MICHEL TALAGRAND

Volume 55

Ergebnisse
der Mathematik
und ihrer
Grenzgebiete
3. Folge

A Series
of Modern
Surveys
in Mathematics

Mean Field Models for Spin Glasses

Volume II: Advanced Replica-Symmetry and Low Temperature

Springer

Ergebnisse der Mathematik und ihrer Grenzgebiete

Volume 55

3. Folge

A Series of Modern Surveys in Mathematics

For further volumes:
www.springer.com/series/728

Michel Talagrand

Mean Field Models for Spin Glasses

Volume II: Advanced Replica-Symmetry and Low Temperature

Springer

Michel Talagrand
Université Paris 6
Institut de mathématiques
UMR 7586 CNRS
Place Jussieu 4
75252 Paris Cedex 05
France
michel@talagrand.net

This volume is the second part of a treatise on Spin Glasses in the series Ergebnisse der Mathematik und ihrer Grenzgebiete. The first part is Vol. 54 of the Ergebnisse series (ISBN 978-3-642-15201-6). The first edition of the treatise appeared as Vol. 46 of the same series (978-3-540-00356-4).

ISSN 0071-1136 Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys in Mathematics
ISBN 978-3-642-22252-8 e-ISBN 978-3-642-22253-5
DOI 10.1007/978-3-642-22253-5
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2011936027

Cover design: deblik

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

To Giorgio Parisi, for the new territories he discovered.

Contents

Part II. Low Temperature

Introduction

Welcome to the second volume of the treatise “Mean fields models for spin glasses”. You certainly do *not* need to have read all of Volume 1 to enjoy the present work. For the low-temperature results of Part II, starting with Chapter 12, only (the beginning of) Chapter 1 is really needed. This is also true for Chapter 11.

In the first part of this volume we continue, at a deeper level, the study of four of the models that were introduced in Volume I. Chapter 8 continues the study of the Shcherbina-Tirozzi model of Chapter 3; Chapter 9 continues the study of the Perceptron model of Chapter 2. Both chapters culminate in the proof of the “Gardner formula” which computes the proportion of the sphere (respectively the discrete cube) that belongs to the intersection of many random half-spaces. Chapters 8 and 9 are somewhat connected. They could in principle be read with only the previous understanding of the corresponding chapter of Volume 1, although we feel that it should help to have also read at least a part of each of Chapters 2 to Chapter 4, where the basic techniques are presented.

Chapter 10 continues and deepens the study of the Hopfield model of Chapter 4. We achieve a good understanding for a larger region of parameters than in Chapter 4 and this understanding is better, as we reach the correct rates of convergence in $1/N$. This chapter can be read independently of Chapters 8 and 9, and in principle with only the knowledge of some of the material of Chapter 4.

Chapter 11 provides an in-depth study of the Sherrington-Kirkpatrick model at high temperature and without external field. As this is a somewhat simpler case than the other models considered in this work, we can look deeper into it. Only (the beginning of) Chapter 1 is a prerequisite from this point on.

In my lecture in the International Congress of Mathematicians in Berlin, 1998, I presented (an earlier form of) some of the results explained here. At the end of the lecture, while I was still panting under the effort, a man (whose name I have mercifully forgotten) came to me, and handed me one of his papers with the following comment “you should read this instead of doing this trivial replica-symmetric stuff”. To him I dedicate these four chapters.

The second part of this volume explores genuine low-temperature results. In Chapter 12 we describe the Ghirlanda-Guerra identities and some rather striking consequences. This chapter can be read without any detailed knowledge of any other material presented so far.

In Chapters 13 and 14 we learn how to prove a celebrated formula of G. Parisi which gives the value of the "limiting free energy" at any temperature for the Sherrington-Kirkpatrick model. A very special case of this formula determines that high-temperature region of this model. We present first this special case in Chapter 13. This seems to require all the important ideas, and these are better explained in this technically simpler setting. Parisi's formula is believed to be only a small part of a very beautiful structure that we call the Parisi Solution. We attempt to describe this structure in Chapter 15 where we also prove as many parts of it as is currently possible. Chapter 15 can be read *without* having read the details of the (difficult) proof of Parisi formula in Chapter 14, and is probably the highlight of this entire work. We also explain what are the remaining (fundamental) questions to be answered before we reach a really satisfactory understanding.

In the final Chapter 16 we study the p-spin interaction model, in a case not covered by the theory of Chapter 14. The approach is based on a clear physical picture of what happens in the phase of "one step of replica-symmetry breaking" and new aspects of the cavity method.

I am very much grateful to Sourav Chatterjee and Albert Hanen who read this entire volume, sometimes in several versions in the most difficult parts, and also to Dmitry Panchenko and Marc Yor who read most of it. Each suggested countlessly many improvements, sometimes correcting serious errors. Special thanks are also due to Wei-Kuo Chen. I claim full responsibility for all the remaining mistakes.

Part I

Advanced Replica-Symmetry

8. The Gardner Formula for the Sphere

8.1 Introduction

In this Chapter we continue the study of the Shcherbina-Tirozzi model of Chapter 3 with Hamiltonian given by (3.1), i.e.

$$-H_{N,M}(\boldsymbol{\sigma}) = \sum_{k\le M} u\biggl(\frac{1}{\sqrt{N}}\sum_{i\le N} g_{i,k}\sigma_i\biggr) + h\sum_{i\le N} g_i\sigma_i - \kappa\|\boldsymbol{\sigma}\|^2 \,. \tag{8.1}$$

As usual we write $S_k = N^{-1/2}\sum_{i\le N} g_{i,k}\sigma_i$. In Theorem 3.3.2 we have succeeded in computing

$$\lim_{\exp u\to \mathbf{1}_{\{x\ge\tau\}}}\biggl\{\lim_{N\to\infty, M/N\to\alpha} \mathsf{E}\frac{1}{N}\log\int \exp(-H_{N,M}(\boldsymbol{\sigma}))\mathrm{d}\boldsymbol{\sigma}\biggr\}\,.$$

Our first goal is to prove that it is possible to exchange the limits in this relation. Writing

$$U_k := \{S_k \ge \tau\} = \biggl\{\sum_{i\le N} g_{i,k}\sigma_i \ge \tau\sqrt{N}\biggr\}\,,$$

we will be able, in Theorem 8.3.1 below, to asymptotically compute the typical value of the quantity

$$\frac{1}{N}\log\int_{\cap_{k\le M}U_k}\exp\biggl(-\kappa\|\boldsymbol{\sigma}\|^2 + h\sum_{i\le N} g_i\sigma_i\biggr)\mathrm{d}\boldsymbol{\sigma}\,. \tag{8.2}$$

In this quantity, let us think of adding the sets U_k one at a time. That is, if we denote by a_M the quantity (8.2), we have

$$a_{M+1} - a_M = \frac{1}{N}\log G(U_{M+1}) = \frac{1}{N}\log G\biggl(\biggl\{\sum_{i\le N} g_{i,M+1}\sigma_i \ge \tau\sqrt{N}\biggr\}\biggr)\,,$$

where G is the probability measure on $\cap_{k\le M}U_k$ with density proportional to $\exp(-\kappa\,\|\boldsymbol{\sigma}\|^2 + h\sum_{i\le N} g_i\,\sigma_i)$. We then see the importance, given a Gibbs measure G, of the quantity

M. Talagrand, *Mean Field Models for Spin Glasses*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys in Mathematics 55, DOI 10.1007/978-3-642-22253-5_1,

$$G(\{\boldsymbol{\sigma}\ ;\ \boldsymbol{\sigma}\cdot\mathbf{g}\geq\tau\sqrt{N}\})\ , \tag{8.3}$$

where $\mathbf{g}$ is a "fresh" Gaussian random vector, i.e. independent of the randomness of G. This quantity should typically not be too small, otherwise it would be unlikely that the quantity (8.2) remains bounded independently of N. For our analysis, we need to know a stronger property, namely that it is extremely rare that the quantity (8.3) is very small. This implies that in the expression (8.2) the influence of each single set U_k is rather moderate. As these sets are independent, we can then expect the quantity (8.2) to have small fluctuations. This is indeed the case, as is shown in Proposition 8.3.6 below, and this is a crucial step in the computation of this quantity (8.2).

8.2 Gaussian Processes

In this section we prove the fundamental fact of this chapter: under rather general conditions for a Gibbs measure G, given a "fresh" random vector $\mathbf{g}$, it is very rare that the set $\{\boldsymbol{\sigma};\boldsymbol{\sigma}\cdot\mathbf{g}\geq\tau\sqrt{N}\}$ has a very small measure for G. The result is stated in Proposition 8.2.6 below. It will be deduced from general properties of Gaussian processes. The main fact is that a jointly Gaussian family $(u_\ell)_{\ell\leq n}$ of r.v.s is "large" as soon as for some number a we have $\mathsf{E}(u_\ell-u_{\ell'})^2\geq a^2$ for $\ell\neq\ell'$; more specifically, such a family is in some sense "as large as a family $(\xi_\ell)_{\ell\leq n}$ of independent Gaussian r.v.s with $\mathsf{E}\xi_\ell^2=a^2$". (Note that $\mathsf{E}(\xi_\ell-\xi_{\ell'})^2=2a^2$, but the factor 2 is unimportant here).

Given a number s, for $\mathbf{x}=(x_\ell)_{\ell\leq n}$ we consider the function

$$F(\mathbf{x})=\log\sum_{\ell\leq n}\exp sx_\ell\ . \tag{8.4}$$

Lemma 8.2.1. *We have*

$$\text{For every } \ell \text{ and } \ell'\neq\ell,\quad \frac{\partial^2F}{\partial x_\ell^2}\geq 0\ ;\ \frac{\partial^2F}{\partial x_\ell\partial x_{\ell'}}\leq 0\ . \tag{8.5}$$

Proof. We define $Z=\sum_{\ell\leq n}\exp sx_\ell$, so that

$$\frac{\partial F}{\partial x_\ell}=s\frac{\exp sx_\ell}{Z}$$

$$\frac{\partial^2F}{\partial x_\ell^2}=s^2\left(\frac{\exp sx_\ell}{Z}-\left(\frac{\exp sx_\ell}{Z}\right)^2\right)\ ,$$

which is ≥ 0 because $\exp sx_\ell/Z\leq 1$. For $\ell\neq\ell'$, we have

$$\frac{\partial^2F}{\partial x_\ell\partial x_{\ell'}}=-s^2\frac{\exp s(x_\ell+x_{\ell'})}{Z^2}\leq 0\ . \qquad \square$$

Proposition 8.2.2. *Consider two jointly Gaussian families* $\mathbf{u} = (u_\ell)_{\ell \le n}$, $\mathbf{v} = (v_\ell)_{\ell \le n}$, *and assume*

$$\forall \ell\,,\ \mathsf{E}u_\ell^2 \ge \mathsf{E}v_\ell^2\ ;\ \forall \ell \ne \ell'\,,\ \mathsf{E}u_\ell u_{\ell'} \le \mathsf{E}v_\ell v_{\ell'}\ . \tag{8.6}$$

Then

$$\mathsf{E}F(\mathbf{u}) \ge \mathsf{E}F(\mathbf{v})\ . \tag{8.7}$$

Proof. Combine (8.5) and Lemma 1.3.1. □

As an application of Proposition 8.2.2, an example of fundamental importance is the case where, for some numbers $b > c > 0$ we have

$$\forall \ell \ne \ell'\,,\ \mathsf{E}u_\ell u_{\ell'} \le c < b \le \mathsf{E}u_\ell^2\ .$$

Then (8.6) holds if $v_\ell = z\sqrt{c} + \xi_\ell\sqrt{b-c}$, where z, ξ_ℓ are independent standard Gaussian r.v.s; indeed $\mathsf{E}v_\ell^2 = b$ and $\mathsf{E}v_\ell v_{\ell'} = c$ if $\ell \ne \ell'$. Thus (8.7) implies

$$\mathsf{E}F(\mathbf{u}) \ge \mathsf{E}F(\mathbf{v}) = \mathsf{E}\log\sum_{\ell\le n}\exp(s\sqrt{b-c}\,\xi_\ell)\ . \tag{8.8}$$

The right-hand side involves only independent r.v.s, so with some work we will be able to find an explicit lower bound for it. This is what is done in the next argument. The proof is elementary but a bit messy. It would be nice to have a clean argument, but it is not certain that such an argument exists: we are working in "the hard direction". The function log is concave and we bound from below the quantity $\mathsf{E}\log Y$ for a certain r.v. Y.

Lemma 8.2.3. *There exists a number* L *with the following property. Consider independent standard Gaussian r.v.s* $(\xi_\ell)_{\ell\le n}$. *For* $L \le s \le \sqrt{\log n}/L$, *we have*

$$\mathsf{E}\log\sum_{\ell\le n}\exp s\xi_\ell \ge \log n + \frac{s^2}{5}\ . \tag{8.9}$$

Proof. We consider a number t to be chosen later, and

$$X = \mathrm{card}\{\ell \le n\ ;\ \xi_\ell \ge t\} = \sum_{\ell\le n}\mathbf{1}_{\{\xi_\ell \ge t\}}\ ,$$

so that $\mathsf{E}X = n\delta$ where $\delta = \mathsf{P}(\xi_\ell \ge t)$. We note that by independence, for $\ell \ne \ell'$ we have $\mathsf{E}\mathbf{1}_{\{\xi_\ell\ge t\}}\mathbf{1}_{\{\xi_{\ell'}\ge t\}} = \delta^2$. Thus

$$\mathsf{E}X^2 = n\delta + n(n-1)\delta^2$$

and therefore

$$n\delta \ge 1 \Rightarrow \mathsf{E}X^2 \le 2n^2\delta^2 = 2(\mathsf{E}X)^2\ .$$

The Paley-Zygmund inequality (A.61) implies that when $n\delta \ge 1$

$$P\left(X \geq \frac{n\delta}{2}\right) \geq \frac{1}{8}\,.$$

Since $\sum_{\ell\leq n} \exp s\xi_\ell \geq X \exp st$, taking logarithms we have $\log \sum_{\ell\leq n} \exp s\xi_\ell \geq st + \log X$, and therefore

$$\mathsf{P}\left(\log \sum_{\ell\leq n} \exp s\xi_\ell \geq st + \log\left(\frac{n\delta}{2}\right)\right) \geq \frac{1}{8}\,. \tag{8.10}$$

We observe (by computing the gradient and using trivial bounds) that the quantity $\log \sum_{\ell\leq n} \exp s\xi_\ell$, when seen as a function of $(\xi_\ell)_{\ell\leq n} \in \mathbb{R}^n$, has a Lipschitz constant s, so that Theorem 1.3.4 implies that for any $u > 0$

$$\mathsf{P}\left(\left|\log \sum_{\ell\leq n} \exp s\xi_\ell - \mathsf{E}\log \sum_{\ell\leq n} \exp s\xi_\ell\right| \geq u\right) \leq 2\exp\left(-\frac{u^2}{4s^2}\right)\,.$$

In particular, taking $u = 4s$, and since $e \geq 2$,

$$\mathsf{P}\left(\log \sum_{\ell\leq n} \exp s\xi_\ell \geq \mathsf{E}\log \sum_{\ell\leq n} \exp s\xi_\ell + 4s\right) < \frac{1}{8}\,,$$

and combining with (8.10), then, when $n\delta \geq 1$, we have

$$\mathsf{E}\log \sum_{\ell\leq n} \exp s\xi_\ell \geq st - 4s + \log\left(\frac{n\delta}{2}\right)\,.$$

We now use the crude fact that $\delta \geq 2\exp(-t^2)$ for $t \geq L$ (see (A.5)), and we choose $t = s/2$ to obtain that the right-hand side of the above inequality is $\geq st - 4s - t^2 + \log n = s^2/4 - 4s + \log n$ and this is $\geq s^2/5 + \log n$ for $s \geq 80$. □

Combining (8.8) and (8.9), we have obtained a lower bound for $F(\mathbf{u})$ under rather general conditions. As made precise in the forthcoming theorem, this bound implies that, typically, many of the terms u_ℓ are large. To understand the statement of this theorem, one should keep in mind the case where the quantities u_ℓ are independent standard Gaussian r.v.s. In that case, we expect that about $n\mathsf{P}(\xi \geq s)$ terms u_ℓ will be $\geq s$. It is unlikely that we can do as well under general conditions, but when the r.v.s u_ℓ are not too much correlated we can still guarantee that about $n\exp(-Ks^2)$ terms u_ℓ will be $\geq s$.

Theorem 8.2.4. *Consider $d > b > c > 0$, and assume moreover that $d \geq 1/(b-c)$. Then there exists a number $K = K(d)$, depending on d only, with the following property: Consider a jointly Gaussian family $\mathbf{u} = (u_\ell)_{\ell\leq n}$, and assume that*

$$\forall \ell \leq n\,,\ \ b \leq \mathsf{E}u_\ell^2 \leq d \tag{8.11}$$

$$\forall \ell \neq \ell'\,,\ \ \mathsf{E}u_\ell u_{\ell'} \leq c\,. \tag{8.12}$$

Then for $K \le s \le \sqrt{\log n}/K$ *we have*

$$\mathsf{P}\left(\operatorname{card}\left\{\ell \le n\,; u_\ell \ge \frac{s}{K}\right\} \le n\exp(-Ks^2)\right) \le K\exp\Bigl(-\frac{s^2}{K}\Bigr)\,. \tag{8.13}$$

The unusual condition $d \ge 1/(b-c)$ allows to use the single parameter d to control both the size of (u_ℓ) and the fact that b and c are not too close, since $b - c \ge 1/d$.

Proof. Let as usual $F(\mathbf{x}) = \log\sum_{\ell\le n}\exp sx_\ell$. We first claim that

$$\forall t > 0\,,\ \mathsf{P}(|F(\mathbf{u}) - \mathsf{E}F(\mathbf{u})| \ge t) \le 2\exp\left(-\frac{t^2}{4ds^2}\right)\,. \tag{8.14}$$

This is a consequence of Theorem 1.3.4. To see this, we note that for $1 \le n$ we can find vectors $\mathbf{z}_\ell \in \mathbb{R}^n$ such that the sequence $(u_\ell)_{\ell\le n}$ has the same distribution as the sequence $(\mathbf{z}_\ell\cdot\mathbf{g})_{\ell\le n}$, where $\mathbf{g}$ is an independent standard Gaussian sequence. (This simple statement is proved in Section A.2, but is not really needed since when we will apply this theorem, the family $(u_\ell)_{\ell\le n}$ will already be given in this form.) We have $\|\mathbf{z}_\ell\|^2 = \mathsf{E}u_\ell^2 \le d$, so that the gradient of the function

$$\mathbf{x} \mapsto \log\sum_{\ell\le n}\exp s\mathbf{z}_\ell\cdot\mathbf{x}$$

has a Lipschitz constant $\le s\sqrt{d}$, which implies (8.14).

Next, using (8.8) and then (8.9) for $s\sqrt{b-c}$ rather than s, we have

$$\mathsf{E}F(\mathbf{u}) \ge \log n + \frac{s^2(b-c)}{5}\,, \tag{8.15}$$

provided s satisfies $L \le s\sqrt{b-c} \le \sqrt{\log n}/L$. We combine (8.15) with (8.14) for $t = s^2(b-c)/10$, to obtain

$$\begin{aligned}\mathsf{P}\Biggl(\sum_{\ell\le n}\exp su_\ell \ge n\exp\frac{s^2(b-c)}{10}\Biggr) &= \mathsf{P}\left(F(\mathbf{u}) \ge \log n + \frac{s^2(b-c)}{10}\right)\\ &\ge 1 - \mathsf{P}\left(F(\mathbf{u}) < \mathsf{E}F(\mathbf{u}) - \frac{s^2(b-c)}{10}\right)\\ &\ge 1 - 2\exp\Bigl(-\frac{s^2(b-c)^2}{Ld}\Bigr)\,. \end{aligned}\tag{8.16}$$

Also, using (A.1) and (8.11),

$$\mathsf{E}\sum_{\ell\le n}\exp 2su_\ell = \sum_{\ell\le n}\exp 2s^2\mathsf{E}u_\ell^2 \le n\exp 2s^2 d$$

which, by Markov inequality (A.2) implies

$$\mathsf{P}\Big(\sum_{\ell \le n} \exp 2su_\ell \le n \exp 3s^2 d\Big) \ge 1 - \exp(-s^2 d) \ . \tag{8.17}$$

Consider the (typical) event Ω given by

$$\Omega = \Big\{\sum_{\ell \le n} \exp su_\ell \ge n \exp \frac{s^2(b-c)}{10}\Big\} \bigcap \Big\{\sum_{\ell \le n} \exp 2su_\ell \le n \exp 3s^2 d\Big\} \ , \tag{8.18}$$

so that combining (8.16) and (8.17) and recalling that $b - c \ge 1/d$, we obtain

$$\begin{aligned}\mathsf{P}(\Omega) &\ge 1 - 2\exp\left(-\frac{s^2(b-c)^2}{Ld}\right) - \exp(-s^2 d) \\ &\ge 1 - 3\exp\left(-\frac{s^2}{K(d)}\right) \ .\end{aligned} \tag{8.19}$$

We now study the generic sequence (u_ℓ) when Ω occurs. Consider the uniform probability space $\mathcal{U}_n = (\{1, \ldots, n\}, \mathsf{P}_0)$, and denote by E_0 expectation with respect to the probability P_0. Consider the r.v. X on the probability space $\mathcal{U}_n$ given by $X(\ell) = \exp su_\ell$, so that $\mathsf{E}_0 X = 1/n \sum_{\ell \le n} \exp su_\ell$ etc. It follows from (8.18) that

$$\mathsf{E}_0 X \ge \exp \frac{s^2(b-c)}{10} \ ; \ \mathsf{E}_0 X^2 \le \exp 3s^2 d \ .$$

Since we assume $s\sqrt{b-c} \ge L$, we may also assume $\mathsf{E}_0 X > 2\exp(s^2(b-c)/20)$ and the Paley-Zygmund inequality

$$\mathsf{P}_0\Big(X \ge \frac{\mathsf{E}_0 X}{2}\Big) \ge \frac{1}{4} \frac{(\mathsf{E}_0 X)^2}{\mathsf{E}_0 X^2} \tag{A.61}$$

implies

$$\mathrm{card}\left\{\ell \le n \ ; u_\ell \ge \frac{s(b-c)}{20}\right\} > n \exp(-3s^2 d) \ .$$

This occurs whenever Ω occurs. Consequently,

$$\mathsf{P}\left(\mathrm{card}\left\{\ell \le n \ ; u_\ell \ge \frac{s(b-c)}{20}\right\} \le n \exp(-3s^2 d)\right) \le 3\exp\Big(-\frac{s^2}{K}\Big) \ ,$$

which proves the theorem. □

Corollary 8.2.5. *Consider numbers b, c, d and a Gaussian family $\mathbf{u}$ as in Theorem 8.2.4. Then there exists a constant $K = K(d)$, depending on d only, with the following property: for every ε with $n^{-1/K} \le \varepsilon \le \exp(-K\tau^2)/K$*

$$\mathsf{P}\big(\mathrm{card}\{\ell \le n \ ; u_\ell \ge \tau\} \le \varepsilon n\big) \le K\varepsilon^{1/K} \ . \tag{8.20}$$

Proof. It follows from (8.13) that for $s \geq K$, $s \geq K\tau$, $s \leq \sqrt{\log n}/K$ we have

$$\mathsf{P}\big(\operatorname{card}\{\ell \leq n\,; u_\ell \geq \tau\} \leq n\exp(-Ks^2)\big) \leq K\exp\Big(-\frac{s^2}{K}\Big)\,.$$

Setting $\varepsilon = \exp(-Ks^2)$, we obtain (8.20). □

The relevance of Theorem 8.2.4 to spin glasses is shown by the following statement, where we recall that $\mathbf{g}$ denotes a standard Gaussian random vector in $\mathbb{R}^N$.

Proposition 8.2.6. *Consider a probability measure G on $\mathbb{R}^N$. Assume that for some numbers $0 \leq c < b < d$, $d \geq 1/(b-c)$ we have*

$$G(\{\boldsymbol{\sigma}\,;\, bN \leq \|\boldsymbol{\sigma}\|^2 \leq dN\}) \geq 1 - \exp\Big(-\frac{N}{d}\Big) \tag{8.21}$$

$$G^{\otimes 2}(\{(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2)\,;\, |\boldsymbol{\sigma}^1 \cdot \boldsymbol{\sigma}^2| \leq cN\}) \geq 1 - \exp\Big(-\frac{N}{d}\Big)\,. \tag{8.22}$$

Then, for any $\tau \geq 0$ the following implication holds,

$$K\exp\Big(-\frac{N}{K}\Big) \leq \varepsilon \leq \frac{1}{K}\exp(-K\tau^2) \;\Rightarrow\; \mathsf{P}\bigg(G\bigg(\bigg\{\boldsymbol{\sigma}\,; \frac{\mathbf{g}\cdot\boldsymbol{\sigma}}{\sqrt{N}} \geq \tau\bigg\}\bigg) \leq \varepsilon\bigg) \leq K\varepsilon^{1/K}, \tag{8.23}$$

where K depends on d only.

As hinted at with the notation, this will be applied when G is a Gibbs measure. The single parameter d is also used to control the size of the exceptional sets (8.21) and (8.22).

Proof. The idea behind this proof is very simple. Consider generic points $\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n$ chosen independently according to G. Then (8.21) and (8.22) show that unless n is very large (or we are very unlucky) the Gaussian r.v.s $u_\ell = \boldsymbol{\sigma}^\ell \cdot \mathbf{g}/\sqrt{N}$ satisfy the hypothesis of Corollary 8.2.5. Therefore, from (8.20), it is likely that a significant proportion of them is $\geq \tau$, which implies that the set $\{\boldsymbol{\sigma}\,; \mathbf{g}\cdot\boldsymbol{\sigma} \geq \tau\sqrt{N}\}$ cannot be very small for G.

To implement this idea, let

$$\begin{aligned} Q_n = \{(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n)\,;\, &\forall \ell \leq n\,, bN \leq \|\boldsymbol{\sigma}^\ell\|^2 \leq dN\,; \\ &\forall\, 1 \leq \ell < \ell' \leq n\,, |\boldsymbol{\sigma}^\ell \cdot \boldsymbol{\sigma}^{\ell'}| \leq cN\}\,, \end{aligned}$$

so that (8.21) and (8.22) imply

$$G^{\otimes n}(Q_n) \geq 1 - n^2\exp\Big(-\frac{N}{d}\Big) \geq \frac{1}{2} \tag{8.24}$$

provided $2n^2 \leq \exp(N/d)$. For $(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n) \in Q_n$ we define the event

$$\Omega(\boldsymbol{\sigma}^1,\ldots,\boldsymbol{\sigma}^n) = \left\{\mathrm{card}\{\ell \le n\ ;\ \boldsymbol{\sigma}^\ell\cdot\mathbf{g} \ge \tau\sqrt{N}\} \le \varepsilon n\right\} .$$

When $n^{-1/K} \le \varepsilon \le \exp(-K\tau^2)/K$, condition (8.20) implies that

$$\mathsf{P}(\Omega(\boldsymbol{\sigma}^1,\ldots,\boldsymbol{\sigma}^n)) \le K\varepsilon^{1/K} ,$$

where K depends on d only. Thus, the expectation of

$$Y = Y(g) = \int_{Q_n} \mathbf{1}_{\{\Omega(\boldsymbol{\sigma}^1,\ldots,\boldsymbol{\sigma}^n)\}} \mathrm{d}G(\boldsymbol{\sigma}^1)\cdots\mathrm{d}G(\boldsymbol{\sigma}^n)$$

is

$$\mathsf{E}Y = \int_{Q_n} \mathsf{P}(\Omega(\boldsymbol{\sigma}^1,\ldots,\boldsymbol{\sigma}^n))\mathrm{d}G(\boldsymbol{\sigma}^1)\cdots\mathrm{d}G(\boldsymbol{\sigma}^n) \le K\varepsilon^{1/K}$$

and Markov's inequality (A.2) implies

$$\mathsf{P}\Big(Y \ge \frac{1}{4}\Big) \le K\varepsilon^{1/K} . \tag{8.25}$$

Now by definition, Y also equals

$$Y = G^{\otimes n}\left(\left\{(\boldsymbol{\sigma}^1,\ldots,\boldsymbol{\sigma}^n)\in Q_n\ ;\ \mathrm{card}\left\{\ell\le n\,;\frac{\boldsymbol{\sigma}^\ell\cdot\mathbf{g}}{\sqrt{N}}\ge\tau\right\}\le n\varepsilon\right\}\right) ,$$

so that, recalling (8.24), we obtain

$$\begin{aligned} G^{\otimes n}&\left(\left\{(\boldsymbol{\sigma}^1,\ldots,\boldsymbol{\sigma}^n)\in Q_n\ ;\ \mathrm{card}\left\{\ell\le n\,;\frac{\boldsymbol{\sigma}^\ell\cdot\mathbf{g}}{\sqrt{N}}\ge\tau\right\}> n\varepsilon\right\}\right)\\ &= G^{\otimes n}(Q_n) - Y \ge \frac{1}{4}\end{aligned}$$

whenever $Y \le 1/4$. In that case,

$$\begin{aligned} &nG\left(\left\{\boldsymbol{\sigma}\ ;\ \frac{\boldsymbol{\sigma}\cdot\mathbf{g}}{\sqrt{N}}\ge\tau\right\}\right)\\ &= \int \mathrm{card}\left\{\ell\le n\ ;\ \frac{\boldsymbol{\sigma}^\ell\cdot\mathbf{g}}{\sqrt{N}}\ge\tau\right\}\mathrm{d}G(\boldsymbol{\sigma}^1)\cdots\mathrm{d}G(\boldsymbol{\sigma}^n)\\ &\ge n\varepsilon G^{\otimes n}\left(\left\{(\boldsymbol{\sigma}^1,\ldots,\boldsymbol{\sigma}^n)\in Q_n\ ;\ \mathrm{card}\left\{\ell\le n\,;\frac{\boldsymbol{\sigma}^\ell\cdot\mathbf{g}}{\sqrt{N}}\ge\tau\right\}>n\varepsilon\right\}\right)\\ &\ge \frac{n\varepsilon}{4} .\end{aligned}$$

Therefore, whenever $Y \le 1/4$, it holds that $G(\{\boldsymbol{\sigma}\ ;\ \boldsymbol{\sigma}\cdot\mathbf{g}\ge\tau\sqrt{N}\}) \ge \varepsilon/4$.

In summary, recalling (8.25) it follows that if $2n^2 \le \exp(N/d)$ and $n^{-1/K} \le \varepsilon \le \exp(-K\tau^2)/K$, then

$$\mathsf{P}\left(G\left(\left\{\boldsymbol{\sigma}\ ;\ \frac{\boldsymbol{\sigma}\cdot\mathbf{g}}{\sqrt{N}}\ge\tau\right\}\right)\le\frac{\varepsilon}{4}\right) \le K\varepsilon^{1/K} .$$

This concludes the proof, by taking n as large as possible so that $2n^2 \le \exp(N/d)$. □

We now show that the hypotheses of Proposition 8.2.6 are easily satisfied by the Gibbs measures considered in Chapter 3.

Theorem 8.2.7. *Consider a concave function $U \le 0$ on $\mathbb{R}^N$, numbers $0 < \kappa_0 < \kappa_1$, numbers $(a_i)_{i\le N}$. Consider a convex set C of $\mathbb{R}^N$. Consider $\kappa_0 \le \kappa \le \kappa_1$ and the probability measure G_C on $\mathbb{R}^N$ given by*

$$\forall B\,,\ G_C(B) = \frac{1}{Z}\int_{B\cap C} \exp\Big(U(\boldsymbol{\sigma}) - \kappa\|\boldsymbol{\sigma}\|^2 + \sum_{i\le N} a_i\sigma_i\Big)\mathrm{d}\boldsymbol{\sigma}$$

where Z is the normalizing factor

$$Z = \int_C \exp\Big(U(\boldsymbol{\sigma}) - \kappa\|\boldsymbol{\sigma}\|^2 + \sum_{i\le N} a_i\sigma_i\Big)\mathrm{d}\boldsymbol{\sigma}\ .$$

Assume that for a certain number a we have

$$Z \ge \exp(-Na) \tag{8.26}$$

$$\sum_{i\le N} a_i^2 \le Na^2\ . \tag{8.27}$$

Then G_C satisfies (8.23) for a number K depending only on a, κ_0 and κ_1.

Proof. In this proof, K denotes a number depending only on κ_0, κ_1 and a, which does not need to be the same at each occurrence. We will prove that conditions (8.21) and (8.22) are satisfied for numbers b, c, d such that $d = K$ $(= K(a, \kappa_0, \kappa_1))$. We denote by $\langle\cdot\rangle$ an average for the probability G_C or its products.

First, recalling that $R_{\ell,\ell'} = N^{-1}\sum_{i\le N}\sigma_i^\ell\sigma_i^{\ell'}$ we prove that

$$\langle R_{1,1}\rangle - \langle R_{1,2}\rangle \ge \frac{1}{K}\ . \tag{8.28}$$

We observe that

$$\langle R_{1,1}\rangle - \langle R_{1,2}\rangle = \frac{1}{2N}\langle\|\boldsymbol{\sigma} - \langle\boldsymbol{\sigma}\rangle\|^2\rangle\ .$$

We shall prove that there exists a number $d' > 0$, depending on a and κ only, such that if B is any ball of radius $d'\sqrt{N}$, we have $G_C(B) \le 1/2$. Using this for the ball centered at $\langle\boldsymbol{\sigma}\rangle$ shows that $\langle\|\boldsymbol{\sigma} - \langle\boldsymbol{\sigma}\rangle\|^2\rangle \ge Nd'^2/2$ and proves (8.28). To bound $G_C(B)$ we simply write, completing the squares and using (8.26) in the last line,

$$\begin{aligned} G_C(B) &\le \frac{1}{Z}\int_B \exp\Big(-\kappa\|\boldsymbol{\sigma}\|^2 + \sum_{i\le N} a_i\sigma_i\Big)\mathrm{d}\boldsymbol{\sigma} \\ &= \frac{1}{Z}\exp\Big(\sum_{i\le N}\frac{a_i^2}{4\kappa}\Big)\int_B \exp\Big(-\kappa\sum_{i\le N}\Big(\sigma_i - \frac{a_i}{2\kappa}\Big)^2\Big)\mathrm{d}\boldsymbol{\sigma} \\ &\le \exp N\Big(a + \frac{a^2}{4\kappa}\Big)\int_{B'}\exp(-\kappa\|\boldsymbol{\sigma}\|^2)\mathrm{d}\boldsymbol{\sigma} \end{aligned}$$

where B' is a ball of radius $d'\sqrt{N}$. The last integral is maximum when B' is centered at 0. (This statement is left as an exercise, since the use of Fubini's theorem allows to reduce it to the case $N = 1$.) We then make a standard computation as follows. Considering $\lambda > 0$, we write

$$\begin{aligned} \int_{B'}\exp(-\kappa\|\boldsymbol{\sigma}\|^2)\mathrm{d}\boldsymbol{\sigma} &\le \exp\lambda d'^2 N \int \exp(-(\lambda+\kappa)\|\boldsymbol{\sigma}\|^2)\mathrm{d}\boldsymbol{\sigma} \\ &\le \big(\exp\lambda d'^2 N\big)\Big(\frac{\pi}{\lambda+\kappa}\Big)^{N/2}, \end{aligned}$$

using (3.23). Therefore, since $\pi \le e^2$ and $\lambda + \kappa \ge \lambda$ we have

$$G_C(B) \le \lambda^{-N/2}\exp N\Big(a + \frac{a^2}{4\kappa} + \lambda d'^2 + 1\Big)\,. \tag{8.29}$$

We then choose e.g. $\lambda = \exp 2(a + a^2/4\kappa + 3)$ and $d' = \lambda^{-1/2}$ to deduce from (8.29) that then $G_C(B)$ is very small.

Next, we prove that

$$\Big\langle \exp\frac{\kappa}{2}\|\boldsymbol{\sigma}\|^2\Big\rangle \le \exp NK\,. \tag{8.30}$$

For this purpose, we simply follow the proof of (3.26), using the bound $1/Z \le \exp NK$ rather than Lemma 3.1.6. Third, we show that for $k \le N$ we have

$$\langle (R_{1,2} - \langle R_{1,2}\rangle)^{2k}\rangle \le \Big(\frac{Kk}{N}\Big)^k \tag{8.31}$$

$$\langle (R_{1,1} - \langle R_{1,1}\rangle)^{2k}\rangle \le \Big(\frac{Kk}{N}\Big)^k\,. \tag{8.32}$$

For this, we simply repeat the arguments of Theorem 3.1.11, replacing B^* by KN throughout. It is in this proof that (8.30) is used.

An inequality such as $\langle f^{2k}\rangle \le (Kk/N)^k$ for $k \le N$ implies an exponentially good control of certain deviations; indeed we write

$$G_C(\{|f| \ge x\}) \le \frac{\langle f^{2k}\rangle}{x^{2k}} \le \Big(\frac{Kk}{x^2 N}\Big)^k\,,$$

and we choose for k the smallest integer $\ge x^2 N/Ke$ to obtain

$$x^2 \le Ke \ \Rightarrow \ G_C(\{|f| \ge x\}) \le \exp\left(-\frac{x^2 N}{Ke}\right) . \tag{8.33}$$

Denoting by K_0 the constant of (8.28) we then take $c = \langle R_{1,2}\rangle + 3/K_0$ and $b = \langle R_{1,1}\rangle - 3/K_0$, so that $1/(b-c) \le 3K_0$. Using (8.33) for $f = R_{1,2} - \langle R_{1,2}\rangle$ we obtain

$$G_C^{\otimes 2}\{(\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2) \ ; \ |R_{1,2} - \langle R_{1,2}\rangle| \ge 3/K_0\} \le \exp\left(-\frac{N}{K}\right) ,$$

so that

$$G_C^{\otimes 2}\{(\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2) \ ; \ |R_{1,2}| \ge c\} \le \exp\left(-\frac{N}{K}\right)$$

which proves that G_C satisfies (8.21). The proof that it satisfies (8.22) is similar. □

8.3 The Gardner Formula for the Gaussian Measure

We recall some notation from Chapter 3. We define

$$\mathcal{N}(x) = \mathsf{P}(z \ge x) ,$$

where z is standard Gaussian; for $0 \le q \le \rho$ we define

$$F(q,\rho) = \alpha \mathsf{E} \log \mathcal{N}\left(\frac{\tau - z\sqrt{q}}{\sqrt{\rho - q}}\right) + \frac{1}{2}\frac{q}{\rho - q} + \frac{1}{2}\log(\rho - q) - \kappa\rho + \frac{h^2}{2}(\rho - q) . \tag{8.34}$$

In Theorem 3.3.1 we proved that when $\tau \ge 0$ and $\alpha < 2$ the system of two equations

$$\frac{\partial F}{\partial q} = \frac{\partial F}{\partial \rho} = 0 \tag{8.35}$$

have a unique solution (q_0, ρ_0). As in Chapter 3 we define

$$\mathrm{RS}_0(\alpha) = F(q_0, \rho_0) , \tag{8.36}$$

and we further define

$$\mathrm{RSG}(\alpha) = F(q_0, \rho_0) + \frac{1}{2}\log(2e\pi) , \tag{8.37}$$

where "RS" stands as usual for "replica-symmetric" and G stands for Gaussian. These quantities also depend on κ and h, but this dependence is kept implicit. We recall that $S_k = N^{-1/2}\sum_{i\le N} g_{i,k}\sigma_i$ and $U_k = \{S_k \ge \tau\}$.

Theorem 8.3.1. *Assume that* $\tau \geq 0$. *Consider* $0 < \kappa_0 < \kappa_1$, $h_0 > 0$, $\alpha_0 < 2$ *and* $\varepsilon > 0$. *Then there is a number* K, *depending on* $\kappa_0, \kappa_1, h_0, \alpha_0$ *and* ε *only, such that if* $N \geq K$, *whenever* $\kappa_0 \leq \kappa \leq \kappa_1$, *and whenever* $h \leq h_0$ *and* $M \leq \alpha_0 N$ *we have*

$$\mathsf{P}\left(\left|\frac{1}{N}\log\int_{\cap_{k\leq M}U_k}\exp\left(-\kappa\,\|\boldsymbol{\sigma}\|^2 + h\sum_{i\leq N} g_i\,\sigma_i\right)\mathrm{d}\boldsymbol{\sigma} - \mathrm{RSG}\left(\frac{M}{N}\right)\right| \geq \varepsilon\right)$$
$$\leq K\exp\left(-\frac{N}{K}\right). \tag{8.38}$$

That is, we have succeeded in computing the typical measure, for the Gaussian measure of density proportional to $\exp(-\kappa\|\boldsymbol{\sigma}\|^2 + h\sum_{i\leq N} g_i\sigma_i)$ of the intersection of the M random half-spaces $(U_k)_{k\leq M}$.

Research Problem 8.3.2. (Level 2) Find out what happens for $\tau < 0$.

The technical difficulty is that we no longer know that the equations (8.34) have a unique solution.

Exercise 8.3.3. Prove that

$$\lim_{N\to\infty}\frac{1}{N}\log\int_{\cap_{k\leq 2N}U_k}\exp\left(-\kappa\,\|\boldsymbol{\sigma}\|^2 + h\sum_{i\leq N} g_i\,\sigma_i\right)\mathrm{d}\boldsymbol{\sigma} = -\infty\,. \tag{8.39}$$

Hint: Reduce first to the case $\tau = 0$. Prove then that $\lim_{\alpha\to 2}\mathrm{RS}_0(\alpha) = -\infty$. For this observe that by Corollary (3.3.12) b) we have $F(\rho_0, q_0) \leq \max_\rho \inf_q F(q,\rho)$, and study the proof of Lemma 8.4.8 below. Then argue as in the few lines following this lemma.

The following might be easier than Problem 8.3.2.

Research Problem 8.3.4. Does (8.39) remain true for $\tau < 0$? And what happens if instead we normalize by N^α for some $\alpha > 1$? If, as is plausible, the limit is then 0, is there a normalization factor that ensures a finite limit?

Research Problem 8.3.5. Find a rate of convergence in Theorem 8.3.1.

For example, one may wonder whether

$$\mathsf{P}\left(\left|\frac{1}{N}\log\int_{\bigcap_{k\leq M}U_k}\exp\left(-\kappa\|\boldsymbol{\sigma}\|^2 + h\sum_{i\leq N} g_i\sigma_i\right)\mathrm{d}\boldsymbol{\sigma} - \mathrm{RSG}\left(\frac{M}{N}\right)\right| \geq \varepsilon(N)\right)$$
$$\to 0 \tag{8.40}$$

whenever $N\varepsilon(N) \to \infty$. Other possible formulations will be given below. None of these formulations is as simple as one might wish, due to the fact that $\bigcap_{k\le M} U_k = \emptyset$ with a positive probability. This makes these formulations a bit awkward, because we cannot compute $\mathsf{E}\log Y$ when $\mathsf{P}(Y=0)>0$. There is however no real difficulty, because the event $\bigcap_{k\le M} U_k = \emptyset$ is so rare (its probability is exponentially small in N) that it has no influence. This is not yet obvious, but follows e.g. from (8.76) below

In order to be able to keep computing expectations, given a number $A \ge 0$ we define

$$\log_A x = \log\max(x, e^{-A}) = \max(\log x, -A)\ ,$$

so that $\log_A 0 = -A$.

The overall scheme of proof of Theorem 8.3.1 is as follows. First, given any $a > 0$ (very large) the fluctuations of the random quantity

$$\frac{1}{N}\log_{aN}\int_{\bigcap_{k\le M} U_k} \exp\Big(-\kappa\|\boldsymbol{\sigma}\|^2 + h\sum_{i\le N} g_i\sigma_i\Big)\mathrm{d}\boldsymbol{\sigma} \tag{8.41}$$

around its expectation are small; namely, for any $t > 0$ it is very rare that the difference between (8.41) and its expectation is $\ge t$.

Second, given $a > 0$, if the concave function u is such that $\exp u$ is a sufficiently good approximation of $\mathbf{1}_{\{x\ge\tau\}}$, the expectation of the quantity (8.41) is nearly

$$\frac{1}{N}\mathsf{E}\log_{aN}\int \exp(-H_{N,M}(\boldsymbol{\sigma}))\mathrm{d}\boldsymbol{\sigma} \tag{8.42}$$

where $H_{N,M}$ is the Hamiltonian (3.1), i.e.

$$-H_{N,M}(\boldsymbol{\sigma}) = \sum_{k\le M} u(S_k) + h\sum_{i\le N} g_i\sigma_i - \kappa\|\boldsymbol{\sigma}\|^2\ . \tag{8.43}$$

Third, given such a function u, the quantity

$$\frac{1}{N}\log\int \exp(-H_{N,M}(\boldsymbol{\sigma}))\mathrm{d}\boldsymbol{\sigma} \tag{8.44}$$

has itself small fluctuations around its mean (observe that the log is not truncated here). By Theorem 3.3.2 this means that

$$\frac{1}{N}\mathsf{E}\log\int \exp(-H_{N,M}(\boldsymbol{\sigma}))\mathrm{d}\boldsymbol{\sigma} \tag{8.45}$$

is nearly $\mathrm{RSG}(M/N)$, so that the quantity (8.44) is nearly always very close to $\mathrm{RSG}(M/N)$. Thus if a is large (say $a \ge -\mathrm{RSG}(M/N)+1$) the quantities (8.45) and (8.42) are nearly the same (i.e. the influence of the truncation is very small), and the quantity (8.42) is nearly $\mathrm{RSG}(M/N)$, so that the quantity (8.41) is nearly always very close to $\mathrm{RSG}(M/N)$; and finally the quantity

$$\frac{1}{N}\log\int_{\bigcap_{k\le M} U_k} \exp\Big(-\kappa\|\boldsymbol{\sigma}\|^2 + h\sum_{i\le N} g_i\sigma_i\Big)\mathrm{d}\boldsymbol{\sigma}$$

itself is nearly always very close to RSG(M/N).

The estimates in the different steps all ultimately rely on Theorem 8.2.7 above.

Proposition 8.3.6. *Consider $\tau \geq 0$. Consider $0 < \kappa_0 < \kappa_1$, $h_0 > 0$, $a > 0$. Then there is a constant K, depending only on $\kappa_0, \kappa_1, h_0, a$ and τ such that, if $\kappa_0 \leq \kappa \leq \kappa_1$ and $h \leq h_0$, whenever $u \leq 0$ is a concave function with*

$$x \geq \tau \;\Rightarrow\; u(x) = 0 \tag{8.46}$$

we have

$$\forall N\,, \forall M \leq 10N\,, \forall t > 0\,, \tag{8.47}$$
$$\mathsf{P}\left(\left|\frac{1}{N}\log_{aN} Z_{N,M} - \mathsf{E}\frac{1}{N}\log_{aN} Z_{N,M}\right| \geq t\right) \leq 2\exp\left(-\frac{N}{K}\min(t^2, t)\right)\,,$$

where

$$\begin{aligned} Z_{N,M} &= \int \exp\Big(\sum_{k\leq M} u(S_k) - \kappa\|\boldsymbol{\sigma}\|^2 + h\sum_{i\leq N} g_i\sigma_i\Big)\mathrm{d}\boldsymbol{\sigma} \\ &= \int \exp(-H_{N,M}(\boldsymbol{\sigma}))\mathrm{d}\boldsymbol{\sigma}\,. \end{aligned}$$

This statement includes the situation where $\exp u(x) = \mathbf{1}_{\{x\geq\tau\}}$, in which case

$$Z_{N,M} = \int_{\bigcap_{k\leq M} U_k} \exp\left(-\kappa\|\boldsymbol{\sigma}\|^2 + h\sum_{i\leq N} g_i\sigma_i\right)\mathrm{d}\boldsymbol{\sigma}\,. \tag{8.48}$$

Lemma 8.3.7. *If $0 \leq x, z \leq 1$, we have*

$$|\log_A x - \log_A z| \leq \left|\log_A\left(\frac{x}{z}\right)\right|\,. \tag{8.49}$$

Proof. Assuming without loss of generality that $z \leq x$, we have $0 \leq z \leq x \leq 1$ so that

$$0 \leq \log_A x - \log_A z \leq -\log_A z \leq A$$

and also

$$|\log_A x - \log_A z| \leq |\log x - \log z| \leq \left|\log\frac{x}{z}\right|\,.$$

Consequently,

$$|\log_A x - \log_A z| \leq \min\left(A, \left|\log\frac{x}{z}\right|\right) = \left|\log_A\left(\frac{x}{z}\right)\right|\,. \qquad \square$$

Lemma 8.3.8. *Consider a r.v. $V \geq 0$ and assume that for certain numbers $C, D \geq 1$, and a we have*

$$\mathsf{P}(V \geq \exp Na) = 0 \tag{8.50}$$

$$D \leq t \leq \frac{1}{C}\exp\Big(\frac{N}{C}\Big) \;\Rightarrow\; \mathsf{P}(V \geq t) \leq Ct^{-1/C} \tag{8.51}$$

Then we can find a number $\lambda > 0$ depending only on a, C and D for which

$$\mathsf{E}V^\lambda \leq (D + 2 + 2C)\;. \tag{8.52}$$

Proof. We start with the formula

$$\mathsf{E}V^\lambda = \lambda \int_0^\infty t^{\lambda-1}\mathsf{P}(V \geq t)\mathrm{d}t\;,$$

and we will show that this implies (8.52). Since $V \leq \exp aN$, setting $t_0 = D$, $t_1 = C^{-1}\exp(N/C)$ and $t_2 = \exp aN$ we have

$$\lambda \int_0^\infty t^{\lambda-1}\mathsf{P}(V \geq t)\mathrm{d}t = \mathrm{I} + \mathrm{II} + \mathrm{III}\;,$$

where

$$\mathrm{I} = \lambda \int_0^{t_0} t^{\lambda-1}\mathsf{P}(V \geq t)\mathrm{d}t \leq \lambda \int_0^{t_0} t^{\lambda-1}\mathrm{d}t = t_0^\lambda\;,$$

$$\mathrm{II} = \lambda \int_{t_0}^{t_1} t^{\lambda-1}\mathsf{P}(V \geq t)\mathrm{d}t \leq C\lambda \int_{t_0}^{t_1} t^{\lambda-1-1/C}\mathrm{d}t\;.$$

$$\mathrm{III} = \lambda \int_{t_1}^{t_2} t^{\lambda-1}\mathsf{P}(V \geq t)\mathrm{d}t \leq \lambda\mathsf{P}(V \geq t_1) \int_{t_1}^{t_2} t^{\lambda-1}\mathrm{d}t \leq Ct_1^{-1/C}t_2^\lambda\;.$$

Since $t_1^{-1/C}t_2^\lambda = C^{1/C}\exp(-N/C + \lambda Na)$ and since $C^{1/C} \leq 2$, when $\lambda a \leq 1/C$, then $\mathrm{III} \leq 2C$. Moreover $\mathrm{II} \leq 2$ for $\lambda \leq 1/2C$ and $\mathrm{I} \leq D$ for $\lambda \leq 1$. □

Proof of Proposition 8.3.6. The term $h\sum_{i \leq N} g_i\sigma_i$ creates only secondary complications, and the most interesting situation is when $h = 0$. For this reason, although we had the nerve to state Theorem 8.3.1 with this term $h\sum_{i \leq N} g_i\sigma_i$, we will prove it only when $h = 0$.

The idea of the proof is to apply Bernstein's inequality (A.41) to a suitable martingale difference sequence. We denote by Ξ_m the σ-algebra generated by the r.v.s $(g_{i,k})$ for $i \leq N$, $k \leq m$. We denote by E^m conditional expectation given Ξ_m, so that $\mathsf{E} = \mathsf{E}^0$. We set

$$W = \int \exp\Big(\sum_{k \leq M} u(S_k) - \kappa\|\boldsymbol{\sigma}\|^2\Big)\mathrm{d}\boldsymbol{\sigma}$$

and we write

$$\frac{1}{N}\log_{aN} W - \frac{1}{N}\mathsf{E}\log_{aN} W = \sum_{m=1}^{M} X_m$$

where

$$X_m = \mathsf{E}^m\left(\frac{1}{N}\log_{aN} W\right) - \mathsf{E}^{m-1}\left(\frac{1}{N}\log_{aN} W\right) . \tag{8.53}$$

Thus X_m is Ξ_m-measurable and $\mathsf{E}^{m-1}X_m = 0$, i.e. (X_m) is a martingale difference sequence. In order to deduce (8.47) from Bernstein's inequality (A.41), we will prove that for each m we have

$$\mathsf{E}^{m-1}\exp\frac{N|X_m|}{K} \le 2 , \tag{8.54}$$

where K depends only on κ_0, κ_1 and a. This suffices since the bound (A.41), when used for $A = K/N$ and M instead of N is

$$2\exp\left(-\min\left(\frac{N^2t^2}{MK}, \frac{Nt}{K}\right)\right) \le 2\exp\left(-\frac{N}{K}\min(t^2, t)\right)$$

for $M \le 10N$. The proof of (8.54) occupies the rest of the argument.

We define

$$W_m = \int \exp\Bigg(\sum_{k\ne m, k\le M} u(S_k) - \kappa\|\boldsymbol{\sigma}\|^2\Bigg)\mathrm{d}\boldsymbol{\sigma} , \tag{8.55}$$

so that W_m does not depend on the r.v.s $(g_{i,m})_{i\le N}$, and since $u \le 0$ we have $W \le W_m$. In the case where $\exp u(x) = \mathbf{1}_{\{x\ge\tau\}}$, the definition of W_m is

$$W_m = \int_{C_m} \exp(-\kappa\|\boldsymbol{\sigma}\|^2)\mathrm{d}\boldsymbol{\sigma} ,$$

where $C_m = \bigcap_{k\ne m,k\le M} U_k$.

Using (8.49) we get

$$|\log_{aN} W - \log_{aN} W_m| \le \left|\log_{aN}\left(\frac{W}{W_m}\right)\right| .$$

Since $W \le W_m$, the left-hand side is 0 for $W_m \le \exp(-aN)$, since then $\log_{aN} W = \log_{aN} W_m = 0$. Therefore,

$$|\log_{aN} W - \log_{aN} W_m| \le Y := \mathbf{1}_{\{W_m\ge\exp(-aN)\}}\left|\log_{aN}\left(\frac{W}{W_m}\right)\right| . \tag{8.56}$$

Since W_m does not depend on the r.v.s $(g_{i,m})_{i\le N}$, we have $\mathsf{E}^m\log_{aN} W_m = \mathsf{E}^{m-1}\log_{aN} W_m$, and thus

$$\begin{aligned}N|X_m| &= |\mathsf{E}^m \log_{aN} W - \mathsf{E}^{m-1} \log_{aN} W| \\ &= |\mathsf{E}^m(\log_{aN} W - \log_{aN} W_m) - \mathsf{E}^{m-1}(\log_{aN} W - \log_{aN} W_m)| \\ &\le \mathsf{E}^m|\log_{aN} W - \log_{aN} W_m| + \mathsf{E}^{m-1}|\log_{aN} W - \log_{aN} W_m| \\ &\le \mathsf{E}^m Y + \mathsf{E}^{m-1} Y .\end{aligned}$$

Given $\lambda > 0$, we have, using Jensen's inequality in the last line,

$$\begin{aligned}\mathsf{E}^{m-1} \exp \lambda N |X_m| &\le \mathsf{E}^{m-1}(\exp(\lambda \mathsf{E}^m Y) \exp(\lambda \mathsf{E}^{m-1} Y)) \\ &= \exp(\lambda \mathsf{E}^{m-1} Y)\mathsf{E}^{m-1}(\exp \lambda \mathsf{E}^m Y) \\ &\le (\mathsf{E}^{m-1} \exp \lambda Y)^2 . \end{aligned} \tag{8.57}$$

Let us denote by E_m integration in the r.v.s $(g_{i,m})_{i \le N}$ only, so that $\mathsf{E}^{m-1} = \mathsf{E}^{m-1} \mathsf{E}_m$, and therefore

$$\mathsf{E}^{m-1} \exp \lambda Y = \mathsf{E}^{m-1} \mathsf{E}_m \exp \lambda Y .$$

To conclude the proof it suffices to show that there exists some $\lambda > 0$ depending only on $\kappa_1, \kappa_2, a, \tau$ for which

$$\mathsf{E}_m \exp \lambda Y \le K . \tag{8.58}$$

Indeed, Hölder's inequality then shows that $\mathsf{E}_m \exp \lambda Y/K \le 2$ and using (8.57) for λ/K rather than for λ proves (8.54).

To prove (8.58) we work at given values of $(g_{i,k})_{k \ne m}$, we write E for E_m and we denote by P the corresponding probability. We can and do assume that $W_m \ge \exp(-Na)$, for otherwise $Y = 0$.

Let us consider the probability measure G on $\mathbb{R}^M$ with density proportional to $\exp(\sum_{k \le M, k \ne m} u(S_k) - \kappa \|\boldsymbol{\sigma}\|^2)$. (In the case where $\exp u(x) = \mathbf{1}_{\{x \ge \tau\}}$, G is simply the probability on C_m with density proportional to $\exp(-\kappa\|\boldsymbol{\sigma}\|^2)$.) Let us write $\mathbf{g}_m = (g_{i,m})_{i \le N}$, and let us observe the identity

$$\frac{W}{W_m} = \int \exp u(S_m) \mathrm{d}G(\boldsymbol{\sigma}) .$$

Therefore, using (8.46), we have

$$\begin{aligned}1 \ge \frac{W}{W_m} = \int \exp u(S_m)\mathrm{d}G(\boldsymbol{\sigma}) &= \int \exp u\left(\frac{\mathbf{g}_m \cdot \boldsymbol{\sigma}}{\sqrt{N}}\right) \mathrm{d}G(\boldsymbol{\sigma}) \\ &\ge G\left(\left\{\frac{\mathbf{g}_m \cdot \boldsymbol{\sigma}}{\sqrt{N}} \ge \tau\right\}\right) . \end{aligned} \tag{8.59}$$

We then use Theorem 8.2.7 with $a_i = 0$, $U(\boldsymbol{\sigma}) = \sum_{k \ne m} u(S_k)$, so that since $Z = W_m \ge \exp(-Na)$, condition (8.26) holds. Therefore (8.23) holds, and together with (8.59) it follows that

$$K\exp\Big(-\frac{N}{K}\Big)\le\varepsilon\le\frac{1}{K}\exp(-K\tau^2)\ \Rightarrow\ \mathsf{P}\left(\frac{W}{W_m}\le\varepsilon\right)\le K\varepsilon^{1/K}\ .$$

Setting $\varepsilon=1/t$, it follows that $V:=\exp Y=\min(\exp aN, W_m/W)$ satisfies the conditions of Lemma 8.3.8 with $D=K\exp(K\tau^2)$ and $C=K$. This proves (8.58). □

Proposition 8.3.9. *Consider* $0<\kappa_0<\kappa_1$, $h_0>0$, $a>0$, $\varepsilon>0$. *Then there is a number* $\varepsilon'>0$ *with the following property. Consider any concave function* $u\le 0$ *that satisfies (8.46) together with*

$$\exp u(\tau-\varepsilon')\le\varepsilon'\ . \tag{8.60}$$

Then, if $M\le 10N$, $\kappa_0<\kappa<\kappa_1$ *and* $h\le h_0$ *we have*

$$\Bigg|\mathsf{E}\frac{1}{N}\log_{Na}\int_{\bigcap_{k\le M}U_k}\exp\Big(-\kappa\|\boldsymbol{\sigma}\|^2+h\sum_{i\le N}g_i\sigma_i\Big)\mathrm{d}\boldsymbol{\sigma}$$
$$-\ \mathsf{E}\frac{1}{N}\log_{Na}\int\exp(-H_{N,M}(\boldsymbol{\sigma}))\mathrm{d}\boldsymbol{\sigma}\Bigg|\le\varepsilon\ , \tag{8.61}$$

where of course $H_{N,M}$ is as in (8.43).

Lemma 8.3.10. *If* $x,y,z\le 1$, *we have*

$$|\log_A xz-\log_A yz|\le|\log_A x-\log_A y|\,\mathbf{1}_{\{z\ge\exp(-A)\}}\ .$$

Proof. The result is obvious if $z<\exp(-A)$, since both sides are 0. So we assume $z\ge\exp(-A)$.

If $\log_A x=-A$, then $\log_A xz=-A$, and the result follows since $-A\le\log_A yz\le\log_A y$. The same argument works if $\log_A y=-A$. And if $\log_A x=\log x$, $\log_A y=\log y$, then

$$|\log_A xz-\log_A yz|\le|\log xz-\log yz|=|\log x-\log y|=|\log_A x-\log_A y|\ .\ \square$$

Lemma 8.3.11. *If* $0<y\le x\le 1$, *then, for any* $c>0$,

$$|\log_A x-\log_A y|\le|\log_A y|\,\mathbf{1}_{\{y\le c\}}+\frac{|x-y|}{c}\ .$$

Proof. Since $\log_A y\le\log_A x\le 0$, we have $|\log_A x-\log_A y|\le|\log_A y|$. If $y>c$, we have $|\log x-\log y|\le|x-y|/c$. □

Proof of Proposition 8.3.9. The idea is simply to replace the factors $\mathbf{1}_{U_k}$ by $\exp S_k$ one at a time. Given $m\le M$, we consider the quantity

$$V_m=\int_{C_m}\exp\Big(\sum_{m\le k\le M}u(S_k)-\kappa\|\boldsymbol{\sigma}\|^2+h\sum_{i\le N}g_i\sigma_i\Big)\mathrm{d}\boldsymbol{\sigma}$$

where

$$C_m = \bigcap_{k<m} U_k \ . \tag{8.62}$$

The left-hand side of (8.61) is thus

$$\begin{aligned} &\left|\mathsf{E}\frac{1}{N}\log_{Na} V_1 - \mathsf{E}\frac{1}{N}\log_{Na} V_{M+1}\right| \\ &\le \frac{1}{N}\sum_{1\le m\le M} |\mathsf{E}\log_{Na} V_{m+1} - \mathsf{E}\log_{Na} V_m| \ . \end{aligned} \tag{8.63}$$

Let us fix m and denote by E_m expectation only in the r.v.s $(g_{i,m})_{i\le N}$. We are going to bound

$$\mathsf{E}_m|\log_{Na} V_{m+1} - \log_{Na} V_m| \ . \tag{8.64}$$

Let

$$Z_m = \int_{C_m} \exp\Big(\sum_{m<k\le M} u(S_k) - \kappa\|\boldsymbol{\sigma}\|^2 + h\sum_{i\le N} g_i\sigma_i\Big)\mathrm{d}\boldsymbol{\sigma} \ ,$$

where C_m is given by (8.62). Let us consider the Gibbs measure G_m on $\mathbb{R}^N$ given by

$$G_m(B) = \frac{1}{Z_m}\int_{C_m\cap B} \exp\Big(\sum_{m<k\le M} u(S_k) - \kappa\|\boldsymbol{\sigma}\|^2 + h\sum_{i\le N} g_i\sigma_i\Big)\mathrm{d}\boldsymbol{\sigma} \ .$$

Denoting by $\langle\cdot\rangle$ an average for G_m, we observe the identities

$$\begin{aligned} V_m &= Z_m\langle\exp u(S_m)\rangle \\ V_{m+1} &= Z_m\langle\mathbf{1}_{U_m}\rangle = Z_m G_m(U_m) = Z_m G_m(\{S_m \ge \tau\}) \ . \end{aligned}$$

Since $u \le 0$, we have $V_m, V_{m+1} \le Z_m$, so that the quantity (8.64) is 0 unless $Z_m \ge \exp(-Na)$, and to bound this quantity we can and do assume that $Z_m \ge \exp(-Na)$. Since $u(x) = 0$ for $x \ge \tau$, we have

$$Y := G_m(\{S_m \ge \tau\}) \le X := \langle\exp u(S_m)\rangle \ ,$$

and using Lemmas 8.3.10 in the first inequality below and 8.3.11 in the second one yields that for any $c > 0$,

$$\begin{aligned} |\log_{Na} V_{m+1} - \log_{Na} V_m| &\le |\log_{Na} X - \log_{Na} Y| \\ &\le |\log_{Na} Y|\mathbf{1}_{\{Y\le c\}} + \frac{1}{c}|X - Y| \ . \end{aligned} \tag{8.65}$$

We appeal to Theorem 8.2.7, noting that (8.26) holds since $Z_m \ge \exp(-Na)$. Therefore (8.23) implies

$$K\exp\left(-\frac{N}{K}\right) \le t \le \frac{1}{K}\exp(-K\tau^2) \ \Rightarrow\ \mathsf{P}(Y \le t) \le Kt^{1/K}\ , \tag{8.66}$$

where P denotes the probability corresponding to expectation E_m. We will then first deduce from (8.66) that we may choose c small enough so that

$$\mathsf{E}_m |\log_{Na} Y| \mathbf{1}_{\{Y \le c\}} \le \varepsilon\ , \tag{8.67}$$

c depending only on $\kappa_0, \kappa_1, h_0, a$ and τ. For this we simply observe that

$$|\log_{Na} Y| = \log V\ ,$$

where

$$V = \min(\exp Na, 1/Y)$$

satisfies the conditions of Lemma 8.3.8 for $C, D = K$. Since $V \ge 1$ we have $\log V < K(\lambda)V^\lambda$ for any λ, and Lemma 8.3.8 implies

$$\mathsf{E}(\log V)^2 = \mathsf{E}(\log_{Na} Y)^2 \le K\ ,$$

so the Cauchy-Schwarz inequality yields

$$\mathsf{E}_m |\log_{Na} Y| \mathbf{1}_{\{Y \le c\}} \le K\mathsf{P}(\{Y \le c\})^{1/2}\ .$$

Together with (8.66) this implies (8.67).

Next we show that $\mathsf{E}_m|X - Y|$ can be made arbitrarily small if ε' in (8.60) is small enough. Using (8.60),

$$0 \le X - Y \le \varepsilon' + \langle \mathbf{1}_{\{\tau - \varepsilon' \le S_m \le \tau\}} \rangle$$

so that

$$\mathsf{E}_m|X - Y| \le \varepsilon' + \langle \mathsf{E}_m \mathbf{1}_{\{\tau - \varepsilon' \le S_m \le \tau\}} \rangle\ . \tag{8.68}$$

If z is a Gaussian r.v., then for $\tau_1 < \tau_2$ we have

$$\begin{aligned} \mathsf{P}(\tau_1 \le z \le \tau_2) &\le \frac{1}{\sqrt{2\pi \mathsf{E}z^2}} \int_{\tau_1}^{\tau_2} \exp\left(-\frac{t^2}{2\mathsf{E}z^2}\right) \mathrm{d}t \\ &\le \frac{L(\tau_2 - \tau_1)}{(\mathsf{E}z^2)^{1/2}}\ . \end{aligned} \tag{8.69}$$

Since S_m is a Gaussian r.v. with $\mathsf{E}_m S_m^2 = \|\boldsymbol{\sigma}\|^2/N$ we deduce from (8.69) that

$$\mathsf{E}_m \mathbf{1}_{\{\tau - \varepsilon' \le S_M \le \tau\}} \le L\varepsilon' \frac{\sqrt{N}}{\|\boldsymbol{\sigma}\|}\ ,$$

and since this quantity is at most one we get

$$\langle \mathsf{E}_m \mathbf{1}_{\{\tau - \varepsilon' \le S_M \le \tau\}} \rangle \le L\sqrt{\varepsilon'} + \left\langle \mathbf{1}_{\{\|\boldsymbol{\sigma}\| \le \sqrt{\varepsilon' N}\}} \right\rangle\ .$$

To control the last term (considering again only the case $h = 0$) we simply write that

$$\left\langle \mathbf{1}_{\{\|\boldsymbol{\sigma}\| \le \sqrt{\varepsilon' N}\}} \right\rangle \le \frac{1}{Z_m} \int_{\|\boldsymbol{\sigma}\| \le \sqrt{\varepsilon' N}} \exp(-\kappa \|\boldsymbol{\sigma}\|^2) \mathrm{d}\boldsymbol{\sigma} \ ,$$

and we check as in (8.29) that the integral is $\le \varepsilon' \exp(-Na)$ provided ε' is small enough (depending only on a, κ_1 and κ_2).

In this manner, if we choose c as in (8.67) and then ε' is small enough, then (8.65) implies

$$\mathsf{E}_m |\log_{Na} V_{m+1} - \log_{Na} V_m| \le 2\varepsilon \ ,$$

so that

$$|\mathsf{E} \log_{Na} V_{m+1} - \mathsf{E} \log_{Na} V_m| \le \mathsf{E}\mathsf{E}_m |\log_{Na} V_{m+1} - \log_{Na} V_m| \le 2\varepsilon \ ,$$

and since $M \le 10N$ the left-hand side of (8.61) is $\le 20\varepsilon$. □

Proof of Theorem 8.3.1. We fix a large enough so that for $\kappa_1 \le \kappa \le \kappa_2$, $h \le h_0$, $\alpha \le \alpha_0$ we have

$$\mathrm{RSG}(\alpha) \ge -a + 1 \ . \tag{8.70}$$

Next, we use Proposition 8.3.9 to find ε' small enough that whenever u is a concave function, $u \le 0$, which satisfies (8.46) and (8.60), then (8.61) holds. Moreover, if ε' is small enough, Theorem 3.3.2 shows that we can fix such a function u, which is moreover four times differentiable with bounded first four derivatives, so that (3.127) holds. Since $\log_A x \ge \log x$, we see that for N large enough, we have, using from (3.127) in the first inequality and (8.70) in the second one,

$$\mathsf{E}\frac{1}{N} \log_{Na} Z_{N,M} \ge \mathsf{E}\frac{1}{N} \log Z_{N,M} \ge \mathrm{RSG}\left(\frac{M}{N}\right) - \varepsilon \ge -a + 1 - \varepsilon \ ,$$

where we write for simplicity $Z_{N,M} = \int \exp(-H_{N,M}(\boldsymbol{\sigma}))\mathrm{d}\boldsymbol{\sigma}$. Assuming, as we may, that $\varepsilon \le 1/4$, so that $N^{-1}\mathsf{E} \log_{Na} Z_{N,M} \ge -a + 3/4$, we obtain from (8.47) that

$$\mathsf{P}(\Omega) \le K \exp\left(-\frac{N}{K}\right) \tag{8.71}$$

where

$$\Omega = \left\{ \frac{1}{N} \log_{Na} Z_{N,M} \le -a \right\} = \{ Z_{N,M} \le \exp(-Na) \} \ . \tag{8.72}$$

Next, we claim that for N large enough we have

$$\left| \mathsf{E}\frac{1}{N} \log_{Na} Z_{N,M} - \mathsf{E}\frac{1}{N} \log Z_{N,M} \right| \le \varepsilon \ . \tag{8.73}$$

Since on Ω we have $\log_{Na} Z_{N,M} = -a$ and on Ω^c we have $\log_{Na} Z_{N,M} = \log Z_{N,M}$, the left-hand side is

$$\left|\mathsf{E}\mathbf{1}_\Omega\Big(a + \frac{1}{N}\log Z_{N,M}\Big)\right| . \tag{8.74}$$

It should be obvious from the bound of Lemma 3.1.6 that $N^{-2}\mathsf{E}(\log Z_{N,M})^2 \le K$, so that

$$\mathsf{E}\Big(a + \frac{1}{N}\log Z_{N,M}\Big)^2 \le K .$$

Using (8.71) and the Cauchy-Schwarz inequality in (8.74) then proves (8.73). Combining (8.73) with (3.127) and (8.61) (and recalling (8.37)) we see that if N is large enough and $M \le \alpha_0 N$ we have

$$\left|\mathsf{E}\frac{1}{N}\log_{Na}\int_{\bigcap_{k\le M} U_k}\exp\Big(-\kappa\|\boldsymbol{\sigma}\|^2 + h\sum_{i\le N} g_i\sigma_i\Big)\mathrm{d}\boldsymbol{\sigma} - \mathrm{RSG}\Big(\frac{M}{N}\Big)\right| \le 3\varepsilon ,$$

and combining with (8.47) for $\exp u(x) = \mathbf{1}_{\{x\ge\tau\}}$ and $t=\varepsilon$ we get

$$\mathsf{P}\left(\left|\frac{1}{N}\log_{Na}\int_{\bigcap_{k\le M} U_k}\exp\Big(-\kappa\|\boldsymbol{\sigma}\|^2 + h\sum_{i\le N} g_i\sigma_i\Big)\mathrm{d}\boldsymbol{\sigma} - \mathrm{RSG}\left(\frac{M}{N}\right)\right| \le 4\varepsilon\right)$$
$$\ge 1 - 2\exp\left(-\frac{N}{K}\right) .$$

However, since $\mathrm{RSG}(\alpha) \ge -a + 1 \ge -a + 4\varepsilon$, when the event on the left occurs, $\log_{Na}$ and log are equal, and this completes the proof. □

8.4 The Gardner Formula for the Sphere

Theorem 8.3.1 computes the Gaussian measure of $\bigcap_{k\le M} U_k$, i.e. the proportion of the Gaussian space that belongs to $\bigcap_{k\le M} U_k$. Let us denote by $\mathbb{S}_N$ the sphere of radius $\sqrt{N}$,

$$\mathbb{S}_N = \{\boldsymbol{\sigma} \in \mathbb{R}^N \ ; \ \|\boldsymbol{\sigma}\|^2 = N\}$$

and by μ_N the uniform measure on $\mathbb{S}_N$. Gardner's formula computes the proportion of $\mathbb{S}_N$ that belongs to $\bigcap_{k\le M} U_k$. Throughout the section we assume that $h = 0$.

Theorem 8.4.1. *(Shcherbina-Tirozzi [86]) Consider $\tau \ge 0$, and define*

$$\alpha(\tau) = \frac{1}{\mathsf{E}\,\max^2(z+\tau, 0)} \tag{8.75}$$

where z is standard Gaussian. Then if $\alpha_0 < \alpha(\tau)$, given $\varepsilon > 0$, for N large enough, we have

$$\frac{M}{N} \le \alpha_0 \Rightarrow \mathsf{P}\left(\left|\frac{1}{N}\log \mu_N\left(\bigcap_{k\le M} U_k\right) - \mathrm{RS}\left(\frac{M}{N}\right)\right| \ge \varepsilon\right) \le \exp\left(-\frac{N}{K}\right) \tag{8.76}$$

where K depends on τ and α_0 only and where

$$\mathrm{RS}(\alpha) = \min_q \left(\alpha\, \mathsf{E}\log \mathcal{N}\left(\frac{\tau - z\sqrt{q}}{\sqrt{1-q}}\right) + \frac{1}{2}\frac{q}{1-q} + \frac{1}{2}\log(1-q)\right) . \tag{8.77}$$

Moreover given any number a (taking possibly very large negative values) we have, for N large enough

$$\frac{M}{N} \ge \alpha(\tau) \Rightarrow \mathsf{P}\left(\frac{1}{N}\log \mu_N\left(\bigcap_{k\le M} U_k\right) \ge a\right) \le \exp\left(-\frac{N}{K}\right) . \tag{8.78}$$

Here as usual $\mathcal{N}(x) = \mathsf{P}(z \ge x)$. The reader will not confuse the definition (8.77) with the definition (8.36).

Research Problem 8.4.2. (Level 1) Extend (8.78) to the case $\tau < 0$, α small.

Research Problem 8.4.3. (Level ≥ 2) Understand what happens for any $\tau < 0$, $\alpha > 0$.

Conjecture 8.4.4. (Level 2) If $M/N \ge \alpha_0 > \alpha(\tau)$, for large N the set $\mathbb{S}_N \cap \bigcap_{k\le M} U_k$ is empty with probability at least $1 - \exp(-N/K)$.

Research Problem 8.4.5. (Level 2) Find the correct rate of convergence in Theorem 8.4.1.

We will deduce Theorem 8.4.1 from Theorem 8.3.1, and we explain first the overall strategy. We define

$$C = C_{N,M} = \bigcap_{k\le M} U_k$$

and

$$F_{N,M}(\kappa) = \frac{1}{N}\log \int_{C_{N,M}} \exp(-\kappa\,\|\boldsymbol{\sigma}\|^2)\,\mathrm{d}\boldsymbol{\sigma} , \tag{8.79}$$

a quantity we have studied in Theorem 8.3.1. It is now clearer to indicate in the notation that the quantity (8.37) depends on κ, so that we denote it now by $\mathrm{RSG}(\alpha,\kappa)$. The content of Theorem 8.3.1 is that $F_{N,M}(\kappa) \simeq \mathrm{RSG}(\alpha,\kappa)$, hence

$$\int_{C_{N,M}} \exp(-\kappa\|\boldsymbol{\sigma}\|^2)\mathrm{d}\boldsymbol{\sigma} \simeq \exp(N\mathrm{RSG}(\alpha,\kappa)) \ . \tag{8.80}$$

Let us denote by $\mathrm{Vol}\,A$ the N-dimensional volume of a subset A of $\mathbb{R}^N$. It follows from (8.80) that for any κ we approximatively have

$$\exp(-N\kappa)\mathrm{Vol}(C_{N,M} \cap \{\|\boldsymbol{\sigma}\| \le \sqrt{N}\}) \le \exp(N\mathrm{RSG}(\alpha,\kappa)) \ ,$$

and thus

$$\frac{1}{N}\log \mathrm{Vol}(C_{M,N} \cap \{\|\boldsymbol{\sigma}\| \le \sqrt{N}\}) \le \kappa + \mathrm{RSG}(\alpha,\kappa) \ .$$

If the function $\kappa \mapsto \mathrm{RSG}(\alpha,\kappa)$ is convex, we can expect that

$$\frac{1}{N}\log \mathrm{Vol}(C_{M,N} \cap \{\|\boldsymbol{\sigma}\| \le \sqrt{N}\}) \simeq \inf_{\kappa}\{\kappa + \mathrm{RSG}(\alpha,\kappa)\} \ . \tag{8.81}$$

This should be natural to the reader familiar with (elementary) large deviation theory; a simple proof will be given below.

We observe from (8.37) and (8.35) that

$$\frac{\partial}{\partial\kappa}\mathrm{RSG}(\alpha,\kappa) = -\rho_0 \ ,$$

and thus the value of κ that achieves the infimum in (8.81) is the value for which $\rho_0 = \rho_0(\alpha,\kappa) = 1$. Interestingly one can show that in Theorem 8.3.1 the Gibbs measure nearly lives on the sphere of radius $\sqrt{N}\rho_0$. Therefore it is quite natural that the case where $\rho_0 = 1$ should give information on the left-hand side of (8.81). Of course, if one can compute the left-hand side of (8.81), one should be able to compute $\mu_N(C_{N,M})$ because in high dimension, only the region very close to the boundary really contributes to the volume of a ball.

We start by performing the calculus necessary to show that the previous program has a chance to be completed. We recall the solutions $q_0 = q_0(\alpha,\kappa)$ and $\rho_0 = \rho_0(\alpha,\kappa)$ of the equations (8.35).

Proposition 8.4.6. *We have*

$$\frac{\partial\rho_0}{\partial\kappa} < 0 \tag{8.82}$$

$$\lim_{\kappa\to\infty}\rho_0(\alpha,\kappa) = \rho(\alpha) := \inf\left\{\rho > 0 \ ; \ \alpha\mathsf{E}\max{}^2\left(z + \frac{\tau}{\sqrt{\rho}}, 0\right) \le 1\right\} \ . \tag{8.83}$$

Proof. Making the change of variable $q = \rho x/(1+x)$, we have $F(q,\rho) = G(x,\rho)$, where $G(x,\rho)$ is given by (3.149). In Proposition 3.3.11 we proved that at every point we have

$$\frac{\partial^2 G}{\partial \rho^2} < 0 \; ; \; \frac{\partial}{\partial x}\left(\frac{x+1}{x}\frac{\partial G}{\partial x}\right) > 0 \; . \tag{8.84}$$

Let us define x_0 by $q_0 = \rho x_0/(1+x_0)$, so that by Theorem 3.3.1 (x_0, ρ_0) is the unique point at which

$$\frac{\partial G}{\partial x}(x_0, \rho_0) = \frac{\partial G}{\partial \rho}(x_0, \rho_0) = 0 \; . \tag{8.85}$$

We obtain from (8.84) that

$$\frac{\partial^2 G}{\partial \rho^2} < 0 \; , \; \frac{\partial^2 G}{\partial x^2} > 0 \tag{8.86}$$

where now, as in the rest of the proof, it is understood that G and its partial derivatives are computed at the arguments x_0 and ρ_0. We write $\rho' = \partial \rho_0/\partial \kappa$, $x' = \partial x_0/\partial \kappa$. Differentiating with respect to κ the relations $\partial G/\partial \rho = 0$ and $\partial G/\partial x = 0$ we get respectively

$$x' \frac{\partial^2 G}{\partial \rho \partial x} + \rho' \frac{\partial^2 G}{\partial \rho^2} + \frac{\partial^2 G}{\partial \kappa \partial \rho} = 0$$
$$x' \frac{\partial^2 G}{\partial x^2} + \rho' \frac{\partial^2 G}{\partial \rho \partial x} + \frac{\partial^2 G}{\partial \kappa \partial x} = 0 \; .$$

Since $\partial^2 G/\partial \kappa \partial \rho = 1$ and $\partial^2 G/\partial \kappa \partial x = 0$, combining these relations to eliminate the mixed derivative $\partial^2 G/\partial \rho \partial x$ we get

$$\rho' = \rho'^2 \frac{\partial^2 G}{\partial \rho^2} - x'^2 \frac{\partial^2 G}{\partial x^2} < 0 \tag{8.87}$$

by (8.86). This proves (8.82).

To prove (8.83), we recall from Lemma 3.3.13 that

$$\frac{\partial G}{\partial \rho}(x_0, \rho_0) = 0 \Rightarrow \frac{\alpha \tau \sqrt{1+x_0}}{\rho_0^{3/2}} \mathsf{E} \mathcal{A}\left(\frac{\tau\sqrt{1+x_0}}{\sqrt{\rho_0}} - z\sqrt{x_0}\right) + \frac{1}{\rho_0} = 2\kappa \tag{8.88}$$

and

$$2\frac{x_0+1}{x_0}\frac{\partial G}{\partial x}(x_0, \rho_0) = -\frac{\alpha}{x_0} \mathsf{E} \mathcal{A}\left(\frac{\tau\sqrt{1+x_0}}{\sqrt{\rho_0}} - z\sqrt{x_0}\right)^2 + 1 = 0 \; , \tag{8.89}$$

where $\mathcal{A}(y) = e^{-y^2/2}/(\sqrt{2\pi}\mathcal{N}(y))$. When $\tau = 0$, (8.88) implies that $\rho_0 = 1/2\kappa$, so (8.83) is obvious in this case because $\rho(\alpha) = 0$ since $\alpha \le 2$. Thus we assume $\tau > 0$, in which case $\rho(\alpha) > 0$ is defined by

$$\alpha \mathsf{E} \max{}^2\left(z + \frac{\tau}{\sqrt{\rho(\alpha)}}, 0\right) = 1 \; ,$$

or, equivalently, changing z into $-z$,

$$\alpha \mathsf{E} \max{}^2\left(\frac{\tau}{\sqrt{\rho(\alpha)}} - z, 0\right) = 1 \; . \tag{8.90}$$

By (8.84) and (8.85) for $x \geq x_0$ we have

$$\frac{x+1}{x}\frac{\partial G}{\partial x}(x, \rho_0) > 0 \; ,$$

so that, by (8.89)

$$\frac{\alpha}{x}\mathsf{E}\mathcal{A}\left(\frac{\tau\sqrt{1+x}}{\sqrt{\rho_0}} - z\sqrt{x}\right)^2 < 1 \; .$$

Since $\mathcal{A} \geq 0$ and $\mathcal{A}(v) \geq v$ by (3.133), we have $\mathcal{A}(v) \geq \max(v, 0)$, and thus

$$\frac{\alpha}{x}\mathsf{E} \max{}^2\left(\frac{\tau\sqrt{1+x}}{\sqrt{\rho_0}} - z\sqrt{x}, 0\right) < 1 \; .$$

Letting $x \to \infty$ implies

$$\alpha \mathsf{E} \max{}^2\left(\frac{\tau}{\sqrt{\rho_0}} - z, 0\right) \leq 1$$

and therefore $\rho_0 \geq \rho(\alpha)$ by (8.90). As $\kappa \to \infty$, the left-hand side of (8.88) must go to ∞. Since $\rho_0 \geq \rho(\alpha) > 0$, we must have $x_0 \to \infty$. Let $\rho_\infty = \lim_{\kappa\to\infty} \rho_0$, that exists by (8.87). From (8.89) we have

$$\frac{\alpha}{x_0}\mathsf{E}\mathcal{A}\left(a(\kappa)\right)^2 = 1 \; , \tag{8.91}$$

where

$$a(\kappa) = \frac{\tau\sqrt{1+x_0}}{\sqrt{\rho_0}} - z\sqrt{x_0} \; .$$

Since

$$\lim_{\kappa\to\infty}\frac{a(\kappa)}{\sqrt{x_0}} = \frac{\tau}{\sqrt{\rho_\infty}} - z \; ,$$

as $\kappa \to \infty$ we have $a(\kappa) \to \infty$ when $\tau/\sqrt{\rho_\infty} - z > 0$, while $a(\kappa) \to -\infty$ when $\tau/\sqrt{\rho_\infty} - z < 0$. Writing (8.91) as

$$\alpha \mathsf{E}\frac{a(\kappa)^2}{x_0}\frac{\mathcal{A}(a(\kappa))^2}{a^2(\kappa)} = 1 \; ,$$

and since $\mathcal{A}(v)/v \to 1$ as $v \to \infty$ and $\mathcal{A}(v)/v \to 0$ as $v \to -\infty$ we get in the limit $\kappa \to \infty$ that

$$\alpha \mathsf{E} \max{}^2\left(\frac{\tau}{\sqrt{\rho_\infty}} - z, 0\right) = 1$$

i.e. $\rho_\infty = \rho(\alpha)$. □

We denote by $\mathrm{RS}_0(\alpha, \kappa)$ the quantity (8.36), to bring out the fact that it depends on κ. We recall that from (8.35) and (8.36) we have

$$\frac{\partial}{\partial \kappa}\mathrm{RS}_0(\alpha, \kappa) = -\rho_0(\alpha, \kappa) \; . \tag{8.92}$$

Lemma 8.4.7. *If $\alpha < \alpha(\tau)$ (given by (8.75)), there is a unique value $\kappa = \kappa(\alpha)$ for which $\rho_0(\alpha, \kappa) = 1$. The function $f(\kappa) = \kappa + \mathrm{RS}_0(\alpha, \kappa)$ attains its minimum at this value of κ, and this minimum is the quantity $\mathrm{RS}(\alpha)$ of (8.77).*

Proof. When $\alpha < \alpha(\tau)$ we have $\rho(\alpha) < 1$ so that by (8.83) we have $\rho_0(\alpha, \kappa) < 1$ for large κ. Also, by (8.88) and since $\mathcal{A} > 0$ we have $\rho_0 \geq 1/2\kappa$ so $\rho_0 > 1$ for $\kappa < 1/2$. By (8.82) there is a unique κ^* for which $\rho_0(\alpha, \kappa^*) = 1$. Since by (8.92) we have $f'(\kappa) = 1 - \rho_0(\alpha, \kappa)$, (8.82) proves that $f''(\kappa) > 0$, so that $f(\kappa)$ attains its minimum at $\kappa = \kappa^*$ and by (8.34) and (8.36) this minimum is $U(q_0)$, where

$$U(q) = \alpha \mathsf{E} \log \mathcal{N}\left(\frac{\tau - z\sqrt{q}}{\sqrt{1-q}}\right) + \frac{1}{2}\frac{q}{1-q} + \frac{1}{2}\log(1-q) \; . \tag{8.93}$$

Since $U(q) = F(1, q) + \kappa$, and since, because $\rho_0 = 1$,

$$\frac{\mathrm{d}U}{\mathrm{d}q}(q_0) = \frac{\partial F}{\partial q}(\rho_0, q_0) = 0 \; ,$$

it follows from the second part of Corollary 3.3.12 that the function $U(q)$ attains its minimum for $q = q_0$, so that $U(q_0)$ is the quantity $\mathrm{RS}(\alpha)$ of (8.77). □

We now settle a side story. We will prove that (8.78) follows from (8.76). This relies on the following.

Lemma 8.4.8. *We have*

$$\lim_{\alpha \to \alpha(\tau)} \mathrm{RS}(\alpha) = -\infty \; . \tag{8.94}$$

Proof. We observe that, since $\mathcal{N}(x) = \mathsf{P}(\xi \geq x) \leq \exp(-\max^2(0, x)/2)$, we have, given any $q < 1$, and using (8.75) in the last line,

$$\begin{aligned}
\mathsf{E} \log \mathcal{N}\left(\frac{\tau - z\sqrt{q}}{\sqrt{1-q}}\right) &\leq -\frac{1}{2(1-q)}\,\mathsf{E}\left(\max^2(\tau - z\sqrt{q}, 0)\right) \\
&= -\frac{q}{2(1-q)}\,\mathsf{E}\left(\max^2\left(z + \frac{\tau}{\sqrt{q}}, 0\right)\right) \\
&\leq -\frac{q}{2(1-q)}\,\mathsf{E}\left(\max^2(z + \tau, 0)\right) \\
&= -\frac{q}{2\alpha(\tau)(1-q)}
\end{aligned}$$

so that

$$\begin{aligned}\mathrm{RS}(\alpha) &\le \alpha\,\mathsf{E}\log\mathcal{N}\left(\frac{\tau - z\sqrt{q}}{\sqrt{1-q}}\right) + \frac{1}{2}\frac{q}{1-q} + \frac{1}{2}\log(1-q)\\ &\le \frac{q}{2(1-q)}\left(1 - \frac{\alpha}{\alpha(\tau)}\right) + \frac{1}{2}\log(1-q)\ .\end{aligned}$$

As $\alpha \to \alpha(\tau)$, we get

$$\limsup_{\alpha\to\alpha(\tau)} \mathrm{RS}(\alpha) \le \frac{1}{2}\log(1-q)\ ,$$

and this proves the result since q is arbitrary. □

To prove that (8.78) follows from (8.76), consider a number a as in (8.78), and $\alpha_1 < \alpha(\tau)$ such that $\mathrm{RS}(\alpha) < a$ for $\alpha_1 \le \alpha < \alpha(\tau)$. Consider α_0 with $\alpha_1 < \alpha_0 < \alpha(\tau)$. Then for $\alpha_1 < M/N < \alpha_0$ and N by large, by (8.76) the right-hand side of (8.78) holds true, and this implies that (8.78) holds true since $\mu_N(C_{N,M})$ decreases as M increases.

Let us denote by Area A the $(N-1)$-dimensional area of a subset of A of $v\mathbb{S}_N$, for any $v > 0$. Thus, for any set A we have

$$\mu_N(A) = \frac{\mathrm{Area}\, A \cap \mathbb{S}_N}{\mathrm{Area}\,\mathbb{S}_N}$$

so that

$$\frac{1}{N}\log\mu_N(A) = \frac{1}{N}\log\mathrm{Area}\,(A\cap\mathbb{S}_N) - \frac{1}{N}\log\mathrm{Area}\,\mathbb{S}_N \tag{8.95}$$

and (classically)

$$\lim_{N\to\infty}\frac{1}{N}\log\mathrm{Area}\,\mathbb{S}_N = \frac{1}{2}\log(2e\pi)\ . \tag{8.96}$$

We recall that $C = C_{N,M} = \bigcap_{k\le M} U_k$. A simple observation is that, since $\tau \ge 0$, if $v \ge 1$ we have

$$vU_k = \{S_k \ge v\tau\} \subset \{S_k \ge \tau\} = U_k\ ,$$

so that $vC \subset C$, and if $v \le 1$ we have $vC \supset C$. Thus

$$v \ge 1 \Rightarrow \mathrm{Area}\,(C\cap v\mathbb{S}_N) \ge \mathrm{Area}\,v(C\cap\mathbb{S}_N) = v^{N-1}\mathrm{Area}\,(C\cap\mathbb{S}_N) \tag{8.97}$$

$$v \le 1 \Rightarrow \mathrm{Area}\,(C\cap v\mathbb{S}_N) \le \mathrm{Area}\,v(C\cap\mathbb{S}_N) = v^{N-1}\mathrm{Area}\,(C\cap\mathbb{S}_N)\ . \tag{8.98}$$

Proof of Theorem 8.4.1. To prove (8.76) we show that given $\alpha < \alpha(\tau)$, given $\varepsilon > 0$, then if $N \to \infty$, $M = \lfloor \alpha N \rfloor$, for N large enough we have

$$\mathsf{P}\left(\left|\frac{1}{N}\log\mu_N\left(\bigcap_{k\le M}U_k\right)-\mathrm{RS}(\alpha)\right|\ge\varepsilon\right)\le\exp\left(-\frac{N}{K}\right). \tag{8.99}$$

The claim that the same value of N achieves (8.76) for all $M \le \alpha_0 N$ follows by a simple argument from the fact that $\mu_N(C_{N,M})$ decreases as M increases.

We recall the content of Theorem 8.3.1: Given $\varepsilon > 0$ and $\kappa > 0$, for N large enough (and $M = \lfloor \alpha N \rfloor$) we have

$$\mathsf{P}\left(|F_{N,M}(\kappa)-\mathrm{RSG}(\alpha,\kappa)|\ge\varepsilon\right)\le\exp\left(-\frac{N}{K}\right). \tag{8.100}$$

Consider $w > 1$, and observe that

$$\int_C \exp(-\kappa\|\boldsymbol{\sigma}\|^2)\,\mathrm{d}\boldsymbol{\sigma}\ge\exp(-\kappa N w^2)\,\mathrm{Vol}(C\cap\{\boldsymbol{\sigma}\,;\ N\le\|\boldsymbol{\sigma}\|^2\le Nw^2\}).$$

Using (8.97) in the second line we get

$$\begin{aligned}\mathrm{Vol}(C\cap\{\boldsymbol{\sigma}\,;\ N\le\|\boldsymbol{\sigma}\|^2\le Nw^2\}) &= \sqrt{N}\int_1^w \mathrm{Area}(C\cap v\mathbb{S}_N)\,\mathrm{d}v\\ &\ge \frac{1}{\sqrt{N}}\,(w^N-1)\,\mathrm{Area}(C\cap\mathbb{S}_N).\end{aligned}$$

Now we take $w = 1 + 1/\kappa N$, so $w^N - 1 \ge 1/\kappa$, and hence

$$\int_C \exp(-\kappa\|\boldsymbol{\sigma}\|^2)\mathrm{d}\boldsymbol{\sigma}\ge\frac{1}{\sqrt{N}}\exp(-\kappa N w^2)\mathrm{Area}(C\cap\mathbb{S}_N).$$

Recalling (8.79) and taking logarithms yields

$$F_{N,M}(\kappa)\ge-\kappa w^2+\frac{1}{N}\log\mathrm{Area}(C\cap\mathbb{S}_N)-\frac{1}{N}\log(\kappa\sqrt{N}).$$

Since $\kappa w^2 = 2/N + 1/\kappa N^2 + \kappa$, we get

$$\frac{1}{N}\log\mathrm{Area}\,(C\cap\mathbb{S}_N)\le F_{N,M}(\kappa)+\kappa+\varepsilon_N, \tag{8.101}$$

where

$$\varepsilon_N=\frac{\log(\kappa\sqrt{N})}{N}+\frac{2}{N}+\frac{1}{\kappa N^2}.$$

Combining with (8.95), we obtain

$$\frac{1}{N}\log\mu_N(C)\le F_{N,M}(\kappa)+\kappa-\frac{1}{N}\log\mathrm{Area}\,\mathbb{S}_N+\varepsilon_N.$$

Consequently,

$$\begin{aligned}
&\mathsf{P}\left(\frac{1}{N}\log\mu_N(C)\geq \mathrm{RS}(\alpha)+\varepsilon\right)\\
&\leq \mathsf{P}\left(F_{N,M}(\kappa)+\kappa-\frac{1}{N}\log \mathrm{Area}\,\mathbb{S}_N+\varepsilon_N\geq \mathrm{RS}(\alpha)+\varepsilon\right)\\
&= \mathsf{P}\left(F_{N,M}(\kappa)\geq \mathrm{RS}(\alpha)-\kappa+\frac{1}{N}\log \mathrm{Area}\,\mathbb{S}_N+\varepsilon-\varepsilon_N\right).
\end{aligned}$$

Let us use this for the value $\kappa=\kappa(\alpha)$ for which $\mathrm{RS}_0(\alpha,\kappa(\alpha))+\kappa(\alpha)=\mathrm{RS}(\alpha)$. We have, setting $\varepsilon'_N=\frac{1}{N}\log\mathrm{Area}\,\mathbb{S}_N-\frac{1}{2}\log(2e\pi)$,

$$\begin{aligned}
\mathrm{RS}(\alpha)-\kappa(\alpha)+\frac{1}{N}\log\mathrm{Area}\,\mathbb{S}_N &= \mathrm{RS}_0(\alpha,\kappa(\alpha))+\frac{1}{N}\log\mathrm{Area}\,\mathbb{S}_N\\
&= \mathrm{RS}_0(\alpha,\kappa(\alpha))+\frac{1}{2}\log(2e\pi)+\varepsilon'_N\\
&= \mathrm{RSG}(\alpha,\kappa(\alpha))+\varepsilon'_N\ .
\end{aligned}$$

Therefore

$$\begin{aligned}
&\mathsf{P}\left(\frac{1}{N}\log\mu_N(C)\geq \mathrm{RS}(\alpha)+\varepsilon\right)\\
&\leq \mathsf{P}\left(F_{N,M}(\kappa(\alpha))\geq \mathrm{RSG}(\alpha,\kappa(\alpha))+\varepsilon-\varepsilon_N+\varepsilon'_N\right),
\end{aligned}$$

and combining with (8.96) and (8.100), we obtain that for every ε and N large we have

$$\mathsf{P}\left(\frac{1}{N}\log\mu_N(C)\geq \mathrm{RS}(\alpha)+\varepsilon\right)\leq \exp\left(-\frac{N}{K}\right)\ .$$

To prove a lower bound on $\mu_N(C)$, consider the function $f(\kappa)=\mathrm{RS}_0(\alpha,\kappa)+\kappa$. This function attains its minimum at $\kappa=\kappa(\alpha)$. Considering $v<1$, for small δ we have

$$f(\kappa(\alpha))<f(\kappa(\alpha)+\delta)\ ;\quad f(\kappa(\alpha)+\delta)>f(\kappa(\alpha)+2\delta)-(1-v^2)\delta\ ,$$

the second statement resulting from the fact that $\lim_{\delta\to 0}(f(\kappa(\alpha)+\delta)-f(\kappa(\alpha))/\delta=0$. Setting $\kappa=\kappa(\alpha)+\delta$ where $\delta>0$ is as above, these inequalities are equivalent to

$$\mathrm{RS}_0(\alpha,\kappa-\delta)-\delta<\mathrm{RS}_0(\alpha,\kappa)\,;\ \mathrm{RS}_0(\alpha,\kappa+\delta)+\delta v^2<\mathrm{RS}_0(\alpha,\kappa)\ . \tag{8.102}$$

We observe the inequalities (recalling (8.79))

$$\begin{aligned}
I_1 &:= \int_{C\cap\{\|\boldsymbol{\sigma}\|\geq\sqrt{N}\}}\exp(-\kappa\,\|\boldsymbol{\sigma}\|^2)\,\mathrm{d}\boldsymbol{\sigma}\\
&\leq\ \exp(-N\delta)\int_C\exp(-(\kappa-\delta)\,\|\boldsymbol{\sigma}\|^2)\,\mathrm{d}\boldsymbol{\sigma}\\
&=\ \exp(N(-\delta+F_{M,N}(\kappa-\delta)))
\end{aligned}$$

$$\begin{aligned} I_2 :&= \int_{C\cap\{\|\boldsymbol{\sigma}\|\le v\sqrt{N}\}} \exp(-\kappa\,\|\boldsymbol{\sigma}\|^2)\,\mathrm{d}\boldsymbol{\sigma} \\ &\le \exp(N\delta v^2)\int_C \exp(-(\kappa+\delta)\,\|\boldsymbol{\sigma}\|^2)\,\mathrm{d}\boldsymbol{\sigma} \\ &= \exp(N(\delta v^2 + F_{M,N}(\kappa+\delta)))\ , \end{aligned} \tag{8.103}$$

and thus

$$\begin{aligned} \int_{C\cap\{v\sqrt{N}\le\|\boldsymbol{\sigma}\|\le\sqrt{N}\}} \exp(-\kappa\,\|\boldsymbol{\sigma}\|^2)\,\mathrm{d}\boldsymbol{\sigma} &\ge \int_C \exp(-\kappa\,\|\boldsymbol{\sigma}\|^2)\,\mathrm{d}\boldsymbol{\sigma} - I_1 - I_2 \\ &\ge \exp N\,F_{M,N}(\kappa) \\ &\quad - \exp N(-\delta + F_{M,N}(\kappa-\delta)) \\ &\quad - \exp N(\delta\,v^2 + F_{M,N}(\kappa+\delta))\ . \end{aligned}$$

The left-hand side is at most

$$\exp(-\kappa\,v^2\,N)\,\mathrm{Vol}(C\cap\{\|\boldsymbol{\sigma}\|\,;\ \|\boldsymbol{\sigma}\|\le\sqrt{N}\})\ .$$

Now, using (8.98),

$$\begin{aligned} \mathrm{Vol}(C\cap\{\|\boldsymbol{\sigma}\|\,;\ \|\boldsymbol{\sigma}\|\le\sqrt{N}\}) &= \sqrt{N}\int_0^1 \mathrm{Area}(C\cap v\,\mathbb{S}_N)\,\mathrm{d}v \\ &\le \sqrt{N}\int_0^1 v^{N-1}\mathrm{Area}(C\cap\mathbb{S}_N)\,\mathrm{d}v \\ &= \frac{1}{\sqrt{N}}\,\mathrm{Area}(C\cap\mathbb{S}_N)\ . \end{aligned}$$

Thus we proved the bound

$$\begin{aligned} \mathrm{Area}(C\cap\mathbb{S}_N) \ge \sqrt{N}\exp(N\kappa v^2)\big(&\exp NF_{M,N}(\kappa) \\ &- \exp N(-\delta + F_{M,N}(\kappa-\delta)) - \exp N(\delta v^2 + F_{M,N}(\kappa+\delta))\big)\ . \end{aligned} \tag{8.104}$$

The first part of (8.102) implies

$$\varepsilon' := \frac{1}{3}\big(\mathrm{RSG}(\alpha,\kappa) - (\mathrm{RSG}(\alpha,\kappa-\delta)-\delta)\big) > 0\ .$$

Using (8.100) for $\varepsilon = \varepsilon'$ shows that for N large enough, with probability $\ge 1 - K\exp(-N/K)$ we have

$$F_{M,N}(\kappa) \ge \mathrm{RSG}(\alpha,\kappa) - \varepsilon'$$

$$F_{M,N}(\kappa-\delta) \le \mathrm{RSG}(\alpha,\kappa-\delta) + \varepsilon'$$

and therefore

$$F_{M,N}(\kappa) - F_{M,N}(\kappa-\delta) \ge \mathrm{RSG}(\alpha,\kappa) - \mathrm{RSG}(\alpha,\kappa-\delta) - 2\varepsilon' = -\delta + \varepsilon'\ .$$

Consequently,

$$F_{M,N}(\kappa) \geq F_{M,N}(\kappa - \delta) - \delta + \varepsilon'$$

and for N large enough we have

$$\exp NF_{M,N}(\kappa) \geq 4 \exp N(-\delta + F_{M,N}(\kappa - \delta)) \; .$$

Proceeding in a similar manner for the other term in (8.104), we obtain that for N large, with probability $\geq 1 - K\exp(-N/K)$,

$$\exp NF_{M,N}(\kappa) \geq 4 \exp N(\delta v^2 + F_{M,N}(\kappa + \delta)) \; ,$$

and therefore

$$\exp NF_{M,N}(\kappa) \geq 2 \exp N(-\delta + F_{M,N}(\kappa - \delta)) + 2 \exp N(\delta v^2 + F_{M,N}(\kappa + \delta)) \; .$$

Therefore (8.104) implies that

$$\mathrm{Area}(C \cap \mathbb{S}_N) \geq \frac{\sqrt{N}}{2} \exp N(F_{M,N}(\kappa) + \kappa v^2) \; ,$$

and thus

$$\frac{1}{N} \log \mathrm{Area}(C \cap \mathbb{S}_N) \geq \kappa v^2 + \mathrm{RSG}(\alpha, \kappa) \; .$$

Lemma 8.4.7 implies that $\mathrm{RS}_0(\alpha, \kappa) + \kappa \geq \mathrm{RS}(\alpha)$, and hence

$$\mathrm{RSG}(\alpha, \kappa) + \kappa \geq \mathrm{RS}(\alpha) + \frac{1}{2} \log(2e\pi) \; .$$

Using (8.95) and (8.96) we obtain that with overwhelming probability, for large N, it holds true that

$$\frac{1}{N} \log \mu_N(C) \geq \mathrm{RS}(\alpha) - 2(1 - v^2)\kappa$$

and this conclude the proof since $v < 1$ is arbitrary. □

8.5 The Bernoulli Model

In this model we replace the independent Gaussian r.v.s $g_{i,k}$ by independent Bernoulli r.v.s $\eta_{i,k}$, $\mathsf{P}(\eta_{i,k} = \pm 1) = 1/2$; so the basic Hamiltonian is now

$$-H_{N,M}(\boldsymbol{\sigma}) = \sum_{k \leq M} u\left(\frac{1}{\sqrt{N}} \sum_{i \leq N} \eta_{i,k}\sigma_i\right) - \kappa \|\boldsymbol{\sigma}\|^2 + h \sum_{i \leq N} g_i \sigma_i \; . \qquad (8.105)$$

One expects the behavior of this model to be very similar to the behavior of the model with the Gaussian r.v.s. Proving this is, however, another matter, in particular because so far we have used specifically Gaussian tools. It is rather unfortunate to find in the published literature papers considering in the introduction Hamiltonians such as (8.105), and then stating that "for simplicity we will consider only the case where the r.v.s $\eta_{i,k}$ are Gaussian". The uninformed reader might form the impression that it is just a matter of a few details to extend the Gaussian results to the Bernoulli case. This is not true, in particular for the Shcherbina-Tirozzi model.

Research Problem 8.5.1. (Level 2) Prove that Proposition 3.1.17 holds true with the Hamiltonian (8.105).

Rather than studying in detail the system with Hamiltonian (8.105), we will carry on the easier program of proving that Theorem 8.3.1, and therefore Theorem 8.4.1, remain true for this Hamiltonian (8.105). This still requires a rather significant amount of work.

The nature of this work is essentially to show that certain (large) families $u_\ell = \sum_{i\le N} a_{i,\ell}\eta_i$ of r.v.s behave somewhat as Gaussian families when the r.v.s η_i are Bernoulli and the coefficients $a_{i,\ell}$ are not too large. The general idea is very familiar, but the results required here are certainly not within the reach of standard tools of probability theory. (In particular because the number of variables (u_ℓ) grows exponentially with N.) A physicist might find that the work of the present section lacks glory (since there is no reason to doubt the result in the first place), but the mathematics involved are rather challenging. In the general direction of comparing large families of r.v.s of the type $u_\ell = \sum_{i\le N} a_{i,\ell}\eta_i$ with Gaussian families of r.v.s, some of the most important and natural questions have been open for a rather long time [107].

We define

$$p^b_{N,M} = \frac{1}{N}\mathsf{E}\log\int \exp(-H_{N,M}(\boldsymbol{\sigma}))\mathrm{d}\boldsymbol{\sigma}$$

where $H_{N,M}$ is the Hamiltonian (8.105) and $p^g_{N,M}$ is the corresponding quantity for the Hamiltonian with Gaussian r.v.s $g_{i,k}$ rather than $\eta_{i,k}$.

Theorem 8.5.2. *Assume that the function u is concave, and*

$$\forall \ell \le 4\,, \qquad |u^{(\ell)}| \le D\,. \tag{8.106}$$

Then

$$|p^b_{N,M} - p^g_{N,M}| \le \frac{M}{N^2} K(D,\kappa,h)\,. \tag{8.107}$$

As a consequence when we assume (8.106), Theorem 3.3.2 remains true for the Bernoulli model (here and in the rest of the section, we refer to the Hamiltonian (8.105) as the Bernoulli model, and to the Hamiltonian (8.1) as the Gaussian model), and we can prove this without studying the Bernoulli model in detail.

We could prove Theorem 8.5.2 by interpolation between the Bernoulli and the Gaussian models, but for the pleasure of change we will use a slightly different method, called Trotter's method in the West (it was invented to prove central limit theorems without using Fourier transform, and it seems to have been discovered much earlier by J. Lindeberg). The basic idea is that to compare

$$\mathsf{E}U(g_1,\dots,g_N) \quad \text{and} \quad \mathsf{E}U(\eta_1,\dots,\eta_N)$$

where $g_1,\dots,g_N$ are independent standard Gaussian r.v.s and $\eta_1,\dots,\eta_N$ are independent Bernoulli r.v.s, we replace the g_k by the η_k one at a time: that is, we define

$$U_k(x) = \mathsf{E}U(g_1,\dots,g_{k-1},x,\eta_{k+1},\dots,\eta_N)\ , \tag{8.108}$$

and we observe that

$$|\mathsf{E}U(g_1,\dots,g_N) - \mathsf{E}U(\eta_1,\dots,\eta_N)| \le \sum_{k\le N} |\mathsf{E}U_k(g_k) - \mathsf{E}U_k(\eta_k)|\ . \tag{8.109}$$

Now, if we define

$$R_k(x) = U_k(x) - U_k(0) - xU_k'(0) - \frac{x^2}{2}U_k''(0) - \frac{x^3}{3!}U_k^{(3)}(0)\ , \tag{8.110}$$

we have

$$\begin{aligned}|\mathsf{E}U_k(g_k) - \mathsf{E}U_k(\eta_k)| &= |\mathsf{E}R_k(g_k) - \mathsf{E}R_k(\eta_k)| \\ &\le \mathsf{E}|R_k(g_k)| + \mathsf{E}|R_k(\eta_k)|\ ,\end{aligned} \tag{8.111}$$

because $\mathsf{E}g_k^\ell = \mathsf{E}\eta_k^\ell$ for $\ell = 0,1,2,3$. Since R_k is a “fourth order term”, we expect the right-hand side of (8.109) to be small provided U does not depend too much on the k-th variable.

To implement the idea of (8.108) and (8.109) in the setting of Theorem 8.5.2, “we replace the variables $g_{i,k}$ by the variables $\eta_{i,k}$ one at a time”. By symmetry we may assume $i = N$, $k = M$. We consider Hamiltonians of the type

$$H_x(\boldsymbol{\sigma}) = \sum_{k\le M} u\Big(\frac{1}{\sqrt{N}}\sum_{i\le N}\xi_{i,k}\sigma_i\Big) - \kappa\|\boldsymbol{\sigma}\|^2 + h\sum_{i\le N} g_i\sigma_i\ , \tag{8.112}$$

where the $\xi_{i,k}$ for $(i,k) \ne (N,M)$ are independent r.v.s that are either Bernoulli or standard Gaussian, and where $\xi_{N,M} = x$. We consider the function

$$V(x) = \frac{1}{N}\mathsf{E}\log\int \exp(-H_x(\boldsymbol{\sigma}))\mathrm{d}\boldsymbol{\sigma}\ .$$

To prove Theorem 8.5.2 it suffices to establish a bound

$$|\mathsf{E}V(g) - \mathsf{E}V(\eta)| \le \frac{K(D)}{N^3}\ , \tag{8.113}$$

where g is a standard Gaussian r.v. and η is a Bernoulli r.v.s. Let us denote by $\langle\cdot\rangle_x$ an average for the Hamiltonian (8.112). We compute $V^{(4)}(x)$. “Each derivation brings out a factor $N^{-1/2}$". Using (8.106) and Hölder's inequality, it is straightforward to reach a bound

$$|V^{(4)}(x)| \le \frac{K(D)}{N^3}\mathsf{E}\langle\sigma_N^4\rangle_x \ . \tag{8.114}$$

The difficulty is to control the last term. For this we must revisit the proof of Lemma 3.2.6. We write

$$S_k = \frac{1}{\sqrt{N}}\sum_{i\le N}\xi_{i,k}\sigma_i \ .$$

We recall that $\boldsymbol{\rho} = (\sigma_1, \dots, \sigma_{N-1})$, and we consider, as in (3.85), the function

$$f(\sigma_N) = \log\int\exp\biggl(\sum_{k\le M}u(S_k) - \kappa\sum_{i\le N-1}\sigma_i^2 + h\sum_{i\le N-1}g_i\sigma_i\biggr)\mathrm{d}\boldsymbol{\rho} \ .$$

Proceeding as in (3.87) it suffices to control $\mathsf{E}\langle\sigma_N\rangle_x^4$, and proceeding as in (3.86) shows that

$$\mathsf{E}\langle\sigma_N\rangle_x^4 \le K(\tau,h)(1+\mathsf{E}f'(0)^4) \ .$$

Now

$$f'(0) = \frac{1}{\sqrt{N}}\sum_{k\le M}\xi_{N,k}\langle u'(S_{k,0})\rangle \ ,$$

where $\langle u'(S_{k,0})\rangle$ are certain Gibbs averages. We recall that $\xi_{N,M} = x$, while for $(i,k) \ne (N,M)$ the r.v $\xi_{i,k}$ is either standard Gaussian or Bernoulli. We gather those of these variables that are Bernoulli, and we use the subgaussian inequality (A.17) to obtain that $\mathsf{E}f'(0)^4 \le K(D)(1+x^4/N^2)$, and finally from (8.114) that

$$|V^{(4)}(x)| \le \frac{K(D)}{N^3}\left(1+\frac{x^4}{N^2}\right) \ .$$

Let us define

$$R(x) = V(x) - V(0) - xV'(0) - \frac{x^2}{2}V''(0) - \frac{x^3}{3!}V^{(3)}(0) \ .$$

Proceeding as in (8.110) we get

$$|\mathsf{E}V(g) - \mathsf{E}V(\eta)| \le \mathsf{E}|R(g)| + \mathsf{E}|R(\eta)| \ .$$

Using the bound

$$|R(x)| \le \sup_{|y|\le|x|}|V^{(4)}(y)| \ ,$$

we then obtain (8.113) and conclude the proof. □

The rest of this section is devoted to adapt the arguments of the previous three sections to prove how to deduce Theorem 8.3.1 from Theorem 8.5.2

in the case of the Bernoulli model. Most of the work is devoted to find an appropriate substitute for Theorem 8.2.7. The mathematics are beautiful, in particular in Proposition 8.5.3 below. Still, this is somewhat special material, reserved for the reader who has already mastered the previous three sections.

An obvious idea is that a r.v. of the type $\sum_{i\le N} a_i\eta_i$ is more likely to look like a Gaussian r.v. when the coefficients a_i are all small. Given a number $\overline{a}>0$ (that remains implicit in the notation) we consider the truncation function

$$\psi(x)=\psi_{\overline{a}}(x):=\max(-\overline{a},\min(x,\overline{a}))\ , \tag{8.115}$$

so that $|\psi(x)|\le\overline{a}$ and

$$|\psi(x)-\psi(y)|\le|x-y|\ . \tag{8.116}$$

We recall the function $F(\mathbf{x})=\log\sum_{\ell\le n}\exp sx_\ell$ of (8.4).

Proposition 8.5.3. *Consider numbers $(a_{i,\ell})_{i\le N,\ell\le n}$ and consider the r.v.s $u_\ell=\sum_{i\le N} a_{i,\ell}\eta_i$, where $(\eta_i)_{i\le N}$ are independent Bernoulli r.v.s. Let $u'_\ell=\sum_{i\le N}\psi(a_{i,\ell})\eta_i$. Then*

$$\mathsf{E}F(\mathbf{u})\ge\mathsf{E}F(\mathbf{u}')\ . \tag{8.117}$$

It will be easier to work with the family $\mathbf{u}'$ than with $\mathbf{u}$ because all the coefficients $\psi(a_{i,\ell})$ are bounded by $\overline{a}$, that will be chosen small.

Proof. Let us fix the numbers $(\eta_i)_{i<N}$, and let us define

$$w_\ell=\exp\left(s\sum_{i<N}a_{i,\ell}\eta_i\right)\ .$$

We will prove that

$$\frac{1}{2}\log\left(\sum_{\ell\le n}w_\ell\exp sa_{N,\ell}\right)+\frac{1}{2}\log\left(\sum_{\ell\le n}w_\ell\exp(-sa_{N,\ell})\right)$$
$$\ge\frac{1}{2}\log\left(\sum_{\ell\le n}w_\ell\exp s\psi(a_{N,\ell})\right)+\frac{1}{2}\log\left(\sum_{\ell\le n}w_\ell\exp(-s\psi(a_{N,\ell}))\right)\ .$$

Taking expectation in the r.v.s $(\eta_i)_{i\le N}$ proves that $\mathsf{E}F(\mathbf{u})$ decreases if we replace the numbers $a_{N,\ell}$ by $\psi(a_{N,\ell})$ and leave the numbers $a_{i,\ell}$ $(i<N)$ unchanged. Iterating the process yields (8.117).

Writing a_ℓ rather than $a_{N,\ell}$, we want to prove that

$$\sum_{\ell\le n}w_\ell\exp sa_\ell\sum_{\ell\le n}w_\ell\exp(-sa_\ell)\ge\sum_{\ell\le n}w_\ell\exp s\psi(a_\ell)\sum_{\ell\le n}w_\ell\exp(-s\psi(a_\ell))\ . \tag{8.118}$$

We will prove that for any r.v. X we have

$$\mathsf{E}\exp sX\mathsf{E}\exp(-sX) \geq \mathsf{E}\exp s\psi(X)\mathsf{E}\exp(-s\psi(X)) . \tag{8.119}$$

Now (8.119) implies (8.118) by considering the probability on $\{1, \cdots, n\}$ that gives a mass proportional to w_ℓ to ℓ and the r.v. X such that $X(\ell) = a_\ell$.

To prove (8.119) let us consider an independent copy $X^\sim$ of X, so that, since the r.v. $X - X^\sim$ is symmetric,

$$\mathsf{E}\exp sX\mathsf{E}\exp(-sX) = \mathsf{E}\exp s(X - X^\sim) = \mathsf{E}\,\mathrm{ch}(X - X^\sim) = \mathsf{E}\,\mathrm{ch}|X - X_s| ,$$

and, using this for $\psi(X)$ rather than X, we get

$$\mathsf{E}\exp s\psi(X)\mathsf{E}\exp(-s\psi(X)) = \mathsf{E}\exp|\psi(X) - \psi(X^\sim)| ,$$

and (8.119) follows from (8.116). □

The following should be compared to Proposition 8.2.2.

Proposition 8.5.4. *Given $\bar{b} > 0$ consider for $\ell \leq n$, $i \leq N$ consider numbers $b_{i,\ell}$ such that $|b_{i,\ell}| \leq \bar{b}$. Define the r.v.s*

$$u_\ell = \sum_{i \leq N} b_{i,\ell}\eta_i ,$$

where $(\eta_i)_{i \leq N}$ are independent Bernoulli r.v.s. Consider a jointly Gaussian family $(v_\ell)_{\ell \leq n}$, and assume that

$$\forall \ell ,\ \sum_{i \leq N} b_{i,\ell}^2 \geq \mathsf{E}v_\ell^2 ; \qquad \forall \ell \neq \ell' ,\ \sum_{i \leq N} b_{i,\ell}b_{i,\ell'} \leq \mathsf{E}v_\ell v_{\ell'} . \tag{8.120}$$

Then, if $\mathbf{u} = (u_\ell)_{\ell \leq n}$ and $\mathbf{v} = (v_\ell)_{\ell \leq n}$, we have

$$\mathsf{E}F(\mathbf{u}) \geq \mathsf{E}F(\mathbf{v}) - LN(s\bar{b})^4 . \tag{8.121}$$

Proof. We set $\mathbf{u}(t) = \sqrt{t}\mathbf{u} + \sqrt{1-t}\mathbf{v}$ and

$$\varphi(t) = \mathsf{E}F(\mathbf{u}(t))$$

so that $\varphi(1) = \mathsf{E}F(\mathbf{u})$ and $\varphi(0) = \mathsf{E}F(\mathbf{v})$, and

$$\varphi'(t) = \frac{1}{2}\sum_{\ell \leq n} \mathsf{E}\left(\frac{u_\ell}{\sqrt{t}} - \frac{v_\ell}{\sqrt{1-t}}\right)\frac{\partial F}{\partial x_\ell}(\mathbf{u}(t)) .$$

We compute this quantity by integrating by parts in v_ℓ, and by using the "approximate integration by parts" of (4.197) for the variables η_i. The "main terms" resulting of this integration by parts are the same as if the r.v.s η_i were Gaussian, and the contribution to $\varphi'(t)$ of these terms and of the terms coming from v_ℓ is ≥ 0 (which is what the proof of Proposition 8.2.2 shows).

The issue is to control the error terms. Let us fix $i \le N$, all the r.v.s η_j, $j \ne i$ and the r.v.s v_ℓ, $\ell \le n$. Consider the function $v(x)$ given by

$$v(x) = \frac{1}{2\sqrt{t}} \sum_{\ell \le n} b_{i,\ell} \frac{\partial F}{\partial x_\ell}(\mathbf{u}(t,x)) \ ,$$

where $\mathbf{u}(t,x)$ means that in the expression $\mathbf{u}(t)$ every occurrence of η_i has been replaced by x. Approximate integration by parts in η_i is the evaluation of $\mathsf{E}\eta_i v(\eta_i)$, and the error term is bounded by

$$\mathsf{E} \sup_{|x| \le 1} |v^{(3)}(x)| \ .$$

Thus it suffices to prove that $\sup_{|x|\le 1} |v^{(3)}(x)| \le L(s\overline{b})^4$, where $\overline{b} = \sup |b_{i,\ell}|$. Computing $v^{(3)}$ makes it obvious that

$$|v^{(3)}(x)| \le \overline{b}^4 \sum_{\ell_1,\ell_2,\ell_3,\ell_4 \le n} \left| \frac{\partial^4 F}{\partial x_{\ell_1} \partial x_{\ell_2} \partial x_{\ell_3} \partial x_{\ell_4}}(\mathbf{u}(t,x)) \right| \ ,$$

and the special form of F ensures that the quadruple sum is bounded by Ls^4. (To understand why, the reader should first check that $\sum_{\ell_1,\ell_2} |\frac{\partial^2 F}{\partial x_{\ell_1} \partial x_{\ell_2}}(\mathbf{x})| \le Ls^2$.) □

Proposition 8.5.5. *Given $d > 0$, there exists a constant $K = K(d)$ with the following property. Consider numbers $0 \le c < b \le d$, $\overline{a} \le d$, with $1/(b-c) \le d$, and the function ψ of (8.115). Consider numbers $(a_{i,\ell})_{i \le N, \ell \le n}$ and assume that*

$$\forall \ell \le n \ , \ \ Nb \le \sum_{i \le N} \psi(a_{i,\ell})^2 \le \sum_{i \le N} a_{i,\ell}^2 \le dN \tag{8.122}$$

$$\forall \ell \ne \ell' \ , \ \ \sum_{i \le N} \psi(a_{i,\ell}) \psi(a_{i,\ell'}) \le cN \ . \tag{8.123}$$

Consider the r.v.s $u_\ell = \sum_{i \le N} a_{i,\ell} \eta_i / \sqrt{N}$, where $(\eta_i)_{i \le N}$ are independent Bernoulli r.v.s. Then if

$$K \le s \le \frac{\sqrt{\log n}}{K} \ ; \qquad Ks^2 \le N \tag{8.124}$$

we have

$$\mathsf{P}\left(\mathrm{card} \left\{ \ell \le n \ ; \ u_\ell \ge \frac{s}{K} \right\} \le n \exp(-Ks^2) \right) \le K \exp\left(-\frac{s^2}{K} \right) \ . \tag{8.125}$$

Proof. Let $u'_\ell = \sum_{i \le N} \psi(a_{i,\ell}) \eta_i / \sqrt{N}$, so that $\mathsf{E}F(\mathbf{u}) \ge \mathsf{E}F(\mathbf{u}')$ by Proposition 8.5.3. Consider a Gaussian family $\mathbf{v}' = (v_\ell)_{\ell \le n}$ such that $v_\ell =$

$z\sqrt{q}+\sqrt{1-q}\xi^\ell$, where as usual z and ξ^ℓ are independent standard Gaussian r.v.s. We apply Proposition 8.5.4 with $b_{i,\ell} = \psi(a_{i,\ell})/\sqrt{N}$ and $\overline{b} = d/\sqrt{N}$ to see that $\mathsf{E}F(\mathbf{u}') \geq \mathsf{E}F(\mathbf{v}') - Ls^4d^4/N$. Lemma 8.2.3 implies that if $K \leq s \leq \sqrt{\log n}/K$, where K depends on d only, then

$$\mathsf{E}F(\mathbf{v}') \geq \log n + \frac{s^2}{K} \; .$$

Combining these inequalities, we see that

$$\mathsf{E}F(\mathbf{u}) \geq \log n + \frac{s^2}{K}$$

provided $K \leq s \leq \sqrt{\log n}/K$ and $Ks^2 \leq N$, where K depends only on d.

If we think of $F(\mathbf{u})$ as a function of the real numbers $(\eta_i)_{i\leq N}$, it has two fundamental properties. First the condition $\sum_{i\leq N} a_{i,\ell}^2 \leq dN$ implies (as in the case of Theorem 8.2.4) that its Lipschitz constant is $\leq ds$. Second, for any number a the set of (η_i) for which $F(\mathbf{u}) \leq a$ is convex. Thus, according to the principle explained in Section 6 of [91], the inequality (8.14) remains valid. The remainder is as in Theorem 8.2.4. □

Proposition 8.5.6. *Under the conditions of Proposition 8.5.5, assume that we are moreover given numbers $(b_i)_{i\leq N}$ with $\sum_{i\leq N} b_i^2 \leq Nd$, and consider the r.v. $u_0 = \sum_{i\leq N} b_i\eta_i/\sqrt{N}$. Then the constant K can be chosen so that under (8.124) we have*

$$\mathsf{P}\left(\mathrm{card}\left\{\ell \leq n\,;\; u_\ell + u_0 \geq \frac{s}{K}\right\} \leq n\exp(-Ks^2)\right) \leq K\exp\left(-\frac{s^2}{K}\right) \; . \tag{8.126}$$

Consequently, for

$$\varepsilon \leq \frac{1}{K}\exp(-K\tau^2)\;;\quad \varepsilon \geq n^{-1/K}\;;\quad \varepsilon \geq \exp\left(-\frac{N}{K}\right) \tag{8.127}$$

we have

$$\mathsf{P}\left(\mathrm{card}\left\{\ell \leq n\,;\; u_\ell + u_0 \geq \frac{s}{K}\right\} \leq \varepsilon n\right) \leq \varepsilon^{1/K} \; . \tag{8.128}$$

Proof. Denoting by K_0 the constant of (8.125), we consider the events

$$\Omega_1 : \mathrm{card}\left\{\ell \leq n\,;\; u_\ell \geq \frac{2s}{K_0}\right\} \leq n\exp(-4K_0s^2)$$

$$\Omega_2 : u_0 \geq -\frac{s}{K_0}$$

$$\Omega \;\; : \mathrm{card}\left\{\ell \leq n\,;\; u_\ell + u_0 \geq \frac{s}{K_0}\right\} \leq n\exp(-4K_0s^2) \; .$$

We will prove that

$$\Omega \subset \Omega_1 \cup \Omega_2^c \; . \tag{8.129}$$

This proves (8.128) since the subgaussian inequality (A.16) implies $\mathsf{P}(\Omega_2^c) = \mathsf{P}(u_0 \leq -s/K_0) \leq \exp(-s^2/2d^2K_0^2)$, and because we control $\mathsf{P}(\Omega_1)$ by (8.125) used for $2s$ rather than s. To prove (8.129) we simply observe that if $u_0 \geq -s/K_0$ we have

$$u_\ell \geq \frac{2s}{K_0} \quad \Rightarrow \quad u_\ell + u_0 \geq \frac{s}{K_0} \; ,$$

and thus

$$\mathrm{card}\left\{\ell \leq n \; ; \; u_\ell \geq \frac{2s}{K_0}\right\} \leq \mathrm{card}\left\{\ell \leq n \; ; \; u_\ell + u_0 \geq \frac{s}{K_0}\right\} \; . \qquad \square$$

Proposition 8.5.7. *Given a number d, there exists a constant $K = K(d)$ with the following properties. Consider numbers $0 \leq c < b \leq d$, $d \geq 1/(b-c)$, $\overline{a} \leq d$. Consider a probability measure G on $\mathbb{R}^N$, and assume that for a certain $\mathbf{b} = (b_i)_{i \leq N} \in \mathbb{R}^N$, the following occurs, where ψ is the function (8.115):*

$$G\biggl(\biggl\{\boldsymbol{\sigma} \; ; \; \sum_{i \leq N} \psi(\sigma_i - b_i)^2 \geq bN\biggr\}\biggr) \geq 1 - \exp\left(-\frac{N}{d}\right) \tag{8.130}$$

$$G\left(\{\boldsymbol{\sigma} \; ; \; \|\boldsymbol{\sigma}\|^2 \leq dN\}\right) \geq 1 - \exp\left(-\frac{N}{d}\right) \tag{8.131}$$

$$G^{\otimes 2}\biggl(\biggl\{(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) \; ; \; \sum_{i \leq N} \psi(\sigma_i^1 - b_i)\psi(\sigma_i^2 - b_i) \leq cN\biggr\}\biggr) \tag{8.132}$$

$$\geq 1 - \exp\left(-\frac{N}{d}\right) \; .$$

Then, for any $\tau \geq 0$, we have

$$K \exp\left(-\frac{N}{K}\right) \leq \varepsilon \leq \frac{1}{K}\exp(-K\tau^2)$$
$$\Rightarrow \mathsf{P}\biggl(G\biggl(\biggl\{\boldsymbol{\sigma} \; ; \; \frac{1}{\sqrt{N}} \sum_{i \leq N} \sigma_i \eta_i \geq \tau\biggr\}\biggr) \leq \varepsilon\biggr) \leq K\varepsilon^{1/K} \; . \tag{8.133}$$

Proof. We follow the proof of Proposition 8.2.6, defining now

$$Q_n = \biggl\{(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n) \; ; \; \forall \ell \leq n \; , \; \|\boldsymbol{\sigma}^\ell\|^2 \leq dN \; , \; bN \leq \sum_{i \leq N} \psi(\sigma_i^\ell - b_i)^2 \; ;$$
$$\forall \ell \neq \ell' \; , \; \sum_{i \leq N} \psi(\sigma_i^\ell - b_i)\psi(\sigma_i^{\ell'} - b_i) \leq cN\biggr\} \; ,$$

and using now (8.128) for $u_\ell = \sum_{i\le N}\eta_i(\sigma_i^\ell - b_i)/\sqrt{N}$, so that $u_0 + u_\ell = \sum_{i\le N}\eta_i\sigma_i^\ell/\sqrt{N}$. Since we consider only values of $n \le \exp(N/K)$, and of $\varepsilon \ge n^{-1/K}$, the restriction $\varepsilon \ge \exp(-N/K)$ of (8.128) is immaterial. □

Before we move on, we need a simple lemma.

Lemma 8.5.8. *Consider a probability measure μ on $\mathbb{R}$ with density proportional to $\exp w(x)$ with respect to Lebesgue measure, where w is a concave function with $w''(x) \le -c$, where $c > 0$. (The case where w is finite only on an interval and is $-\infty$ elsewhere is permitted.) Assume that $\int x\mathrm{d}\mu(x) = 0$. Then, if $\overline{a} \ge L/\sqrt{c}$ and if X is a r.v. of law μ,*

$$\mathsf{E}(\psi(X) - \mathsf{E}\psi(X))^2 \ge \frac{1}{L}\mathsf{E}(X - \mathsf{E}X)^2 = \frac{1}{L}\mathsf{E}X^2 \tag{8.134}$$

where ψ is given by (8.115).

In words, we can witness a proportion of the variance of X by looking at $\psi(X)$ instead of X. It is because of the requirement $\int x\mathrm{d}\mu(x) = 0$ that the vector $\mathbf{b}$ is needed in Proposition 8.5.7. This Proposition will be used with $\mathbf{b} = \langle\boldsymbol{\sigma}\rangle$, and (8.134) will be used to prove that

$$\left\langle\big(\psi(\sigma_i - b_i) - \langle\psi(\sigma_i - b_i)\rangle\big)^2\right\rangle \ge \frac{1}{L}\langle(\sigma_i - \langle\sigma_i\rangle)^2\rangle\ .$$

Proof. Consider the unique point x^* where $w(x^*)$ is maximum; without loss of generality we assume that $w(x^*) = 0$, so that $w(x) \le 0$ for all x. Define now $x_1 < x^* < x_2$ by $w(x_1) = w(x_2) = -1$; thus

$$\int \exp w(x)\mathrm{d}x \ge \int_{x_1}^{x_2} \exp w(x)\mathrm{d}x \ge \frac{1}{e}(x_2 - x_1)\ . \tag{8.135}$$

We prove that

$$\int (x - x^*)^2 \exp w(x)\mathrm{d}x \le L(x_2 - x_1)^3\ . \tag{8.136}$$

Indeed, by concavity of w, for $x \ge x_2$ we have $w(x) \le -(x - x^*)/(x_2 - x^*)$, and thus, since $w(x) \le 0$,

$$\begin{aligned}
&\int_{x^*}^{\infty} (x - x^*)^2 \exp w(x)\mathrm{d}x \\
&\le \int_{x^*}^{x_2} (x - x^*)^2 \exp w(x)\mathrm{d}x + \int_{x_2}^{\infty} (x - x^*)^2 \exp w(x)\mathrm{d}x \\
&\le (x_2 - x^*)^3 + \int_{x_2}^{\infty} (x - x^*)^2 \exp\left(-\frac{x - x^*}{x_2 - x^*}\right)\mathrm{d}x \\
&\le L(x_2 - x^*)^3\ ,
\end{aligned}$$

and proceeding similarly for the other half proves (8.136). Comparing (8.136) and (8.135) we get

$$\int (x - x^*)^2 \mathrm{d}\mu(x) \le L(x_2 - x_1)^2 \le \frac{L}{c} \tag{8.137}$$

because $(x_2 - x_1)^2 \le L/c$ since $w''(x) \le -c$. Since $\int x \mathrm{d}\mu(x) = 0$, we have

$$x^{*2} = \left(\int (x - x^*) \mathrm{d}\mu(x) \right)^2 \le \int (x - x^*)^2 \mathrm{d}\mu(x) \le \frac{L}{c} ,$$

so that $|x^*| \le L/\sqrt{c}$ and since $x_1 \le x^* \le x_2$ and $x_2 - x_1 \le L/\sqrt{c}$ we have $|x_1|, |x_2| \le L/\sqrt{c}$. Therefore $|x_1|, |x_2| \le \overline{a}$ provided $\overline{a} \ge L/\sqrt{c}$. Next, we prove that when $|x_1|, |x_2| \le \overline{a}$, then

$$\mathsf{E}(\psi(X) - \mathsf{E}\psi(X))^2 \ge \frac{1}{L}(x_2 - x_1)^2 . \tag{8.138}$$

First we notice that, as in (8.136), we have

$$\int \exp w(x) \mathrm{d}x \le L(x_2 - x_1) .$$

Also, for $[a, b] \subset [x_1, x_2]$, and since $w(x) \ge -1$ on $[a, b]$, it holds

$$\int_a^b \exp w(x) \mathrm{d}(x) \ge \frac{1}{e}(b - a) ,$$

and therefore

$$\mu([a, b]) = \frac{\int_a^b \exp w(x) \mathrm{d}x}{\int \exp w(x) \mathrm{d}x} \ge \frac{b - a}{L(x_2 - x_1)} ,$$

which implies (8.138). Finally, (8.137) implies that

$$\mathsf{E}(X - \mathsf{E}X)^2 \le \mathsf{E}(X - x^*)^2 \le L(x_2 - x_1)^2 ,$$

and comparing with (8.138) completes the argument. □

Theorem 8.5.9. *Consider a concave function $U(\boldsymbol{\sigma})$ on $\mathbb{R}^N$ with $U \le 0$, numbers $0 < \kappa_0 < \kappa_1$, numbers $(a_i)_{i \le N}$ and a convex set C of $\mathbb{R}^N$. Consider κ with $\kappa_0 < \kappa < \kappa_1$ and the probability measure G_C on $\mathbb{R}^N$ given by*

$$\forall B , \quad G_C(B) = \frac{1}{Z} \int_{B \cap C} \exp\Big(U(\boldsymbol{\sigma}) - \kappa \|\boldsymbol{\sigma}\|^2 + \sum_{i \le N} a_i \sigma_i \Big) \mathrm{d}\boldsymbol{\sigma} ,$$

where Z is the normalizing factor. Assume that for a certain number a we have

$$Z \ge \exp(-Na) ; \qquad \sum_{i \le N} a_i^2 \le Na^2 . \tag{8.139}$$

Then we may find a number K, depending only on a, κ_0 and κ_1 such that G_C satisfies (8.133).

Proof. We shall show that conditions (8.130) to (8.133) are satisfied for numbers $\overline{a}$, b, c with $c \le b \le d$, $\overline{a} \le d$, $1/(b-c) \le d$, $\mathbf{b} = \int \boldsymbol{\sigma} \mathrm{d}G_C(\boldsymbol{\sigma})$, where d depends only on a, κ_0 and κ_1. Define

$$T_{1,1}(\boldsymbol{\sigma}) = \frac{1}{N} \sum_{i \le N} \psi(\sigma_i - b_i)^2 \; ; \quad T_{1,2}(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) = \frac{1}{N} \sum_{i \le N} \psi(\sigma_i^1 - b_i)\psi(\sigma_i^2 - b_i) \; , \tag{8.140}$$

so that, denoting by $\langle \cdot \rangle$ an average for G_C or its products, we have

$$\langle T_{1,1}(\boldsymbol{\sigma}) \rangle - \langle T_{1,2}(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) \rangle = \frac{1}{2N} \left\langle \sum_{i \le N} (\psi(\sigma_i - b_i) - \langle \psi(\sigma_i - b_i) \rangle)^2 \right\rangle . \tag{8.141}$$

We first prove that

$$\langle T_{1,1}(\boldsymbol{\sigma}) \rangle - \langle T_{1,2}(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) \rangle \ge \frac{1}{K} \; , \tag{8.142}$$

where K depends only on a, κ_1, κ_2. In the proof of Theorem 8.2.7 we have shown that

$$\langle R_{1,1} \rangle - \langle R_{1,2} \rangle = \frac{1}{2N} \sum_{i \le N} \langle (\sigma_i - \langle \sigma_i \rangle)^2 \rangle \ge \frac{1}{K} \; . \tag{8.143}$$

In Lemma 3.2.5 we have shown that the law of $\sigma_i - b_i$ under G_C is of the type considered in Lemma 8.5.8, with $c = 2\kappa$. Thus, if $\overline{a} \ge L/\sqrt{\kappa_1}$, we deduce from (8.134) that

$$\left\langle \big(\psi(\sigma_i - b_i) - \langle \psi(\sigma_i - b_i) \rangle \big)^2 \right\rangle \ge \frac{1}{K} \langle (\sigma_i - \langle \sigma_i \rangle)^2 \rangle$$

and therefore (8.141) and (8.143) imply (8.142).

It is very simple to see that the Lipschitz constant of the functions (8.140) is $\le L\overline{a}/\sqrt{N}$. Thus (3.16) yields

$$\int \exp \frac{N\kappa}{L\overline{a}^2} (T_{1,1}(\boldsymbol{\sigma}) - \langle T_{1,1} \rangle)^2 \mathrm{d}G_C(\boldsymbol{\sigma}) \le 4 \; ,$$

and therefore

$$G_C(\{\boldsymbol{\sigma} \; ; \; |T_{1,1}(\boldsymbol{\sigma}) - \langle T_{1,1} \rangle| \ge t\}) \le L \exp\left(-\frac{Nt^2}{K\overline{a}^2} \right) \; .$$

We observe that moreover the measure $G_C^{\otimes 2}$ on $\mathbb{R}^{2N}$ corresponds to the Hamiltonian $H_{N,M}(\boldsymbol{\sigma}^1) + H_{N,M}(\boldsymbol{\sigma}^2)$, which satisfies (3.4) on $\mathbb{R}^{2N}$. Therefore as previously (3.16) implies that

$$G_C^{\otimes 2}(\{(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) \; ; \; |T_{1,2}(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) - \langle T_{1,2} \rangle| \ge t\}) \le L \exp\left(-\frac{Nt^2}{K\overline{a}^2} \right) \; .$$

We then continue as in the proof of Theorem 8.2.7. □

Even though it was harder to prove, Theorem 8.5.9 is a perfect substitute for Theorem 8.2.7. We leave to the reader the task to prove that, once we have obtained Theorem 8.5.9, the proof of Theorem 8.3.1 carries through with essentially no changes in the case of the Bernoulli model, with the exception of the proof of Proposition 8.3.9 from (8.68) on. The difference there is that $S_m = N^{-1/2}\sum_{i\le N}\eta_{i,m}\sigma_i$ is no longer Gaussian. What is required to make the proof work is to show that given ε, we can find $\varepsilon' > 0$ and N_0 such that for $N \ge N_0$ and each m we have

$$\langle \mathsf{E}_m \mathbf{1}_{\{\tau-\varepsilon' \le S_m \le \tau\}} \rangle \le \varepsilon \; .$$

First, we prove that given ε' there is N_0 such that if $N \ge N_0$, given numbers a_i with $|a_i| \le N^{1/4}$, then for each number x,

$$\mathsf{P}\left(\sum_{i\le N}\frac{a_i\eta_i}{\sqrt{N}} \in [x-\varepsilon', x]\right) \le L\varepsilon'\sqrt{\frac{N}{\sum_{i\le N} a_i^2}} + 2\varepsilon' \; . \tag{8.144}$$

Since the left-hand side is ≤ 1, it suffices to consider the case where $\sum_{i\le N} a_i^2 \ge \varepsilon'^2 N$. As we detail now, this statement follows from the one-dimensional central limit theorem. Indeed, if $c = \sum_{i\le N} a_i^2/N$, the r.v. $X = c^{-1/2}\sum_{i\le N} a_i\eta_i/\sqrt{N}$ satisfies $\mathsf{E}X^2 = 1$, and since $c^{-1/2}|a_i|/\sqrt{N} \le N^{-1/4}/\varepsilon'$, for large N we have for each x that

$$|\mathsf{P}(X \le x) - \mathsf{P}(g \le x)| \le \varepsilon' \; ,$$

where g is standard Gaussian. Therefore $|\mathsf{P}(cX \le x) - \mathsf{P}(cg \le x)| \le \varepsilon'$, and by (8.69) it holds

$$\begin{aligned}\mathsf{P}(cX \in [x-\varepsilon', x]) &\le \mathsf{P}(cg \in [x-\varepsilon', x]) + 2\varepsilon' \\ &\le \frac{L\varepsilon'}{c} + 2\varepsilon' \le L\varepsilon'\sqrt{\frac{N}{\sum_{i\le N} a_i^2}} + 2\varepsilon' \; .\end{aligned}$$

Next, let

$$S'_m = \frac{1}{\sqrt{N}}\sum_{|\sigma_i|\le N^{1/4}} \sigma_i\eta_i \; ,$$

where the summation is taken over all i for which $|\sigma_i| \le N^{1/4}$, and let $S''_m = S_m - S'_m$ be the sum of the other terms. Observe that S'_m and S''_m are independent r.v.s. Since

$$\tau - \varepsilon' \le S_m \le \tau \quad \Rightarrow \quad S'_m \in [\tau - S''_m - \varepsilon', \tau - S''_m] \; ,$$

we deduce from (8.144) that for $N \ge N_0$ we have

$$\mathsf{E}_m \mathbf{1}_{\{\tau-\varepsilon' \le S_m \le \tau\}} \le L\varepsilon' \frac{\sqrt{N}}{a(\boldsymbol{\sigma})} + 2\varepsilon',$$

where $a(\boldsymbol{\sigma})^2 = \sum_{i \le N} \sigma_i^2 \mathbf{1}_{\{|\sigma_i| \le N^{1/4}\}}$. Continuing as in the proof of Proposition 8.3.9, it suffices to prove that if ε' is small enough we have

$$\int_{a(\boldsymbol{\sigma}) \le \sqrt{\varepsilon' N}} \exp(-\kappa \|\boldsymbol{\sigma}\|^2) \mathrm{d}\boldsymbol{\sigma} \le \varepsilon' \exp(-Na) \; .$$

This uses standard methods: Denoting by γ is the probability with density proportional to $\exp(-\kappa\|\boldsymbol{\sigma}\|^2)$, we bound $\gamma(\{a(\boldsymbol{\sigma}) \le x\})$ by

$$\inf_\lambda \exp \lambda x \int \exp(-\lambda a(\boldsymbol{\sigma})) \mathrm{d}\gamma(\boldsymbol{\sigma}) \; ;$$

the computations are a bit tedious.

9. The Gardner Formula for the Discrete Cube

9.1 Overview

This chapter continues the work of Chapter 2. We study the Hamiltonian

$$-H_{N,M}(\boldsymbol{\sigma}) = \sum_{k\le M} u(S_k)\ ; \quad S_k = S_k(\boldsymbol{\sigma}) = \frac{1}{\sqrt{N}}\sum_{i\le N} g_{i,k}\sigma_i\ . \tag{9.1}$$

We are concerned mainly with the case where $\exp u(x)$ is nearly $\mathbf{1}_{\{x\ge\tau\}}$ for a certain number τ (which is fixed once and for all). We assume the following

$$u \le 0\ ; \quad x \ge \tau \quad \Rightarrow \quad u(x) = 0\ . \tag{9.2}$$

Since it is very desirable that u be differentiable, we assume that $u^{(3)}$ exists, and that for a certain number D

$$1 \le \ell \le 3 \quad \Rightarrow \quad |u^{(\ell)}| \le D\ . \tag{9.3}$$

The difference between the work of Chapter 2 and the work we are going to present is that the dependence on D of our estimates will be much weaker; every occurrence of D in the estimates will now be multiplied by an exponentially small factor $\exp(-N/L)$. This will allow to have D depend on N. The overall content of the present chapter is that there exists $\alpha(\tau) > 0$ such that if $M/N \le \alpha(\tau)$ and (9.3) holds for $D = \exp(N/L)$ then we understand very well the system governed by the Hamiltonian (9.1). The very weak requirement (9.3) for $D = \exp(N/L)$ allows to find (given N and M) a function u satisfying this requirement and for which $\exp u(x)$ is a very good approximation of $\mathbf{1}_{\{x\ge\tau\}}$. It is worth repeating this. We will approximate the function $\mathbf{1}_{\{x\ge\tau\}}$ by a function u which varies with N. What makes the argument work is that condition (9.3) for $D = \exp(N/L)$ becomes very weak for large N.

At this point it is probably wise to make explicit a rather important difference between the way we look at spin glasses and traditional statistical mechanics. In spin glasses, there is no "limiting system" as $N \to \infty$, and the object under study is really the system considered for a given large value of N. With this in mind, it is quite natural to try to approximate the function $\mathbf{1}_{\{x\ge\tau\}}$ by a function u that depends on the situation under study, i.e. on N.

M. Talagrand, *Mean Field Models for Spin Glasses*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys in Mathematics 55, DOI 10.1007/978-3-642-22253-5_2,

The central difference between the situation of Chapter 8 and the present situation is that we no longer have a magic proof of the fact that $R_{1,2} \simeq \langle R_{1,2} \rangle$, and we will have to work very hard to prove that $R_{1,2} \simeq$ Const. (On the other hand, the fact that the spins are bounded removes several minor - yet irritating - sources of complications.)

The overall approach is the same as in Chapter 2, and it would be very helpful for the reader to have Sections 2.2 and 2.3 at hand while proceeding. We use the smart path of Section 2.1, and we attempt to show that the terms I and II of Proposition 2.2.2 nearly cancel out. This is done though the "cavity in M" method of Section 2.3; what we need is a better estimate than Lemma 2.3.2 provides. In the remainder of this section we try to outline the general strategy that will achieve this. Since we describe the overall structure of the approach, we do not recall the definitions of the various quantities involved in complete detail, as these details are irrelevant now and will be given in due time. For the time being, we recall that the average $\langle \cdot \rangle_{t,\sim}$ corresponds to the Hamiltonian (2.30) (i.e. when M has been replaced by $M-1$), while $\nu_{t,v}$ is given by the formula

$$\nu_{t,v}(f) = \mathsf{E} \frac{\left\langle f \exp \sum_{\ell \le n} u(S_v^\ell) \right\rangle_{t,\sim}}{\langle \mathsf{E}_\xi \exp u(S_v^1) \rangle_{t,\sim}^n} ,$$

where E_ξ denotes expectation "in all the r.v.s labeled ξ". At first sight the above formula differs from the formula (2.35). This is simply because in (2.35) we made the convention that the expectation E_ξ is built-in the bracket $\langle \cdot \rangle_{t,\sim}$, while in the present chapter we find it more economical to write explicitly this expectation instead of constantly reminding the reader of this convention.

Given a function f on Σ_N^4, and $B_v \equiv 1$ or $B_v \equiv u'(S_v^1)u'(S_v^2)$, we want to bound $\frac{\mathrm{d}}{\mathrm{d}v} \nu_{t,v}(B_v f)$. After differentiation and integration by parts, this quantity is a sum of terms of the type

$$\nu_{t,v}(f(R_{1,2}^t - q)A) \tag{9.4}$$

where A is a monomial in the quantities $u'(S_v^\ell)$, $u''(S_v^\ell)$, $u'''(S_v^\ell)$ and where $R_{1,2}^t = N^{-1} \left(\sum_{i<N} \sigma_i^1 \sigma_i^2 + t\sigma_N^1 \sigma_N^2 \right)$. Of course it does not matter that we have $R_{1,2}^t$ rather than $R_{1,2}$. The problem is that A might take huge values, because the derivatives of u can be very large (which could not happen in Section 2.3) and we have to show that somehow these huge values cancel out. With the notation of Section 2.2 we have

$$\begin{aligned} \nu_{t,v}(f(R_{1,2}^t - q)A) &= \mathsf{E} \frac{\left\langle f(R_{1,2}^t - q) \mathsf{E}_\xi A \exp \sum_{\ell \le n} u(S_v^\ell) \right\rangle_{t,\sim}}{\langle \mathsf{E}_\xi \exp u(S_v^1) \rangle_{t,\sim}^n} \\ &= \mathsf{E}\,\mathsf{E}' \frac{\left\langle f(R_{1,2}^t - q) \mathsf{E}_\xi A \exp \sum_{\ell \le n} u(S_v^\ell) \right\rangle_{t,\sim}}{\langle \mathsf{E}_\xi \exp u(S_v^1) \rangle_{t,\sim}^n} , \end{aligned} \tag{9.5}$$

where E' denotes expectation only in the randomness of the S_v^ℓ. This randomness is independent of the randomness of $\langle \cdot \rangle_{t,\sim}$. We then separate the numerator and the denominator using the Cauchy-Schwarz inequality

$$\mathsf{E}' \left| \frac{\langle f(R^t_{\ell,\ell'} - q)\mathsf{E}_\xi A \exp \sum_{\ell \le n} u(S^\ell_v)\rangle_{t,\sim}}{\langle \mathsf{E}_\xi \exp u(S^1_v)\rangle^n_{t,\sim}} \right| \le \mathrm{I} \times \mathrm{II} \tag{9.6}$$

where

$$\mathrm{I} = \left(\mathsf{E}' \Big\langle f(R^t_{\ell,\ell'} - q)\mathsf{E}_\xi A \exp \sum_{\ell \le n} u(S^\ell_v) \Big\rangle^2_{t,\sim} \right)^{1/2} \tag{9.7}$$

and

$$\mathrm{II} = \left(\mathsf{E}' \frac{1}{\langle \mathsf{E}_\xi \exp u(S^1_v)\rangle^{2n}_{t,\sim}} \right)^{1/2} . \tag{9.8}$$

We will bound both terms separately. To bound the denominator in (9.8) from below we cannot do better than

$$\langle \mathsf{E}_\xi \exp u(S^1_v)\rangle_{t,\sim} \ge \langle \mathsf{E}_\xi \mathbf{1}_{\{S^1_v \ge \tau\}}\rangle_{t,\sim} .$$

This quantity is closely connected (in particular when $v = 1$) to the quantity $\langle \mathbf{1}_{\{S_M \ge \tau\}}\rangle_{t,\sim}$, a random variable for which we have obtained the estimate (8.23). This estimate is however insufficient, even if we consider only the case $n = 6$. Indeed, given a random variable $X \ge 0$, to obtain the integrability of X^{-12}, it does not suffice to know that $\mathsf{P}(X \le \varepsilon) \le \varepsilon^{1/L}$, we need something like $\mathsf{P}(X \le \varepsilon) \le \varepsilon^a$ for $a > 12$. So we will have to improve on the estimate (8.23), and this will be the purpose of Section 9.3.

To control the term (9.7), if $A = A((S^\ell_v)_{\ell \le n})$, let $A' = A((S^{\ell+n}_v)_{\ell \le n})$ and define a replicated version f' of f similarly. Then

$$\begin{aligned} &\mathsf{E}' \Big\langle f(R^t_{1,2} - q)\mathsf{E}_\xi A \exp \sum_{\ell \le n} u(S^\ell_v) \Big\rangle^2_{t,\sim} \\ &= \Big\langle f f'(R^t_{1,2} - q)(R^t_{n+1,n+2} - q)\mathsf{E}'\mathsf{E}_\xi A A' \exp \sum_{\ell \le 2n} u(S^\ell_v) \Big\rangle_{t,\sim} . \end{aligned} \tag{9.9}$$

To control this quantity, we will prove the following. There is an exponentially small set of configurations $(\boldsymbol{\sigma}^1, \dots, \boldsymbol{\sigma}^{2n})$ such that, outside this set, we have

$$\left| \mathsf{E}'\mathsf{E}_\xi A A' \exp \sum_{\ell \le 2n} u(S^\ell_v) \right| \le L .$$

The reason for this is simply that when there is enough independence among the r.v.s $(S^\ell_v)_{\ell \le 2n}$, one can eliminate the derivatives of u occurring in A and A' through integration by parts (and these were the cause for A to be large). On the exceptionally small set of configurations we use (9.3) to control $|A|$. In this manner we will prove that the quantity (9.9) is at most

$$\begin{aligned} & L\langle|f||f'||R^t_{1,2}-q||R^t_{\ell+n,\ell'+n}-q|\rangle_{t,\sim}+\mathcal{R} \\ & = L\langle|f||R^t_{1,2}-q|\rangle^2_{t,\sim}+\mathcal{R} \end{aligned}$$

where $\mathcal{R}$ is exponentially small. Therefore

$$\mathrm{I} \le L\langle|f||R^t_{1,2}-q|\rangle_{t,\sim}+\mathcal{R}$$

which (modulo the fact that we have $\langle\cdot\rangle_{t,\sim}$ rather than $\langle\cdot\rangle_t$) is very much what we are looking for. We should also point out that it does not work to use the Cauchy-Schwarz inequality on the whole of E in (9.5); this would yield a bound $\mathsf{E}\langle|f|^2|R^t_{1,2}-q|^2\rangle^{1/2}$, which is useless.

Learning how to perform integration by parts will occupy Section 9.4.

There is a further complication. Each of the two bounds previously described needs the knowledge that the average $\langle\cdot\rangle_{t,\sim}$ is not pathological. We know how to prove this when $Z_{N,M}=\sum_{\boldsymbol{\sigma}}\exp(-H_{N,M}(\boldsymbol{\sigma}))$ is not too small. We will prove a priori that this is the case with overwhelming probability, provided $\alpha=M/N$ is not too large.

Once these obstacles are overcome, we can recover the results of Section 2.4 when $L\alpha\exp L\tau^2\le 1$, but this time under the much less stringent conditions (9.3) with $D=\exp(N/L)$. This will be done in Section 9.5. The main estimate is obtained in Proposition 9.5.5. Roughly speaking, this Proposition replaces the estimate

$$\left|\frac{\mathrm{d}}{\mathrm{d}v}\nu_{t,v}(B_vf)\right|\le K(D)\left(\nu_{t,v}(|f|^{\tau_1})^{1/\tau_1}\nu_{t,v}(|R_{1,2}-q|^{\tau_2})^{1/\tau_2}+\frac{1}{N}\nu_{t,v}(|f|)\right)$$

of Lemma 2.3.2 by the estimate

$$\begin{aligned} \left|\frac{\mathrm{d}}{\mathrm{d}v}\nu_{t,v}(B_vf)\right| \le L\exp L\tau^2\Bigg(&\nu_{t,v}(|f|^{\tau_1})^{1/\tau_1}\nu_{t,v}(|R_{1,2}-q|^{\tau_2})^{1/\tau_2} \\ &+\frac{1}{N}\nu_{t,v}(|f|)+\max|f|D^4\exp\left(-\frac{N}{L}\right)\Bigg)\,. \end{aligned}$$

From this, we will be able, in Section 9.6, to deduce the Gardner formula for the cube when $L\alpha\exp L\tau^2\le 1$ by repeating (in a simpler manner) the arguments of Section 8.3. We will also, in Section 9.8, show the surprising fact that, in the end, the differentiability of u is largely irrelevant. In the remainder of the chapter, we will prove a central limit theorem for the overlaps, and we will investigate the Bernoulli model, when the Gaussian randomness is replaced by coin-flipping randomness.

9.2 A Priori Estimates

We already have the tools to prove that $Z_{N,M}$ is typically not too small. This will be done in Theorem 9.2.3 below. We start by a simple observation.

Lemma 9.2.1. *Consider a probability measure G on Σ_N; assume that G has a density proportional to $W \le 1$ with respect to the uniform measure on Σ_N, and assume that for a certain number $t > 0$ we have*

$$Z := \sum_{\boldsymbol{\sigma}} W(\boldsymbol{\sigma}) \ge 2^N \exp\left(-\frac{Nt^2}{8}\right) . \tag{9.10}$$

Then we have

$$G^{\otimes 2}\big(\{(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2)\ ;\ R_{1,2} \ge t\}\big) \le \exp\left(-\frac{Nt^2}{4}\right) . \tag{9.11}$$

Proof. Using that $W \le 1$ in the second line and (9.10) in the third line, we obtain

$$\begin{aligned} G^{\otimes 2}(\{(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2)\ ;\ R_{1,2} \ge t\}) &= \frac{1}{Z^2} \sum_{R_{1,2} \ge t} W(\boldsymbol{\sigma}^1) W(\boldsymbol{\sigma}^2) \\ &\le \frac{1}{Z^2} \mathrm{card}\{(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2)\ ;\ R_{1,2} \ge t\} \\ &\le 2^{-2N} \exp\left(\frac{Nt^2}{4}\right) \mathrm{card}\{(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2)\ ;\ R_{1,2} \ge t\} . \end{aligned}$$

Now, if $(\eta_i)_{i \le N}$ are independent Bernoulli r.v.s,

$$\begin{aligned} 2^{-2N} \mathrm{card}\{(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2)\ ;\ R_{1,2} > t\} &= \mathsf{P}\Big(\sum_{i \le N} \eta_i \ge tN\Big) \\ &\le \exp\left(-\frac{Nt^2}{2}\right) \end{aligned}$$

by the subgaussian inequality (A.16) used for $a_i = 1$. □

Lemma 9.2.2. *There exists a number L and a number $\lambda_0 > 0$ with the following property. Consider a probability measure G on Σ_N; assume that G has a density proportional to $W \le 1$ with respect to the uniform measure on Σ_N, and assume that*

$$Z := \sum_{\boldsymbol{\sigma}} W(\boldsymbol{\sigma}) \ge 2^N \exp\left(-\frac{N}{32}\right) . \tag{9.12}$$

Then, for independent standard normal r.v.s $(g_i)_{i \le N}$ we have

$$\begin{aligned} & L \exp\left(-\frac{N}{L}\right) \le \varepsilon \le \frac{1}{L} \exp(-L\tau^2) \\ \Rightarrow \quad & \mathsf{P}\bigg(G\bigg(\bigg\{\boldsymbol{\sigma}\ ;\ \frac{1}{\sqrt{N}} \sum_{i \le N} g_i \sigma_i \ge \tau\bigg\}\bigg) \le \varepsilon\bigg) \le \varepsilon^{1/L} . \end{aligned} \tag{9.13}$$

Moreover the r.v.

$$V = \frac{1}{\max\Big(\exp(-N/32), G\big(\{\boldsymbol{\sigma}\ ;\ \frac{1}{\sqrt{N}}\sum_{i\le N} g_i\sigma_i \ge \tau\}\big)\Big)} \tag{9.14}$$

satisfies

$$\mathsf{E} V^{\lambda_0} \le L \exp L\tau^2\ . \tag{9.15}$$

Of course the value 1/32 is just a convenient choice. We write

$$a = \frac{1}{32}\ ;\ b^* = \exp(-aN) = \exp(-N/32)$$

throughout this chapter.

Proof. From (9.12) and Lemma 9.2.1, we see that (9.11) holds for $t = 1/2$. Thus (9.13) follows from Proposition 8.2.6 used for $b = 1$ and $c = 1/2$, $d = 32$. The r.v. V satisfies

$$t > \exp aN \quad \Rightarrow \quad \mathsf{P}(V > t) = 0$$
$$L\exp(L\tau^2) \le t \le \frac{1}{L}\exp\left(\frac{N}{L}\right) \quad \Rightarrow \quad \mathsf{P}(V > t) \le Lt^{-1/L}\ ,$$

and the conclusion follows from Lemma 8.3.8. □

We recall that S_k is defined in (9.1) and we state the main result of this section.

Theorem 9.2.3. *There exists a number L with the following property. If $b \ge b^* = \exp(-aN)$, then*

$$\mathsf{P}\big(\mathrm{card}\{\boldsymbol{\sigma}\ ;\ \forall k \le M\ ,\ S_k(\boldsymbol{\sigma}) \ge \tau\} \le b2^N\big) \le b^{1/L}\exp(LM(1+\tau^2))\ . \tag{9.16}$$

This inequality is of interest only for $b \le 1$ so the larger the value of L, the weaker the inequality. If we take $b = b^* = \exp(-N/32)$, the right-hand side is $\exp(L_1 M(1+\tau^2) - N/L_1)$, which is exponentially small as soon as $2L_1^2\alpha(1+\tau^2) \le 1$. This might be the place to remind the reader that by L we always denote a number, that does not depend on any parameter whatsoever, but that need not be the same at each occurrence. With this convention, the short-hand way to write the previous claim is that "when $b = b^* = \exp(-N/32)$, the right-hand side of (9.16) is exponentially small when $L\alpha(1+\tau^2) \le 1$" The reader will then understand by herself that the constant L occurring in this inequality is a new number that depends only on the (different) number L occurring in the right-hand side of (9.16).

Since $u(x) = 0$ for $x \ge \tau$, we have

$$Z_{N,M} \ge \mathrm{card}\{\boldsymbol{\sigma}\ ;\ \forall k \le M\ ,\ S_k(\boldsymbol{\sigma}) \ge \tau\}\ ,$$

and the previous result shows that $Z_{N,M}$ is typically $\ge 2^N\exp(-aN)$ when $L\alpha(1+\tau^2) \le 1$. Therefore in that case the Gibbs measure typically satisfies (9.13).

Research Problem 9.2.4. (Level 2) Can the main results of this chapter be proved under the condition $L\alpha(1+\tau^2) \le 1$ rather than under the condition $L\alpha \exp L\tau^2 \le 1$?

Apparently solving this problem requires finding a different approach.

Research Problem 9.2.5. (Level 2) Extend the results of this section to the case of the Hamiltonian

$$H_{M,N}(\boldsymbol{\sigma}) = \sum_{k\le M} u(S_k) + h\sum_{i\le N} \sigma_i \tag{9.17}$$

where h is large.

The point of this problem is that the influence of a large external field will make $R_{1,2}$ typically close to one, while our arguments constantly require that "$R_{1,2}$ be typically small", so the solution of this problem is also likely to require a different approach. Also one often gets the feeling (but maybe this has no basis) that adding an external field can only improve matters.

Let us also note that it should be obvious to the reader, once she understands our arguments, that for $\tau \le 0$, the condition $L\alpha \le 1$ suffices.

Proof of Theorem 9.2.3. We set

$$V_M = 2^{-N}\mathrm{card}\{\boldsymbol{\sigma}\ ;\ \forall k \le M\ ,\ S_k(\boldsymbol{\sigma}) \ge \tau\}$$

so that

$$V_M \le V_{M-1} \le 1\ .$$

Let us denote by G the probability measure on Σ_N of density $W(\boldsymbol{\sigma}) = \mathbf{1}_{\bigcap_{k\le M-1} U_k}(\boldsymbol{\sigma})$ with respect to the uniform measure on Σ_N. It satisfies the condition (9.12) of Lemma 9.2.2 provided $V_{M-1} \ge b^*$. Also,

$$\frac{V_M}{V_{M-1}} = G(\{\boldsymbol{\sigma}\ ;\ S_M(\boldsymbol{\sigma}) \ge \tau\})\ .$$

It then follows from (9.15) that if E_M denotes expectation only with respect to the r.v.s $g_{i,M}$, we have

$$V_{M-1} \ge b^* \quad \Rightarrow \quad \mathsf{E}_M \frac{1}{\max{(b^*, V_M/V_{M-1})}^{\lambda_0}} \le L \exp L\tau^2 \le \exp L(\tau^2+1)\ . \tag{9.18}$$

When $V_{M-1} \ge b^*$ we further have, since $V_{M-1} \le 1$,

$$\begin{aligned}\max(b^*, V_M) &= V_{M-1} \max\left(\frac{b^*}{V_{M-1}}, \frac{V_M}{V_{M-1}}\right) \\ &\ge \max(b^*, V_{M-1}) \max\left(b^*, \frac{V_M}{V_{M-1}}\right)\end{aligned}$$

and combining with (9.18) yields

$$\mathsf{E}_M \frac{1}{\max(b^*, V_M)^{\lambda_0}} \le \frac{1}{\max(b^*, V_{M-1})^{\lambda_0}} \exp L(1+\tau^2) \,.$$

This relation remains true when $V_{M-1} \le b^*$ because then the left-hand side is $\le b^{*-\lambda_0}$ while the right-hand side is $\ge b^{*-\lambda_0}$. Iteration of this relation yields

$$\mathsf{E} \frac{1}{\max(b^*, V_M)^{\lambda_0}} \le \exp LM(1+\tau^2)$$

so that, if $b \ge b^*$,

$$\mathsf{P}(V_M \le b) b^{-\lambda_0} \le \exp LM(1+\tau^2) \,. \qquad \square$$

Throughout the chapter we use the notation

$$U_k = \{S_k \ge \tau\} \,.$$

Later it will be of fundamental importance that the r.v.

$$\operatorname{card}\{\boldsymbol{\sigma} \,;\, \forall k \le M \,,\, S_k(\boldsymbol{\sigma}) \ge \tau\} = \operatorname{card} \bigcap_{k \le M} U_k$$

has small fluctuations. Since the argument is close in spirit to the previous one, we present it now, but the result itself will not be used before Section 9.6. We recall that $a = 1/32$ and from Chapter 8 the notation

$$\log_A(x) = \max(-A, \log x) \,. \tag{9.19}$$

Proposition 9.2.6. *There exists a number L with the following property. Consider any function u satisfying (9.2) and let*

$$Z = Z(u) = Z_{N,M}(u) = \sum_{\boldsymbol{\sigma}} \exp \sum_{k \le M} u(S_k(\boldsymbol{\sigma})) \,. \tag{9.20}$$

Then for each $t > 0$ we have

$$\mathsf{P}\left(\left| \frac{1}{N} \log_{aN}(2^{-N} Z) - \frac{1}{N} \mathsf{E} \log_{aN}(2^{-N} Z) \right| \ge t \right) \le 2 \exp\left(-\frac{1}{L} \min\left(\frac{N^2 t^2}{M(1+\tau^2)}, \frac{Nt}{1+\tau^2} \right) \right) \,. \tag{9.21}$$

This result includes the case $\exp u(x) = \mathbf{1}_{\{x \ge \tau\}}$.

Proof. We follow the proof of Proposition 8.3.6. Denoting Ξ_m the σ-algebra generated by the r.v.s $(g_{i,k})$ for $i \le N$, $k \le m$ and by E^m the conditional expectation given Ξ_m, we write

$$\frac{1}{N}\log_{aN}(2^{-N}Z) - \frac{1}{N}\mathsf{E}\log_{aN}(2^{-N}Z) = \sum_{m=1}^{M} X_m$$

where

$$X_m = \frac{1}{N}\mathsf{E}^m \log_{aN}(2^{-N}Z) - \frac{1}{N}\mathsf{E}^{m-1}\log_{aN}(2^{-N}Z) \,. \tag{9.22}$$

Using Bernstein's inequality for martingale difference sequences (A.41), it suffices to prove that

$$\mathsf{E}\exp\frac{|X_m|}{L(1+\tau^2)} \le 2 \,. \tag{9.23}$$

Let us define

$$Z_m = \sum_{\boldsymbol{\sigma}} \exp \sum_{k\ne m} u(S_k(\boldsymbol{\sigma}))$$

and

$$Y = \mathbf{1}_{\{Z_m \ge b^*\}} \log_{aN}\frac{Z_m}{Z} \,.$$

Denoting by E_m expectation in the r.v.s $(g_{i,m})_{i\le N}$ only, it suffices (as in (8.58)) to prove that

$$\mathsf{E}_m \exp 2\lambda Y \le 2 \tag{9.24}$$

for $\lambda = 1/L(1+\tau^2)$. To prove this we may assume $Z_m \ge b^*$, for otherwise $Y = 0$. The probability measure G on Σ_N with density proportional to $W = \exp\big(\sum_{k\ne m} u(S_k)\big)$ then satisfies the conditions of Lemma 9.2.2, and thus, by (9.15) we have

$$\mathsf{E}_m \frac{1}{\max(b^*, G(U_m))^{\lambda_0}} \le L\exp L\tau^2 \,. \tag{9.25}$$

Now

$$Z_m G(U_m) = \sum_{\boldsymbol{\sigma}\in U_m} \exp\Bigg(\sum_{k\ne m} u(S_k(\boldsymbol{\sigma}))\Bigg) \le \sum_{\boldsymbol{\sigma}} \exp\Bigg(\sum_{k\le M} u(S_k(\boldsymbol{\sigma}))\Bigg) = Z$$

because $u(S_m) = 0$ on U_m. Thus, using in the last equality that $Y = \log_{aN} Z_m/Z = \max(aN, Z_m/Z)$, we get

$$\frac{1}{\max(b^*, G(U_m))} = \min\left(\exp aN, \frac{1}{G(U_m)}\right) \ge \min\left(\exp aN, \frac{Z_m}{Z}\right) = \exp Y$$

and (9.25) implies

$$\mathsf{E}_m \exp \lambda_0 Y \le L\exp L\tau^2 \le \exp L_2(1+\tau^2) \,,$$

from which (9.24) follows through Hölder's inequality for $\lambda = \lambda_0/2L_2(1+\tau^2)$.
□

9.3 Gaussian Processes

The goal of this section is to bound the quantity (9.8). It should help the reader to look again at Section 8.2, up to Proposition 8.2.6. The arguments here are similar, just a bit more elaborate.

Theorem 9.3.1. *There exists a number L with the following property. Consider $0 < c \le 1/2$ and a jointly Gaussian family $(w_\ell)_{\ell \le n}$; assume that $\mathsf{E}w_\ell^2 = 1$ and that*

$$\forall \ell \ne \ell' \,, \quad \mathsf{E}w_\ell w_{\ell'} \le c \,. \tag{9.26}$$

Then if $nc \ge 2$ and $L \le s \le \sqrt{\log(nc/2)}/L$ we have

$$\mathsf{P}\big(\mathrm{card}\{\ell \le n \,;\, w_\ell \ge s\} \le n\exp(-Ls^2)\big) \le L\exp\left(-\frac{s^2}{Lc}\right) \,. \tag{9.27}$$

The point of (9.27) is that if we set $\varepsilon = \exp(-Ls^2)$ the bound is $L\varepsilon^{1/Lc}$, and the exponent will be large for c small.

The proof relies on an elementary geometrical lemma.

Lemma 9.3.2. *Consider a number $c > 0$, and vectors $(\mathbf{x}_\ell)_{\ell \le n}$ in a Hilbert space. Assume that $\|\mathbf{x}_\ell\| \le 1$ and $\mathbf{x}_\ell \cdot \mathbf{x}_{\ell'} \le c$ whenever $\ell \ne \ell'$. Then, for any vector $\mathbf{x}$ we have*

$$\mathrm{card}\{\ell \,;\, \mathbf{x}\cdot\mathbf{x}_\ell \ge \|\mathbf{x}\|\sqrt{2c}\} \le \frac{1}{c} \,.$$

Proof. Assume that $\mathbf{x}\cdot\mathbf{x}_\ell \ge \|\mathbf{x}\|\sqrt{2c}$ for $\ell \le k$. Then

$$k\|\mathbf{x}\|\sqrt{2c} \le \mathbf{x}\cdot\Big(\sum_{\ell\le k}\mathbf{x}_\ell\Big) \le \|\mathbf{x}\|\Big\|\sum_{\ell\le k}\mathbf{x}_\ell\Big\| \,,$$

and

$$\Big\|\sum_{\ell\le k}\mathbf{x}_\ell\Big\|^2 = \sum_{\ell\le k}\|\mathbf{x}_\ell\|^2 + \sum_{\ell\ne\ell'}\mathbf{x}_\ell\cdot\mathbf{x}_{\ell'} \le k + ck(k-1) \,.$$

Thus

$$k\sqrt{2c} \le \sqrt{k + ck(k-1)} \le \sqrt{k}\sqrt{1+ck} \,,$$

so that $2ck \le 1 + ck$ i.e. $k \le 1/c$. □

The following useful fact is a consequence of Theorem 1.3.4. We denote by $\mathbf{g} = (g_i)_{i\le N}$ a standard Gaussian vector.

Lemma 9.3.3. *Consider a closed subset B of $\mathbb{R}^N$, and*

$$d(\mathbf{x}, B) = \inf\{d(\mathbf{x},\mathbf{y}) \,;\, \mathbf{y} \in B\} \,,$$

the Euclidean distance from $\mathbf{x}$ to B. Then for $t > 0$, we have

$$\mathsf{P}\left(d(\mathbf{g}, B) \ge t + 2\sqrt{\log\frac{2}{\mathsf{P}(\mathbf{g}\in B)}}\right) \le 2\exp\left(-\frac{t^2}{4}\right) \,. \tag{9.28}$$

Proof. The function $F(\mathbf{x}) = d(\mathbf{x}, B)$ satisfies (1.45) with $A = 1$, so that for all $t > 0$ by (1.46) we have

$$\mathsf{P}(|d(\mathbf{g}, B) - \mathsf{E}d(\mathbf{g}, B)| \geq t) \leq 2\exp\left(-\frac{t^2}{4}\right) . \tag{9.29}$$

If $t = \mathsf{E}d(\mathbf{g}, B)$, then

$$\mathsf{P}(\mathbf{g} \in B) \leq \mathsf{P}(|d(\mathbf{g}, B) - \mathsf{E}d(\mathbf{g}, B)| \geq t) ,$$

and combining with (9.29) we get

$$t = \mathsf{E}d(\mathbf{g}, B) \Rightarrow \mathsf{P}(\mathbf{g} \in B) \leq 2\exp\left(-\frac{t^2}{4}\right) ,$$

so that

$$\mathsf{E}d(\mathbf{g}, B) \leq 2\sqrt{\log \frac{2}{\mathsf{P}(\mathbf{g} \in B)}} ,$$

and combining with (9.29) gives (9.28). □

Proof of Theorem 9.3.1. We consider vectors $\mathbf{x}_\ell$ in $\mathbb{R}^N$ such that the sequence $(w_\ell)_{\ell \leq n}$ has the same law as $(\mathbf{x}_\ell \cdot \mathbf{g})_{\ell \leq n}$. (The existence of these vectors is proved in Section A.2 but will be obvious in the situation where we will apply the lemma.) Using (8.13) with $b = 1$, $c = 1/2$, $d = 2$ (and changing s into sL) yields that for $L \leq s \leq \sqrt{\log n}/L$ we have

$$\mathsf{P}\big(\mathrm{card}\{\ell \leq n\ ;\ w_\ell \geq s\} > n\exp(-Ls^2)\big) \geq 1 - L\exp\left(-\frac{s^2}{L}\right)$$

i.e. if

$$B = \left\{\mathbf{x} \in \mathbb{R}^N\ ;\ \mathrm{card}\{\ell \leq n\ ;\ \mathbf{x} \cdot \mathbf{x}_\ell \geq s\} > n\exp(-Ls^2)\right\} , \tag{9.30}$$

then

$$\mathsf{P}(\mathbf{g} \in B) \geq 1 - L\exp\left(-\frac{s^2}{L}\right) .$$

Consequently, there exists a large enough constant L_3 such that for $s \geq L_3$ we have $\mathsf{P}(\mathbf{g} \in B) \geq 1/2$. (Of course, according to our conventions about the meaning of the symbol L, we should simply say that $s \geq L$ implies that $\mathsf{P}(\mathbf{g} \in B) \geq 1/2$.) It then follows from (9.28) that for $t > 0$ we have

$$\mathsf{P}(d(\mathbf{g}, B) \geq t + 4) \leq 2\exp\left(-\frac{t^2}{4}\right) ,$$

and setting $v = t + 4$, it follows that for $t > 0$, it holds

$$\mathsf{P}(d(\mathbf{g}, B) \geq v) \leq L\exp\left(-\frac{v^2}{8}\right) . \tag{9.31}$$

Let $B_v = \{\mathbf{x}\,;\ d(\mathbf{x}, B) \le v\}$, so that (9.31) implies

$$\mathsf{P}(\mathbf{g} \in B_v) \ge 1 - L\exp\left(-\frac{v^2}{8}\right)\ . \tag{9.32}$$

By definition of B_v, for $\mathbf{g} \in B_v$ we can find $\mathbf{g}' \in B$ with $\|\mathbf{g} - \mathbf{g}'\| \le v$. We note that $\mathbf{x}_\ell \cdot \mathbf{x}_{\ell'} = \mathsf{E} w_\ell w_{\ell'} \le c$ for $\ell \ne \ell'$ so that by Lemma 9.3.2 we have

$$\mathrm{card}\big\{\ell \le n\,;\ (\mathbf{g}' - \mathbf{g}) \cdot \mathbf{x}_\ell \ge v\sqrt{2c}\big\} \le \frac{1}{c}\ .$$

On the other hand, since $\mathbf{g}' \in B$, recalling the definition (9.30) of B we have

$$\mathrm{card}\{\ell \le n\,;\ \mathbf{g}' \cdot \mathbf{x}_\ell \ge s\} \ge n\exp(-Ls^2)$$

and thus

$$\mathrm{card}\{\ell \le n\,;\ \mathbf{g} \cdot \mathbf{x}_\ell \ge s - v\sqrt{2c}\} \ge n\exp(-Ls^2) - \frac{1}{c}$$

because $\mathbf{g}' \cdot \mathbf{x}_\ell \ge s$ and $(\mathbf{g}' - \mathbf{g}) \cdot \mathbf{x}_\ell < v\sqrt{2c}$ imply $\mathbf{g} \cdot \mathbf{x}_\ell \ge s - v\sqrt{2c}$. Taking $v = s/2\sqrt{2c}$, we have shown that

$$\mathbf{g} \in B_v \quad \Rightarrow \quad \mathrm{card}\left\{\ell \le n\,;\ \mathbf{g} \cdot \mathbf{x}_\ell \ge \frac{s}{2}\right\} \ge n\exp(-Ls^2) - \frac{1}{c} \ge n\exp(-Ls^2)$$

provided $s \le \sqrt{\log(nc/2)}/L$. Combining with (9.32) this completes the proof. □

Corollary 9.3.4. *There exists a number L and a number $\overline{c} > 0$ with the following property. For a jointly Gaussian family $(w_\ell)_{\ell \le n}$ with $\mathsf{E} w_\ell^2 = 1$ and*

$$\ell \ne \ell' \quad \Rightarrow \quad \mathsf{E} w_\ell w_{\ell'} \le \overline{c}\ ,$$

then for any number $\tau \ge 0$ and

$$Ln^{-1/L} \le \varepsilon \le \frac{1}{L}\exp(-L\tau^2)$$

we have

$$\mathsf{P}\big(\mathrm{card}\{\ell \le n\,;\ w_\ell \ge \tau\} \le \varepsilon n\big) \le L\varepsilon^{24}\ . \tag{9.33}$$

Proof. From (9.27) we obtain

$$\mathsf{P}\big(\mathrm{card}\{\ell \le n\,;\ w_\ell \ge \tau\} < n\exp(-Ls^2)\big) \le L\exp\left(-\frac{s^2}{L\overline{c}}\right)$$

provided $s \ge \tau$, $s \ge L$, $s \le \sqrt{\log(n\overline{c}/2)}/L$. Letting $\varepsilon = \exp(-Ls^2)$ we have

$$L\exp\left(-\frac{s^2}{L\overline{c}}\right) = L\varepsilon^{1/L_4\overline{c}} = L\varepsilon^{24}$$

if $\overline{c} = 1/24L_4$. This completes the proof. □

The meaning of the quantity $\overline{c}$ remains as above in the remainder of this chapter.

Proposition 9.3.5. *There exists a number L with the following property. Consider a probability measure G on Σ_N, and a family $(w(\boldsymbol{\sigma}))_{\boldsymbol{\sigma}\in\Sigma_N}$ of jointly Gaussian r.v.s such that $\mathsf{E}w^2(\boldsymbol{\sigma})=1$ and*

$$G^{\otimes 2}\big(\{(\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2)\ ;\ \mathsf{E}w(\boldsymbol{\sigma}^1)w(\boldsymbol{\sigma}^2)>\overline{c}\}\big)\le 32\exp\left(-\frac{N}{d}\right) \tag{9.34}$$

for a certain number d. Then for any number $\tau\ge 0$ we have

$$L\exp\left(-\frac{N}{Ld}\right)\le\varepsilon\le\frac{1}{L}\exp(-L\tau^2)\ \Rightarrow\ \mathsf{P}\big(G(\{\boldsymbol{\sigma}\ ;\ w(\boldsymbol{\sigma})\ge\tau\})\le\varepsilon\big)\le L\varepsilon^{24}\ . \tag{9.35}$$

Proof. We copy the proof of Proposition 8.2.6, using now Corollary 9.3.4 instead of Corollary 8.2.5. Let

$$Q_n=\{(\boldsymbol{\sigma}^1,\ldots,\boldsymbol{\sigma}^n)\ ;\ \forall\ell\ne\ell'\ ,\ \mathsf{E}w(\boldsymbol{\sigma}^\ell)w(\boldsymbol{\sigma}^{\ell'})\le\overline{c}\}\ ,$$

so that since there are at most $n(n-1)/2\le n^2/2$ choices for ℓ and ℓ' it follows from (9.34) that we have

$$32n^2\exp\left(-\frac{N}{d}\right)\le 1\quad\Rightarrow\quad G^{\otimes n}(Q_n)\ge\frac{1}{2}\ .$$

For $(\boldsymbol{\sigma}^1,\ldots,\boldsymbol{\sigma}^n)\in Q_n$, consider the event

$$\Omega(\boldsymbol{\sigma}^1,\ldots,\boldsymbol{\sigma}^n)=\{\mathrm{card}\{\ell\le n\ ;\ w(\boldsymbol{\sigma}^\ell)\ge\tau\}\le\varepsilon n\}\ ,$$

so that Corollary 9.3.4 implies

$$Ln^{-1/L}\le\varepsilon\le\frac{1}{L}\exp(-L\tau^2)\quad\Rightarrow\quad\mathsf{P}(\Omega(\boldsymbol{\sigma}^1,\ldots,\boldsymbol{\sigma}^n))\le L\varepsilon^{24}\ .$$

Thus if we define

$$Y=\int_{Q_n}\mathbf{1}_{\{\Omega(\boldsymbol{\sigma}^1,\ldots,\boldsymbol{\sigma}^n)\}}\mathrm{d}G(\boldsymbol{\sigma}^1)\cdots\mathrm{d}G(\boldsymbol{\sigma}^n)\ ,$$

we have

$$\mathsf{E}Y=\int_{Q_n}\mathsf{P}(\Omega(\boldsymbol{\sigma}^1,\ldots,\boldsymbol{\sigma}^n))\mathrm{d}G(\boldsymbol{\sigma}^1)\cdots\mathrm{d}G(\boldsymbol{\sigma}^n)\le L\varepsilon^{24}\ ,$$

and therefore by Markov's inequality $\mathsf{P}(Y\ge 1/4)\le L\varepsilon^{24}$. Now

$$Y=G^{\otimes n}(\{(\boldsymbol{\sigma}^1,\ldots,\boldsymbol{\sigma}^n)\in Q_n\ ;\ \mathrm{card}\{\ell\le n\ ;\ w(\boldsymbol{\sigma}^\ell)\ge\tau\}\le n\varepsilon\})$$

so that

$$\begin{aligned} Y \le \frac{1}{4} \Rightarrow G^{\otimes n}(\{(\boldsymbol{\sigma}^1, \dots, \boldsymbol{\sigma}^n) \in Q_n \; ; \; \mathrm{card}\{\ell \le n \; ; \; w(\boldsymbol{\sigma}^\ell) \ge \tau\} > n\varepsilon\}) \\ = G^{\otimes n}(Q_n) - Y \ge \frac{1}{2} - \frac{1}{4} = \frac{1}{4} \; . \end{aligned}$$

In that case,

$$\begin{aligned} nG(\{\boldsymbol{\sigma} \; ; \; w(\boldsymbol{\sigma}) \ge \tau\}) &= \int \mathrm{card}\{\ell \le n \; ; \; w(\boldsymbol{\sigma}^\ell) \ge \tau\} \mathrm{d}G(\boldsymbol{\sigma}^1) \cdots \mathrm{d}G(\boldsymbol{\sigma}^n) \\ &\ge n\varepsilon G^{\otimes n}\big(\{(\boldsymbol{\sigma}^1, \dots, \boldsymbol{\sigma}^n) \in Q_n \; ; \; \mathrm{card}\{\ell \le n \; ; \\ &\quad w(\boldsymbol{\sigma}^\ell) \ge \tau\} > n\varepsilon\}\big) \\ &\ge \frac{n\varepsilon}{4} \; , \end{aligned}$$

so that we have proved that

$$Y \le \frac{1}{4} \quad \Rightarrow \quad G(\{\boldsymbol{\sigma} \; ; \; w(\boldsymbol{\sigma}) \ge \tau\}) \ge \frac{\varepsilon}{4} \; ,$$

and therefore

$$G(\{\boldsymbol{\sigma} \; ; \; w(\boldsymbol{\sigma}) \ge \tau\}) < \frac{\varepsilon}{4} \quad \Rightarrow \quad Y > \frac{1}{4} \; .$$

In conclusion, if $32n^2 \le \exp(N/d)$ and ε satisfies

$$Ln^{-1/L} \le \varepsilon \le \frac{1}{L} \exp(-L\tau^2)$$

we have

$$\mathsf{P}(G(\{\boldsymbol{\sigma} \; ; \; w(\boldsymbol{\sigma}) \ge \tau\}) < \varepsilon/4) \le L\varepsilon^{24} \; .$$

We conclude by taking n as large as possible. □

Corollary 9.3.6. *There exists a constant L with the following property. Consider a probability measure G on Σ_N and a family $(w(\boldsymbol{\sigma}))$ as in Proposition 9.3.5. Then if $b = L\exp(-N/Ld)$, for any number $\tau \ge 0$ we have*

$$\mathsf{E} \frac{1}{\max(b, G(\{\boldsymbol{\sigma} \; ; \; w(\boldsymbol{\sigma}) \ge \tau\}))^{12}} \le L \exp L\tau^2 \; . \tag{9.36}$$

Proof. We define $b = L\exp(-N/Ld)$ where L is the constant of (9.35). Let $Y = \max(b, G(\{\boldsymbol{\sigma} \; ; \; w(\boldsymbol{\sigma}) \ge \tau\}))$, so that by (9.35) we have

$$\varepsilon \le \varepsilon_0 := \frac{1}{L} \exp(-L\tau^2) \quad \Rightarrow \quad \mathsf{P}(Y \le \varepsilon) \le L\varepsilon^{24} \; ,$$

because $\mathsf{P}(Y \le \varepsilon) = 0$ if $\varepsilon < c$. We use that

$$\begin{aligned} \mathsf{E} \frac{1}{Y^{12}} &= 12 \int_0^\infty \mathsf{P}(Y \le \varepsilon) \varepsilon^{-13} \mathrm{d}\varepsilon \\ &\le L \int_0^{\varepsilon_0} \varepsilon^{11} \mathrm{d}\varepsilon + 12 \int_{\varepsilon_0}^\infty \varepsilon^{-13} \mathrm{d}\varepsilon \\ &\le L\varepsilon_0^{12} + \varepsilon_0^{-12} \le L \exp L\tau^2 \; , \end{aligned}$$

and this completes the proof. □

We recall the definition of S_k given in (9.1). If we define $w(\boldsymbol{\sigma}) = S_k(\boldsymbol{\sigma})$, then $R_{1,2} = \mathsf{E} S_k(\boldsymbol{\sigma}^1) S_k(\boldsymbol{\sigma}^2) = \mathsf{E} w(\boldsymbol{\sigma}^1) w(\boldsymbol{\sigma}^2)$. Here is a simple situation where (9.34) is satisfied in this case.

Proposition 9.3.7. *Assume that*

$$\operatorname{card}\big\{\boldsymbol{\sigma} \in \Sigma_N \ ; \ \forall k \le M-1 \ , \ S_k \ge \tau\big\} \ge 2^N \exp\left(-\frac{N\overline{c}^2}{8}\right) . \tag{9.37}$$

Consider the Gibbs measure G with Hamiltonian $-\sum_{k \le M-1} u(S_k(\boldsymbol{\sigma}))$. Then

$$G^{\otimes 2}\big(\{(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) \ ; \ R_{1,2} > \overline{c}\}\big) \le \exp\left(-\frac{N\overline{c}^2}{4}\right) . \tag{9.38}$$

Proof. Use Lemma 9.2.1 with $t = \overline{c}$ and $W(\boldsymbol{\sigma}) = \exp \sum_{k \le M-1} u(S_k(\boldsymbol{\sigma}))$. □

We must now take care of some (tedious and unsurprising) details in order to be able to apply the above principles to our interpolating Hamiltonians. We recall the notation

$$S_k^0(\boldsymbol{\sigma}) = \frac{1}{\sqrt{N}} \sum_{i<N} g_{i,k}\sigma_i = S_k(\boldsymbol{\sigma}) - \frac{g_{N,k}}{\sqrt{N}}$$

$$S_{k,t}(\boldsymbol{\sigma}) = S_k^0(\boldsymbol{\sigma}) + \sqrt{\frac{t}{N}} g_{N,k}\sigma_N + \sqrt{\frac{1-t}{N}} \xi_k \sigma_N \ ,$$

where ξ_k are independent standard Gaussian random variables (independent of all the other r.v.s already introduced). It is of course almost certain that replacing S_k by $S_{k,t}$ in the interpolating Hamiltonian cannot really change anything, but we must nonetheless check this. This occupies the rest of this section.

Lemma 9.3.8. *If we have $L\alpha(1 + tau^2) \le 1$ for a large enough constant L then the following two events*

$$\operatorname{card}\big\{\boldsymbol{\sigma}; \ \forall k < M \ , \ S_k(\boldsymbol{\sigma}) \ge \tau + 3\big\} \ge 2^N \exp\left(-\frac{N\overline{c}^2}{16}\right) \ ; \tag{9.39}$$

$$\forall k < M \ , \qquad |g_{N,k}| \le \sqrt{N} \ . \tag{9.40}$$

occur with probability $\ge 1 - L\exp(-N/L)$.

Proof. We use Theorem 9.2.3 with $\tau + 3$ instead of τ and $b = \exp(-N\overline{c}^2/16)$. □

Lemma 9.3.9. *If $N \geq 10$, the following holds true. Let us assume that (9.39) and (9.40) hold true. Then, for any number y we have*

$$\mathsf{E}_\xi \sum_{\boldsymbol{\sigma}} \exp\Bigl(\sum_{k<M} u(S_{k,t}(\boldsymbol{\sigma})) + \sigma_N y\Bigr) \geq 2^{N-1} \exp\left(-\frac{N\overline{c}^2}{16}\right) \operatorname{ch} y \;, \tag{9.41}$$

where E_ξ denotes expectation in the r.v.s ξ_k.

Proof. Let

$$A = \bigl\{\boldsymbol{\sigma} \;;\; \forall k < M \;,\; S_k^0(\boldsymbol{\sigma}) \geq \tau + 2\bigr\} \;.$$

Let us assume that

$$\forall k < M \;, \qquad |\xi_k| \leq \sqrt{N} \;. \tag{9.42}$$

Then using that $|g_{N,k}| \leq \sqrt{N}$ in the first line and that $|g_{N,k}| \leq \sqrt{N}$ and $|\xi_k| \leq \sqrt{N}$ in the second line yields

$$S_k^0(\boldsymbol{\sigma}) \geq S_k(\boldsymbol{\sigma}) - 1 \tag{9.43}$$

$$S_{k,t}(\boldsymbol{\sigma}) \geq S_k^0(\boldsymbol{\sigma}) - (\sqrt{t} + \sqrt{1-t}) \geq S_k^0(\boldsymbol{\sigma}) - 2. \tag{9.44}$$

Since $u(x) = 0$ for $x \geq \tau$ we have $u(S_{k,t}(\boldsymbol{\sigma})) = 0$ if $S_k^0(\boldsymbol{\sigma}) \geq \tau + 2$. Consequently for $\boldsymbol{\sigma} \in A$ we have

$$\exp\Bigl(\sum_{k<M} u(S_{k,t}(\boldsymbol{\sigma})) + \sigma_N y\Bigr) = \exp \sigma_N y \;,$$

so that

$$\begin{aligned} Z := \sum_{\boldsymbol{\sigma}} \exp\Bigl(\sum_{k<M} u(S_{k,t}(\boldsymbol{\sigma})) + \sigma_N y\Bigr) &\geq \sum_{A} \exp\Bigl(\sum_{k<M} u(S_{k,t}(\boldsymbol{\sigma})) + \sigma_N y\Bigr) \\ &= \sum_{A} \exp \sigma_N y = \operatorname{card} A \operatorname{ch} y \;, \end{aligned} \tag{9.45}$$

because the set A is invariant under the transformation $(\sigma_1, \ldots, \sigma_{N-1}, \sigma_N) \mapsto (\sigma_1, \ldots, \sigma_{N-1}, -\sigma_N)$. Also, (9.39) and (9.43) imply that

$$\operatorname{card} A \geq 2^N \exp(-N\overline{c}^2/16) \;,$$

and thus by (9.45), under (9.42) we have $Z \geq 2^N \exp(-N\overline{c}^2/16)\operatorname{ch} y$. Since ξ_k is standard Gaussian, we have $\mathsf{P}(|\xi_k| \geq \sqrt{N}) \leq 2\exp(-N/2)$, so that for $N \geq 10$ the event Ω described by (9.42) occurs with probability $\geq 1/2$ and this completes the proof since

$$\mathsf{E}_\xi Z \geq \mathsf{E}_\xi(\mathbf{1}_\Omega Z) \geq \mathsf{P}(\Omega) 2^N \exp(-N\overline{c}^2/16)\operatorname{ch} y \;. \qquad \square$$

We recall the notation $R_{1,2}^t = N^{-1}\bigl(\sum_{i<N} \sigma_i^1 \sigma_i^2 + t\sigma_N^1 \sigma_N^2\bigr)$.

Proposition 9.3.10. *For $N \geq 10$ the following occurs. Assume (9.39) and (9.40), and consider the measure G on Σ_N given by*

$$\int f \mathrm{d}G = \frac{\sum_{\boldsymbol{\sigma}} f(\boldsymbol{\sigma}) \mathsf{E}_\xi \exp\left(\sum_{k<M} u(S_{k,t}(\boldsymbol{\sigma})) + \sigma_N y\right)}{\sum_{\boldsymbol{\sigma}} \mathsf{E}_\xi \exp\left(\sum_{k<M} u(S_{k,t}(\boldsymbol{\sigma})) + \sigma_N y\right)} . \tag{9.46}$$

Then

$$G^{\otimes 2}\left(\left\{(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) \; ; \; |R^t_{1,2}| \geq \left(1 - \frac{1-t}{N}\right) \overline{c}\right\}\right) \leq 32 \exp\left(-\frac{N\overline{c}^2}{8}\right) . \tag{9.47}$$

Proof. Let $R^-_{1,2} = N^{-1}\sum_{i<N} \sigma^1_i \sigma^2_i$, so that (for $N \geq 10$)

$$|R^t_{1,2}| > \left(1 - \frac{1-t}{N}\right) \overline{c} \;\; \Rightarrow \;\; |R^-_{1,2}| > \frac{3\overline{c}}{4} .$$

Therefore if Z is as in (9.45), we have

$$\begin{aligned} & G^{\otimes 2}\left(\left\{(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) \; ; \; |R^t_{1,2}| \geq \left(1 - \frac{1-t}{N}\right) \overline{c}\right\}\right) \\ &= \frac{1}{(\mathsf{E}_\xi Z)^2} \sum_{|R^t_{1,2}| > (1-(1-t)/N)\overline{c}} \exp y(\sigma^1_N + \sigma^2_N) \\ &\leq \frac{1}{(\mathsf{E}_\xi Z)^2} \sum_{|R^-_{1,2}| > 3\overline{c}/4} \exp y(\sigma^1_N + \sigma^2_N) \\ &\leq \frac{1}{(\mathsf{E}_\xi Z)^2} \mathrm{ch}^2 y \, \mathrm{card}\left\{(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) \; ; \; |R^-_{1,2}| > \frac{3\overline{c}}{4}\right\} , \end{aligned}$$

because the condition $|R^-_{1,2}| \geq 3\overline{c}/4$ does not depend on the value of σ^1_N and σ^2_N. Now (A.18) implies

$$\begin{aligned} \mathrm{card}\left\{(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) \; ; \; |R^-_{1,2}| > \frac{3\overline{c}}{4}\right\} &\leq 2^{2N+1} \exp\left(-\frac{9N\overline{c}^2}{32}\right) \\ &\leq 2^{2N+1} \exp\left(-\frac{N\overline{c}^2}{4}\right) , \end{aligned}$$

and we conclude using (9.41). □

The following will allow to control the term (9.8).

Proposition 9.3.11. *There exists a number $q_0 > 0$ and a number L with the following properties. Assume that (9.39) and (9.40) hold. Consider the probability measure G given by (9.46) and denote by $\langle\cdot\rangle$ an average for G. Consider any number $0 \leq q \leq q_0$. Consider independent standard Gaussian r.v.s z, ξ', and set*

$$w(\boldsymbol{\sigma}) = \sqrt{v} S_{M,t}(\boldsymbol{\sigma}) + \sqrt{1-v}(z\sqrt{q} + \xi' \sqrt{1-q}) \ . \tag{9.48}$$

Denote by E' *expectation in the r.v.s* $g_{i,M}$ *and* z*. Then, for* $\overline{b} = L\exp(-N/L)$ *we have*

$$\mathsf{E}' \frac{1}{\max(\overline{b}, \mathsf{E}_\xi \langle \exp u(w(\boldsymbol{\sigma}))\rangle)^{12}} \le L \exp L\tau^2 \ , \tag{9.49}$$

where E_ξ *denotes expectation in* ξ' *and* ξ_M*.*

Proof. Be begin the proof by a few observations. Let us denote by E_g expectation in the r.v.s $g_{i,M}$ only (given z). Let

$$w'(\boldsymbol{\sigma}) = \frac{1}{\sqrt{1-(1-t)/N}} \left(\frac{1}{\sqrt{N}} \sum_{i<N} g_{i,M}\sigma_i + \sqrt{\frac{t}{N}} g_{N,M}\sigma_N \right) \ .$$

The purpose of the factor $1/\sqrt{1-(1-t)/N}$ is to ensure that $\mathsf{E}_g w'(\boldsymbol{\sigma})^2 = 1$ in order to apply Proposition 9.3.5. We have

$$\mathsf{E}_g w'(\boldsymbol{\sigma}^1) w'(\boldsymbol{\sigma}^2) = \frac{1}{1-(1-t)/N} R^t_{1,2} \ .$$

It follows from Proposition 9.3.10 that the family $w'(\boldsymbol{\sigma})$ satisfies the conditions of Proposition 9.3.5 for the probability measure G given by (9.46) and for the value $d = 8/\overline{c}^2$. Since d is a universal constant, the number b in (9.36) is of the type $L\exp(-N/L)$ and will from now on be denoted by $\overline{b}$. Therefore by (9.36) for any number $\tau' \ge 0$ we have

$$\mathsf{E}_g \frac{1}{\max(b, G(\{\boldsymbol{\sigma}\ ;\ w'(\boldsymbol{\sigma}) \ge \tau'\}))^{12}} \le L \exp L\tau'^2 \ . \tag{9.50}$$

We now start the proof of (9.49). We observe that

$$w(\boldsymbol{\sigma}) = \sqrt{v}\sqrt{1 - \frac{1-t}{N}} w'(\boldsymbol{\sigma}) + \sqrt{v}\sqrt{\frac{1-t}{N}} \xi_M + \sqrt{1-v}(z\sqrt{q} + \xi'\sqrt{1-q}). \tag{9.51}$$

Case 1 We have $v \ge 1/2$. We set

$$d = \frac{1}{2}\mathcal{N}\left(-\frac{z\sqrt{q}}{\sqrt{1-q}}\right)$$

where $\mathcal{N}(s) = \mathsf{P}(\xi' \ge s)$. Let us define

$$\tau' = \frac{\tau}{\sqrt{v}\sqrt{1-(1-t)/N}} \ .$$

Then (9.51) implies

$$\xi' \ge -\frac{z\sqrt{q}}{\sqrt{1-q}}\ ;\ \ \xi_M \ge 0\ ,\ \ w'(\boldsymbol{\sigma}) \ge \tau' \ \ \Rightarrow\ \ w(\boldsymbol{\sigma}) \ge \tau \ \Rightarrow\ \exp u(w(\boldsymbol{\sigma})) = 1\ .$$

Therefore, if $w'(\boldsymbol{\sigma}) \geq \tau'$, we have

$$\mathsf{E}_\xi \exp u(w(\boldsymbol{\sigma})) \geq \mathsf{P}\left(\xi' \geq -\frac{z\sqrt{q}}{\sqrt{1-q}}, \xi_M \geq 0\right) = d \ ,$$

and thus

$$\mathsf{E}_\xi \langle \exp u(w(\boldsymbol{\sigma}))\rangle = \langle \mathsf{E}_\xi \exp u(w(\boldsymbol{\sigma}))\rangle \geq dG(\{w'(\boldsymbol{\sigma}) \geq \tau'\}) \ , \tag{9.52}$$

so that (9.52) yields, using (9.50),

$$\begin{aligned}\mathsf{E}_g \frac{1}{\max(\overline{b}, \mathsf{E}_\xi \langle \exp u(w(\boldsymbol{\sigma}))\rangle)^{12}} &\leq \frac{1}{d^{12}} \mathsf{E}_g \frac{1}{\max(\overline{b}, G(\{w'(\boldsymbol{\sigma}) \geq \tau'\}))^{12}} \\ &\leq \frac{L}{d^{12}} \exp L\tau'^2 \\ &\leq \frac{L}{d^{12}} \exp L\tau^2 \ , \end{aligned} \tag{9.53}$$

since $\tau' \leq 2\tau$. Now, using the rough estimate $\mathcal{N}(s) \geq \exp(-s^2)/L$, we have

$$\frac{1}{d^{12}} = 2^{12}\mathcal{N}\left(-\frac{z\sqrt{q}}{\sqrt{1-q}}\right)^{-12} \leq L \exp\left(\frac{12 z^2 q}{1-q}\right) \ ,$$

so that if $q \leq q_0$ we have $\mathsf{E}d^{-12} \leq L$ and taking expectation in z in (9.53) the result follows.

Case 2 We have $v \leq 1/2$. Then (9.51) implies

$$\xi' \geq -\frac{z\sqrt{q}}{\sqrt{1-q}} + 2\tau \ ; \quad \xi_M \geq 0 \ , \quad w'(\boldsymbol{\sigma}) \geq 0 \quad \Rightarrow \quad w(\boldsymbol{\sigma}) \geq \tau \ ,$$

and thus (9.52) holds now for

$$\tau' = 0 \ , \quad d = \mathcal{N}\left(-\frac{z\sqrt{q}}{\sqrt{1-q}} + 2\tau\right) \ ,$$

and we proceed as before, using that

$$\mathsf{E} \exp 12\left(2\tau - \frac{z\sqrt{q}}{\sqrt{1-q}}\right)^2 \leq L \exp L\tau^2$$

if $q \leq q_0$. □

9.4 Integration by Parts

If g is a standard Gaussian r.v. and U is a smooth function (of moderate growth), the size of $\mathsf{E}U'(g)$ is governed by the size of U rather than by the size of U' because, by integration by parts,

$$|\mathsf{E}U'(g)| = |\mathsf{E}gU(g)| \leq \mathsf{E}|g||U(g)| \leq L \sup |U| \ .$$

More generally, we have the following elementary fact.

Lemma 9.4.1. *Consider independent standard Gaussian r.v.s* $h_1, \ldots, h_n$, *a smooth function* V *of* n *variables, and integers* $k_1, \ldots, k_n$. *Let* $k = \sum_{i \le n} k_i$. *Then*

$$\left| \mathsf{E} \frac{\partial^k V}{\partial x_1^{k_1} \cdots \partial x_n^{k_n}} (h_1, \ldots, h_n) \right| \le C(k_1, \ldots, k_n) \sup |V| \,, \tag{9.54}$$

where the number $C(k_1, \ldots, k_n)$ *depends only on* $k_1, \ldots, k_n$. *In fact, more generally, if* $\ell_1, \ldots, \ell_n$ *are integers* ≥ 0 *then*

$$\left| \mathsf{E} h_1^{\ell_1} \cdots h_n^{\ell_n} \frac{\partial^k V}{\partial x_1^{k_1} \cdots \partial x_n^{k_n}} (h_1, \ldots, h_n) \right| \le C(k_1, \ldots, k_n, \ell_1, \ldots, \ell_n) \sup |V| \,. \tag{9.55}$$

Proof. The proof goes by induction over k. For $k = 0$ the result is obvious. Assuming that (9.55) has been proved for $k - 1$ (and all values of $\ell_1, \ldots, \ell_n, k_1, \ldots, k_n$ such that $\sum_{i \le n} k_i = k - 1$) we prove it for k. we may and do assume that $k_1 \ge 1$, and we simply write, using integration by parts in h_1, that

$$\begin{aligned} &\mathsf{E} h_1^{\ell_1} \cdots h_n^{\ell_n} \frac{\partial^k V}{\partial x_1^{k_1} \cdots \partial x_n^{k_n}} (h_1, \ldots, h_n) \\ &= \mathsf{E} h_1^{\ell_1 + 1} \cdots h_n^{\ell_n} \frac{\partial^{k-1} V}{\partial x_1^{k_1 - 1} \partial x_2^{k_2} \cdots \partial x_n^{k_n}} (h_1, \ldots, h_n) \\ &- \ell_1 \mathsf{E} h_1^{\ell_1 - 1} \cdots h_n^{\ell_n} \frac{\partial^{k-1} V}{\partial x_1^{k_1 - 1} \partial x_2^{k_2} \cdots \partial x_n^{k_n}} (h_1, \ldots, h_n) \,, \end{aligned}$$

and the proof is complete. □

We proved (9.54) when $h_1, \ldots, h_n$ are independent. Certainly (9.54) will not hold without any condition on $h_1, \ldots, h_n$. For example, at the opposite from the independence situation consider the pair (h, h), a function f of one variable and $U(x_1, x_2) = f(x_1 - x_2)$, so that

$$\frac{\partial^2 U}{\partial x_1 \partial x_2} (x_1, x_2) = -f''(x_1 - x_2)$$

and

$$\mathsf{E} \frac{\partial^2 U}{\partial x_1 \partial x_2} (h, h) = -f''(0)$$

is certainly not controlled by $\sup |f|$. Still, it turns out that (9.54) will hold provided there is enough independence between the r.v.s $h_1, \ldots, h_n$. To see this, assume that there is a linear invertible operator $T : \mathbb{R}^n \to \mathbb{R}^n$ such that

$$T(h_1, \ldots, h_n) = (w_1, \ldots, w_n) \,, \tag{9.56}$$

where the sequence $(w_1, \ldots, w_n)$ consists of independent standard Gaussian r.v.s. That is, T is given by a invertible matrix $(a_{\ell,\ell'})$, and (9.56) means that $w_\ell = \sum_{\ell' \leq n} a_{\ell,\ell'} h_{\ell'}$. Consider the function $V = U \circ T^{-1}$ of n variables, so that $U = V \circ T$. Each term

$$\mathsf{E}\frac{\partial^k U}{\partial x_1^{k_1} \cdots \partial x_n^{k_n}}(h_1, \ldots, h_n)$$

is a linear combination of terms

$$\mathsf{E}\frac{\partial^k V}{\partial x_1^{\ell_1} \cdots \partial x_n^{\ell_n}}(T(h_1, \ldots, h_n)) \tag{9.57}$$

where $\ell_1 + \ldots + \ell_n = k$. The coefficients of this linear combination are determined by the coefficients of the matrix T. Using (9.56) and (9.54), each term (9.57) is controlled by $\sup |V| = \sup |U|$. Thus

$$\left| \mathsf{E}\frac{\partial^k U}{\partial x_1^{k_1} \cdots \partial x_n^{k_n}}(h_1, \ldots, h_n) \right| \leq C(k_1, \ldots, k_n, \|T\|) \sup |U| \tag{9.58}$$

where $\|T\|$ is the size of the largest coefficient of the matrix T and the quantity $C(k_1, \ldots, k_n, \|T\|)$ depends only on $k_1, \ldots, k_n$ and $\|T\|$.

Here is a simple condition under which one can control $\|T\|$.

Definition 9.4.2. *A jointly Gaussian sequence* $(h_1, \ldots, h_n)$ *is* widely spread *if for each* $\ell \leq n$ *we have* $\mathsf{E}h_\ell^2 \leq 1$ *and there exists a Gaussian r.v.* z_ℓ *with* $\mathsf{E}z_\ell^2 \leq 1$*,* $\mathsf{E}z_\ell h_\ell \geq 1/8$ *and* $\mathsf{E}z_\ell h_{\ell'} = 0$ *for* $\ell \neq \ell'$*.*

Of course here we assume that the whole family $(h_1, \ldots, h_n, z_1, \ldots, z_n)$ is jointly Gaussian. Equivalently, we may assume that the r.v.s z_ℓ belong to the linear span of $h_1, \ldots, h_n$. The choice of the constant $1/8$ is quite arbitrary.

It often helps to think in geometrical terms. This is the case here: consider the space W of linear combinations $h = \sum_{\ell \leq n} a_\ell h_\ell$ provided with the scalar product $(h, h') = \mathsf{E}hh'$. Given $\ell \leq n$, consider the linear span W_ℓ of $h_1, \ldots, h_{\ell-1}, h_{\ell+1}, \ldots, h_n$. Then

$$\sup\bigl\{(z, h_\ell)\ ;\ z \in W\ ;\ \|z\|^2 = 1\ ;\ \forall \ell \neq \ell'\ ,\ (z, h_{\ell'}) = 0\bigr\}$$

is the distance from h_ℓ to W_ℓ. So, the sequence $(h_1, \ldots, h_n)$ is widely spread if and only if for each ℓ this distance is $\geq 1/8$.

When

$$w = \sum_{\ell' \leq n} a_{\ell'} h_{\ell'}\ ,$$

and if z_ℓ is as provided by the hypothesis of Definition 9.4.2, i.e. $\mathsf{E}z_\ell^2 \leq 1$, $\mathsf{E}z_\ell h_\ell \geq 1/8$ and $\mathsf{E}z_\ell h_{\ell'} = 0$ for $\ell \neq \ell'$, we have

$$\|w\| \geq |(z_\ell, w)| = |a_\ell(z_\ell, h_\ell)| \geq \left|\frac{a_\ell}{8}\right| ,$$

so that $|a_\ell| \leq 8\|w\|$. It should be obvious that W is n-dimensional. Consider any orthonormal basis $w_1, \ldots, w_n$ of W, so that the sequence $w_1, \ldots, w_n$ is i.i.d. standard normal. We have just shown that the matrix of the map T such that (9.56) holds satisfies $\|T\| \leq 8$. Thus we have proved the following.

Proposition 9.4.3. *If the sequence $(h_\ell)_{\ell \leq n}$ is widely spread then*

$$\left|\mathsf{E}\frac{\partial^k U}{\partial x_1^{k_1} \cdots \partial x_n^{k_n}}(h_1, \ldots, h_n)\right| \leq C(k) \sup |U| , \tag{9.59}$$

where $C(k)$ depends only on $k = k_1 + \cdots + k_n$.

We now show that widely spread sequences occur naturally.

Proposition 9.4.4. *Consider a probability measure G on $\Sigma_N = \{-1, 1\}^N$ and assume that*

$$\forall \mathbf{x} \in \mathbb{R}^N , \quad G\bigg(\bigg\{\boldsymbol{\sigma} \; ; \sum_{i \leq N-1} |\sigma_i - x_i|^2 \leq \frac{N}{16}\bigg\}\bigg) \leq 4 \exp\left(-\frac{N}{32}\right) . \tag{9.60}$$

For $\boldsymbol{\sigma}$ in Σ_N, let $h(\boldsymbol{\sigma}) = N^{-1/2}\sum_{i<N} g_i \sigma_i$. Then

$$\begin{aligned} &G^{\otimes n}(\{(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n) \; ; \; (h(\boldsymbol{\sigma}^1), \ldots, h(\boldsymbol{\sigma}^n)) \,\text{is widely spread}\,\}) \\ &\geq 1 - L^n \exp\left(-\frac{N}{32}\right) . \end{aligned} \tag{9.61}$$

Proof. As a first step, given $\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^{n-1} \in \Sigma_N$ we show that $G(A) \geq 1 - L^n \exp(-N/32)$, where

$$A = \big\{\boldsymbol{\sigma} \; ; \; \exists z, \; \mathsf{E}z^2 = 1 \; , \; \mathsf{E}zh(\boldsymbol{\sigma}) \geq 1/8 \; ; \; \forall \ell \leq n-1 \; , \; \mathsf{E}zh(\boldsymbol{\sigma}^\ell) = 0\big\} ,$$

and where z is a Gaussian r.v. that belongs to the linear span of $(g_i)_{i \leq N}$. To prove this statement we consider the space $\mathbb{R}^{N-1}$ provided with the dot product $(\mathbf{x}, \mathbf{y}) = N^{-1}\sum_{i \leq N-1} x_i y_i$ and the associated distance. The condition $\boldsymbol{\sigma} \in A^c$ means that $\boldsymbol{\rho} = (\sigma_1, \ldots, \sigma_{N-1})$ is at distance $< 1/8$ from the linear span W of $\boldsymbol{\rho}^1, \ldots, \boldsymbol{\rho}^{n-1}$. According to Proposition A.7.1 we may find a subset F of W with card$F \leq L^n$ such that any point of the unit ball of W is within distance $1/8$ of F. Then if the distance of $\boldsymbol{\rho}$ to W is $\leq 1/8$, since $\boldsymbol{\rho}$ is of norm ≤ 1, $\boldsymbol{\rho}$ is within distance $1/8$ of the unit ball of W, so is within distance $\leq 1/4$ of F. Thus

$$G(A^c) \leq G\bigg(\bigg\{\boldsymbol{\sigma} \; ; \; \exists \mathbf{x} \in F \; ; \sum_{i \leq N-1} |\sigma_i - x_i|^2 \leq \frac{N}{16}\bigg\}\bigg) \leq L^n \exp\left(-\frac{N}{32}\right)$$

by (9.60) and this completes the proof that $G(A) \geq 1 - L^n \exp(-N/32)$. We then use Fubini Theorem to obtain that if

$$B_n = \left\{\boldsymbol{\sigma}^1, \cdots \boldsymbol{\sigma}^n \ ; \ \exists z, \ \mathsf{E}z^2 = 1, \ \mathsf{E}zh(\boldsymbol{\sigma}^n) \geq 1/8, \ \forall \ell \leq n-1, \ \mathsf{E}zh(\boldsymbol{\sigma}^\ell) = 0\right\} ,$$

then

$$G^{\otimes n}(B_n) \geq 1 - L^n \exp\left(-\frac{N}{32}\right) ,$$

and therefore

$$\begin{aligned}&G^{\otimes n}(\{(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n) \ ; \ (h(\boldsymbol{\sigma}^1), \ldots, h(\boldsymbol{\sigma}^n)) \text{ is widely spread}\})\\ &\geq 1 - nL^n \exp\left(-\frac{N}{32}\right) ,\end{aligned}$$

which completes the proof. □

Condition (9.60) itself occurs naturally, as the following shows.

Proposition 9.4.5. *Assume (9.39) and (9.40), and consider G as in (9.46). Then G satisfies (9.60).*

Proof. If Z is as in (9.45), then

$$G\left(\left\{\boldsymbol{\sigma} \ ; \ \sum_{i \leq N-1} (\sigma_i - x_i)^2 \leq \frac{N}{16}\right\}\right) \leq \frac{1}{\mathsf{E}_\xi Z} \sum_B \exp y\sigma_N , \tag{9.62}$$

where the summation is over the set

$$B = \left\{\boldsymbol{\sigma} \ ; \ \sum_{i \leq N-1} (\sigma_i - x_i)^2 \leq \frac{N}{16}\right\} .$$

Since B does not depend on the last coordinate, we have

$$\sum_B \exp y\sigma_N = \operatorname{ch} y \operatorname{card} B ,$$

and by (9.41) the right-hand side of (9.62) is $\leq \exp(N\overline{c}^2/16)2^{-N+1}\operatorname{card}B$. Next we proceed to bound $\operatorname{card}B$. Consider $\lambda = 1/2$, so that $\exp(-\lambda) \leq 1 - \lambda/2$ and $1 + \exp(-\lambda) \leq 2(1 - \lambda/4)$. Since for each i either $|1 - x_i| \geq 1$ or $|1 + x_i| \geq 1$, we have

$$\begin{aligned}&\sum_{\boldsymbol{\sigma}} \exp\left(-\lambda \sum_{i \leq N-1} (x_i - \sigma_i)^2\right)\\ &= 2 \prod_{i \leq N-1} (\exp(-\lambda(1 + x_i)^2) + \exp(-\lambda(1 - x_i)^2))\\ &\leq 2(1 + \exp(-\lambda))^{N-1}\\ &\leq 2^N \left(1 - \frac{\lambda}{4}\right)^{N-1}\\ &\leq 2^N \exp\left(-\frac{\lambda}{4}(N-1)\right)\end{aligned}$$

so that

$$\operatorname{card} B \exp\left(-\frac{\lambda N}{16}\right) \le \sum_{\boldsymbol{\sigma}\in B} \exp\Bigl(-\lambda \sum_{i\le N-1}(\sigma_i - x_i)^2\Bigr) \le 2^N \exp\left(-\frac{\lambda}{4}(N-1)\right)$$

i.e., since $\lambda = 1/2$,

$$\operatorname{card} B \le 2^{N+1} \exp\left(-\frac{N}{16}\right) .$$

Since we may assume $\overline{c} \le 1/2$, we have

$$\exp(N\overline{c}^2/16)2^{-N+1}\operatorname{card} B \le L \exp(-N/32)$$

and the result follows. □

Our final technical result will allow us to deal with r.v.s such as in (9.48).

Proposition 9.4.6. *Assume that the sequence $(h_\ell)_{\ell\le n}$ is widely spread. Consider a number $q \le 1/2$ and Gaussian r.v.s h'_ℓ, z, ξ^ℓ. We assume that the r.v.s (h'_ℓ) are independent of the r.v.s (h_ℓ), and that the r.v.s z, ξ^ℓ are independent of the r.v.s h_ℓ and h'_ℓ. Then the sequence*

$$w_\ell = \sqrt{1-v}(z\sqrt{q} + \sqrt{1-q}\xi^\ell) + \sqrt{v}(h_\ell + h'_\ell)$$

is widely spread.

Proof. Since the sequence h_ℓ is widely spread, by definition, for $\ell \le n$ there exists a Gaussian r.v. z_ℓ with $\mathsf{E} z_\ell^2 = 1$, $\mathsf{E} z_\ell h_\ell \ge 1/8$, $\mathsf{E} z_\ell h_{\ell'} = 0$ if $\ell \ne \ell'$. we may assume that $\mathsf{E} z_\ell \xi^{\ell'} = \mathsf{E} z_\ell h'_{\ell'} = 0$ for each ℓ'. The Gaussian r.v.

$$g_\ell = \sqrt{1-v}\xi^\ell + \sqrt{v} z_\ell$$

satisfies $\mathsf{E} g_\ell^2 = 1$, $\mathsf{E} g_\ell w_\ell = (1-v)\sqrt{1-q} + v\mathsf{E} z_\ell h_\ell \ge 1/8$ and $\mathsf{E} g_\ell w_{\ell'} = 0$ if $\ell \ne \ell'$. □

9.5 The Replica Symmetric Solution

We have built the tools necessary to accomplish the program outlined in Section 9.1, and now we will perform the steps of this program in detail. We recall the number τ of (9.2).

Theorem 9.5.1. *There exists a number L with the following property. Consider a function u that satisfies (9.2), and assume that*

$$\forall \ell,\ 1 \le \ell \le 5\ ,\quad |u^{(\ell)}| \le \exp\left(\frac{N}{L}\right)\ . \tag{9.63}$$

Consider α with

$$L\alpha \exp L\tau^2 \le 1\ . \tag{9.64}$$

Then, if z and ξ are independent standard normal r.v.s, the system of equations with unknown (r, p)

$$q = \mathsf{E}\mathrm{th}^2(z\sqrt{r})\ ;\qquad r = \frac{\alpha}{1-q}\mathsf{E}\left(\frac{\mathsf{E}_\xi \xi \exp u(\theta)}{\mathsf{E}_\xi \exp u(\theta)}\right)^2\ , \tag{9.65}$$

where $\theta = z\sqrt{q} + \xi\sqrt{1-q}$ has a unique solution. Moreover, consider the system with Hamiltonian (9.1). Then if $\alpha = M/N$ satisfies (9.64) and if q is as in (9.65) we have

$$\nu\big((R_{1,2} - q)^2\big) \le \frac{L}{N}\ . \tag{9.66}$$

The control of the first 5 derivatives in (9.63) is assumed as a blanket assumption for further use. The reader can check that to prove (9.66) it would suffice to control the first three derivatives.

Let us first study the system of equations (9.65). To compare with the equations (2.66) we recall that by integration by parts we have

$$\widehat{r}(q) := \frac{1}{1-q}\mathsf{E}\left(\frac{\mathsf{E}_\xi \xi \exp u(\theta)}{\mathsf{E}_\xi \exp u(\theta)}\right)^2 = \mathsf{E}\left(\frac{\mathsf{E}_\xi u'(\theta) \exp u(\theta)}{\mathsf{E}_\xi \exp u(\theta)}\right)^2\ . \tag{9.67}$$

Let us define

$$Y = \frac{\tau - z\sqrt{q}}{\sqrt{1-q}}\ .$$

We will prove first that

$$\left(\frac{\mathsf{E}_\xi \xi \exp u(\theta)}{\mathsf{E}_\xi \exp u(\theta)}\right)^2 \le L(Y^2 + 1)\ . \tag{9.68}$$

Since $u \le 0$ and

$$\xi \ge Y \Rightarrow \theta \ge \tau \Rightarrow u(\theta) \ge 0\ ,$$

we have

$$\mathsf{E}_\xi \exp u(\theta) \ge \mathsf{P}_\xi(\xi \ge Y)\ , \tag{9.69}$$

denoting by P_ξ the probability corresponding to E_ξ. Thus (9.68) is obvious when $Y \le 1$, since $|\mathsf{E}_\xi \exp u(\theta)| \le \mathsf{E}|\xi| \le L$ and $\mathsf{E}_\xi \exp u(\theta) \ge 1/L$. When $Y \ge 1$, it holds

$$\begin{aligned}|\mathsf{E}_\xi \xi \exp u(\theta)| &\le \mathsf{E}_\xi \mathbf{1}_{\{|\xi|\le Y\}} |\xi| \exp u(\theta) + \mathsf{E}_\xi \mathbf{1}_{\{|\xi|>Y\}} |\xi| \exp u(\theta) \\ &\le Y \mathsf{E}_\xi \exp u(\theta) + \mathsf{E}_\xi \mathbf{1}_{\{|\xi|>Y\}} |\xi| \ .\end{aligned}$$

Therefore

$$\left| \frac{\mathsf{E}_\xi \xi \exp u(\theta)}{\mathsf{E}_\xi \exp u(\theta)} \right| \le Y + \frac{E_\xi \mathbf{1}_{\{|\xi|>Y\}} |\xi|}{\mathsf{E}_\xi \exp u(\theta)} \ . \tag{9.70}$$

We observe that

$$\mathsf{E}_\xi \mathbf{1}_{\{|\xi|>Y\}} |\xi| = \frac{2}{\sqrt{2\pi}} \int_Y^\infty x \exp\Big(-\frac{x^2}{2}\Big) \mathrm{d}x = \frac{2}{\sqrt{2\pi}} \exp\Big(-\frac{Y^2}{2}\Big) \ ,$$

and that, by (3.136) and (9.69) we have

$$\mathsf{E}_\xi \exp u(\theta) \ge \frac{Y}{1+Y^2} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{Y^2}{2}\right) \ ,$$

so combining with (9.70) we obtain

$$\left| \frac{\mathsf{E}_\xi \xi \exp u(\theta)}{\mathsf{E}_\xi \exp u(\theta)} \right| \le Y + L\frac{Y^2+1}{Y} \le LY + 1 \ .$$

This implies (9.68) and thus, going back to (9.67), we have

$$\widehat{r}(q) \le \frac{L(1+\tau^2)}{(1-q)^2} \ .$$

Therefore

$$q \le \frac{1}{2} \quad \Rightarrow \quad \widehat{r}(q) \le L(1+\tau^2) \ ,$$

so that if the constant in (9.64) is large enough then

$$q \le \frac{1}{2} \quad \Rightarrow \quad \alpha\widehat{r}(q) \le \frac{1}{2}$$

and since $\mathsf{E}\mathrm{th}^2(z\sqrt{r}) \le \mathsf{E} z^2 r \le r$ the continuous function

$$q \mapsto \psi(q) = \mathsf{E}\mathrm{th}^2\big(z\sqrt{\alpha\widehat{r}(q)}\big) \tag{9.71}$$

maps the interval $[0, 1/2]$ into itself; so the equation $q = \psi(q)$ has a solution.

To show that this solution is unique, one simply works harder along the same lines to prove that $|\psi'| < 1$. There is no point however to complete the details, since our argument will show that (9.66) holds for any solution of (9.65), and that therefore this solution is unique.

We turn to the proof of (9.66). We fix once and for all a solution (q, r) of the equations (9.65) and recalling (9.67) we set $\widehat{r} = \widehat{r}(q)$. As we explained in Section 9.1, the key to the results of the present chapter is a better estimate than (2.40) when using the "cavity in M" method; and we turn to this now. We think of t as fixed, and given a function f on 4 replicas, we recall that

$$\nu_{t,v}(fu'(S_v^\ell)u'(S_v^{\ell'})) = \mathsf{E}\frac{\left\langle f\mathsf{E}_\xi u'(S_v^\ell)u'(S_v^{\ell'})\exp\sum_{m\le 4}u(S_v^m)\right\rangle_{t,\sim}}{\langle \mathsf{E}_\xi \exp(S_v^1)\rangle_{t,\sim}^4} \ , \quad (9.72)$$

where E_ξ denotes expectation in the randomness of the variables ξ^ℓ, ξ_M^ℓ, where

$$S_v^\ell = S_v(\boldsymbol{\sigma}^\ell, \xi_M^\ell) = \sqrt{v}S_{M,t}(\boldsymbol{\sigma}^\ell, \xi_M^\ell) + \sqrt{1-v}(z\sqrt{q} + \xi^\ell\sqrt{1-q}) \quad (9.73)$$

$$S_{M,t}(\boldsymbol{\sigma}, \xi) = \sum_{i<N-1}\frac{1}{\sqrt{N}}g_{i,M}\sigma_i + \sqrt{\frac{t}{N}}g_{N,M}\sigma_N + \sqrt{\frac{1-t}{N}}\xi \ ,$$

and where $\langle\cdot\rangle_{t,\sim}$ is the Gibbs average corresponding to the Hamiltonian (2.30).

Proposition 9.5.2. *Consider a function f on Σ_N^4, and*

$$\varphi(v) = \nu_{t,v}(fu'(S_v^\ell)u'(S_v^{\ell'})) \quad \text{or} \quad \varphi(v) = \nu_{t,v}(f) \ . \quad (9.74)$$

Assume that D is as in (9.3), and that (9.64) holds. Then

$$|\varphi'(v)| \le L\exp L\tau^2\Bigg(\sum_{\ell_1,\ell_2\le 6, \ell_1\ne\ell_2}\mathsf{E}\langle|f||R_{\ell_1,\ell_2} - q|\rangle_{t,\sim}$$
$$+\frac{1}{N}\mathsf{E}\langle|f|\rangle_\sim + \max|f|D^4\exp\left(-\frac{N}{L}\right)\Bigg) \ . \quad (9.75)$$

Proof. We observe the very important fact that, if $\ell \ne \ell'$, we have

$$u'(x_\ell)u'(x_{\ell'})\exp\sum_{m\le 4}u(x_m) = \frac{\partial^2}{\partial x_\ell \partial x_{\ell'}}\exp\sum_{m\le 4}u(x_m) \ . \quad (9.76)$$

To compute $\varphi'(v)$ we differentiate the relation (9.72) and we integrate by parts in all the Gaussian r.v.s occurring in S_v^ℓ. We recall the notation

$$R_{1,2}^t = \frac{1}{N}\sum_{i\le N-1}\sigma_i^1\sigma_i^2 + \frac{t}{N}\sigma_N^1\sigma_N^2 \ .$$

Setting $S_v^{\ell\prime} = \partial S_v^\ell/\partial v$, (a quantity that should not be confused with $S_v^{\ell'}$) we see that $\mathsf{E}S_v^{\ell\prime}S_v^{\ell'} = (R_{\ell,\ell'}^t - q)/2$. We have explained in great detail in the proof of Lemma 2.3.2 how to compute $\varphi'(v)$ using integration by parts. This argument shows that $\varphi'(v)$ is a linear combination of terms

$$\mathsf{E}\frac{\langle f(R_{\ell_1,\ell_2}^t - q)\mathsf{E}_\xi V\rangle_{t,\sim}}{\langle \mathsf{E}_\xi \exp u(S_v^1)\rangle_{t,\sim}^6} \ , \quad (9.77)$$

where

$$V = V(S_v^1, \ldots, S_v^6)$$

and

$$V(x_1, \ldots, x_6) = \frac{\partial^k}{\partial x_1^{k_1} \cdots \partial x_6^{k_6}} \exp \sum_{m \le 6} u(x_m) \ ,$$

for integers $k_1, \ldots, k_6$ with $k = \sum_{m \le 6} k_m \le 4$, and $k_m \le 3$. Specifically, $k = 2$ when $\varphi(v) = \nu_{t,v}(f)$ and $k = 4$ when $\varphi(v) = \nu_{t,v}(fu'(S_v^\ell)u'(S_v^{\ell'}))$.

Consider the (exceptional) event Ω that (9.39) or (9.40) fail. Using Lemma 9.3.8 we obtain that if the constant in (9.64) is large enough, then

$$\mathsf{P}(\Omega) \le L \exp\left(-\frac{N}{L}\right) .$$

On the other hand, since $V \le LD^4 \exp(\sum_{m \le 6} u(S_v^m))$, we have

$$\langle \mathsf{E}_\xi V \rangle_{t,\sim} \le D^4 \langle \mathsf{E}_\xi \exp \sum_{m \le 6} u(S_v^m) \rangle_{t,\sim} = \langle \mathsf{E}_\xi \exp u(S_v^1) \rangle_{t,\sim}^6 \ ,$$

so that

$$\left| \frac{\langle f(R_{\ell_1,\ell_2}^t - q) \mathsf{E}_\xi V \rangle_{t,\sim}}{\langle \mathsf{E}_\xi \exp u(S_v^1) \rangle_{t,\sim}^6} \right| \le 2 \max |f| \frac{\langle \mathsf{E}_\xi V \rangle_{t,\sim}}{\langle \mathsf{E}_\xi \exp u(S_v^1) \rangle_{t,\sim}^6} \le L \max |f| D^4 \ . \tag{9.78}$$

Therefore

$$\mathsf{E} \mathbf{1}_\Omega \left| \frac{\langle f(R_{\ell_1,\ell_2}^t - q) \mathsf{E}_\xi V \rangle_{t,\sim}}{\langle \mathsf{E}_\xi \exp u(S_v^1) \rangle_{t,\sim}^6} \right| \le L \max |f| D^4 \exp\left(-\frac{N}{L}\right) . \tag{9.79}$$

This controls what happens on the exceptional event Ω and we turn to the control of what happens on the "generic" event Ω^c. Let us denote by E' the expectation in the randomness of z and of the $g_{i,M}$, $i \le N$. This randomness is independent of Ω so that

$$\mathsf{E} \mathbf{1}_{\Omega^c} \left| \frac{\langle f(R_{\ell_1,\ell_2}^t - q) \mathsf{E}_\xi V \rangle_{t,\sim}}{\langle \mathsf{E}_\xi \exp u(S_v^1) \rangle_{t,\sim}^6} \right| = \mathsf{E} \mathbf{1}_{\Omega^c} \mathsf{E}' \left| \frac{\langle f(R_{\ell_1,\ell_2}^t - q) \mathsf{E}_\xi V \rangle_{t,\sim}}{\langle \mathsf{E}_\xi \exp u(S_v^1) \rangle_{t,\sim}^6} \right| . \tag{9.80}$$

Comparing (9.73) and (9.48) we obtain from (9.49) that

$$\mathbf{1}_{\Omega^c} \mathsf{E}' \frac{1}{\max(\overline{b}, \langle \mathsf{E}_\xi \exp u(S_v^1) \rangle_{t,\sim})^{12}} \le L \exp L\tau^2 \ . \tag{9.81}$$

In particular if $\Omega' = \Omega^c \cap \{\langle \mathsf{E}_\xi \exp u(S_v^1) \rangle_{t,\sim} \le \overline{b}\}$, we have (with the obvious notation that P' denotes the probability corresponding to E')

$$\mathsf{P}'(\Omega') \le L\overline{b}^{12} \exp L\tau^2 \ , \tag{9.82}$$

so that since $\overline{b}$ is exponentially small in N, Ω' is another exceptional event. We observe that, using (9.78) and (9.82), we have

$$\mathbf{1}_{\Omega^c}\mathsf{E}'\mathbf{1}_{\Omega'}\left|\frac{\langle f(R^t_{\ell_1,\ell_2}-q)\mathsf{E}_\xi V\rangle_{t,\sim}}{\langle \mathsf{E}_\xi \exp u(S_v^1)\rangle^6_{t,\sim}}\right| \le L\max|f|\overline{b}^{12}D^4\exp L\tau^2$$
$$\le \max|f|D^4\exp L\tau^2\exp\left(-\frac{N}{L}\right) .$$

Having controlled what happens on the exceptional event Ω' we turn to the control of what happens on the generic event Ω'^c. We note that (9.81) implies

$$\mathbf{1}_{\Omega^c}\mathsf{E}'\mathbf{1}_{\Omega'^c}\frac{1}{\langle \mathsf{E}_\xi \exp u(S_v^1)\rangle^{12}_{t,\sim}} \le \mathbf{1}_{\Omega^c}\mathsf{E}'\frac{1}{\max(\overline{b},\langle \mathsf{E}_\xi \exp u(S_v^1)\rangle_{t,\sim})^{12}}$$
$$\le L\exp L\tau^2 .$$

Combining this with the Cauchy-Schwarz inequality we get

$$\mathbf{1}_{\Omega^c}\mathsf{E}'\mathbf{1}_{\Omega'^c}\left|\frac{\langle f(R^t_{\ell_1,\ell_2}-q)\mathsf{E}_\xi V\rangle_{t,\sim}}{\langle \mathsf{E}_\xi \exp u(S_v^1)\rangle^6_{t,\sim}}\right|$$
$$\le \mathbf{1}_{\Omega^c}L\exp L\tau^2(\mathbf{1}_{\Omega^c}\mathsf{E}'\langle f(R^t_{\ell_1,\ell_2}-q)\mathsf{E}_\xi V\rangle^2_{t,\sim})^{1/2} . \tag{9.83}$$

The remainder of the proof consists in controlling the expectation of the quantity (9.83). This is the main argument. We consider a replicated copy f' of f; that is, if $f=f(\boldsymbol{\sigma}^1,\ldots,\boldsymbol{\sigma}^6)$, we set $f'(\boldsymbol{\sigma}^1,\ldots,\boldsymbol{\sigma}^{12})=f(\boldsymbol{\sigma}^7,\ldots,\boldsymbol{\sigma}^{12})$ and we consider

$$f^\sim = ff'(R^t_{\ell_1,\ell_2}-q)(R^t_{\ell_1+6,\ell_2+6}-q) .$$

Thus

$$\langle f(R^t_{\ell_1,\ell_2}-q)\mathsf{E}_\xi V\rangle^2_{t,\sim} = \langle f^\sim \mathsf{E}_\xi W\rangle_{t,\sim} , \tag{9.84}$$

where

$$W(S_v^1,\ldots,S_v^{12}) = V(S_v^1,\ldots,S_v^6)V(S_v^7,\ldots,S_v^{12}) .$$

In particular W is of the type

$$W(x^1,\ldots,x^{12}) = \frac{\partial^k}{\partial x_1^{k_1}\cdots\partial x_{12}^{k_{12}}}\exp\sum_{m\le 12}u(x_m)$$

for integers $k_1,\ldots,k_{12}$ with $\sum_{m\le 12}k_m\le 8$ (and $k_m\le 3$). From (9.84) we have

$$\mathbf{1}_{\Omega^c}\mathsf{E}'\langle f(R^t_{\ell_1,\ell_2}-q)\mathsf{E}_\xi V\rangle^2_{t,\sim} = \mathbf{1}_{\Omega^c}\langle f^\sim \mathsf{E}'\mathsf{E}_\xi W\rangle_{t,\sim} = \mathrm{I}+\mathrm{II} , \tag{9.85}$$

where I is the contribution to $\langle\cdot\rangle_{t,\sim}$ of all the configurations for which the sequence $S_v^1,\ldots,S_v^{12}$ is widely spread, and II is the contribution of the other configurations. That is, if

$$A = \{(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^{12}) \; ; \; S_v^1, \ldots, S_v^{12} \text{ is widely spread}\} ,$$

$$\mathrm{I} = \mathbf{1}_{\Omega^c} \langle \mathbf{1}_A f^\sim \mathsf{E}' \mathsf{E}_\xi W \rangle_{t,\sim} ,$$

$$\mathrm{II} = \mathbf{1}_{\Omega^c} \langle \mathbf{1}_{A^c} f^\sim \mathsf{E}' \mathsf{E}_\xi W \rangle_{t,\sim} .$$

We use Proposition 9.4.3 with $U(x_1, \ldots, x_{12}) = \exp \sum_{m \le 12} u(x_m)$, so that $0 \le U \le 1$ and we obtain

$$\mathrm{I} \le L \langle |f^\sim| \rangle_{t,\sim} \le L \langle |f| |R^t_{\ell_1,\ell_2} - q| \rangle^2_{t,\sim} .$$

The essential point of the proof is that only the bound for U, and not the much larger bound for W occurs here.

We recall that by definition of the event Ω, conditions (9.39) and (9.40) hold on Ω^c, and the probability G on Σ_N corresponding to the averages $\langle \cdot \rangle_{t,\sim}$ is of the type (9.46). Propositions 9.4.5 and 9.4.4 then imply that the set of configurations $(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^{12})$ for which the sequence $h_1, \cdots, h_{12}$ is *not* widely spread is exponentially small for G, where

$$h_\ell = h(\boldsymbol{\sigma}^\ell) = \frac{1}{\sqrt{N}} \sum_{i<N} g_{i,M} \sigma_i^\ell .$$

We then use Proposition 9.4.6 with

$$h'_\ell = \sqrt{\frac{t}{N}} g_{N,M} \sigma_N^\ell + \sqrt{\frac{1-t}{N}} \xi_M^\ell$$

to obtain that the set A^c of configurations $(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^{12})$ for which the sequence $S_v^1, \ldots, S_v^{12}$ is *not* widely spread is exponentially small, and since $|W| \le L D^8 \exp \sum_{m \le 12} u(x_m)$ we get

$$\mathrm{II} \le L \max |f| \exp\left(-\frac{N}{L}\right) D^8 ,$$

so that (9.85) implies

$$\mathbf{1}_{\Omega^c} (\mathsf{E}' \langle f (R^t_{\ell_1,\ell_2} - q) \mathsf{E}_\xi V \rangle^2_{t,\sim})^{1/2} \le L \langle |f| |R^t_{\ell_1,\ell_2} - q| \rangle_{t,\sim} + L \max |f| \exp\left(-\frac{N}{L}\right) D^4 . \quad (9.86)$$

Finally, we write

$$\langle |f| |R^t_{\ell_1,\ell_2} - q| \rangle_{t,\sim} \le \langle |f| |R_{\ell_1,\ell_2} - q| \rangle_{t,\sim} + \frac{1}{N} \langle |f| \rangle_{t,\sim} ,$$

and we combine this estimate with the previous ones to conclude the proof of (9.75). □

We can now fix the constant L of (9.63) once and for all so that (9.75) becomes

$$|\varphi'(v)| \le L \exp L\tau^2 \Bigg(\sum_{\ell_1,\ell_2 \le 6, \ell_1 \ne \ell_2} \mathsf{E}\langle |f||R_{\ell_1,\ell_2} - q|\rangle_{t,\sim}$$
$$+ \frac{1}{N}\mathsf{E}\langle |f|\rangle_{t,\sim} + \max |f| \exp\left(-\frac{N}{L}\right)\Bigg) . \tag{9.87}$$

To obtain the estimate (9.87) is the main effort in proving Theorem 9.5.1. However we would like however to have $\langle\cdot\rangle_t$ rather than $\langle\cdot\rangle_{t,\sim}$ occurring on the right-hand side, and we now learn how to compare these.

Lemma 9.5.3. *If $L\alpha(1+\tau^2) \le 1$ and $f \ge 0$ is a function on Σ_N^8, we have*

$$\mathsf{E}\langle f\rangle_{t,\sim} \le L \exp L\tau^2 \left(\nu_t(f) + (\max f) \exp\left(-\frac{N}{L}\right)\right) . \tag{9.88}$$

Proof. Let us denote by E' expectation in the r.v.s $g_{i,M}$, $i \le N$. Since $u \le 0$, we have

$$\nu_t(f) = \mathsf{E}\frac{\mathsf{E}_\xi \langle f \exp \sum_{\ell \le 8} u(S^\ell_{M,t})\rangle_{t,\sim}}{\mathsf{E}_\xi \langle \exp u(S^1_{M,t})\rangle^8_{t,\sim}} \ge \mathsf{E}\Big\langle f \mathsf{E}'\mathsf{E}_\xi \exp \sum_{\ell\le 8} u(S^\ell_{M,t})\Big\rangle_{t,\sim} . \tag{9.89}$$

Consider a number $d > 0$, to be determined later, and

$$A = \{(\boldsymbol{\sigma}^1, \dots, \boldsymbol{\sigma}^8) \; ; \; \forall \ell \ne \ell' \; , \; |R_{\ell,\ell'}| \le d\} .$$

In Lemma 9.5.4 below we show that we can choose d (which is a universal constant independent of any other parameter) so that

$$(\boldsymbol{\sigma}^1, \dots, \boldsymbol{\sigma}^8) \in A \quad \Rightarrow \quad \mathsf{E}'\mathsf{E}_\xi \exp \sum_{\ell \le 8} u(S^\ell_{M,t}) \ge \frac{1}{L}\exp(-L\tau^2) . \tag{9.90}$$

Thus

$$\mathsf{E}\Big\langle f \mathsf{E}'\mathsf{E}_\xi \exp \sum_{\ell\le 8} u(S^\ell_{M,t})\Big\rangle_{t,\sim} \ge \frac{1}{L}\exp(-L\tau^2)\mathsf{E}\langle \mathbf{1}_A f\rangle_{t,\sim}$$
$$\ge \frac{1}{L}\exp(-L\tau^2)(\mathsf{E}\langle f\rangle_{t,\sim} - (\max f)\mathsf{E}\langle \mathbf{1}_{A^c}\rangle_{t,\sim}) .$$

Since d is a universal constant we may and do assume that $\overline{c} \le d$. It then follows from Lemma 9.3.8 and Proposition 9.3.10 that if $L\alpha(1+\tau^2) \le 1$ (and using (9.47) for $t = 0$) we have $\mathsf{E}\langle \mathbf{1}_{A^c}\rangle_{t,\sim} \le L\exp(-N/L)$. This concludes the proof, modulo the proof of (9.90), which is given in the next lemma. □

Lemma 9.5.4. *There exists a number $d > 0$ with the following property. If we consider Gaussian r.v.s $(w_\ell)_{\ell \le 8}$, such that $\mathsf{E}w_\ell^2 = 1$, $|\mathsf{E}w_\ell w_{\ell'}| \le d$ for $\ell \ne \ell'$, then for any value of τ we have*

$$\mathsf{P}(\forall \ell \le 8 \; , \; w_\ell \ge \tau) \ge \frac{1}{L}\exp(-L\tau^2) . \tag{9.91}$$

When applied to the case $w_\ell = S^\ell_{M,t}$, this proves (9.90) since $u(x) = 0$ for $x \geq \tau$.

Proof. It should be obvious that one can choose $d > 0$ so that the hypothesis on (w_ℓ) implies that we can find i.i.d. Gaussian r.v.s $(v_\ell)_{\ell \leq 8}$ with

$$w_\ell = \sum_{\ell' \leq 8} a_{\ell,\ell'} v_{\ell'}$$

where for each ℓ we have $|1 - a_{\ell,\ell}| + \sum_{\ell \neq \ell'} |a_{\ell,\ell'}| \leq 1/3$. Consequently,

$$w_\ell \geq v_\ell - \frac{1}{3} \max_{\ell'} |v_{\ell'}| \, . \tag{9.92}$$

To prove (9.91) we may and do assume that $\tau \geq 1$. Then on the event

$$\forall \ell \leq 8 \, , \quad 2\tau \leq v_\ell \leq 3\tau \, , \tag{9.93}$$

we have $w_\ell \geq \tau$ by (9.92); and the event (9.93) is of probability greater than or equal to $(1/L)\exp(-L\tau^2)$. □

Let us summarize what we have proved.

Proposition 9.5.5. *Under (9.63), and if* $L\alpha(1+\tau^2) \leq 1$, *for any function* f *on* Σ_N^4 *and if either* $\varphi(v) = \nu_{t,v}(fu'(S^\ell_v)u'(S^{\ell'}_v))$ *or* $\varphi(v) = \nu_{t,v}(f)$ *then whenever* $1/\tau_1 + 1/\tau_2 = 1$ *and* $0 \leq v \leq 1$ *we have*

$$|\varphi'(v)| \leq L \exp L\tau^2 \bigg((\nu_t(|f|^{\tau_1})^{1/\tau_1} \nu_t(|R_{1,2} - q|^{\tau_2})^{1/\tau_2} \\ + \frac{1}{N}\nu_t(|f|) + \max |f| \exp\left(-\frac{N}{L}\right)\bigg) \, . \tag{9.94}$$

Proof. We combine (9.87) and (9.88) and we use Hölder's inequality. □

Research Problem 9.5.6. Is it true that (9.94) holds with a term $L(1+\tau^2)$ rather than $L \exp L\tau^2$?

Corollary 9.5.7. *Under (9.63) and if* $L\alpha(1+\tau^2) \leq 1$, *for any function* f *on* Σ_N^4, *whenever* $1/\tau_1 + 1/\tau_2 = 1$ *we have*

$$|\nu_t(fu'(S^\ell_{M,t})u'(S^{\ell'}_{M,t})) - \widehat{r}\nu_t(f)| \\ \leq L \exp L\tau^2 \bigg((\nu_t(|f|^{\tau_1})^{1/\tau_1} \nu_t(|R_{1,2} - q|^{\tau_2})^{1/\tau_2} \\ + \frac{1}{N}\nu_t(|f|) + \max |f| \exp\left(-\frac{N}{L}\right)\bigg) \, . \tag{9.95}$$

Proof. If $\mathcal{B}$ denotes the right-hand side of (9.94) then

$$|\varphi(0) - \varphi(1)| \leq \mathcal{B} \,. \tag{9.96}$$

Now $\varphi(1) = \nu_t(fu'(S^{\ell}_{M,t})u'(S^{\ell'}_{M,t}))$ and using (2.38) we see that $\varphi(0) = \widehat{r}\mathsf{E}\langle f\rangle_{t,\sim}$, so that

$$\left|\nu_t(fu'(S^{\ell}_{M,t})u'(S^{\ell'}_{M,t})) - \widehat{r}\mathsf{E}\langle f\rangle_{t,\sim}\right| \leq \mathcal{B} \,. \tag{9.97}$$

Using again (9.96) in the case $\varphi(v) = \nu_{t,v}(f)$ yields $|\nu_t(f) - \mathsf{E}\langle f\rangle_{t,\sim}| \leq \mathcal{B}$, and combining with (9.97) finishes the proof. □

Proposition 9.5.8. *Under (9.63), and if $L\alpha(1+\tau^2) \leq 1$, for any function f on Σ_N^2, whenever $1/\tau_1 + 1/\tau_2 = 1$ we have*

$$\begin{aligned}|\nu_t'(f)| \leq L\alpha \exp L\tau^2 \Big(&\nu_t(|f|^{\tau_1})^{1/\tau_1}\nu_t(|R_{1,2}-q|^{\tau_2})^{1/\tau_2} \\ &+ \frac{1}{N}\nu_t(|f|) + \max|f| \exp\left(-\frac{N}{L}\right)\Big) \,.\end{aligned} \tag{9.98}$$

Proof. Combine (9.95) with (2.23). □

Lemma 9.5.9. *Assume $L\alpha \exp L\tau^2 \leq 1$. Consider a function f on Σ_N^2, $f \geq 0$. Then*

$$\forall t\,, \quad \nu_t(f) \leq 2\nu(f) + L\max|f|\exp\left(-\frac{N}{L}\right) \,. \tag{9.99}$$

Proof. Using (9.98) for $\tau_1 = 1$, $\tau_2 = \infty$ we obtain

$$|\nu_t'(f)| \leq L\alpha \exp L\tau^2 \left(\nu_t(f) + \max|f|\exp\left(-\frac{N}{L}\right)\right) \tag{9.100}$$

and we integrate using Lemma A.11.1. □

Now it is straightforward to check that one can prove (9.66) by following the steps of the proof of (2.67). Theorem 9.5.1 is proved.

Proposition 9.5.10. *Under the conditions of Theorem 9.5.1 we actually have*

$$\forall k \geq 1\,, \quad \nu\big((R_{1,2}-q)^{2k}\big) \leq \left(\frac{Lk}{N}\right)^k \,. \tag{9.101}$$

Proof. We copy the proof of Theorem 2.5.1. In (2.96)we get an extra term $L\max|f|\exp(-N/L) \leq 2^{2k}\exp(-N/L)$. Now, for $x > 0$ we have

$$(ax)^x \geq \exp\left(-\frac{1}{ae}\right)$$

so that

$$\left(\frac{L_0 k}{4N}\right)^k \geq \exp\left(-\frac{4N}{L_0 e}\right)$$

and if L_0 is large enough we have

$$2^{2k} \exp\left(-\frac{N}{L}\right) \leq \left(\frac{L_0 k}{N}\right)^k$$

and the proof of Theorem 2.5.1 carries forward with no other changes. □

We recall that q and r are defined as in (9.65). We recall the notations (2.11) and (2.72):

$$p_{N,M}(u) = \frac{1}{N} \mathsf{E} \log \sum_{\boldsymbol{\sigma}} \exp(-H_{N,M}(\boldsymbol{\sigma}))$$

$$p(u) = -\frac{r}{2}(1-q) + \mathsf{E}\log(2\mathrm{ch}(z\sqrt{r})) + \alpha \mathsf{E} \log \mathsf{E}_\xi \exp u(z\sqrt{q} + \xi\sqrt{1-q}). \tag{9.102}$$

Theorem 9.5.11. *Under the conditions of Theorem 9.5.1 we have*

$$|p_{N,M}(u) - p(u)| \leq \frac{L}{N} . \tag{9.103}$$

The proof follows the approach of the second proof of Theorem 2.4.2. We recall the identity

$$p_{N,M+1}(u) - p_{N,M}(u) = \frac{1}{N} \mathsf{E} \log \left\langle \exp u\left(\frac{1}{\sqrt{N}} \sum_{i \leq N} g_{i,M+1}\sigma_i\right)\right\rangle .$$

We define

$$S_s^\ell = \sqrt{\frac{s}{N}} \sum_{i\leq N} g_{i,M+1}\sigma_i + \sqrt{1-s}\theta^\ell \ ; \quad \theta^\ell = z\sqrt{q} + \xi^\ell\sqrt{1-q} \ ,$$

$$\varphi(s) = \frac{1}{N} \mathsf{E} \log \mathsf{E}_\xi \langle \exp u(S_s^1)\rangle \ .$$

The excuse for using the same notation here and in (9.104) below as in (9.73) is of course that they serve the same purpose. By (2.89) we have

$$\varphi'(s) = -\frac{1}{2} \mathsf{E} \frac{\langle (R_{1,2} - q) u'(S_s^1) u'(S_s^2) \exp(u(S_s^1) + u(S_s^2))\rangle}{(\mathsf{E}_\xi \langle \exp u(S_s^1)\rangle)^2} . \tag{9.104}$$

As in the proof of Theorem 2.4.2 one needs to control $|\varphi'(0)|$ and $|\varphi''(s)|$.

Since $\varphi'(0) = -(\widehat{r}/2)\nu(R_{1,2} - q)$, we have $|\varphi'(0)| \leq L/N$ as a consequence of the next Lemma, that we will prove when we study central limit theorems in Section 9.7.

Lemma 9.5.12. *Under the conditions of Theorem 9.5.1 we have*

$$|\nu(R_{1,2} - q)| \leq \frac{L}{N} .$$

To control φ'', we compute it from (9.104) using integration by parts; this brings a new factor $(R_{\ell,\ell'} - q)$ in each resulting term. To bound the resulting quantity is not obvious a priori, because the denominator can be small, and the derivatives of u can be huge. But we simply repeat the steps of the proof of Proposition 9.5.2: we separate the numerator from the denominator using the Cauchy-Schwarz inequality, we integrate by parts, and so on. The proof is quite simpler than that of Proposition 9.5.2, because we do not have to be concerned with the pesky interpolating averages $\langle\cdot\rangle_{t,\sim}$. The reader who really likes to understand the previous techniques should carry out the detail of the proofs, as suggested by the following exercise.

Exercise 9.5.13. If $L\alpha \exp L\tau^2 \leq 1$, prove the inequality

$$|\varphi''(s)| \leq L \exp L\tau^2 \left(\nu((R_{1,2} - q)^2) + L \exp\left(-\frac{N}{L}\right)\right) . \tag{9.105}$$

Combining (9.105) and (9.66), we have

$$|\varphi''(s)| \leq \frac{L}{N} \exp L\tau^2 ,$$

and since $|\varphi'(0)| \leq L/N$ by Lemma 9.78 we have reached the bound

$$\left| p_{N,M+1}(u) - p_{N,M}(u) - \frac{1}{N} \mathsf{E} \log \mathsf{E}_\xi \exp u(\theta) \right| \leq \frac{L}{N} \exp L\tau^2 ,$$

and as in Theorem 2.4.2, summation over M (and the fact that $M \leq LN \exp L\tau^2$) yields (9.103).

9.6 The Gardner Formula for the Discrete Cube

We recall the notation

$$U_k = \{\boldsymbol{\sigma} \; ; \; S_k(\boldsymbol{\sigma}) \geq \tau\} \; ; \quad \mathcal{N}(x) = \mathsf{P}(\xi \geq x) \; ; \quad \mathcal{A}(x) = \frac{1}{\sqrt{2\pi}} \frac{e^{-x^2/2}}{\mathcal{N}(x)} .$$

In the case $\exp u(x) = \mathbf{1}_{\{x \geq \tau\}}$ the equations (9.65) become

$$q = \mathsf{E}\mathrm{th}^2(z\sqrt{r}) \; ; \qquad r = \frac{\alpha}{1-q} \mathsf{E}\mathcal{A}\left(\frac{\tau - z\sqrt{q}}{\sqrt{1-q}}\right)^2 , \tag{9.106}$$

where z is a standard Gaussian r.v.

Theorem 9.6.1. *There is a constant L with the following property. Consider τ and M with $L\alpha \exp L\tau^2 \le 1$. Then for $t \ge 0$ we have*

$$\mathsf{P}\left(\left|\frac{1}{N}\log\left(2^{-N}\mathrm{card}\bigcap_{k\le M} U_k\right) - \mathrm{RS}(\alpha)\right| \ge t\right) \le L\exp\left(-\frac{1}{L}\min\left(N, \frac{Nt}{1+\tau^2}, \frac{N^2t^2}{M(1+\tau^2)^2}\right)\right) , \tag{9.107}$$

where

$$\mathrm{RS}(\alpha) = -\frac{r}{2}(1-q) + \mathsf{E}\log\mathrm{ch}(z\sqrt{q}) + \alpha\mathsf{E}\log\mathcal{N}\left(\frac{\tau - z\sqrt{q}}{\sqrt{1-q}}\right) , \tag{9.108}$$

where $\alpha = M/N$ and q and r are solutions of the equations (9.106).

If one does not care about the dependence on τ^2 (which is unlikely to be sharp anyway) one can simplify (9.107) as

$$\mathsf{P}\left(\left|\frac{1}{N}\log\left(2^{-N}\mathrm{card}\bigcap_{k\le M} U_k\right) - \mathrm{RS}(\alpha)\right| \ge t\right) \le K\exp\left(-\frac{Nt^2}{K}\right)$$

for $t \le 1$, where K depends on τ only.

The existence of a solution to the equations (9.106) where $L\alpha \exp L\tau^2 \le 1$ was actually obtained in the proof of Theorem 9.5.1, because this part of the argument never used (9.63), so it remains valid in the case $\exp u(x) = \mathbf{1}_{\{x\ge\tau\}}$. A bit of extra work would show that these solutions are unique.

The following is a rather weak consequence of Theorem 9.6.1. The main motivation for proving it is that it was announced in Section 2.1.

Corollary 9.6.2. *When $\tau = 0$ there exists $\alpha_0 < 1$ such that*

$$\mathsf{P}\left(\bigcap_{k\le\alpha_0 N} U_k = \emptyset\right) \ge 1 - L\exp\left(-\frac{N}{L}\right) . \tag{9.109}$$

Proof. The difficulty is that Theorem 9.6.1 holds only for $\alpha \le 1/L$ while we are trying to prove something for α close to 1. As a first step we prove that for $\alpha > 0$ we have

$$\mathrm{RS}(\alpha) < -\alpha\log 2 . \tag{9.110}$$

Let us denote by $F(q,r,\alpha)$ the left-hand side of (9.108) when we consider q, r and α as unrelated variables. Then the conditions (9.106) mean that

$$\frac{\partial F}{\partial q}(q,r,\alpha) = \frac{\partial F}{\partial r}(q,r,\alpha) = 0 ,$$

so that

$$\mathrm{RS}'(\alpha) := \frac{\mathrm{d}}{\mathrm{d}\alpha}\mathrm{RS}(\alpha) = \mathsf{E}\log\mathcal{N}\left(\frac{-z\sqrt{q}}{\sqrt{1-q}}\right) .$$

Now, Jensen's inequality implies

$$\mathsf{E}\log\mathcal{N}\left(\frac{-z\sqrt{q}}{\sqrt{1-q}}\right) \le \log \mathsf{E}\mathcal{N}\left(\frac{-z\sqrt{q}}{\sqrt{1-q}}\right) , \tag{9.111}$$

and since $\mathcal{N}(x)+\mathcal{N}(-x) = 1$, the expectation in the right-hand side of (9.111) is equal to $1/2$. There cannot be equality in (9.111) unless $q = 0$, and this does not occur for $\alpha > 0$ as is apparent from the equations (9.108). Thus we have proved that $\mathrm{RS}'(\alpha) < -\log 2$ for $\alpha > 0$ and (9.110) follows.

Consequently, we can find α_1 small enough so that Theorem 9.6.1 holds for α_1 while we have

$$\mathrm{RS}(\alpha_1) = -(\alpha_1 + 4\delta)\log 2 , \tag{9.112}$$

where $\delta > 0$. We are now going to prove that if $\alpha_0 = 1 - \delta$ then

$$\mathsf{P}\left(\mathrm{card}\left(\bigcap_{k\le\alpha N} U_k\right) \le 4\times 2^{-\delta N}\right) \ge 1 - L\exp\left(-\frac{N}{L}\right) . \tag{9.113}$$

This implies (9.109) since

$$\mathrm{card}\left(\bigcap_{k\le\alpha_0 N} U_k\right) < 1 \Rightarrow \mathrm{card}\left(\bigcap_{k\le\alpha_0 N} U_k\right) = 0 .$$

To prove (9.113) we first observe that by Theorem 9.6.1, for any integer M with $LM\exp L\tau^2 \le N$ it holds

$$\mathsf{P}\left(\mathrm{card}\bigcap_{k\le M} U_k \le 2^{N(1+\delta/2)}\exp N\mathrm{RS}\left(\frac{M}{N}\right)\right) \ge 1 - L\exp\left(-\frac{N}{L}\right) .$$

Thus if $M = \lceil \alpha_1 N\rceil$ and

$$V = \bigcap_{k\le M} U_k ,$$

it follows from (9.112) that

$$\mathsf{P}(\mathrm{card}V \le 2^{N(1-\alpha_1-3\delta)}) \ge 1 - L\exp\left(-\frac{N}{L}\right) . \tag{9.114}$$

Moreover given any set V and any integer $M' > M$ we have

$$\mathsf{E}\mathrm{card}\left(V\cap\bigcap_{M<k\le M'} U_k\right) = 2^{M-M'}\mathrm{card}V ,$$

because any point of V has a 50% chance to belong to each set U_k. Therefore by Markov's inequality we have

$$\mathsf{P}\left(\operatorname{card}\left(V\cap\bigcap_{M<k\le M'}U_k\right)\le 2^{M-M'+\delta N}\operatorname{card}V\right)\ge 1-L\exp\left(-\frac{N}{L}\right),$$

and if we combine with (9.114) we see that

$$\mathsf{P}\left(\operatorname{card}\left(\bigcap_{k\le M'}U_k\right)\le 2^{M-M'-(2\delta+\alpha_1)N}\right)\ge 1-L\exp\left(-\frac{N}{L}\right).$$

Taking $M'=\lfloor\alpha_0 N\rfloor$ this proves (9.113) since $M\le\alpha_1 N+1$ and $M'\ge N\alpha_0-1$. □

Exercise 9.6.3. Prove from (9.108) that $\mathrm{RS}(0)=0$ and $\mathrm{RS}'(0)=\mathcal{N}(\tau)$. Offer an intuitive explanation for this fact.

As in Chapter 8, the key to Theorem 9.6.1 is that the fluctuations of the random quantity $\log\left(2^{-N}\operatorname{card}\bigcap_{k\le M}U_k\right)$ are very small, so that it will be sufficient to compute its expectation (after suitable truncation), using Theorem 9.5.1.

Proof of Theorem 9.6.1. We will use a function u that satisfies (9.2), but such that $\exp u$ approximates well the function $\mathbf{1}_{\{x\ge\tau\}}$. We will require that, for a certain number τ' depending on N, with $\tau'<\tau$ we have

$$x\le\tau'\quad\Rightarrow\quad u(x)=-N\ .$$

In order to be able to use Theorem 9.5.1, we want u to satisfy (9.63), and yet $\tau-\tau'$ to be as small as possible. It is obvious from scaling arguments that u can be found with $|u^{(\ell)}|\le NL(\tau-\tau')^{-\ell}$ for $1\le\ell\le 5$, so that we may achieve (9.63) with

$$\tau-\tau'\le L_1\exp\left(-\frac{N}{L_1}\right) \tag{9.115}$$

for a certain number L_1.

Let us define

$$V=2^{-N}\operatorname{card}\{\boldsymbol{\sigma}\in\Sigma_N\ ;\ \exists k\le M\ ;\ \tau'\le S_k(\boldsymbol{\sigma})\le\tau\}\ .$$

This r.v. is very small, since, if g is standard Gaussian r.v.

$$\mathsf{E}V\le M\mathsf{P}(\tau'\le\ g\le\tau)\le M(\tau-\tau')\le L_2\exp\left(-\frac{N}{L_2}\right)$$

since $M\le N$. In particular, if we consider the event

$$\Omega_1=\left\{\ V\le\exp\left(-\frac{N}{2L_2}\right)\right\},$$

Markov's inequality implies

$$\mathsf{P}(\Omega_1) \geq 1 - L_2 \exp\left(-\frac{N}{2L_2}\right) .$$

Consider also the event

$$\Omega_2 = \left\{2^{-N}\mathrm{card} \bigcap_{k\leq M} U_k \geq \exp\left(-\frac{N}{2L_2}\right)\right\} .$$

We use (9.16) with $b = \exp(-N/2L_2)$ to obtain that

$$L\alpha(1+\tau^2) \leq 1 \quad \Rightarrow \quad \mathsf{P}(\Omega_2) \geq 1 - L\exp\left(-\frac{N}{L}\right) . \tag{9.116}$$

Let us define

$$Z(u) = \sum_{\boldsymbol{\sigma}} \exp \sum_{k\leq M} u(S_k(\boldsymbol{\sigma})) ,$$

and prove the inequality

$$2^{-N}\mathrm{card} \bigcap_{k\leq M} U_k \leq 2^{-N} Z(u) \leq 2^{-N}\mathrm{card}\left(\bigcap_{k\leq M} U_k\right) + V + e^{-N} . \tag{9.117}$$

The left-hand side inequality is obvious since $u(x) = 0$ for $x \geq \tau$, so that

$$\boldsymbol{\sigma} \in \bigcap_{k\leq M} U_k \quad \Rightarrow \quad \exp \sum_{k\leq M} u(S_k(\boldsymbol{\sigma})) = 1 .$$

To prove the second inequality we consider the sets

$$A = \bigcap_{k\leq M} U_k \ ; \qquad B = \{\boldsymbol{\sigma} \ ; \ \exists k \leq M \ , \ \tau' \leq S_k(\boldsymbol{\sigma}) \leq \tau\}$$

$$C = \{\boldsymbol{\sigma} \ ; \ \exists k \leq M \ , \ S_k(\boldsymbol{\sigma}) < \tau'\}$$

so that $\Sigma_N \subset A \cup B \cup C$. For any $\boldsymbol{\sigma}$ we have that $\exp \sum_{k\leq M} u(S_k(\boldsymbol{\sigma})) \leq 1$ since $u \leq 0$; moreover if $\boldsymbol{\sigma} \in C$ we have $\exp \sum_{k\leq M} u(S_k(\boldsymbol{\sigma})) \leq e^{-N}$ because $u \leq 0$ and $u(x) = -N$ if $x < \tau'$. Therefore

$$\sum_{\boldsymbol{\sigma}} \exp \sum_{k\leq M} u(S_k(\boldsymbol{\sigma})) \leq \mathrm{card}A + \mathrm{card}B + 2^N e^{-N}$$

and this proves (9.117) since $V = 2^{-N}\mathrm{card}B$.

The point of introducing the events Ω_1 and Ω_2 is that on the event $\Omega_1 \cap \Omega_2$ we have

$$V \leq \exp\left(-\frac{N}{2L_2}\right) \leq 2^{-N}\mathrm{card} \bigcap_{k\leq M} U_k ,$$

and, since without loss of generality we may assume that $L_2 \geq 16$ the right-hand side inequality above implies that, recalling the notation $a = 1/32$ of (9.19),

$$e^{-N} \leq \exp(-aN) \leq 2^{-N} \mathrm{card} \bigcap_{k \leq M} U_k \,.$$

Therefore, using (9.117), on $\Omega_1 \cap \Omega_2$ we have

$$\exp(-aN) \leq 2^{-N} \mathrm{card} \bigcap_{k \leq M} U_k \leq 2^{-N} Z(u) \leq 3 \Big(2^{-N} \mathrm{card} \bigcap_{k \leq M} U_k \Big) \,, \tag{9.118}$$

and then

$$\Big| \log(2^{-N} Z(u)) - \log \Big(2^{-N} \mathrm{card} \bigcap_{k \leq M} U_k \Big) \Big| \leq 2 \,.$$

Recalling the notation (9.19), by (9.118) we also have $\log_{aN}(2^{-N} Z(u)) = \log(2^{-N} Z(u))$ on $\Omega_1 \cap \Omega_2$ and thus using (9.21) in the last line before (9.119) below,

$$\begin{aligned}
&\mathsf{P}\Big(\Big| \frac{1}{N} \log \Big(2^{-N} \mathrm{card} \bigcap_{k \leq M} U_k \Big) - \frac{1}{N} \mathsf{E} \log_{aN}(2^{-N} Z(u)) \Big| \geq t + \frac{2}{N} \Big) \\
&\leq \mathsf{P}\Big(\Big| \frac{1}{N} \log(2^{-N} Z(u)) - \frac{1}{N} \mathsf{E} \log_{aN}(2^{-N} Z(u)) \Big| \geq t \Big) \\
&\leq \mathsf{P}(\Omega_1^c \cup \Omega_2^c) + \mathsf{P}\Big(\Big| \frac{1}{N} \log_{aN}(2^{-N} Z(u)) - \frac{1}{N} \mathsf{E} \log_{aN}(2^{-N} Z(u)) \Big| \geq t \Big) \\
&\leq L \exp\Big(-\frac{N}{L} \Big) + 2 \exp\Big(-\frac{1}{L} \min\Big(\frac{N^2 t^2}{M(1+\tau^2)^2}, \frac{Nt}{1+\tau^2} \Big) \Big) \,.
\end{aligned} \tag{9.119}$$

Also, since $\sum_{k \leq M} u(S_k) \geq -NM$, we have $\log(2^{-N} Z(u)) \geq -MN$ and

$$\begin{aligned}
\Big| \frac{1}{N} \mathsf{E} \log_{aN}(2^{-N} Z(u)) - \frac{1}{N} \mathsf{E} \log(2^{-N} Z(u)) \Big| &\leq M \mathsf{P}(2^{-N} Z(u) \leq \exp(-Na)) \\
&\leq L \exp\Big(-\frac{N}{L} \Big) \leq \frac{L}{N}
\end{aligned}$$

using (9.16). Finally, Theorem 9.5.11 implies

$$\Big| \frac{1}{N} \mathsf{E} \log(2^{-N} Z(u)) - p(u) + \log 2 \Big| \leq \frac{L}{N} \,. \tag{9.120}$$

Therefore from (9.119) we get

$$\mathsf{P}\left(\left|\frac{1}{N}\log\left(2^{-N}\mathrm{card}\bigcap_{k\le N}U_k\right)-p(u)+\log 2\right|\ge t+\frac{L}{N}\right)$$
$$\le L\exp\left(-\frac{1}{L}\min\left(N,\frac{Nt}{1+\tau^2},\frac{N^2t^2}{M(1+\tau^2)^2}\right)\right)\ . \quad (9.121)$$

Recalling the definition (9.102) of $p(u)$ we observe that quantities $p(u)-\log 2$ and $\mathrm{RS}(\alpha)$ are computed by the same procedure, that is applied to the function u in the case of $p(u)$ and to the function $\mathbf{1}_{\{x\ge\tau\}}$ in the case of $\mathrm{RS}(\alpha)$. Therefore we expect that these quantities are exponentially close to each other. However proving this rigorously is no fun, one has to perform the tedious estimates required to prove that the function ψ of (9.71) satisfies $|\psi'|\le 1/2$, after which it is not so difficult to see that the unique solution of the equations (9.65) depends smoothly on the parameters. A simpler way to proceed is to fix a solution (q,r) of the equations (9.106) and to define

$$r'=\frac{\alpha}{1-q}\mathsf{E}\left(\frac{\mathsf{E}_\xi\xi\exp u(\theta)}{\mathsf{E}_\xi\exp u(\theta)}\right)^2$$

where $\theta=z\sqrt{q}+\xi\sqrt{1-q}$, so that $r'-r$ is very small, and $q-\mathsf{E}\mathrm{th}^2(z\sqrt{r'})$ is very small.

Then, nothing needs to be changed to the proof of (9.104) if one uses the values (q,r') rather than a solution of the equations (9.65), so that instead of (9.121) one obtains directly

$$\left|\frac{1}{N}\mathsf{E}\log(2^{-N}Z(u))-\mathrm{RS}(\alpha)\right|\le\frac{L}{N}\ ,$$

and then as in (9.121) one gets

$$\mathsf{P}\left(\left|\frac{1}{N}\log\left(2^{-N}\mathrm{card}\bigcap_{k\le N}U_k\right)-\mathrm{RS}(\alpha)\right|\ge t+\frac{L}{N}\right)$$
$$\le L\exp\left(-\frac{1}{L}\min\left(N,\frac{Nt}{1+\tau^2},\frac{N^2t^2}{M(1+\tau^2)^2}\right)\right)$$

from which (9.107) follows. □

Should the reader find none of the two above arguments above convincing, another possibility is to look for a refund of this book. One may also observe that the rate L/N in (9.120) is not critical since the right-hand side of (9.107) becomes small only for t about $1/\sqrt{N}$.

Theorem 9.6.1 is much more precise than Theorem 8.4.1. This suggests the following.

Research Problem 9.6.4. (Level 1) For α small, improve Theorem 8.4.1 to a statement as precise as (9.107).

Of course, (see Research Problem 8.3.5) the case of $\alpha\le\alpha_0<2$ is even more interesting, but it is no longer level 1.

9.7 Higher Order Expansion and Central Limit Theorems

The main result of this section is probably Theorem 9.7.12 below. The basic idea is simple, and has been used many times. If $\varphi(v)$ is given by (9.74) (reproduced in (9.123) just below) then (9.95) is a consequence of the inequality $|\varphi(1) - \varphi(0)| \le \sup|\varphi'(v)|$. Instead of the "first order expansion" we will use a "second order expansion", $|\varphi(1) - \varphi(0) - \varphi'(0)| \le \sup|\varphi''(v)|$. If we roughly describe the action of taking the derivatives of φ as "bringing out a factor $R_{1,2} - q$" then we expect that taking the second derivative "brings out another such factor" and increases accuracy by a factor $N^{-1/2}$.

We recall the notation S_v^ℓ of (9.73).

Proposition 9.7.1. *There exists a number L with the following property. Consider a function u that satisfies (9.2), as well as*

$$\forall \ell\,,\ 1 \le \ell \le 5\,,\quad |u^{(\ell)}| \le \exp\frac{N}{L}\,. \tag{9.122}$$

Consider a function f on Σ_N^4, consider $\ell \ne \ell' \le 4$ and consider either

$$\varphi(v) = \nu_{t,v}(fu'(S_v^\ell)u'(S_v^{\ell'})) \quad\text{or}\quad \varphi(v) = \nu_{t,v}(f)\,. \tag{9.123}$$

Then if $L\alpha(1+\tau^2) \le 1$, we have

$$|\varphi''(v)| \le (L\exp L\tau^2)\Big(\nu_t(f^2)^{1/2}\nu_t\big((R_{1,2}-q)^4\big)^{1/2} + \frac{1}{N}\nu_t(|f|) + \max|f|\exp\Big(-\frac{N}{L}\Big)\Big)\,. \tag{9.124}$$

Proof. We differentiate twice the relation (9.72), which brings out a second factor of the type $(R^t_{\ell_1,\ell_2} - q)$ in each term. We then repeat the proof of Proposition 9.5.5 with the following small difference: we need to control a few more replicas, since $\varphi''(v)$ depends on 8 replicas while $\varphi'(v)$ depends only on 6 replicas. □

Corollary 9.7.2. *Under the conditions of Proposition 9.7.1, and if moreover $L\alpha\exp L\tau^2 \le 1$, we have*

$$|\varphi''(v)| \le L\exp L\tau^2\left(\frac{1}{N}\nu(f^2)^{1/2} + \max|f|\exp\Big(-\frac{N}{L}\Big)\right)\,. \tag{9.125}$$

Proof. This is a consequence of (9.124). Using (9.101), we have $\nu((R_{1,2} - q)^4)^{1/2} \le L/N$, and using (9.99) we have $\nu_t((R_{1,2}-q)^4)^{1/2} \le L/N$ and

$$\nu_t(f^2) \le L\nu(f^2) + L\max(f^2)\exp(-N/L)\,. \qquad □$$

At this stage we realize that we are facing a nasty unsolved technical problem. While controlling "a few more replicas" (i.e. any finite number rather than 4 only) by increasing the value of the number L of Proposition 9.5.2 is easy, we do not know how to control all the replicas at the same time. Here is the precise version of the problem.

Research Problem 9.7.3. Given $\tau \geq 0$, does there exists a number L and a number $K_0(\tau)$ depending on τ only such that if u satisfies (9.2) and (9.122), then for $MK_0(\tau) \leq N$, any n and any function f on Σ_N^n, we have

$$|\varphi''(v)| \leq K(\tau, n)\left(\frac{1}{N}\nu(f^2)^{1/2} + \max(|f|)\exp\left(-\frac{N}{L}\right)\right) , \tag{9.126}$$

where $K(\tau, n)$ depends only on τ and n?

There is nothing specific about the second derivative here. It is probably the same problem to ask whether we have

$$|\varphi'(v)| \leq K(\tau, n)\left(\frac{1}{\sqrt{N}}\nu(f^2)^{1/2} + \max(|f|)\exp\left(-\frac{N}{L}\right)\right) . \tag{9.127}$$

The difficulty lies with our method of "separating the numerators from the denominators" using the Cauchy-Schwarz inequality. When we work with n replicas rather that 6 replicas, we have to replace the exponent 12 by $2n$ in (9.81). It is not difficult using Theorem 9.3.1 to see that given n, we can control $|\varphi''(v)|$ as in (9.126) for all f on Σ_N^n, provided $\alpha K(\tau, n) \leq 1$ but $K(\tau, n) \to \infty$ as $n \to \infty$. It is reasonable to think that the previous research problem is closely related to Research Problem 9.2.4. Here is a less technical question.

Research Problem 9.7.4. Can one prove a central limit theorem in the spirit of Theorem 1.10.1 under a condition of the type $\alpha K(\tau) \leq 1$?

More precisely, we would like to prove such a theorem where $O(k)$ denotes a quantity $A = A_{N,u}$ with $N^{k/2}A$ bounded independently of N *and* of the choice of u satisfying (9.2) (and maybe a mild condition such as (9.122)).

At present we know how to prove a central limit theorem for all overlaps only when assuming that u is "bounded from below independently of N".

Proposition 9.7.5. *There exists a constant L with the following property. Assume as usual that u satisfies (9.2) and (9.122), and moreover that*

$$-C \leq u \leq 0 . \tag{9.128}$$

Then for each n and each function f on Σ_N^n, we have, with the notation (9.123), that, whenever $1/\tau_1 + 1/\tau_2 = 1$, and $L\alpha \exp L\tau^2 \leq 1$

$$|\varphi'(v)| \le K(C,n,\tau)\Big(\nu_t(|f|^{\tau_1})^{1/\tau_1}\nu_t(|R_{1,2}-q|^{\tau_2})^{1/\tau_2} + \frac{1}{N}\nu_t(|f|) + \max|f|\exp\left(-\frac{N}{L}\right)\Big) \tag{9.129}$$

$$|\varphi''(v)| \le K(C,n,\tau)\left(\frac{1}{N}\nu_t(f^2)^{1/2} + \max|f|\exp\left(-\frac{N}{L}\right)\right) , \tag{9.130}$$

where of course $K(C,n,\tau)$ depends only on C, n and τ, and not N, u, or anything else.

Proof. The proof follows the lines of Proposition 9.5.2, but is much simpler. We use that $\exp u(S_v^1) \ge \exp(-C)$ rather than (9.81) and

$$\nu_t(f) \ge \mathsf{E}_\xi\left\langle f \exp\sum_{\ell\le n} u(S^\ell_{M,t})\right\rangle_{t,\sim} \ge \exp(-nC)\langle f\rangle_{t,\sim}$$

rather than (9.89). Only minor changes are needed for the remainder of the proof. □

Corollary 9.7.6. *Under the conditions of Proposition 9.7.5, for any function $f \ge 0$ on Σ_N^n we have*

$$\nu_t(f) \le K(C,n,\tau)\left(\nu(f) + \max|f|\exp\left(-\frac{N}{L}\right)\right) . \tag{9.131}$$

Proof. We copy the proof of (9.99), using now (9.129) instead of (9.94) to obtain

$$|\nu_t'(f)| \le K(C,n,\tau)\left(\nu(f) + \max|f|\exp\left(-\frac{N}{L}\right)\right) ,$$

and we integrate as usual. □

Proving central limit theorems requires the explicit computation of $\varphi'(0)$.

Lemma 9.7.7. *Assume that f is a function on Σ_N^n, and that, with our usual notation,*

$$\varphi(v) = \nu_{t,v}(fu'(S_v^\ell)u'(S_v^{\ell'})) , \tag{9.132}$$

where $\ell \ne \ell'$. Then

$$\varphi(0) = \widehat{r}\mathsf{E}\langle f\rangle_{t,\sim} \tag{9.133}$$

$$\begin{aligned}\varphi'(0) = &\sum_{1\le\ell_1<\ell_2\le n} c(\ell_1,\ell_2;\ell,\ell')\mathsf{E}\langle f(R^t_{\ell_1,\ell_2}-q)\rangle_{t,\sim} \\ &- n\sum_{\ell_1\le n} c(\ell_1,n+1;\ell,\ell')\mathsf{E}\langle f(R^t_{\ell_1,n+1}-q)\rangle_{t,\sim} \\ &+ \frac{n(n+1)}{2}c(n+1,n+2;\ell,\ell')\mathsf{E}\langle f(R^t_{n+1,n+2}-q)\rangle_{t,\sim}\end{aligned} \tag{9.134}$$

where

$$c(\ell_1,\ell_2;\ell,\ell') = c\big(\mathrm{card}(\{\ell_1,\ell_2\}\cap\{\ell,\ell'\})\big) \ ,$$

$$c(0) = \mathsf{E}\left(\frac{\mathsf{E}_\xi U'(\theta)}{\mathsf{E}_\xi U(\theta)}\right)^4 \ ; \ c(1) = \mathsf{E}\frac{\mathsf{E}_\xi U''(\theta)(\mathsf{E}_\xi U'(\theta))^2}{(\mathsf{E}_\xi U(\theta))^3}$$

$$c(2) = \mathsf{E}\left(\frac{\mathsf{E}_\xi U''(\theta)}{\mathsf{E}_\xi U(\theta)}\right)^2 \ ,$$

for $U(x) = \exp u(x)$ *and, as usual,* $\theta = z\sqrt{q} + \xi\sqrt{1-q}$.

Moreover if now

$$\varphi(v) = \nu_{t,v}(f) \tag{9.135}$$

then we obtain $\varphi(0) = \mathsf{E}\langle f\rangle_{t,\sim}$ *and*

$$\begin{aligned}\varphi'(0) = \widehat{r}\Big(&\sum_{1\le \ell_1<\ell_2\le n} \mathsf{E}\langle f(R^t_{\ell_1,\ell_2}-q)\rangle_{t,\sim} \\ &- n\sum_{\ell_1\le n} \mathsf{E}\langle f(R^t_{\ell_1,n+1}-q)\rangle_{t,\sim} \\ &+ \frac{n(n+1)}{2}\mathsf{E}\langle f(R^t_{n+1,n+2}-q)\rangle_{t,\sim}\Big) \ .\end{aligned} \tag{9.136}$$

Proof. The proof of (9.133) is done in the course of the proof of Proposition 2.3.5. To prove (9.134) we proceed as follows. We observe that $U'(x) = u'(x)U(x)$, and as in (9.72) we write

$$\nu_{t,v}(fu'(S_v^\ell)u'(S_v^{\ell'})) = \mathsf{E}\frac{\big\langle f\mathsf{E}_\xi \prod_{r\le n} U^{(k(r))}(S_v^r)\big\rangle_{t,\sim}}{\langle \mathsf{E}_\xi \exp(S_v^1)\rangle^n_{t,\sim}} \ , \tag{9.137}$$

where $k(r) = 1$ if either $r = \ell$ or $r = \ell'$, and $k(r) = 0$ otherwise. Let us define

$$W_v(\ell_1,\ell_2) = \mathsf{E}\frac{\big\langle f(R^t_{\ell_1,\ell_2}-q)\mathsf{E}_\xi \prod_{r\le n+2} U^{(k(r,\ell_1,\ell_2))}(S_v^r)\big\rangle_{t,\sim}}{\langle \mathsf{E}_\xi \exp(S_v^1)\rangle^n_{t,\sim}} \ ,$$

were $k(r,\ell_1,\ell_2) = k(r)+1$ if $r = \ell_1$ or $r = \ell_2$, and $k(r,\ell_1,\ell_2) = k(r)$ otherwise. Then differentiation of (9.137) and integration by parts as we have learned to do in the proof of Lemma 2.3.2 yield the formula

$$\varphi'(v) = \sum_{1\le\ell_1<\ell_2\le n} W_v(\ell_1,\ell_2) - n\sum_{\ell_1\le n} W_v(\ell_1,n+1) + \frac{n(n+1)}{2}W_v(n+1,n+2)$$

Proceeding as in the proof of (9.133) we then obtain that

$$W_0(\ell_1,\ell_2) = c(\ell_1,\ell_2,\ell,\ell')\mathsf{E}\langle f(R^t_{\ell_1,\ell_2}-q)\rangle_{t,\sim} \ .$$

The proof of (9.136) is similar but simpler, taking now $k(r) = 0$ in (9.137). □

We do not really like to deal with the averages $\mathsf{E}\langle\cdot\rangle_{t,\sim}$ in (9.134), and would rather deal with ν_t instead. Relating these two averages is made easier by the small factor $R^t_{\ell_1,\ell_2} - q$. There, and everywhere else in this section, there are really two situations we can handle. Either a small number n of replicas is involved (say $n \le 8$), or else we assume that u bounded from below as in (9.128) and we can control any number of replicas. To lighten the exposition, we will state only the results in the second case, but we will remember when we need to prove Lemma 9.5.12 that it is then not needed to assume that u is bounded from below.

Lemma 9.7.8. *Under the conditions of Proposition 9.7.5, for a function f on Σ_N^n we have*

$$\begin{aligned} &|\nu_t(f(R_{\ell_1,\ell_2} - q)) - \mathsf{E}\langle f(R^t_{\ell_1,\ell_2} - q)\rangle_{t,\sim}| \\ &\le K\left(\frac{1}{N}\nu(f^2)^{1/2} + \max(|f|)\exp\left(-\frac{N}{L}\right)\right) . \end{aligned} \tag{9.138}$$

Here and below, K **is permitted to depend on** C, n, τ**, but not on** N **nor** f**.**

Proof. Let $\varphi(v) = \nu_{t,v}(f(R^t_{\ell_1,\ell_2} - q))$, so that $\varphi(1) = \nu_t(f(R^t_{\ell_1,\ell_2} - q))$, $\varphi(0) = \mathsf{E}\langle f(R^t_{\ell_1,\ell_2} - q)\rangle_{t,\sim}$. Since

$$\left|\nu_t(f(R_{\ell_1,\ell_1} - q)) - \nu_t(f(R^t_{\ell_1,\ell_2} - q))\right| \le \frac{\nu_t(|f|)}{N} ,$$

we get

$$|\nu_t(f(R_{\ell_1,\ell_2} - q)) - \mathsf{E}\langle f(R^t_{\ell_1,\ell_2} - q)\rangle_{t,\sim}| \le \sup_v |\varphi'(v)| + \frac{1}{N}\nu_t(|f|) .$$

Next we use (9.129) for $\tau_2 = 4$, $\tau_1 = 4/3$, and $f(R^t_{\ell_1,\ell_2} - q)$ rather than f, to get

$$\begin{aligned} |\varphi'(v)| \le K\Big(&\nu_t\big(|f(R^t_{\ell,\ell'} - q)|^{4/3}\big)^{3/4}\nu_t\big((R_{1,2} - q)^4\big)^{1/4} \\ &+ \frac{1}{N}\nu_t(|f|) + \max|f|\exp\left(-\frac{N}{L}\right)\Big) . \end{aligned}$$

Then we use again Hölder's inequality to get

$$\nu_t\big(|f|^{4/3}|R^t_{\ell_1,\ell_2} - q|^{4/3}\big) \le \nu_t(f^2)^{2/3}\nu_t\big((R^t_{\ell_1,\ell_2} - q)^4\big)^{1/3} ,$$

and the Cauchy-Schwarz inequality to obtain that $\nu_t(|f|) \le \nu_t(f^2)^{1/2}$. Finally we use Corollary 9.7.6 to replace ν_t by ν and (9.101) to see that $\nu((R_{\ell_1,\ell_2} -$

$q)^4) \le K/N^2$ and $\nu((R^t_{\ell_1,\ell_2}-q)^4) \le K/N^2$. Combining these estimates yields the result. □

Another nice feature is that we can change the value of t in the term $\nu_t(f(R_{\ell_1,\ell_2}-q))$ without creating a large error.

Lemma 9.7.9. *Under the conditions of Proposition 9.7.5, for a function f on Σ_N^n and $0 \le t \le 1$, we have*

$$\left|\nu_t'(f(R_{\ell_1,\ell_2}-q))\right| \le K\left(\frac{1}{N}\nu(f^2)^{1/2} + \max(|f|)\exp\left(-\frac{N}{L}\right)\right). \tag{9.139}$$

As a consequence, if $0 \le t, t' \le 1$ then

$$\begin{aligned}\left|\nu_t(f(R_{\ell_1,\ell_2}-q)) - \nu_{t'}(f(R_{\ell_1,\ell_2}-q))\right| \le K\bigg(&\frac{1}{N}\nu(f^2)^{1/2} \\ &+ \max(|f|)\exp\left(-\frac{N}{L}\right)\bigg).\end{aligned} \tag{9.140}$$

This will allow us in particular to replace the computation of $\nu_t(\cdot)$ by that of $\nu_0(\cdot)$, for which one can take advantage of the decoupling of the last spin.

Proof. Using (9.129) as in Corollary 9.5.7 implies that for a function f on Σ_N^n it holds

$$\begin{aligned}|\nu_t(fu'(S^\ell_{M,t})u'(S^{\ell'}_{M,t})) - \widehat{r}\nu_t(f)| \le K\bigg(&\nu_t(|f|^{\tau_1})^{1/\tau_2}\nu_t(|R_{1,2}-q|^{\tau_2})^{1/\tau_2} \\ &+\frac{1}{N}\nu_t(|f|) + \max|f|\exp\left(-\frac{N}{L}\right)\bigg).\end{aligned} \tag{9.141}$$

Using this for $f(R_{\ell_1,\ell_2}-q)$ rather than f and using Hölder's inequality as in Lemma 9.7.8 we get

$$\begin{aligned}&\left|\nu_t\big(f(R_{\ell_1,\ell_2}-q)u'(S^\ell_{M,t})u'(S^{\ell'}_{M,t})\big) - \widehat{r}\nu_t(f(R_{\ell_1,\ell_2}-q))\right| \\ &\le K\left(\frac{1}{N}\nu_t(f^2) + \max|f|\exp\left(-\frac{N}{L}\right)\right).\end{aligned}$$

Combining with Proposition 2.2.2 and (9.131) we can bound $|\nu_t'(f(R_{\ell_1,\ell_2}-q))|$ as in (9.139). □

In the proof of a central limit theorem for the overlap, there is the aspect of controlling the error terms, and the matter of handling the algebra, which are rather distinct. In order to illustrate the basic procedure before we get into algebraic complications, we prove Lemma 9.5.12 (as was promised when this lemma was stated). The method of proof should certainly not come as a surprise. Throughout the proof $O(2)$ denotes quantity A such that $N|A|$ remains bounded independently of N.

Proof of Lemma 9.5.12. We have $\nu(R_{1,2}-q)=\nu(f)$ for $f=\varepsilon_1\varepsilon_2-q$. We have $\nu_0(f)=0$, and we compute $\nu_t'(f)$ using (2.23). For each term of the type $\nu_t(f\varepsilon_\ell\varepsilon_{\ell'}u'(S^\ell_{M,t})u'(S^{\ell'}_{M,t}))$ we consider the function

$$\varphi(v)=\nu_{t,v}\big(f\varepsilon_\ell\varepsilon_{\ell'}u'(S^\ell_v)u'(S^{\ell'}_v)\big)\ ,$$

so that by (9.130) we have

$$\varphi(1)=\varphi(0)+\varphi'(0)+O(2)\ . \tag{9.142}$$

Using (9.133) and (9.134) we see that $\varphi(0)=\widehat{r}\mathsf{E}\langle f\varepsilon_\ell\varepsilon_{\ell'}\rangle_{t,\sim}$ and that $\varphi'(0)$ is a linear combination of terms of the type $\mathsf{E}\langle\varepsilon_\ell\varepsilon_{\ell'}f(R^t_{\ell_1,\ell_2}-q)\rangle_{t,\sim}$. The coefficients of these terms are all the type $\pm c(j)$ for $j=0,1,2$. Using Lemmas 9.7.8 and 9.7.9 we get successively

$$\begin{aligned}\mathsf{E}\langle\varepsilon_\ell\varepsilon_{\ell'}f(R^t_{\ell_1,\ell_2}-q)\rangle_{t,\sim}&=\nu_t\big(\varepsilon_\ell\varepsilon_{\ell'}(\varepsilon_1\varepsilon_2-q)(R_{\ell_1,\ell_2}-q)\big)+O(2)\\&=\nu_0\big(\varepsilon_\ell\varepsilon_{\ell'}(\varepsilon_1\varepsilon_2-q)(R_{\ell_1,\ell_2}-q)\big)+O(2)\\&=\nu_0\big(\varepsilon_\ell\varepsilon_{\ell'}(\varepsilon_1\varepsilon_2-q)(R^-_{\ell_1,\ell_2}-q)\big)+O(2)\\&=\nu_0(\varepsilon_\ell\varepsilon_{\ell'}(\varepsilon_1\varepsilon_2-q))\nu_0(R^-_{\ell_1,\ell_2}-q)+O(2)\\&=\nu_0(\varepsilon_\ell\varepsilon_{\ell'}(\varepsilon_1\varepsilon_2-q))\nu_0(R_{\ell_1,\ell_2}-q)+O(2)\\&=\nu_0(\varepsilon_\ell\varepsilon_{\ell'}(\varepsilon_1\varepsilon_2-q))\nu(R_{1,2}-q)+O(2)\ .\end{aligned}\tag{9.143}$$

In this manner we obtain from (9.142) that

$$\alpha\nu_v\big(f\varepsilon_\ell\varepsilon_{\ell'}u'(S^\ell_v)u'(S^{\ell'}_v)\big)=\alpha\widehat{r}\mathsf{E}\langle f\varepsilon_\ell\varepsilon_{\ell'}\rangle_{t,\sim}+A\nu(R_{1,2}-q)+O(2)\ ,$$

where A is a sum of terms of the type $\pm\alpha c(j)\nu_0(\varepsilon_\ell\varepsilon_{\ell'}(\varepsilon_1\varepsilon_2-q))$.

We collect the terms in (2.23). We associate each quantity $\alpha\widehat{r}\mathsf{E}\langle f\varepsilon_\ell\varepsilon_{\ell'}\rangle_{t,\sim}$ with the corresponding term in (2.25), and we get

$$\nu_t'(f)=\mathrm{I}+\mathrm{II}+O(2)\ ,$$

where I is a sum of terms of the kind $\pm r(\mathsf{E}\langle\varepsilon_\ell\varepsilon_{\ell'}f\rangle_{t,\sim}-\nu_t(\varepsilon_\ell\varepsilon_{\ell'}f))$ and $\mathrm{II}=A\nu(R_{1,2}-q)$, with A a sum of terms of the type $\pm\alpha c(j)\nu_0(\varepsilon_\ell\varepsilon_{\ell'}(\varepsilon_1\varepsilon_2-q))$.

To control the term I we will control separately each difference

$$\mathsf{E}\langle\varepsilon_\ell\varepsilon_{\ell'}f\rangle_{t,\sim}-\nu_t(\varepsilon_\ell\varepsilon_{\ell'}f)\ .$$

For each such difference we consider the function $\varphi(v)=\nu_{t,v}(f\varepsilon_\ell\varepsilon_{\ell'})$ and we use (9.142) again, together with the second part of Lemma 9.7.7, so $\varphi(0)=\mathsf{E}\langle\varepsilon_\ell\varepsilon_{\ell'}f\rangle_{t,\sim}$ and $\varphi'(0)$ is given by (9.136). Proceeding as in (9.143) if follows from $\nu_0(\varepsilon_1\varepsilon_2-q)=0$ that

$$\mathsf{E}\langle f(R^t_{\ell_1,\ell_2}-q)\rangle_{t,\sim}=\nu_0(\varepsilon_1\varepsilon_2-q)\nu(R_{1,2}-q)+O(2)=O(2)\ ,$$

so that $\mathrm{I}=O(2)$.

In conclusion we have shown that

$$\nu_t'(f) = A\nu(R_{1,2} - q) + O(2) \ ,$$

and since $\nu_0(f) = 0$,

$$\nu_1(f) = \nu(R_{1,2} - q) = A\nu(R_{1,2} - q) + O(2) \ .$$

This implies that $\nu(R_{1,2} - q) = O(2)$ (the desired result) provided we can show that $|A| \le 1/2$. To prove Lemma 9.5.12 we need to know this is the case under a condition of the form $L\alpha \exp L\tau^2 \le 1$. It suffices to show that the quantities $\widehat{r}, c(0), c(1), c(2)$ are all $\le \exp L\tau^2$. We have obtained this bound for $\widehat{r}$ in the proof of the uniqueness of the equations (9.65); the case of $c(j)$ is entirely similar. □

We now turn to the "algebra". We consider the numbers

$$d(0) = \widehat{q} - q^2 \ ; \quad d(1) = q - q^2 \ ; \quad d(2) = 1 - q^2 \ ,$$

where $q = \mathsf{E}\mathrm{th}^2(z\sqrt{r})$ and $\widehat{q} = \mathsf{E}\mathrm{th}^4(z\sqrt{r})$. We observe the formula

$$\nu_0((\varepsilon_1\varepsilon_2 - q)\varepsilon_1\varepsilon_2) = d(\mathrm{card}\{1,2\} \cap \{\ell, \ell'\}) \ . \tag{9.144}$$

We consider the numbers

$$b_0(2) = \ c(2)\, d(2) - 4\, c(1)\, d(1) + 3\, c(0)\, d(0) \tag{9.145}$$

$$\begin{aligned} b_0(1) = \ & c(1)\, d(2) + c(2)\, d(1) - 2\, c(1)\, d(1) - 3\, c(0)\, d(1) - 3\, c(1)\, d(0) \\ & + 6\, c(0)\, d(0) \end{aligned} \tag{9.146}$$

$$\begin{aligned} b_0(0) = \ & c(2)\, d(0) + c(0)\, d(2) + 4\, c(1)\, d(1) - 8\, c(1)\, d(0) - 8\, c(0)\, d(1) \\ & + 10\, c(0)\, d(0) \end{aligned} \tag{9.147}$$

and finally

$$b(2) = \alpha b_0(2) - \alpha\widehat{r}^2(d(2) - 4d(1) + 3d(0)) \tag{9.148}$$

$$b(1) = \alpha b_0(1) - \alpha\widehat{r}^2(d(2) - 4d(1) + 3d(0)) \tag{9.149}$$

$$b(0) = \alpha b_0(0) - \alpha\widehat{r}^2(d(2) - 4d(1) + 3d(0)) \ . \tag{9.150}$$

Despite the fact that the previous formulas look a bit complicated, there definitely exists some structure (that is not entirely elucidated). The next lemma seems to indicate that, somewhere, we take the product of two operators. Clarifying what really happens seems related to Research Problem 1.8.3.

Lemma 9.7.10. *We have*

$$b(2)-2\,b(1)+b(0)=\alpha\,(c(2)-2\,c(1)+c(0))(d(2)-2\,d(1)+d(0)) \quad (9.151)$$

$$b(2)-4\,b(1)+3\,b(0)=\alpha\,(c(2)-4\,c(1)+3\,c(0))(d(2)-4\,d(1)+3\,d(0))\;. \quad (9.152)$$

Proof. Straightforward algebra. □

One should also mention that $d(2)-4\,d(1)+3\,d(0)$ is the quantity $1-4\,q+3\,\widehat{q}$, that already occurred in Section 1.8 (see e.g. (1.235)).

Theorem 9.7.11. *There exists a number L with the following property. Assume that u satisfies (9.2), (9.122) and (9.128). Then if $L\alpha\exp L\tau^2\le 1$, given a function f^- on Σ_{N-1}^n, which is a product of k functions of the type $R^-_{\ell,\ell'}-q$, $\ell,\ell'\le n$, and given integers $x,y\le n$, we have*

$$\begin{aligned}\nu((\varepsilon_x\varepsilon_y-q)f^-)=&\sum_{1\le\ell<\ell'\le n}b(\ell,\ell';x,y)\nu(f^-(R_{\ell,\ell'}-q))\\&-n\sum_{\ell\le n}b(\ell,n+1;x,y)\nu(f^-(R_{\ell,n+1}-q))\\&+\frac{n(n+1)}{2}b(0)\nu(f^-(R_{n+1,n+2}-q))\\&+O(k+2)\,, \end{aligned}\quad (9.153)$$

where $b(\ell,\ell';x,y)=b(\mathrm{card}(\{\ell,\ell'\}\cap\{x,y\}))$ and $b(j)$, $j=0,1,2$ are given by (9.148) to (9.150).

Here, $O(k+2)$ is a quantity B such that $|B|\le K(C,n,\tau)N^{-(k+2)/2}$, when $K(C,n,\tau)$ is independent of n (and in fact also of the choice of u).

Once this theorem is proved, we can copy the proof of Theorem 1.10.1 in the present setting. Repeating the computations of Section 1.8, the values of A,B, and C are now given by

$$\begin{aligned}A^2&=\frac{1-2q+\widehat{q}}{N(1-(b(2)-2b(1)+b(0)))}\\B^2&=\frac{1}{1-(b(2)-4b(1)+3b(0))}\left(\frac{1}{N}(q-\widehat{q})+(b(1)-b(0))A^2\right)\\C^2&=\frac{1}{1-(b(2)-4b(1)+3b(0))}\left(\frac{1}{N}(\widehat{q}-q^2)+b(0)A^2+(4b(1)-6b(0))B^2\right).\end{aligned}$$

Theorem 9.7.12. *There exists a number L with the following property. Assume that u satisfies (9.2), (9.122) and (9.128). Then if $L\alpha\exp L\tau^2\le 1$, the following occurs with the values of A,B,C given above. Consider an integer n. For $1\le\ell<\ell'\le n$ consider integers $k(\ell,\ell')$. For $1\le\ell\le n$ consider* *integers* $k(\ell)$. *Set* $k_1 = \sum_{1\le\ell<\ell'\le n} k(\ell,\ell')$, $k_2 = \sum_{1\le\ell\le n} k(\ell)$, *consider an integer* k_3 *and set* $k = k_1 + k_2 + k_3$. *Then*

$$\nu\Big(\prod_{1\le\ell<\ell'\le n} T_{\ell,\ell'}^{k(\ell,\ell')} \prod_{1\le\ell\le n} T_\ell^{k(\ell)} T^{k_3}\Big)$$
$$= \prod_{1\le\ell<\ell'\le n} a(k(\ell,\ell')) \prod_{1\le\ell\le n} a(k(\ell)) a(k_3) A^{k_1} B^{k_2} C^{k_3} + O(k+1)\,.$$

So, it remains only to prove (9.153). The scheme of proof is as follows. we may assume $x = 1$, $y = 2$, we use that $\nu_0((\varepsilon_1\varepsilon_2 - q)f^-) = 0$, so that

$$\nu((\varepsilon_1\varepsilon_2 - q)f^-) = \int_0^1 \nu_t'((\varepsilon_1\varepsilon_2 - q)f^-)\mathrm{d}t\,.$$

We compute $\nu_t'((\varepsilon_1\varepsilon_2 - q)f^-)$ using (2.23). For each term of the type

$$\nu_t\big((\varepsilon_1\varepsilon_2 - q)f^-\varepsilon_\ell\varepsilon_{\ell'}u'(S^\ell_{M,t})u'(S^{\ell'}_{M,t})\big)\,,$$

we consider the function

$$\varphi(v) = \nu_{t,v}\big((\varepsilon_1\varepsilon_2 - q)f^-\varepsilon_\ell\varepsilon_{\ell'}u'(S^\ell_v)u'(S^{\ell'}_v)\big)\,.$$

We know from (9.101) that $\nu((R^-_{1,2} - q)^{2k}) = O(2k)$, so that from (9.99) we have $\nu_t((R^-_{1,2} - q)^{2k}) = O(2k)$ (uniformly in t) and we know through (9.130) that $\varphi''(v) = O(k+2)$, so

$$\varphi(1) = \varphi(0) + \varphi'(0) + O(k+2)\,. \tag{9.154}$$

We compute $\varphi'(0)$ using (9.134). According to Lemmas 9.7.8 and 9.7.9, within error $O(k+2)$ we may replace the terms

$$\mathsf{E}\langle(\varepsilon_1\varepsilon_2 - q)f^-\varepsilon_\ell\varepsilon_{\ell'}(R^t_{\ell_1,\ell_2} - q)\rangle_{t,\sim}$$

by

$$\nu_0\big((\varepsilon_1\varepsilon_2 - q)f^-\varepsilon_\ell\varepsilon_{\ell'}(R_{\ell_1,\ell_2} - q)\big)$$
$$= \nu_0(\varepsilon_\ell\varepsilon_{\ell'}(\varepsilon_1\varepsilon_2 - q))\nu_0(f^-(R_{\ell_1,\ell_2} - q))$$

and we may in turn replace $\nu_0(f^-(R_{\ell_1,\ell_2} - q))$ by $\nu(f^-(R_{\ell_1,\ell_2} - q))$ using Lemma 9.7.9 again. The terms $\varphi(0) = \widehat{r}\mathsf{E}\langle(\varepsilon_1\varepsilon_2 - q)\varepsilon_\ell\varepsilon_{\ell'}f^-\rangle_{t,\sim}$ regroup with the corresponding terms of the quantity II of (2.23). To compute the difference

$$\mathsf{E}\langle(\varepsilon_1\varepsilon_2 - q)\varepsilon_\ell\varepsilon_{\ell'}f^-\rangle_{t,\sim} - \nu_t((\varepsilon_1\varepsilon_2 - q)\varepsilon_\ell\varepsilon_{\ell'}f^-)\,,$$

we introduce the function $\varphi(v) = \nu_{t,v}((\varepsilon_1\varepsilon_2 - q)\varepsilon_\ell\varepsilon_{\ell'}f^-)$, we use (9.154) and (9.136) to compute $\varphi'(0)$. Within an error $O(k+2)$ we reach as previously that this difference is a sum of terms

$$\pm\nu_0((\varepsilon_1\varepsilon_2 - q)\varepsilon_\ell\varepsilon_{\ell'})\nu(f^-(R_{\ell_1,\ell_2} - q)) \; .$$

By this procedure we have obtained that within an error $O(k+2)$, $\nu((\varepsilon_1\varepsilon_2 - q)f^-)$ is a certain sum of terms of the type $C_{\ell_1,\ell_2}\nu(f^-(R_{\ell_1,\ell_2} - q))$; and to complete the proof it remains to perform the algebra: the computation of these coefficients C_{ℓ_1,ℓ_2}. This computation will require carefully counting terms in certain formulas. This could look tedious, unless of course one keeps marveling, as one should, about why the relations we find can be at all true. For the computation of the terms C_{ℓ_1,ℓ_2}, it helps to use the quantities of (1.226), i.e.

$$T_{\ell,\ell'} = \frac{(\boldsymbol{\sigma}^\ell - \mathbf{b})\cdot(\boldsymbol{\sigma}^{\ell'} - \mathbf{b})}{N} \; , \quad T_\ell = \frac{(\boldsymbol{\sigma}^\ell - \mathbf{b})\cdot\mathbf{b}}{N} \; , \quad T = \frac{\mathbf{b}\cdot\mathbf{b}}{N} - q \; ,$$

for $\mathbf{b} = (\langle\sigma_i\rangle)_{i\le N}$. We start with a general identity.

Lemma 9.7.13. *Consider numbers $a(0), a(1)$ and $a(2)$. Given two integers $\ell, \ell' \le n$ we define*

$$a(\ell_1,\ell_2) = a(\mathrm{card}\{\ell_1,\ell_2\}\cap\{\ell,\ell'\}) \; .$$

Then for any function f on Σ_N^n we have the identity

$$\begin{aligned}&\sum_{1\le\ell_1<\ell_2\le n} a(\ell_1,\ell_2)\nu(f(R_{\ell_1,\ell_2} - q))\\ &- n\sum_{\ell_1\le n} a(\ell_1,n+1)\nu(f(R_{\ell_1,n+1} - q))\\ &+ \frac{n(n+1)}{2}a(n+1,n+2)\nu(f(R_{n+1,n+2} - q))\\ &= \sum_{\ell_1<\ell_2} a(\ell_1,\ell_2)\nu(fT_{\ell_1,\ell_2}) + \sum_{\ell_1} a_1(\ell_1)\nu(fT_{\ell_1})\\ &+ (a(2) - 4a(1) + 3a(0))\nu(fT)\end{aligned} \tag{9.155}$$

where

$$a_1(\ell_1) = \begin{cases} 2a(1) - 3a(0) & \text{if } \ell_1 \notin \{\ell,\ell'\} \\ a(2) - 2a(1) & \text{if } \ell_1 \in \{\ell,\ell'\}. \end{cases}$$

The reader observes that the range of summation need not be specified for ℓ_1 and ℓ_2 in the right hand side of (9.155), because $\nu(fT_{\ell_1})$ is zero if f does not depend on replica ℓ_1, and similarly for $\nu fT_{\ell_1,\ell_2}$.

Proof. We substitute the relation

$$R_{\ell_1,\ell_2} - q = T_{\ell_1,\ell_2} + T_{\ell_1} + T_{\ell_2} + T \tag{9.156}$$

in each of the terms in the left-hand side of the seeked relation (9.155), and we simply count how many times each term occurs in order to get the right-hand side of (9.155). This is straightforward but requires a bit of patience. The coefficient of $\nu(fT)$ is

$$\sum_{1\le\ell_1<\ell_2\le n} a(\ell_1,\ell_2) - n\sum_{\ell_1\le n} a(\ell_1,n+1) + \frac{n(n+1)}{2}a(0)\ . \tag{9.157}$$

In the first sum above, one term exactly is $a(2)$. There are $(n-2)(n-3)/2$ terms for which $\{\ell_1,\ell_2\}\cap\{\ell,\ell'\}=\emptyset$, and which are equal to $a(0)$. Since the sum has $n(n-1)/2$ terms, there are exactly

$$\frac{n(n-1)}{2} - \frac{(n-2)(n-3)}{2} - 1 = 2n-4$$

terms for which $\mathrm{card}(\{\ell_1,\ell_2\}\cap\{\ell,\ell'\})=1$, and which are equal to $a(1)$. The second sum of (9.157) is $2a(1)+(n-2)a(0)$ so that (9.157) is

$$\begin{aligned}&a(2)-4a(1)+a(0)\left(\frac{(n-2)(n-3)}{2} - n(n-2) + \frac{n(n+1)}{2}\right)\\ &= a(2)-4a(1)+3a(0)\ .\end{aligned}$$

To compute the coefficient of $\nu(fT_{\ell_1})$, we may assume $\ell_1\le n$, for otherwise $\nu(fT_{\ell_1})=0$ since $\langle fT_{\ell_1}\rangle=0$. This coefficient is then

$$\sum_{\ell_2\le n,\ell_2\ne\ell_1} a(\ell_1,\ell_2) - na(\ell_1,n+1)\ .$$

When $\ell_1\notin\{\ell,\ell'\}$, this is

$$2a(1)+(n-3)a(0)-na(0) = 2a(1)-3a(0) = a_1(\ell_1)\ ,$$

while if $\ell_1\in\{\ell,\ell'\}$ this is

$$a(2)+(n-2)a(1)-na(1) = a(2)-2a(1) = a_1(\ell_1)\ . \qquad \square$$

Lemma 9.7.14. *Under the conditions of Proposition 9.7.5, if f^- is a product of k functions of the type $R^-_{\ell,\ell'}-q$, $\ell,\ell'\le n$ and*

$$f' = \varepsilon_\ell\varepsilon_{\ell'}(\varepsilon_1\varepsilon_2-q)f^-\ ,$$

we have

$$\begin{aligned}\nu_t(f'u'(S^\ell_{M,t})u'(S^{\ell'}_{M,t})) &= \widehat{r}\mathsf{E}\langle f'\rangle_{t,\sim}\\ &+ \sum_{\ell_1<\ell_2} c(\ell,\ell';\ell_1,\ell_2)\nu(f'T_{\ell_1,\ell_2})\\ &+ \sum_{\ell_1} c(\ell_1;\ell,\ell')\nu(f'T_{\ell_1})\\ &+ (c(2)-4c(1)+3c(0))\nu(f'T)\\ &+ O(k+2)\end{aligned} \tag{9.158}$$

where

$$c(\ell_1;\ell,\ell') = \begin{cases} 2c(1) - 3c(0) \text{ if } \ell_1 \notin \{\ell,\ell'\} \\ c(2) - 2c(1) \text{ if } \ell_1 \in \{\ell,\ell'\}. \end{cases} \tag{9.159}$$

We also have

$$\nu_t(f'u'(S^\ell_{M,t})u'(S^{\ell'}_{M,t})) = \widehat{r}\mathsf{E}\langle f'\rangle_{t,\sim} + A_{\ell,\ell'} + O(k+2) \tag{9.160}$$

where

$$A_{\ell,\ell'} = d(\ell,\ell')\bigg(\sum_{\ell_1<\ell_2} c(\ell,\ell';\ell_1,\ell_2)\nu(f^-T_{\ell_1,\ell_2}) + \sum_{\ell_1} c(\ell_1;\ell,\ell')\nu(f^-T_{\ell_1}) + (c(2) - 4c(1) + 3c(0))\nu(f^-T)\bigg) \tag{9.161}$$

for $d(\ell,\ell') = d(1,2;\ell,\ell') = d(\mathrm{card}(\{1,2\}\cap\{\ell,\ell'\}))$.

Proof. We first show that (9.158) implies (9.160). For this we simply write (using Lemma 9.7.9) that

$$\begin{aligned}\nu(f'T_{\ell_1,\ell_2}) &= \nu((\varepsilon_1\varepsilon_2 - q)\varepsilon_\ell\varepsilon_{\ell'}f^-T_{\ell_1,\ell_2}) \\ &= \nu_0((\varepsilon_1\varepsilon_2 - q)\varepsilon_\ell\varepsilon_{\ell'}f^-T_{\ell_1,\ell_2}) + O(k+2) \\ &= \nu_0((\varepsilon_1\varepsilon_2 - q)\varepsilon_\ell\varepsilon_{\ell'})\nu_0(f^-T_{\ell_1,\ell_2}) + O(k+2) \\ &= d(\ell,\ell')\nu_0(f^-T_{\ell_1,\ell_2}) + O(k+2) \\ &= d(\ell,\ell')\nu(f^-T_{\ell_1,\ell_2}) + O(k+2)\ .\end{aligned}$$

Passing from the second to the third line goes via Lemma 2.2.1, using as an intermediate step if one wishes that $T_{\ell_1,\ell_2} = T^-_{\ell_1,\ell_2} + (\sigma^{\ell_1}_N - \langle\sigma_N\rangle)(\sigma^{\ell_2}_N - \langle\sigma_N\rangle)/N$, where $T^-_{\ell_1,\ell_2}$ does not depend on the last spins. In a similar manner we get

$$\nu(f'T_{\ell_1}) = d(\ell,\ell')\nu(f^-T_{\ell_1}) + O(k+2)\ ,$$

and

$$\nu(f'T) = d(\ell,\ell')\nu(f^-T) + O(k+2)\ .$$

Substituting these relations in (9.158) proves (9.160).

To prove (9.158) we deduce from Lemmas 9.7.7and 9.7.8 (handling the error terms as already explained) that

$$\begin{aligned}\nu_t(f'u'(S^\ell_{M,t})u'(S^{\ell'}_{M,t})) = \widehat{r}\mathsf{E}\langle f'\rangle_{t,\sim} &+ \sum_{1\le\ell_1<\ell_2\le n} c(\ell_1,\ell_2;\ell,\ell')\nu(f'(R_{\ell_1,\ell_2} - q)) \\ &- n\sum_{\ell_1\le n} c(\ell_1,n+1;\ell,\ell')\nu(f'(R_{\ell_1,n+1} - q)) \\ &+ \frac{n(n+1)}{2}c(n+1,n+2;\ell,\ell')\nu(f'(R_{n+1,n+2} - q)) \\ &+ O(k+2)\ .\end{aligned}$$

We then use Lemma 9.7.13 with $a(j) = c(j)$ to get the result. □

Lemma 9.7.15. *Under the conditions of Proposition 9.7.5, if f^- is a product of k functions of the type $R^-_{\ell,\ell'} - q$, $\ell, \ell' \le n$, and if $f' = \varepsilon_\ell \varepsilon_{\ell'}(\varepsilon_1\varepsilon_2 - q)f^-$, we have*

$$\nu_t(f') = \mathsf{E}\langle f'\rangle_{t,\sim} + B_{\ell,\ell'} + O(k+2) \,, \tag{9.162}$$

where

$$B_{\ell,\ell'} = \widehat{r}d(\ell,\ell')\bigg(\sum_{\ell_1<\ell_2} \nu(f^- T_{\ell_1,\ell_2}) - \sum_{\ell_1} \nu(f^- T_{\ell_1})\bigg) \,. \tag{9.163}$$

Proof. This is entirely similar to Lemma 9.7.14, using (9.136) rather than (9.134). In fact, there is no need to reproduce the computation since the right-hand side of (9.136) is obtained from (9.134) simply by replacing everywhere each of the terms $c(0)$, $c(1)$ and $c(2)$ by $\widehat{r}$. □

Recalling the numbers $b_0(j)$ of (9.145) to (9.147), we define

$$b_0(\ell_1,\ell_2) = b_0(\mathrm{card}\{\ell_1,\ell_2\} \cap \{1,2\}) \,.$$

Lemma 9.7.16. *Let $A_{\ell,\ell'}$ and $B_{\ell,\ell'}$ be given by (9.161) and (9.163) respectively. Then the following identities hold:*

$$\begin{aligned}
&\sum_{1\le\ell<\ell'\le n} A_{\ell,\ell'} - n\sum_{\ell\le n} A_{\ell,n+1} + \frac{n(n+1)}{2}A_{n+1,n+2} \\
&= \sum_{1\le\ell_1<\ell_2\le n} b_0(\ell_1,\ell_2)\nu(f^-(R_{\ell_1,\ell_2} - q)) \\
&- n\sum_{\ell_1\le n} b_0(\ell_1,n+1)\nu(f^-(R_{\ell_1,n+1} - q)) \\
&+ \frac{n(n+1)}{2} b_0(0)\nu(f^-(R_{n+1,n+2} - q))
\end{aligned} \tag{9.164}$$

and

$$\begin{aligned}
&\sum_{1\le\ell<\ell'\le n} B_{\ell,\ell'} - n\sum_{\ell\le n} B_{\ell,n+1} + \frac{n(n+1)}{2}B_{n+1,n+2} \\
&= \widehat{r}(d(2) - 4d(1) + 3d(0))\bigg(\sum_{1\le\ell_1<\ell_2\le n} \nu(f^-(R_{\ell_1,\ell_2} - q)) \\
&- \sum_{\ell_1\le n} \nu(f^-(R_{\ell_1,n+1} - q)) + \frac{n(n+1)}{2}\nu(f^-(R_{n+1,n+2} - q))\bigg).
\end{aligned} \tag{9.165}$$

Proof. We prove (9.164) first. We use Lemma 9.7.13 with $a(j) = b_0(j)$ to see that the right-hand side of this quantity is

$$\sum_{\ell_1<\ell_2} b_0(\ell_1,\ell_2)\nu(f^-T_{\ell_1,\ell_2}) + \sum_{\ell_1} b_1(\ell_1)\nu(f^-T_{\ell_1})$$
$$+ (b_0(2) - 4b_0(1) + 3b_0(0))\nu(f^-T) \;, \tag{9.166}$$

where

$$b_1(\ell) = \begin{cases} 2b_0(1) - 3b_0(0) \text{ if } \ell_1 \notin \{1,2\} \\ b_0(2) - 2b_0(1) \quad \text{if } \ell_1 \in \{1,2\}. \end{cases}$$

We will show that the coefficients of $\nu(f^-T_{\ell_1,\ell_2})$, $\nu(f^-T_{\ell_1})$ and $\nu(f^-T)$ are the same in (9.166) and in the left-hand side of (9.164). That is, recalling (9.161), that $d(\ell,\ell') = d(\text{card}(\{\ell,\ell'\} \cap \{1,2\})$ and (9.159), we have to prove the relations

$$\sum_{1\le\ell<\ell'\le n} c(\ell,\ell';\ell_1,\ell_2)d(\ell,\ell') - n\sum_{\ell\le n} c(\ell,n+1;\ell_1,\ell_2)d(\ell,n+1)$$
$$+ \frac{n(n+1)}{2}c(n+1,n+2;\ell_1,\ell_2)d(n+1,n+2) = b_0(\ell_1,\ell_2) \;; \tag{9.167}$$

$$\sum_{1\le\ell<\ell'\le n} c(\ell_1;\ell,\ell')d(\ell,\ell') - n\sum_{\ell\le n} c(\ell_1;\ell,n+1)d(\ell,n+1)$$
$$+ \frac{n(n+1)}{2}c(\ell_1;n+1,n+2)d(n+1,n+2) = b_1(\ell_1) \;; \tag{9.168}$$

$$(c(2) - 4c(1) + 3c(0))\Bigg(\sum_{1\le\ell<\ell'\le n} d(\ell,\ell') - n\sum_{\ell\le n} d(\ell,n+1)$$
$$+ \frac{n(n+1)}{2}d(n+1,n+2)\Bigg) = b_0(2) - 4b_0(1) + 3b_0(0) \;. \tag{9.169}$$

To prove these relations we may assume that $n > \ell_1, \ell_2$. This is because in (2.23) we may increase n if we wish, since the extra terms this creates cancel out. The proof is completely straightforward, but it requires real patience. The impatient reader may jump ahead directly to the proof of Theorem 9.7.11.

Proof of (9.167). *Case 1: $\{\ell_1,\ell_2\} = \{1,2\}$.*

There are respectively

$$1 \;\;;\;\; 2n-4 \;\;;\;\; \frac{(n-2)(n-3)}{2}$$

choices of $1 \le \ell < \ell' \le n$ for which $\text{card}(\{\ell,\ell'\} \cap \{1,2\}) = 2$, 1, or 0. Therefore the term $\sum_{1\le\ell<\ell'\le n}$ in the left-hand side of (9.167) is

$$c(2)d(2) + c(1)d(1)(2n-4) + c(0)d(0)\left(\frac{(n-2)(n-3)}{2}\right) .$$

There are respectively 2 and $n-2$ choices of $\ell \le n$ for which $\mathrm{card}(\{\ell, n+1\} \cap \{1,2\}) = 1$ or 0, and the term $\sum_{\ell \le n}$ in the left-hand side of (9.167) is

$$2c(1)d(1) + (n-2)c(0)d(0)\ .$$

Therefore the left-hand side of (9.167) is

$$\begin{aligned}
&c(2)d(2) + c(1)d(1)(2n-4) + c(0)d(0)\left(\frac{(n-2)(n-3)}{2}\right)\\
&- n(2c(1)d(1) + (n-2)c(0)d(0)) + \frac{n(n+1)}{2}c(0)d(0)\\
&= c(2)d(2) - 4c(1)d(1) + 3c(0)d(0) = b_0(2) = b_0(1,2) = b_0(\ell_1, \ell_2)\ .
\end{aligned}$$

Case 2: $\mathrm{card}(\{1,2\} \cap \{\ell_1, \ell_2\}) = 1$.

Without loss of generality we assume $\ell_1 = 1$, $\ell_2 = 3$. The sum $\sum_{1 \le \ell < \ell' \le n}$ in the left-hand side of (9.167) is best computed by first calculating the sum over ℓ' for $\ell = 1, 2, 3$. This sum is as follows.

If $\ell = 1$:

$$c(1)d(2) + c(2)d(1) + (n-3)c(1)d(1)\ ,$$

corresponding respectively to the case $\ell' = 2$, $\ell' = 3$, $\ell' \ge 4$.

If $\ell = 2$:

$$c(1)d(1) + (n-3)c(0)d(1)\ ,$$

corresponding respectively to $\ell' = 3$, $\ell' \ge 4$.

If $\ell = 3$:

$$(n-3)c(1)d(0)\ .$$

Moreover, the sum $\sum_{4 \le \ell < \ell' \le n}$ is

$$\frac{(n-3)(n-4)}{2}c(0)d(0)\ .$$

The sum $\sum_{\ell \le n}$ is

$$c(1)d(1) + c(0)d(1) + c(1)d(0) + (n-3)c(0)d(0)\ ,$$

the terms corresponding of course to the cases $\ell = 1$, $\ell = 2$, $\ell = 3$, $\ell \ge 4$. Collecting the terms and using that

$$\frac{(n-3)(n-4)}{2} - n(n-3) + \frac{n(n+1)}{2} = 6$$

to compute the coefficient of $c(0)\,d(0)$, we get a total contribution of

$$c(2)d(1) + c(1)d(2) - 2c(1)d(1) - 3c(0)d(1) - 3c(1)d(0) + 6c(0)d(0)\ ,$$

and this is $b_0(1) = b_0(\ell_1, \ell_2)$.

Case 3: $\{\ell_1, \ell_2\} \cap \{1, 2\} = \emptyset$.

We first compute the sum $\sum_{1 \le \ell < \ell' \le n}$ in the left-hand side of (9.167). There are 6 pairs $1 \le \ell < \ell' \le n$ such that $\{\ell, \ell'\} \subset \{1, 2, \ell_1, \ell_2\}$, for a total contribution of

$$c(2)d(0) + c(0)d(2) + 4c(1)d(1) \; .$$

There are

$$\frac{n(n-1)}{2} - \frac{(n-4)(n-5)}{2} - 6 = 4n - 16$$

choices of $1 \le \ell < \ell' \le n$ for which $\operatorname{card}(\{\ell, \ell'\} \cap \{1, 2, \ell_1, \ell_2\}) = 1$, for a total contribution of

$$(2n-8)(c(0)d(1) + c(1)d(0)) \; .$$

Thus the sum $\sum_{1 \le \ell < \ell' \le n}$ in the left-hand side of (9.167) is

$$\begin{aligned} &c(2)d(0) + c(0)d(2) + 4c(1)d(1) + (2n-8)(c(0)d(1) + c(1)d(0)) \\ &+ \frac{(n-4)(n-5)}{2} c(0)d(0) \; . \end{aligned}$$

Next we compute the sum $\sum_{\ell \le n}$ in the left-hand side of (9.167). We distinguish the cases where $\ell = 1, 2, \ell_1, \ell_2$ to obtain that this sum is

$$2c(1)d(0) + 2c(0)d(1) + (n-4)c(0)d(0) \; .$$

Thus the left-hand side of (9.167) is

$$\begin{aligned} &c(2)d(0) + c(0)d(2) + 4c(1)d(1) + (2n-8)(c(0)d(1) + c(1)d(0)) \\ &+ \frac{(n-4)(n-5)}{2} c(0)d(0) - n(2c(1)d(0) + 2c(0)d(1) + (n-4)c(0)d(0)) \\ &+ \frac{n(n+1)}{2} c(0)d(0) \; . \end{aligned}$$

Using that

$$\frac{(n-4)(n-5)}{2} - n(n-4) + \frac{n(n+1)}{2} = 10 \; ,$$

this is

$$c(2)d(0) + c(0)d(2) + 4c(1)d(1) - 8(c(0)d(1) + c(1)d(0)) + 10c(0)d(0) \; ,$$

which is $b_0(0) = b_)(\ell_1, \ell_2)$, and we have proved (9.167).

Proof of (9.168). We set $c'(1) = c(2) - 2c(1)$, $c'(0) = 2c(1) - 3c(0)$, so that by (9.159) we have $c(\ell_1; \ell, \ell') = c'(\operatorname{card}(\{\ell_1\} \cap \{\ell, \ell'\}))$.

Case 1: $\ell_1 \notin \{1, 2\}$.

Without loss of generality we assume that $\ell_1 = 3$. To compute the sum $\sum_{1\le\ell<\ell'\le n}$ of the left-hand side of (9.168), we again compute first the sum over ℓ' for $\ell = 1, 2, 3$.

If $\ell = 1$:

$$c'(0)d(2) + c'(1)d(1) + (n-3)c'(0)d(1) \ ,$$

corresponding respectively to the values $\ell' = 2$, $\ell' = 3$, $\ell' \ge 4$.

If $\ell = 2$:

$$c'(1)d(1) + (n-3)c'(0)d(1) \ ,$$

corresponding respectively to the values $\ell' = 3$ and $\ell' \ge 4$.

If $\ell = 3$:

$$(n-3)c'(1)d(0) \ .$$

Moreover the sum $\sum_{4\le\ell<\ell'\le n}$ is

$$\frac{(n-3)(n-4)}{2}c'(0)d(0) \ .$$

The sum $\sum_{\ell\le n}$ is

$$2c'(0)d(1) + c'(1)d(0) + (n-3)c'(0)d(0).$$

The first term is the contribution of the values $\ell = 1, 2$, the second term is the contribution of the values $\ell = 3$ and the third term is the contribution of the values $\ell \ge 4$. Collecting the terms we find a total contribution of

$$2c'(1)d(1) + c'(0)d(2) - 3c'(1)d(0) - 6c'(0)d(1) + 6c'(0)d(0) \ .$$

Substituting the values of $c'(1)$ and $c'(0)$, algebra yields the following expression

$$\begin{aligned} &2(c(2)d(1) + c(1)d(2)) - 3(c(2)d(0) + c(0)d(2)) - 16c(1)d(1) \\ &+ 18(c(1)d(0) + c(0)d(1)) - 18c(0)d(0) \end{aligned}$$

and this is indeed $2b_0(1) - 3b_0(0) = b_1(\ell_1)$.

Case 2: $\boxed{\ell_1 \in \{1, 2\}}$.

Without loss of generality we assume that $\ell_1 = 1$. The contribution of the sum $\sum_{1\le\ell<\ell'\le n}$ for the various values of ℓ is as follows.

If $\ell = 1$:

$$c'(1)d(2) + (n-2)c'(1)d(1) \ ,$$

corresponding to the terms $\ell' = 2$ and $\ell' \ge 3$.

If $\ell = 2$:

$$(n-2)c'(0)d(1) \ .$$

The sum of the contributions for $3 \le \ell \le n$ is

$$\frac{(n-2)(n-3)}{2} c'(0) d(0) \; .$$

The sum $\sum_{\ell \le n}$ is

$$c'(1)d(1) + c'(0)d(1) + (n-2)c'(0)d(0) \; ,$$

corresponding to the cases $\ell = 1$, $\ell = 2$, and $\ell \ge 3$.

Collecting the terms, the total contribution is

$$c'(1)d(2) - 2c'(1)d(1) - 2c'(0)d(1) + 3c'(0)d(0) \; ,$$

and substituting the values of $c'(1), c'(0)$ this is indeed $b_0(2) - 2b_0(1) = b_1(\ell_1)$. We have proved (9.168).

Proof of (9.169). It is simpler than the previous one. As in (9.151) we have

$$b_0(2) - 2\, b_0(1) + b_0(0) = (c(2) - 2\, c(1) + c(0))(d(2) - 2\, d(1) + d(0)) \; ,$$

so it suffices to show that

$$\sum_{1 \le \ell < \ell' \le n} d(\ell, \ell') - n \sum_{\ell \le n} d(\ell, n+1) + \frac{n(n+1)}{2} d(n+1, n+2)$$
$$= d(2) - 4d(1) + 3d(0) \; .$$

The computation has been done many times and is left to the reader.

We have proved (9.169), (9.168) and (9.167). Therefore we have proved (9.164).

To prove (9.165), we simply notice that we obtain $B_{\ell,\ell'}$ from $A_{\ell,\ell'}$ by replacing each of the quantities $c(2), c(1)$ and $c(0)$ by $\widehat{r}$: this replaces each of the quantities $b_0(2), b_0(1)$ and $b_0(0)$ by $\widehat{r}(d(2) - 4d(1) + 3d(0))$. □

Proof of Theorem 9.7.11. We apply (2.23) to the function $f = (\varepsilon_1\varepsilon_2 - q)f^-$. We apply (9.160) to each term $\nu_t(\varepsilon_\ell\varepsilon_{\ell'}(\varepsilon_1\varepsilon_2 - q)u'(S^\ell_{M,t})u'(S^{\ell'}_{M,t})f^-)$. Setting

$$D_{\ell,\ell'} = \mathsf{E}\langle(\varepsilon_1\varepsilon_2 - q)\varepsilon_\ell\varepsilon_{\ell'} f^-\rangle_{t,\sim} - \nu_t((\varepsilon_1\varepsilon_2 - q)\varepsilon_\ell\varepsilon_{\ell'} f^-) \; ,$$

and, combining with the contribution of the term II of (2.23) we get

$$\begin{aligned}
\nu_t'((\varepsilon_1\varepsilon_2 - q)f^-) = \alpha \Bigg(& \sum_{1 \le \ell < \ell' \le n} A_{\ell,\ell'} - n \sum_{\ell \le n} A_{\ell,n+1} + \frac{n(n+1)}{2} A_{n+1,n+2} \Bigg) \\
& + r \Bigg(\sum_{1 \le \ell < \ell' \le n} D_{\ell,\ell'} - n \sum_{\ell \le n} D_{\ell,n+1} + \frac{n(n+1)}{2} D_{n+1,n+2} \Bigg) \\
& + O(k+2) \; .
\end{aligned}$$

By (9.162) we have $D_{\ell,\ell'} = -B_{\ell,\ell'} + O(k+2)$, and combining with (9.164) and (9.165) proves (9.153). □

9.8 An Approximation Procedure

In the previous sections we worked with functions u such that

$$\forall \ell\ ,\ 1 \le \ell \le 5\ ,\quad |u^{(\ell)}| \le \exp\left(\frac{N}{L}\right)\ . \tag{9.170}$$

In the present section we show that this condition is not really necessary. The method consists of showing that, given a function u (that does not take excessively large values) we can find a function v that satisfies (9.170), for which the Gibbs measures associated to u and v are nearly identical. This shows that in the end the differentiability of u is really irrelevant and it makes one wonder whether we have used the correct approach. Throughout the section we assume (9.2).

Let us give an example of what can be achieved.

Theorem 9.8.1. *There exists a number L with the following property. Assume that the measurable function u satisfies*

$$-\frac{N}{L} \le u \le 0\ ,$$

and consider the solution q of the equations (9.65). Then if $L\alpha \exp L\tau^2 \le 1$ we have

$$\forall k \ge 1\ ,\quad \nu\big((R_{1,2} - q)^{2k}\big) \le \left(\frac{Lk}{N}\right)^k\ .$$

Again, no differentiability of u whatsoever is necessary here.

We now describe the basic approximation procedure. We assume that

$$-D \le u \le 0\ , \tag{9.171}$$

and we will specify D later. Let us consider a (very small) number b. This parameter controls the quality of our approximation of u.

By scaling arguments, there exists a function ζ supported by the interval $[-b, b]$, with $\zeta \ge 0$, and such that

$$\int \zeta(x)\mathrm{d}x = 1$$

$$\forall \ell\ ,\ 0 \le \ell \le 5\ ,\quad |\zeta^{(\ell)}| \le \frac{L}{b^{\ell+1}}\ . \tag{9.172}$$

We define the function v by

$$\exp v(x) = \zeta * \exp u(x) = \int \exp u(t)\zeta(x-t)\mathrm{d}t\ , \tag{9.173}$$

so that $-D \le v \le 0$. Moreover, if $u(x) = 0$ for $x \ge \tau$, we have

$$x \geq \tau + b \quad \Rightarrow \quad v(x) = 0 \ . \tag{9.174}$$

We claim that

$$\forall \ell \ , \ 0 \leq \ell \leq 5 \ , \quad |v^{(\ell)}| \leq L \frac{\exp \ell D}{b^\ell} \ . \tag{9.175}$$

This follows from (9.173) and computation of the derivatives of v. For example, $v' \exp v = \zeta' * \exp u$, and since ζ' is supported by $[-b, b]$ and $|\zeta'| \leq L/b^2$,

$$|\zeta' * \exp u| \leq \int |\zeta'(x)| \mathrm{d}x \leq 2b \frac{L}{b^2} \ .$$

Lemma 9.8.2. *For any number x and any Gaussian r.v. g we have*

$$|\mathsf{E} \exp u(g + x) - \mathsf{E} \exp v(g + x)| \leq \frac{b}{\sqrt{\mathsf{E} g^2}} \ . \tag{9.176}$$

Proof. The function

$$W(x) = \mathsf{E} \exp u(g + x)$$

satisfies (using integration by parts in the second line),

$$W'(x) = \mathsf{E} u'(g + x) \exp u(g + x) = \frac{1}{\mathsf{E} g^2} \mathsf{E} g \exp u(g + x)$$

so that

$$|W'(x)| \leq \frac{\mathsf{E}|g|}{\mathsf{E} g^2} \leq \frac{1}{\sqrt{\mathsf{E} g^2}}$$

and thus $|W(x) - W(x - t)| \leq b/(\mathsf{E} g^2)^{1/2}$ for $|t| \leq b$. Now, the left-hand side of (9.176) is

$$\begin{aligned} |W(x) - \zeta * W(x)| &= \left| W(x) - \int W(x - t) \zeta(t) \mathrm{d}t \right| \\ &= \left| \int (W(x) - W(x - t)) \zeta(t) \mathrm{d}t \right| \\ &\leq \frac{b}{\sqrt{\mathsf{E} g^2}} \int \zeta(t) \mathrm{d}t = \frac{b}{\sqrt{\mathsf{E} g^2}} \ , \end{aligned}$$

and this completes the proof. □

Lemma 9.8.3. *For any subset A of Σ_N we have*

$$\begin{aligned} &\mathsf{E} \left(\sum_{\boldsymbol{\sigma} \in A} \left(\exp \sum_{k \leq M} u(S_k(\boldsymbol{\sigma})) - \exp \sum_{k \leq M} v(S_k(\boldsymbol{\sigma})) \right) \right)^2 \\ &\leq 2 \mathrm{card} A + b^2 N M^2 (\mathrm{card} A)^2 \ . \end{aligned} \tag{9.177}$$

Proof. The left-hand side of (9.177) is

$$\sum_{\boldsymbol{\sigma},\boldsymbol{\sigma}'\in A} \mathsf{E}B(\boldsymbol{\sigma})B(\boldsymbol{\sigma}') \tag{9.178}$$

where

$$B(\boldsymbol{\sigma}) = \exp\sum_{k\le M} u(S_k(\boldsymbol{\sigma})) - \exp\sum_{k\le M} v(S_k(\boldsymbol{\sigma}))\ .$$

In the sum (9.178) we bound separately the terms for which $\boldsymbol{\sigma} = \pm\boldsymbol{\sigma}'$. For these, we use the trivial bound $|B(\boldsymbol{\sigma})| \le 2$. Next, consider a pair $(\boldsymbol{\sigma},\boldsymbol{\sigma}')$ with $\boldsymbol{\sigma} \ne \pm\boldsymbol{\sigma}'$. Since $\boldsymbol{\sigma} \ne -\boldsymbol{\sigma}'$, there exists a coordinate i with $\sigma_i \ne -\sigma'_i$, i.e. $\sigma_i = \sigma'_i$. Since $\boldsymbol{\sigma} \ne \boldsymbol{\sigma}'$, there exists a coordinate j with $\sigma_j = -\sigma'_j$. Without loss of generality, we assume that $\sigma_1 = \sigma'_1$, $\sigma_2 = -\sigma'_2$. Let us denote by E_0 integration in the variables $g_{i,k}$, $k \le M$, $i = 1, 2$, all other r.v.s being fixed.

The key observation is that all the variables of the type

$$\sigma_1\, g_{1,k} + \sigma_2\, g_{2,k} \text{ and } \sigma'_1\, g_{1,k} + \sigma'_2\, g_{2,k} = \sigma_1\, g_{1,k} - \sigma_2\, g_{2,k}$$

are independent as $k \le M$, so that, under E_0, the r.v.s $B(\boldsymbol{\sigma})$ and $B(\boldsymbol{\sigma}')$ are independent, and

$$\mathsf{E}_0(B(\boldsymbol{\sigma})\, B(\boldsymbol{\sigma}')) = \mathsf{E}_0 B(\boldsymbol{\sigma})\, \mathsf{E}_0 B(\boldsymbol{\sigma}')\,.$$

Now

$$\mathsf{E}_0 B(\boldsymbol{\sigma}) = \prod_{k\le M} X_k - \prod_{k\le M} Y_k\ , \tag{9.179}$$

where

$$X_k = \mathsf{E}_0 \exp u(S_k(\boldsymbol{\sigma}))\,;\ Y_k = \mathsf{E}_0 \exp v(S_k(\boldsymbol{\sigma}))\ .$$

We use Lemma 9.8.2 with

$$g = \frac{1}{\sqrt{N}}(\sigma_1 g_{1,k} + \sigma_2 g_{2,k})$$

$$x = \frac{1}{\sqrt{N}}\sum_{i\ge 3}\sigma_i g_{i,k}$$

and we obtain

$$|X_k - Y_k| \le b\sqrt{N}\ . \tag{9.180}$$

Next, we use that for numbers $(x_k)_{k\le M}$, $(y_k)_{k\le M}$, if $|x_k|, |y_k| \le 1$, then

$$\left|\prod_{k\le M} x_k - \prod_{k\le M} y_k\right| \le \sum_{k\le M} |x_k - y_k|\,, \tag{9.181}$$

to deduce from (9.179) and (9.180) that

$$|\mathsf{E}_0 B(\boldsymbol{\sigma})| \le b\sqrt{N}M \ .$$

The same inequality holds for $\boldsymbol{\sigma}'$ rather than $\boldsymbol{\sigma}$, and thus

$$\sum \mathsf{E}(B(\boldsymbol{\sigma})B(\boldsymbol{\sigma}')) \le b^2 N M^2 (\mathrm{card} A)^2 \ ,$$

where the summation is over $\boldsymbol{\sigma} \ne \pm\boldsymbol{\sigma}'$, $\boldsymbol{\sigma}, \boldsymbol{\sigma}' \in A$. This finishes the proof. □

The Cauchy-Schwarz inequality and (9.177) imply the following.

Corollary 9.8.4. *We have*

$$\mathsf{E}\left(\left|\sum_{\boldsymbol{\sigma}\in A}\left(\exp\sum_{k\le M} u(S_k(\boldsymbol{\sigma})) - \exp\sum_{k\le M} v(S_k(\boldsymbol{\sigma}))\right)\right|\right) \le 2\sqrt{\mathrm{card} A} + b\sqrt{N}M\mathrm{card} A \ . \tag{9.182}$$

We use the notation $\langle\cdot\rangle_u$ and $\langle\cdot\rangle_v$ to distinguish the Gibbs measures associated to u and v.

Corollary 9.8.5. *Assume that $u(x) = 0$ for $x \ge \tau$. Given any subset I of $\{1, \ldots, N\}$ and $d \ge \exp(-N/32)$, we have*

$$\mathsf{E}\left|\left\langle\prod_{i\in I}\sigma_i\right\rangle_u - \left\langle\prod_{i\in I}\sigma_i\right\rangle_v\right| \le L d^{1/L} \exp LM(1+\tau^2) + \frac{L}{d}2^{-N/2} + \frac{Lb}{d}\sqrt{N}M \ . \tag{9.183}$$

Proof. Consider the set

$$A = \left\{\boldsymbol{\sigma} \ ; \ \prod_{i\in I}\sigma_i = 1\right\}$$

so that

$$\left\langle\prod_{i\in I}\sigma_i\right\rangle_u = \langle \mathbf{1}_A\rangle_u - \langle \mathbf{1}_{A^c}\rangle_u = 2\langle \mathbf{1}_A\rangle_u - 1$$

and

$$\left|\left\langle\prod_{i\in I}\sigma_i\right\rangle_u - \left\langle\prod_{i\in I}\sigma_i\right\rangle_v\right| \le 2|\langle \mathbf{1}_A\rangle_u - \langle \mathbf{1}_A\rangle_v| \ .$$

We define

$$S_u = \sum_{\boldsymbol{\sigma}\in A}\exp\sum_{k\le M} u(S_k) \ ; \quad Z_u = \sum_{\boldsymbol{\sigma}}\exp\sum_{k\le M} u(S_k) \tag{9.184}$$

(and similarly we define S_v and Z_v). Thus (9.182) yields

$$\mathsf{E}|S_u - S_v| \leq 2^{N/2+1} + b\sqrt{N}M2^N \tag{9.185}$$

$$\mathsf{E}|Z_u - Z_v| \leq 2^{N/2+1} + b\sqrt{N}M2^N \; . \tag{9.186}$$

Since $\langle \mathbf{1}_A \rangle_u = S_u/Z_u$, we have $\langle \mathbf{1}_A \rangle_v = S_v/Z_v$, and

$$\begin{aligned} |\langle \mathbf{1}_A \rangle_u - \langle \mathbf{1}_A \rangle_v| &= \left| \frac{S_u}{Z_u} - \frac{S_v}{Z_v} \right| = \left| \frac{S_u - S_v}{Z_u} + \frac{S_v(Z_v - Z_u)}{Z_u Z_v} \right| \\ &\leq \frac{|S_u - S_v|}{Z_u} + \frac{|Z_u - Z_v|}{Z_u} \end{aligned} \tag{9.187}$$

since $S_v \leq Z_v$. Consider the event $\Omega = \{Z_u \leq d2^N\}$. It follows from (9.16) that $\mathsf{P}(\Omega) \leq d^{1/L} \exp LM(1+\tau^2)$. Then (9.187) implies

$$|\langle \mathbf{1}_A \rangle_u - \langle \mathbf{1}_A \rangle_v| \leq \mathbf{1}_\Omega + \frac{2^{-N}}{d}(|S_u - S_v| + |Z_u - Z_v|) \; .$$

Taking expectations and using (9.185) and (9.186) completes the proof. □

Proposition 9.8.6. *There exists a constant L such that if $L\alpha \exp L\tau^2 \leq 1$ and the function u satisfies $u(x) = 0$ for $x \geq \tau$ and*

$$-D = -\frac{N}{L} \leq u \leq 0 \; ,$$

then given any subset I of $\{1, \ldots, N\}$ we have

$$\mathsf{E}\left| \left\langle \prod_{i \in I} \sigma_i \right\rangle_u - \left\langle \prod_{i \in I} \sigma_i \right\rangle_v \right| \leq L \exp\left(-\frac{L}{N} \right) . \tag{9.188}$$

Proof. If L_1 denotes the constant in (9.170) we assume

$$D = \frac{N}{10L_1} \; , \tag{9.189}$$

so that, if $b = L_2 \exp(-N/10L_1)$ where L_2 is large enough, (9.175) proves that the function v satisfies (9.170). We then choose $d = \exp(-N/20L_1)$ so that if $L'M(1+\tau^2) \leq N$ for L' large enough the bound in (9.183) is of the type $L\exp(-N/L)$. □

Lemma 9.8.7. *Under the conditions of Proposition 9.8.6, the following occurs. Consider for $1 \leq \ell < \ell' < n$ integers $k(\ell, \ell')$. Then*

$$\mathsf{E}\left| \left\langle \prod_{1 \leq \ell < \ell' \leq n} R_{\ell,\ell'}^{k(\ell,\ell')} \right\rangle_u - \left\langle \prod_{1 \leq \ell < \ell' \leq n} R_{\ell,\ell'}^{k(\ell,\ell')} \right\rangle_v \right| \leq Ln \exp\left(-\frac{N}{L} \right) . \tag{9.190}$$

The surprising part of this result is that here we study functions on n replicas; one would think that having to deal with the quantity Z_u of (9.184) occurring at power n in denominators will create trouble; the content of the lemma is that this is not the case.

Proof. We write

$$R_{\ell,\ell'} = \frac{1}{N}\sum_{i\le N}\sigma_i^\ell\,\sigma_i^{\ell'}$$

so that

$$R_{\ell,\ell'}^{k(\ell,\ell')} = \frac{1}{N^{k(\ell,\ell')}}\sum\prod_{k\le k(\ell,\ell')}\sigma^\ell_{i(\ell,\ell',k)}\,\sigma^{\ell'}_{i(\ell,\ell',k)}$$

where the summation is over all choices of $i(\ell,\ell',1),\dots,i(\ell,\ell',k(\ell,\ell'))$ of integers smaller than or equal to N. Thus

$$\prod_{\ell<\ell'}R_{\ell,\ell'}^{k(\ell,\ell')} = \frac{1}{N^{\overline{k}}}\sum\prod_{\ell<\ell'}\prod_{k\le k(\ell,\ell')}\sigma^\ell_{i(\ell,\ell',k)}\,\sigma^{\ell'}_{i(\ell,\ell',k)} \tag{9.191}$$

where $\overline{k} = \sum_{1\le\ell<\ell'\le n}k(\ell,\ell')$ and where the sum is over the $N^{\overline{k}}$ choices of indices $1\le i(\ell,\ell',k)\le N$ for $1\le\ell<\ell'\le n$, $k\le k(\ell,\ell')$. A product of any collection of spins σ_i^ℓ, $\ell\le n$, $i\le N$ (each of them occurring possibly several times) is of the type $\prod_{\ell\le n}\prod_{i\in I(\ell)}\sigma_i^\ell$ (where $I(\ell)$ is a certain subset of $\{1,\dots,N\}$). This is simply because $(\sigma_i^\ell)^2=1$. This is in particular the case of the double product in (9.191). Therefore

$$\Big\langle\prod_{\ell<\ell'}R_{\ell,\ell'}^{k(\ell,\ell')}\Big\rangle_u = \frac{1}{N^{\overline{k}}}\sum\prod_{\ell\le n}\Big\langle\prod_{i\in I(\ell)}\sigma_i^\ell\Big\rangle_u\,,$$

where the sum contains $N^{\overline{k}}$ terms. Thus to obtain (9.190) it suffices to bound the quantities

$$\mathsf{E}\Big|\prod_{\ell\le n}\Big\langle\prod_{i\in I(\ell)}\sigma_i^\ell\Big\rangle_u - \prod_{\ell\le n}\Big\langle\prod_{i\in I(\ell)}\sigma_i^\ell\Big\rangle_v\Big|\,.$$

This follows from (9.181) and (9.183). □

Proof of Theorem 9.8.1. Let (q_u,r_u) be a solution of the equations (9.65) for u. First, we show that the pair (q_u,r), where

$$r = \alpha\mathsf{E}\left(\frac{\mathsf{E}_\xi v'(\theta)\exp v(\theta)}{\mathsf{E}_\xi\exp v(\theta)}\right)^2$$

(and where $\theta = z\sqrt{q}+\xi\sqrt{1-q}$) is very close to be a solution of the equations (9.65) for v, and thus $\nu_v((R_{1,2}-q)^{2k})\le(Lk/N)^k$ by (9.101), from which the result is deduced for ν_u by expanding the power $(R_{1,2}-q)^{2k}$ and using (9.190) on each term. The details are straightforward. □

9.9 The Bernoulli Model

In the Bernoulli model the Gaussian r.v.s $g_{i,k}$ are replaced by independent random signs $\eta_{i,k}$ and the Hamiltonian is now

$$-H_{N,M}(\boldsymbol{\sigma}) = \sum_{k\le M} u\Big(\frac{1}{\sqrt{N}}\sum_{i\le N}\eta_{i,k}\sigma_i\Big) . \tag{9.192}$$

Throughout this section ν refers to this Hamiltonian. This model is harder to study than the Gaussian model, because we cannot use special Gaussian tools, such as integration by parts or more importantly Lemma 9.3.3. We must replace integration by parts by "approximate integration by parts" (as defined in Section 4.6, equation (4.198)). The error terms introduced by approximate integration by parts depend on the size of the derivatives of u. In order to be able to say anything at all about the structure of the Gibbs measure we essentially need to control the size of these derivatives uniformly over N. On the other hand, the problem of approximating $p_{N,M}(u)$ is easier: one can expect that it will suffices to approximate u by a smooth function (independent of N), for which one can understand the system for large N.

Theorem 9.9.1. *Assume that the function u of (9.192) satisfies (9.2), and, moreover*

$$\forall \ell\ ,\ 1\le \ell\le 5\ ,\quad |u^{(\ell)}|\le D\ .$$

Then there is a number $N(D)$ and a number L such that if $L\alpha\exp L\tau^2\le 1$ and $N\ge N(D)$ we have (9.66), i.e.

$$\nu\big((R_{1,2}-q)^2\big)\le \frac{L}{N}$$

and (9.103), where q is solution of the equations (9.65).

On the other hand, if u is not differentiable, we know very little about Gibbs' measure.

Research Problem 9.9.2. (Level 2) Assume that $u(x) = -\beta\mathbf{1}_{\{x\le 0\}}$. If β and α are small enough, is it true that

$$\lim_{N\to\infty}\lim_{M/N\to\alpha}\nu\big((R_{1,2}-q)^2\big)=0 \tag{9.193}$$

where q is given by the equations (9.65)? And is it true that

$$\sup_N N\nu\big((R_{1,2}-q)^2\big)<\infty$$

if $\alpha = M/N\le\alpha_0$ (small enough)?

This problem, and (9.193) in particular, is a very good case for the "what else could happen?" argument. It illustrates well the substantial gap between heuristic arguments (however convincing) and mathematical proofs.

We turn to the proof of Theorem 9.9.1. This proof is obtained by suitably modifying the proof of Theorem 9.5.1, so that this theorem must be mastered first before attempting to read the next two pages. The reader interested only in Gardner's formula should skip to Theorem 9.9.4 below.

Proposition 9.9.3. *The conditions of Proposition 2.2.2 imply*

$$\nu_t'(f) = \mathrm{I} + \mathrm{II} + \mathcal{R} \ , \tag{9.194}$$

where I *and* II *are as in Proposition 2.2.2 and*

$$|\mathcal{R}| \le \frac{\alpha}{N} K(n, D) \nu_t(|f|) \ . \tag{9.195}$$

Proof. We repeat the proof of Proposition 2.2.2, replacing integration by parts by "approximate integration by parts" as defined in (4.197). The main terms are the same as in the Gaussian case, and create the term I + II. The issue is to prove that the error term satisfies (9.195). This error term occurs when performing "approximate integration by parts" in the term III of (2.28) (with $\eta_{N,M}$ instead of g_M). This is as simple as can be, and we have already dealt with a more complicated situation in Chapter 4 when we introduced this method of approximate integration by parts. Still, there is no harm in repeating the argument in the case of the typical term

$$\nu_t(\eta_M \varepsilon_\ell u'(S^\ell_{M,t}) f) = \mathsf{E}\eta_M \langle \varepsilon_\ell u'(S^\ell_{M,t}) f \rangle_t \ , \tag{9.196}$$

where $\eta_M = \eta_{N,M}$. We consider the function $v_\ell(x)$, obtained by replacing each occurrence of η_M in the explicit expression of $\langle \varepsilon_\ell \, u'(S^\ell_{M,t}) \, f \rangle_t$ by x, and assuming all the other r.v.s $\eta_{i,k}$ fixed. Since each occurrence of η_M in the explicit expression of $\langle \varepsilon_\ell u'(S^\ell_{M,t}) f \rangle_t$ is multiplied by a factor $\sqrt{t/N}$, each derivation brings out this factor, and it should be obvious that

$$|v_\ell'''(x)| \le \frac{t^{3/2}}{N^{3/2}} K(n) D^4 \langle |f| \rangle_{t,x} \ , \tag{9.197}$$

where $\langle \cdot \rangle_{t,x}$ means that in the explicit expression of $\langle \cdot \rangle_t$ each occurrence of η_M is replaced by x (so that $v_\ell(x) = \langle \varepsilon_\ell \, u'(S^\ell_{M,t}) \, f \rangle_{t,x}$). We want to relate $\langle |f| \rangle_{t,x}$ with $\langle |f| \rangle_t$. For this we simply observe that

$$\left| \frac{\partial}{\partial x} \langle |f| \rangle_{t,x} \right| \le \frac{K(n) D}{\sqrt{N}} \, \langle |f| \rangle_{t,x} \ ,$$

so that, by integration

$$\sup_{|x| \le 1} \langle |f| \rangle_{t,x} \le K(n, D) \langle |f| \rangle_{t,1} \ .$$

Since

$$\langle |f| \rangle_{t,1} \le 2 \frac{\langle |f| \rangle_{t,1} + \langle |f| \rangle_{t,-1}}{2} ,$$

and since expectation averages η_M over ± 1, we get

$$\mathsf{E} \frac{\langle |f| \rangle_{t,1} + \langle |f| \rangle_{t,-1}}{2} = \nu_t(|f|) ,$$

so that $\mathsf{E}\langle |f| \rangle_{t,1} \le 2\nu_t(|f|)$ and therefore

$$\mathsf{E} \sup_{|x| \le 1} \langle |f| \rangle_{t,x} \le K(n, D) \nu_t(|f|) . \tag{9.198}$$

Combining with (9.197) we get

$$\mathsf{E} \sup_{|x| \le 1} |v_\ell'''(x)| \le \frac{t^{3/2}}{N^{3/2}} K(n, D) \nu_t(|f|) .$$

It follows from (4.197) and (4.199) that the error term occurring in the approximate integration by parts of the quantity (9.196) is at most

$$\frac{t^{3/2}}{N^{3/2}} K(n, D) \nu_t(|f|) .$$

Despite the coefficient $\sqrt{N/t}$ in the definition of the term III of (2.28) this implies that the error term created while performing approximate integration by parts in this term satisfies (9.195), and this completes the proof. □

Proof of Theorem 9.9.1. The proof of (9.66) is based on the computation of $\nu_t'(f)$ using Corollary 9.5.7. We face the problem to evaluate

$$\nu_t(f u'(S_{M,t}^\ell) u'(S_{M,t}^{\ell'})) = \mathsf{E} \frac{\left\langle f \mathsf{E}_\xi u'(S_{M,t}^\ell) u'(S_{M,t}^{\ell'}) \exp \sum_{m \le 4} u(S_{M,t}^m) \right\rangle_{t,\sim}}{\langle \mathsf{E}_\xi \exp u(S_{M,t}^1) \rangle_{t,\sim}^4} , \tag{9.199}$$

where now

$$S_{M,t}^\ell = \frac{1}{\sqrt{N}} \sum_{i<N} \eta_i \sigma_i^\ell + \frac{\sqrt{t}}{\sqrt{N}} \eta_N \sigma_N^\ell + \frac{\sqrt{1-t}}{\sqrt{N}} \xi_M$$

for $\eta_i = \eta_{i,M}$. For this purpose we simply compare the quantity (9.199) with the similar quantity where $S_{M,t}^\ell$ is replaced by its "Gaussian version"

$$S_{M,t}^{g,\ell} = \frac{1}{\sqrt{N}} \sum_{i<N} g_i \sigma_i^\ell + \frac{\sqrt{t}}{\sqrt{N}} g_N \sigma_N^\ell + \frac{\sqrt{1-t}}{\sqrt{N}} \xi_M , \tag{9.200}$$

with $(g_i)_{i<N}$ independent standard Gaussian r.v.s. For this we use Trotter's method described in the proof of Theorem 8.5.2. If we fix the randomness in $\langle \cdot \rangle_{t,\sim}$ and think of the quantity

$$\frac{\left\langle f \mathsf{E}_\xi u'(S^\ell_{M,t}) u'(S^{\ell'}_{M,t}) \exp \sum_{m\le 4} u(S^m_{M,t})\right\rangle_{t,\sim}}{\langle \mathsf{E}_\xi \exp u(S^1_{M,t})\rangle^4_{t,\sim}}$$

as a function $U(\eta_1,\dots,\eta_N)$, it is immediate that the fourth order partial derivatives of U are bounded by a quantity of the type

$$\frac{K(D)}{N^2}\langle |f|\rangle_{t,\sim} \ ,$$

so the error made while replacing $S^\ell_{M,t}$ by $S^{g,\ell}_{M,t}$ is at most $K(D)\langle |f|\rangle_{t,\sim}/N$. Thus, within this error term, it suffices to study the right-hand side of (9.199) when the quantities $S^\ell_{M,t}$ are replaced by the quantities $S^{g,\ell}_{M,t}$. This study has been done in Sections 9.3 to 9.5, and we leave it to the reader to check that the conclusion of Corollary 9.5.7 remains valid with the extra error term $K(D)\langle |f|\rangle_{t,\sim}/N$, and that the proof of (9.66) carries through. □

We now turn to the proof of Gardner's formula.

Theorem 9.9.4. *There exists a constant L with the following property. Consider a number τ and $\varepsilon > 0$. Then there is a number $N(\varepsilon,\tau)$ such that if $N > N(\varepsilon,\tau)$ and $LM \exp L\tau^2 \le N$, we have*

$$\mathsf{P}\left(\left|\frac{1}{N}\log\left(2^{-N}\mathrm{card}\bigcap_{k\le M} U_k\right) - \mathrm{RS}(\alpha)\right| \ge \varepsilon\right)$$
$$\le L\exp\left(-\frac{1}{L}\min\left(\frac{N^2\varepsilon^2}{M(1+\tau^2)}, \frac{N\varepsilon}{1+\tau^2}\right)\right) \ . \tag{9.201}$$

Here $U_k = \{S_k \ge \tau\} = \left\{\sum_{i\le N}\eta_{i,k}\sigma_i \ge \tau\sqrt{N}\right\}$, $\alpha = M/N$ and $\mathrm{RS}(\alpha)$ is given by (9.108).

In words, as expected, the Gardner formula is the same as in the Gaussian case, but we do not know how to prove as good a convergence rate.

Research Problem 9.9.5. (Level 2) Find the rate at which the convergence takes place in (9.201). For example, how fast does the median of $N^{-1}\log\left(2^{-N}\mathrm{card}\left(\bigcap_{k\le M} U_k\right)\right)$ converge to its limit?

The first major ingredient to the proof of Theorem 9.9.4 is the following.

Proposition 9.9.6. *There exists a constant L with the following property. Consider a probability measure G on Σ_N, and assume that*

$$G^{\otimes 2}\left(\left\{(\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2)\ ;\ R_{1,2}\le\frac{1}{2}\right\}\right) \ge 1-\exp\left(-\frac{N}{16}\right) \ . \tag{9.202}$$

Then for any $\tau \ge 0$, we have

$$L \exp\left(-\frac{N}{L}\right) \le \varepsilon \le \frac{1}{L}\exp(-L\tau^2) \quad \Rightarrow$$
$$\mathsf{P}\left(G\left(\left\{\boldsymbol{\sigma}\ ;\ \frac{1}{\sqrt{N}}\sum_{i\le N}\sigma_i\eta_i \ge \tau\right\}\right) \le \varepsilon\right) \le \varepsilon^{1/L}\ . \tag{9.203}$$

Proof. This is a special case of Proposition 8.5.7, when $\mathbf{b} = \mathbf{0}$, $a = 1$, so $\psi(\sigma_i) = \sigma_i$ for $\sigma_i = \pm 1$. □

Let us point out that the special case of Proposition 8.5.7 used above is much easier than the general case. It does not require in particular Propositions 8.5.3 or 8.5.6.

The second major ingredient to the proof of Theorem 9.9.4 is a weak form of (9.103). There is a simple approach to this result, that does not require a detailed study of the system with Hamiltonian (9.106), and in particular does not require Theorem 9.9.1. It is to use Trotter's method directly on the quantity

$$p^b_{N,M}(u) = \frac{1}{N}\mathsf{E}\log\sum_{\boldsymbol{\sigma}}\exp\sum_{k\le M} u(S_k)\ .$$

If we think of the right-hand side as a function $U(\eta_{i,k})$ of the variables $\eta_{i,k}$, all forth order derivatives of U are bounded by $K(D)N^{-3}$ (each differentiation brings an extra factor $N^{-1/2}$) and Trotter's method immediately implies that (with obvious notation)

$$|p^b_{N,M}(u) - p^g_{N,M}(u)| \le \frac{M}{N^2}K(D)\ , \tag{9.204}$$

which is not as good as (9.103) but will be sufficient for our purposes. The third major ingredient of the proof of Theorem 9.9.4 is the following, where $a = 1/32$.

Proposition 9.9.7. *Consider $\varepsilon > 0$. Then there is a number $\varepsilon' > 0$ with the following property. Consider a function u satisfying (9.2) and*

$$\exp u(\tau - \varepsilon') \le \varepsilon'\ . \tag{9.205}$$

Then, for N large enough and any M with $LM\exp L\tau^2 \le N$, we have

$$\left|\mathsf{E}\frac{1}{N}\log_{aN}\left(2^{-N}\sum_{\boldsymbol{\sigma}}\exp\sum_{k\le M}u(S_k)\right) - \mathsf{E}\frac{1}{N}\log_{aN}\left(2^{-N}\mathrm{card}\bigcap_{k\le M}U_k\right)\right| \le \varepsilon\ . \tag{9.206}$$

Proof. The proof is very similar to that of Proposition 8.3.9, but slightly simpler. We repeat the argument for the convenience of the reader. Let $C_m = \bigcap_{k<m} U_k$ and

$$V_m = 2^{-N} \sum_{\boldsymbol{\sigma} \in C_m} \exp\Bigl(\sum_{m \le k \le M} u(S_k) \Bigr)$$

so that the left-hand side of (9.206) is

$$\begin{aligned} &\frac{1}{N} |\mathsf{E} \log_{aN} V_1 - \log_{aN} V_{M+1}| \\ \le &\frac{1}{N} \sum_{1 \le m \le M} |\mathsf{E} \log_{aN} V_{m+1} - \mathsf{E} \log_{aN} V_m| \,. \end{aligned}$$

Let us fix m and denote by E_m expectation only on the r.v.s $(\eta_{i,m})_{i \le N}$. We are going to bound

$$|\mathsf{E}_m \log_{aN} V_{m+1} - \mathsf{E}_m \log_{aN} V_m| \,. \tag{9.207}$$

Let

$$Z_m = \sum_{\boldsymbol{\sigma} \in C_m} \exp\Bigl(\sum_{m < k \le M} u(S_k) \Bigr) \,.$$

Let us consider the probability measure G on Σ_N given by

$$G(B) = \frac{1}{Z_m} \sum_{\boldsymbol{\sigma} \in B \cap C_m} \exp\Bigl(\sum_{m < k \le M} u(S_k) \Bigr) \,.$$

Denoting by $\langle \cdot \rangle$ an average for G, we observe the identities

$$\begin{aligned} V_m &= 2^{-N} Z_m \langle \exp u(S_m) \rangle \\ V_{m+1} &= 2^{-N} Z_m \langle \mathbf{1}_{\{S_m \ge \tau\}} \rangle = 2^{-N} Z_m G(\{S_m \ge \tau\}) \,. \end{aligned}$$

Since $u(x) = 0$ for $x \ge \tau$, we have

$$Y := \langle \mathbf{1}_{U_m} \rangle = G(\{S_m \ge \tau\}) \le X := \langle \exp u(S_m) \rangle \,,$$

and using Lemmas 8.3.10 and 8.3.11 we see that for any $c > 0$ we have

$$|\log_{aN} V_{m+1} - \log_{aN} V_m| \le |\log_{aN} Y| \mathbf{1}_{\{Y \le c\}} + \frac{1}{c} |X - Y| \,. \tag{9.208}$$

Since $V_m, V_{m+1} \le 2^{-N} Z_m$, the left-hand side is zero unless $2^{-N} Z_m \ge \exp(-aN)$, so we may assume that this is the case. Then (9.202) holds by Lemma 9.2.1, so Proposition 9.9.6 implies

$$L \exp\left(-\frac{N}{L} \right) \le t \le \frac{1}{L} \exp(-L\tau^2) \quad \Rightarrow \quad \mathsf{P}(Y \le t) \le L t^{1/L} \,,$$

where P is the probability in the r.v.s $\eta_{i,k}$ only. It is then straightforward to see that one can choose $c > 0$ (depending on ε and τ only) with

$$\mathsf{E}_m|\log_{aN} Y|\mathbf{1}_{\{Y\le c\}} \le \frac{\varepsilon}{2} \,.$$

All that remains to prove is that $\mathsf{E}_m|X-Y|$ can be made $\le \varepsilon c/2$ if ε' in (9.205) is small enough. Now we observe that $X=Y=1$ if $S_M \ge \tau$, while $0\le X\le \varepsilon'$ and $Y=0$ if $S_M \le \tau-\varepsilon'$, and thus

$$0 \le X - Y \le \varepsilon' + \langle \mathbf{1}_{\{\tau-\varepsilon'\le S_M\le\tau\}}\rangle \,,$$

so that

$$\mathsf{E}_m|X-Y| \le \varepsilon' + \mathsf{P}\Big(\tau-\varepsilon' \le \frac{1}{\sqrt{N}}\sum_{i\le N}\eta_i \le \tau\Big)$$

where $(\eta_i)_{i\le N}$ are random signs. For large N, by the central limit theorem, the right-hand side is $\le \varepsilon' + 2\mathsf{P}(\tau-\varepsilon'\le g\le\tau) \le 3\varepsilon'$. □

Proof of Theorem 9.9.4. Without loss of generality we may assume that $\varepsilon \le a/10$. We consider ε' given by Proposition 9.9.7, and we find a function u that satisfies (9.2), (9.205), and

$$|\mathrm{RS}(\alpha) - (p(u) - \log 2)| \le \varepsilon \tag{9.209}$$

(where $\mathrm{RS}(\alpha)$ is given by (9.108) and $p(u)$ by (9.102)), and such that for a certain number D we have $|u^{(\ell)}| \le D$ for $0\le \ell\le 5$. It follows from (9.204) and Theorem 2.4.2 that for $N \ge N(D,\tau)$ and $L\alpha\exp L\tau^2 \le 1$ we have

$$\left|\frac{1}{N}\mathsf{E}\log(2^{-N}Z) - (p(u)-\log 2)\right| \le \varepsilon \tag{9.210}$$

where $Z = \sum_{\boldsymbol{\sigma}} \exp\sum_{k\le M} u(S_k)$. Thus

$$\frac{1}{N}\mathsf{E}\log_{aN}(2^{-N}Z) \ge \frac{1}{N}\mathsf{E}\log(2^{-N}Z) \ge p(u) - \log 2 - \varepsilon \,.$$

Since $\mathrm{RS}(0)=0$, without loss of generality we may assume that

$$\mathrm{RS}(\alpha) \ge -a/4 \tag{9.211}$$

because this is the case for α small, and since we assume $L\alpha\exp L\tau^2 \le 1$ for a large enough constant L. Since $\varepsilon \le a/10$, (9.209) implies that $p(u)-\log 2-\varepsilon \ge -a/2$, and therefore

$$\frac{1}{N}\mathsf{E}\log_{aN}(2^{-N}Z) \ge -\frac{a}{2} \,.$$

Since $\log_{aN}(x) = -aN$ for $x \le \exp(-aN)$, we deduce from (9.21) by taking $t = a/2$ that

$$\begin{aligned}
\mathsf{P}(2^{-N}Z \le \exp(-aN)) &\le \mathsf{P}\Big(\frac{1}{N}\log_{aN}(2^{-N}Z) \le -a\Big) \\
&\le \mathsf{P}\left(\left|\frac{1}{N}\log_{aN}(2^{-N}Z) - \mathsf{E}\frac{1}{N}\log_{aN}(2^{-N}Z)\right| \ge \frac{a}{2}\right) \\
&\le L\exp\Big(-\frac{N}{L(1+\tau^2)}\Big) \,.
\end{aligned}$$

Since $2^{-N}Z \geq \exp(-MD)$ we obtain that for N large enough

$$\begin{aligned}\frac{1}{N}\mathsf{E}(\mathbf{1}_{\{2^{-N}Z\leq\exp(-aN)\}}\log_{aN}(2^{-N}Z)) &\leq \frac{LMD}{N}\exp\left(-\frac{N}{L(1+\tau^2)}\right)\\ &\leq \varepsilon\end{aligned}$$

and thus

$$\left|\frac{1}{N}\mathsf{E}\log(2^{-N}Z)-\frac{1}{N}\mathsf{E}\log_{aN}(2^{-N}Z)\right|\leq\varepsilon\ .$$

Combining with (9.210) and (9.209) yields

$$\left|\frac{1}{N}\mathsf{E}\log_{aN}(2^{-N}Z)-\mathrm{RS}(\alpha)\right|\leq 3\varepsilon$$

and using (9.206) implies

$$\left|\frac{1}{N}\mathsf{E}\log_{aN}\left(2^{-N}\mathrm{card}\bigcap_{k\leq M}U_k\right)-\mathrm{RS}(\alpha)\right|\leq 4\varepsilon\ . \tag{9.212}$$

Recalling (9.211) and since $5\varepsilon\leq a/2$ we then get

$$\begin{aligned}&\left|\frac{1}{N}\log_{aN}\left(2^{-N}\mathrm{card}\bigcap_{k\leq M}U_k\right)-\mathrm{RS}(\alpha)\right|<5\varepsilon\\ \Rightarrow\ &\frac{1}{N}\log_{aN}\left(2^{-N}\mathrm{card}\bigcap_{k\leq M}U_k\right)>-a\\ \Rightarrow\ &\frac{1}{N}\log_{aN}\left(2^{-N}\mathrm{card}\bigcap_{k\leq M}U_k\right)=\frac{1}{N}\log\left(2^{-N}\mathrm{card}\bigcap_{k\leq M}U_k\right)\\ \Rightarrow\ &\left|\frac{1}{N}\log\left(2^{-N}\mathrm{card}\bigcap_{k\leq M}U_k\right)-\mathrm{RS}(\alpha)\right|<5\varepsilon\end{aligned}$$

and therefore, using (9.212) in the third line,

$$\begin{aligned}&\left|\frac{1}{N}\log\left(2^{-N}\mathrm{card}\bigcap_{k\leq M}U_k\right)-\mathrm{RS}(\alpha)\right|\geq 5\varepsilon\\ \Rightarrow\ &\left|\frac{1}{N}\log_{aN}\left(2^{-N}\mathrm{card}\bigcap_{k\leq M}U_k\right)-\mathrm{RS}(\alpha)\right|\geq 5\varepsilon\\ \Rightarrow\ &\left|\frac{1}{N}\log_{aN}\left(2^{-N}\mathrm{card}\bigcap_{k\leq M}U_k\right)-\frac{1}{N}\mathsf{E}\log_{aN}\left(2^{-N}\mathrm{card}\bigcap_{k\leq M}U_k\right)\right|\geq\varepsilon\ .\end{aligned}$$

The probability of the event above is exponentially small by Proposition 9.2.6, which remains valid in the present case, because this is the case of (9.13), as is proved by (9.203). □

10. The Hopfield Model

10.1 Introduction

In Chapter 4 we investigated the Hopfield model, with Hamiltonian

$$-H_{N,M}(\boldsymbol{\sigma}) = \frac{N\beta}{2} \sum_{k \le M} m_k^2(\boldsymbol{\sigma}) + hNm_1(\boldsymbol{\sigma}) \ ,$$

where $m_k(\boldsymbol{\sigma}) = N^{-1} \sum_{i \le N} \eta_{i,k}\sigma_i$, $\eta_{i,1} = 1$ and $(\eta_{i,k})_{i \le N, k \ge 2}$ are independent random signs. In the present chapter we revisit this model. We will be able to analyze it through the cavity method in a larger region of parameters than in Chapter 4, and this analysis we will be more accurate, getting the correct rates of convergence. This seemed impossible to achieve by the Bovier-Gayrard method of Section 4.4. To illustrate the gain in accuracy over the results of Chapter 4, one of the statements of Theorem 10.7.1 below (one of the main results of this chapter) is that in a certain region of parameters we have (for a certain number r)

$$\nu\left(\left(\frac{1}{N} \sum_{2 \le k \le M} m_k(\boldsymbol{\sigma}^1) m_k(\boldsymbol{\sigma}^2) - r\right)^2\right) \le \frac{K}{N} \ ,$$

while the method of Chapter 4 can reach only a bound of $K/\sqrt{N}$. In the last section of the chapter we will also investigate a non-trivial variation of the model. We will not use all the material of Chapter 4, but it would still be useful for the reader to have this chapter at hand. The reader will find many instances of notions introduced there, such as the expression "with overwhelming probability", which, once again, means with a probability $\ge 1 - K\exp(-N/K)$ where K does not depend on N.

The author considers the present chapter as one of the highlights of this volume, and, in fact, of his entire working lifetime. A number of subtle ideas interlock to yield very precise results. Studying this chapter may require some effort, but the reward should be ample.

We recall that for $\beta > 1$ and $h > 0$ there is a unique number m^* given by $m^* = \mathrm{th}(\beta m^* + h)$. Given a number L_0, let us define the *admissible region* of parameters (α, β, h) as the region

M. Talagrand, *Mean Field Models for Spin Glasses*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys in Mathematics 55, DOI 10.1007/978-3-642-22253-5_3,

$$\left\{(\alpha,\beta,h)\;;\;h>0\;,\;1<\beta\le 2\;,\;\alpha\le\frac{m^{*4}}{L_0}\right\}$$
$$\bigcup\left\{(\alpha,\beta,h)\;;\;h>0\;,\;\beta\ge 2\;,\;\alpha\le\frac{1}{L_0\log\beta}\right\}\;. \tag{10.1}$$

Of course the number 2 does not have any special meaning. The point of the previous definition is that for β close to 1, values of α up to order m^{*4} are permitted, while for β large, values of α up to order $1/\log\beta$ are permitted. We will show that once the number L_0 has been chosen large enough, then for (α,β,h) in the admissible region (10.1), we can analyze the Hopfield model precisely as $N\to\infty$ and $M/N\to\alpha$.

Given β, the challenge is to control the model for values of α as large as possible. The admissible region (10.1) provides a control for $0<\alpha<\alpha_0(\beta)$, where $\alpha_0(\beta)=m^{*4}/L_0$ if $\beta\le 2$ and $\alpha_0(\beta)=1/(L_0\log\beta)$ if $\beta\ge 2$. We show that in this region the model exhibits "replica-symmetric behavior". According to the heuristic work of the physicists (see e.g. [14]) the admissible region has the correct shape (at least in the limiting case $h=0$). More precisely, there exists a number L such that when $\alpha\ge L\alpha_0(\beta)$, the model does not exhibit "replica-symmetric behavior". Proving that the model exhibits replica-symmetric behavior for $\alpha\le\alpha(\beta)$ where $\alpha^*(\beta)$ is the largest possible value such that this is true is likely to be a very difficult problem. And describing what happens for $\alpha>\alpha^*(\beta)$ should be even more difficult.

The reader observes that by definition we have $h>0$ in the admissible region.

It seems useful to stress that through the entire chapter, we think of the number L_0 defining the admissible region as a parameter, that can be adjusted as large as we wish. In our computations *all dependences on* L_0 *are explicit*. In particular, when we write "L" for a universal constant, this means that the value of L has been determined through calculations that *do not* depend on the value of L_0.

It is useful to note right away that in the admissible region (10.1) we have

$$\alpha\le\frac{Lm^{*4}}{L_0}\;, \tag{10.2}$$

as is obvious by considering separately the cases $\beta\le 2$ and $\beta>2$.

10.2 The Replica-Symmetric Equations

Defining as usual $\alpha=M/N$, in Section 4.5 we have put forward the following equations, where as usual z denotes a standard normal r.v.

$$q=\mathsf{E}\mathrm{th}^2(\beta(z\sqrt{r}+\mu)+h) \tag{10.3}$$
$$\mu=\mathsf{E}\mathrm{th}(\beta(z\sqrt{r}+\mu)+h) \tag{10.4}$$
$$r=\frac{\alpha q}{(1-\beta(1-q))^2}\;. \tag{10.5}$$

They are called the "replica-symmetric equations". The goal of this section is to study this system of equations for suitable values of α, β and h. To lighten the exposition we will say "the replica-symmetric equations" when we actually mean "the system consisting of the replica-symmetric equations".

Theorem 10.2.1. *There exist numbers L_1 and L_2 with the following property: If the number L_0 has been chosen large enough, then for (α, β, h) in the admissible region (10.1), the replica-symmetric equations (10.3) to (10.5) have a unique solution (q, μ, r) such that*

$$|q - m^{*2}| \le \frac{m^{*2}}{L_1 \beta^2} \; ; \qquad |\mu - m^*| \le \frac{m^*}{L_2} \; . \tag{10.6}$$

We do not know whether there exist solutions to the replica-symmetric equations (10.3) to (10.5) that do not satisfy (10.6). The uniqueness of the solutions is not important for our approach, but their existence is quite essential.

To lighten notation we write

$$Y = \beta(z\sqrt{q} + \mu) + h \; ,$$

so that equations (10.3) and (10.4) become respectively $q = \mathsf{E}\,\mathrm{th}^2 Y$ and $\mu = \mathsf{E}\,\mathrm{th} Y$. Fixing $h > 0$ once and for all, we consider the functions

$$\Phi(r, \mu) = \mathsf{E}\,\mathrm{th}^2 Y \tag{10.7}$$

$$\Psi(r, \mu) = \mathsf{E}\,\mathrm{th} Y \tag{10.8}$$

$$r(q) = \frac{\alpha q}{(1 - \beta(1 - q))^2} \; , \tag{10.9}$$

so that the equations (10.3) to (10.5) now read

$$q = \Phi(r(q), \mu) \quad ; \quad \mu = \Psi(r(q), \mu) \quad ; \quad r = r(q) \; .$$

The principle of the proof is to show that for a certain distance d on the domain (10.6) the map

$$(q, \mu) \mapsto (\Phi(r(q), \mu), \Psi(r(q), \mu))$$

is a contraction, hence it has a unique fixed point. It does not seem possible to achieve better than an estimate of the type

$$d(\Phi(r(q), \mu), \Psi(r(q), \mu)) \le \left(1 - \frac{m^{*2}}{L}\right) d(q, \mu) \; , \tag{10.10}$$

where, as usual, L denotes a universal constant. When β is close to 1 (and h close to 0), m^{*2} is small, and the "contraction coefficient" $1 - m^{*2}/L$ in (10.10) is close to 1. For this reason, we must prove all our estimates with errors at most of the type m^{*2}/L (rather than, say m^*/L), and this requires some care. The proof of Theorem 10.2.1 is however completely elementary, and gives a new dimension to the word "tedious". The author is aware that it might be difficult for the reader to get the motivation to go line by line through these computations. To ensure that this tedium does not discourage the reader, the proof of Theorem 10.2.1 has been moved to the last section of the chapter, despite the fact that at times it does contain a first version of several of the ideas that are subsequently needed.

From now on, given (α, β, h) in the admissible region, (q, μ, r) will always denote the unique solution of the replica-symmetric equations (10.3) to (10.5) that satisfies (10.6).

We will need further information on these numbers. We list this information now. It will be proved only in the last section of this chapter, while proving Theorem 10.2.1. The relevance of these inequalities will become clear only later. Of course the numbers 9 and 10 are only convenient choices.

Lemma 10.2.2. *In the admissible region (and if L_0 is large enough) we have*

$$|\mu - m^*| \leq \frac{Lm^*}{L_0\beta^{10}} \tag{10.11}$$

$$|q - m^{*2}| \leq \frac{Lm^{*2}}{L_0\beta^{10}} \tag{10.12}$$

$$|r - \alpha| \leq \frac{L}{\beta^9} \quad ; \quad |r - \alpha m^{*2}| \leq \frac{L}{\beta^9} \tag{10.13}$$

$$r \leq \frac{L\alpha}{m^{*2}} \tag{10.14}$$

$$1 - \beta(1-q) \geq \frac{m^{*2}}{L} \tag{10.15}$$

$$\frac{q}{1-\beta(1-q)} \leq L \tag{10.16}$$

$$\beta(1 - 2\mu^2 + q) \leq 1 - \frac{m^{*2}}{L} \, . \tag{10.17}$$

10.3 Localization on Balls with Random Centers

The main result of this section, Theorem 10.3.1 below, is one of the central ideas of this chapter. Like the Bovier-Gayrard localization Theorem 4.3.2, it provides some "a priori" information about the Gibbs measure. This information is a key technical result for all the subsequent work. This is little difficult to explain in detail now, but should become clearer after its first critical use in Theorem 10.4.1 below. The proof of Theorem 10.3.1 is somewhat in the spirit of the proof of the Bovier-Gayrard localization theorem. It does not use the cavity method, and can be skipped by the reader who is interested only in following the main story.

We recall the notation $\mathbf{m}(\boldsymbol{\sigma}) = (m_k(\boldsymbol{\sigma}))_{k\le M} \in \mathbb{R}^M$, and that $\|\cdot\|$ denotes the Euclidean norm on $\mathbb{R}^M$. We define

$$\mathbf{c} = \left(m^* N^{-1} \sum_{i\le N} \eta_{i,k}\right)_{k\le M} , \tag{10.18}$$

and we recall that $\alpha = M/N$. We recall also that in Chapter 4 (see Definition 4.2.3 and the discussion afterwards) we defined a *negligible set* $A \subset \Sigma_N$ to be a set A for which $\mathsf{E}G(A) \le K \exp(-N/K)$ where K depends only on β and h.

Theorem 10.3.1. *If $\beta > 2$ and $h > 0$, then for $\alpha \le 1/(L_0 \log \beta)$ (and L_0 large enough) the set*

$$A = \left\{\boldsymbol{\sigma}\ ;\ \|\mathbf{m}(\boldsymbol{\sigma}) - \mathbf{c}\| \ge \frac{L}{\beta^{10}}\right\}$$

is negligible.

Of course, there is nothing specific about the power 10. Let us recall the notation $\boldsymbol{\theta} = \mathbf{c} - m^*\mathbf{e}_1$, so that $\boldsymbol{\theta} = (\theta_k)_{k\le M}$, where $\theta_1 = 0$ and $\theta_k = m^* N^{-1} \sum_{i\le N} \eta_{i,k}$ for $k \ge 2$.

It turns out that the length of $\boldsymbol{\theta}$ is nearly constant. Indeed,

$$\|\boldsymbol{\theta}\|^2 = \frac{m^{*2}}{N^2} \sum_{2\le k\le M} \Big(\sum_{i\le N} \eta_{i,k}\Big)^2 = m^{*2}\frac{M-1}{N} + \sum_{2\le k\le M} Y_k\ ,$$

where $Y_k = N^{-2}\sum_{i\ne j} \eta_{i,k}\eta_{j,k}$. If follows from a result of C. Borell [20] that $\mathsf{E}\exp(N|Y|/L) \le 2$ and Bernstein's inequality (A.34) implies that given any $\varepsilon > 0$,

$$\text{with overwhelming probability}\ \ \big|\|\boldsymbol{\theta}\|^2 - m^*\sqrt{\alpha}\big| \le \varepsilon\ . \tag{10.19}$$

The content of Theorem 10.3.1 is that the relevant values of $\mathbf{m}(\boldsymbol{\sigma})$ are contained in a ball of small radius centered at the random point $\mathbf{c} = m^*\mathbf{e}_1 + \boldsymbol{\theta}$, and the radius of that ball is often much smaller than $\|\boldsymbol{\theta}\|$. (The Bovier-Gayrard localization theorem, Theorem 4.3.2, provides essentially the best possible result as far as localization on a ball of fixed radius is concerned.)

Let us spell out a specific consequence of Theorem 10.3.1 that will be critical in the proof of Theorem 10.4.1 below: typically we have $|m_1(\boldsymbol{\sigma}) - m^*| \le L/\beta^{10}$. This certainly is less precise than the fact (that we will prove later after considerable work) that $m_1(\boldsymbol{\sigma}) \simeq \mu$ (where μ is obtained through the replica-symmetric equations (10.3) to (10.5))), but it is a key step in obtaining this later result.

Let us also mention that it is possible (with essentially the same proof) to obtain a version of Theorem 10.3.1 that is true for $h = 0$, by considering the set

$$\left\{\boldsymbol{\sigma}\ ;\ \|\mathbf{m}(\boldsymbol{\sigma}) - \mathbf{c}\| \geq \frac{L}{\beta^{10}}\ ,\ \|\mathbf{m}(\boldsymbol{\sigma}) - m^*\mathbf{e}_1\| \leq \frac{m^*}{2}\right\}$$

instead of A.

A first observation is that to prove Theorem 10.3.1 we may assume that $\beta \geq \beta_0$, where β_0 is a number which does not depend on any parameter. This is because by Lemma 4.5.7 there exists a constant L^* such that the set $\{\boldsymbol{\sigma}\ ;\ \|\mathbf{m}(\boldsymbol{\sigma}) - \mathbf{c}\| \leq L^*\}$ is negligible, and $L/\beta^{10} \geq L^*$ if $\beta \leq \beta_0$ and $L \geq \beta_0^{10} L^*$. The key to this theorem is the following.

Proposition 10.3.2. *If $\beta > \beta_0$ and $h > 0$, then for $\alpha \leq 1/(L_0 \log \beta)$ and L_0 large enough the set*

$$A_1 = \left\{\boldsymbol{\sigma} \in \Sigma_N\ ;\ m_1(\boldsymbol{\sigma}) \leq m^* - \frac{2}{\beta^{10}}\right\}$$

is negligible.

Of course the term $2\beta^{-10}$ is just a convenient choice. The reason why Proposition 10.3.2 implies Theorem 10.3.1 is the following simple but unexpected fact.

Lemma 10.3.3. *With overwhelming probability we have*

$$\forall \boldsymbol{\sigma}\ ,\quad m_1(\boldsymbol{\sigma}) \geq m^* - \frac{2}{\beta^{20}} \quad \Rightarrow \quad \|\mathbf{m}(\boldsymbol{\sigma}) - \mathbf{c}\| \leq \frac{L}{\beta^{10}}\ .$$

To formulate things differently, this means that with overwhelming probability the set

$$\left\{\boldsymbol{\sigma}\ ;\ m_1(\boldsymbol{\sigma}) \geq m^* - \frac{2}{\beta^{20}}\ ,\ \|\mathbf{m}(\boldsymbol{\sigma}) - \mathbf{c}\| \leq \frac{L}{\beta^{10}}\right\}$$

is empty. Before proving this we state a simple fact.

Lemma 10.3.4. *Consider numbers $(\rho_i)_{i \leq N}$. Then*

$$\mathsf{P}\left(\sum_{2 \leq k \leq M}\left(\sum_{i \leq N} \eta_{i,k}\,\rho_i\right)^2 \geq t\right) \leq 2^{M/2} \exp\left(-\frac{t}{4\sum_{i \leq N} \rho_i^2}\right)\ . \qquad (10.20)$$

Proof. We use (A.19) to see that

$$\mathsf{E} \exp \frac{\left(\sum_{i \leq N} \eta_{i,k}\,\rho_i\right)^2}{4\sum_{i \leq N} \rho_i^2} \leq \sqrt{2}$$

so that, by independence,

$$\mathsf{E}\exp \frac{\sum_{2\le k\le M}\left(\sum_{i\le N}\eta_{i,k}\,\rho_i\right)^2}{4\sum_{i\le N}\rho_i^2} \le 2^{M/2}\ ,$$

and the result follows by Chebyshev's inequality. □

Proof of Lemma 10.3.3. Let

$$U = \left\{\boldsymbol{\sigma}\ ;\ m_1(\boldsymbol{\sigma}) \ge m^* - \frac{2}{\beta^{20}}\right\}\ .$$

For $\boldsymbol{\sigma}\in U$, let $\rho_i(\boldsymbol{\sigma}) = \sigma_i - m^*$, so that

$$\begin{aligned}\sum_{i\le N}\rho_i^2(\boldsymbol{\sigma}) &= N\Big(1 - \frac{2m^*}{N}\sum_{i\le N}\sigma_i + m^{*2}\Big)\\ &= N(1-2m^* m_1(\boldsymbol{\sigma}) + m^{*2})\\ &\le N\left(1 - m^{*2} + \frac{4m^*}{\beta^{20}}\right) \le \frac{LN}{\beta^{20}}\ ,\end{aligned}$$

using that $m^{*2}\ge 1 - L\beta^{-20}$ from (4.37). Using Lemma 10.3.4, for $t>0$ and any $\boldsymbol{\sigma}\in U$ we have, since $M\le N$,

$$\mathsf{P}\Bigg(\sum_{2\le k\le M}\Big(\sum_{i\le N}\eta_{i,k}\rho_i(\boldsymbol{\sigma})\Big)^2 \ge t\Bigg) \le 2^{M/2}\exp\left(-\frac{t\beta^{20}}{LN}\right) \le \exp\left(N - \frac{t\beta^{20}}{LN}\right)\ .$$

Taking $t = L'N^2\beta^{-20}$ for L' large enough (and recalling that $M\le N$), we obtain with overwhelming probability

$$\forall \boldsymbol{\sigma}\in U\ ,\quad \sum_{2\le k\le M}\Big(\sum_{i\le N}\eta_{i,k}\rho_i(\boldsymbol{\sigma})\Big)^2 \le \frac{LN^2}{\beta^{20}}\ .$$

Now, for $k\ge 2$, let us define $c_k = \theta_k = m^* N^{-1}\sum_{i\le N}\eta_{i,k}$, so that $\sum_{i\le N}\eta_{i,k}\rho_i(\boldsymbol{\sigma}) = m_k(\boldsymbol{\sigma}) - c_k$ and hence

$$\sum_{2\le k\le M}\Big(\sum_{i\le N}\eta_{i,k}\rho_i(\boldsymbol{\sigma})\Big)^2 = N^2\sum_{2\le k\le M}(m_k(\boldsymbol{\sigma}) - c_k)^2\ .$$

Consequently, with overwhelming probability,

$$\begin{aligned}\forall\boldsymbol{\sigma}\in U\ ,\quad \|\mathbf{m}(\boldsymbol{\sigma}) - \mathbf{c}\|^2 &= |m_1(\boldsymbol{\sigma}) - m^*|^2 + \sum_{2\le k\le M}(m_k(\boldsymbol{\sigma}) - c_k)^2\\ &\le |m_1(\boldsymbol{\sigma}) - m^*|^2 + \frac{L}{\beta^{20}}\ .\end{aligned} \tag{10.21}$$

Now for $\boldsymbol{\sigma} \in U$, we have

$$m^* - \frac{2}{\beta^{20}} \le m_1(\boldsymbol{\sigma}) \le 1$$

so that, using again that $1 - m^* \le 2\beta^{-20}$ by (4.37) we obtain

$$|m_1(\boldsymbol{\sigma}) - m^*| \le \max\left(\frac{2}{\beta^{20}}, 1 - m^*\right) \le \frac{L}{\beta^{20}} ,$$

and, combining with (10.21), this completes the proof. □

It remains to prove Proposition 10.3.2. For this we will need some tools from Section 4.2, and we recall the appropriate notation. We denote by G' the probability on $\mathbb{R}^M$ which is the image of Gibbs' measure G under the map $\boldsymbol{\sigma} \mapsto \mathbf{m}(\boldsymbol{\sigma})$. We denote by γ the Gaussian measure on $\mathbb{R}^M$ of density proportional to $\exp(-\beta N \|\mathbf{z}\|^2/2)$ and by $\overline{G} = G' * g$ the convolution of G' and γ.

Proposition 10.3.5. *If $\beta > \beta_0$, then when (α, β, h) belongs to the admissible region (10.1), and when the constant L_0 has been chosen large enough, the set*

$$B = \left\{\mathbf{z} \in \mathbb{R}^M \ ; \ z_1 \le m^* - \frac{1}{\beta^{20}}\right\} \tag{10.22}$$

is negligible for $\overline{G}$, that is $\mathsf{E}\overline{G}(B) \le K \exp(-N/K)$ where K depends only on β and h.

Given β and h, the set of values of α such that (α, β, h) belongs to the admissible region is an interval $[0, \alpha_0(\beta)]$, and the bound for $\mathsf{E}\overline{G}(B)$ is uniform for $\alpha = M/N$ in this interval. This feature is shared by many statements in the chapter.

Proof of Proposition 10.3.2. Consider the sets

$$B_1 = \left\{\mathbf{z} \in \mathbb{R}^M \ ; \ z_1 \le m^* - \frac{2}{\beta^{20}}\right\}$$
$$B_2 = \left\{\mathbf{z} \in \mathbb{R}^M \ ; \ |z_1| \le \frac{1}{\beta^{20}}\right\} .$$

Then, recalling (10.22), for $\mathbf{z} \in B_1$ and $\mathbf{z}' \in B_2$ we have $\mathbf{z} + \mathbf{z}' \in B$ so that, since by definition of convolution, $\overline{G}$ is the image of $G' \otimes \gamma$ under the map $(\mathbf{z}, \mathbf{z}') \mapsto \mathbf{z} + \mathbf{z}'$, we have

$$\overline{G}(B) \ge G'(B_1)\gamma(B_2) .$$

Now B_2 is "one-dimensional" so $\gamma(B_2) = \mathsf{P}(|z| \le \beta^{-20})$ where z is Gaussian with $\mathsf{E}z^2 = 1/(\beta N)$. This quantity is bounded below independently of N or M. Therefore we have $G'(B_1) \le K\overline{G}(B)$. By definition of G' we have $G'(B_1) = G(A_1)$, so $\mathsf{E}G(A_1) \le K\mathsf{E}\overline{G}(B)$, and moreover $\mathsf{E}\overline{G}(B) \le K\exp(-N/K)$ by Proposition 10.3.5. □

We consider the function

$$\xi(x) = \log \mathrm{ch}(\beta x + \beta m^* + h) - \log \mathrm{ch}(\beta m^* + h) - \beta x m^* \,. \tag{10.23}$$

To begin the proof of Proposition 10.3.5, we start with the following rather natural bound for $\overline{G}(C_k)$, where we recall that $\boldsymbol{\eta}_i = (\eta_{i,k})_{k \le M}$.

Lemma 10.3.6. *For a subset C of $\mathbb{R}^M$ we have*

$$\overline{G}(C+\mathbf{c}) \le \exp\left(\frac{1}{2}\sum_{i\le N}\xi(2\boldsymbol{\eta}_i\cdot\boldsymbol{\theta})\right)\int_C \exp\left(\frac{1}{2}\sum_{i\le N}\xi(2\boldsymbol{\eta}_i\cdot\mathbf{w})\right)\mathrm{d}\gamma(\mathbf{w}) \,. \tag{10.24}$$

Proof. We recall that by Lemma 4.2.1 the probability $\overline{G}$ has a density

$$W 2^N Z_{N,M}^{-1} \exp\psi(\mathbf{z})$$

with respect to Lebesgue measure, where

$$\psi(\mathbf{z}) = -\frac{N\beta}{2}\|\mathbf{z}\|^2 + \sum_{i\le N}\log \mathrm{ch}(\beta\boldsymbol{\eta}_i\cdot\mathbf{z} + h) \,, \tag{10.25}$$

and

$$W = \left(\frac{N\beta}{2\pi}\right)^{M/2} \,, \qquad Z_{N,M} = \sum_{\boldsymbol{\sigma}}\exp(-H_{N,M}(\boldsymbol{\sigma})) \,.$$

Proposition 4.3.4 implies that

$$Z_{N,M} \ge 2^N \exp\left(Nb^* + \frac{N\beta}{2}\|\boldsymbol{\theta}\|^2\right) \,,$$

where $b^* = \log \mathrm{ch}(\beta m^* + h) - \beta m^{*2}/2$, so that for a subset D of $\mathbb{R}^M$ we have

$$\overline{G}(D) \le W\exp\left(-Nb^* - \frac{N\beta}{2}\|\boldsymbol{\theta}\|^2\right)\int_D \exp\psi(\mathbf{z})\mathrm{d}\mathbf{z} \,,$$

and hence for a subset C of $\mathbb{R}^M$ we have, setting $\mathbf{z} = \mathbf{w} + \mathbf{c}$,

$$\overline{G}(C+\mathbf{c}) \le W\exp\left(-Nb^* - \frac{N\beta}{2}\|\boldsymbol{\theta}\|^2\right)\int_C \exp\psi(\mathbf{w}+\mathbf{c})\mathrm{d}\mathbf{w} \,. \tag{10.26}$$

Now since $\mathbf{c} = m^*\mathbf{e}_1 + \boldsymbol{\theta}$ and $\eta_{i,1} = 1$ we have

$$\boldsymbol{\eta}_i\cdot\mathbf{c} = m^* + \boldsymbol{\eta}_i\cdot\boldsymbol{\theta}$$

so that

$$\beta\boldsymbol{\eta}_i \cdot (\mathbf{w}+\mathbf{c}) + h = \beta m^* + h + \beta\boldsymbol{\eta}_i \cdot (\mathbf{w}+\boldsymbol{\theta}) \ ,$$

and by definition of the function ξ (see (10.23)) we get

$$\begin{aligned}\log \operatorname{ch}(\beta\boldsymbol{\eta}_i \cdot (\mathbf{w}+\mathbf{c}) + h) = \log \operatorname{ch}(\beta m^* + h) + \beta m^* \boldsymbol{\eta}_i \cdot (\mathbf{w}+\boldsymbol{\theta}) \\ + \xi(\boldsymbol{\eta}_i \cdot (\mathbf{w}+\boldsymbol{\theta})) \ .\end{aligned}$$

Thus

$$\begin{aligned}\psi(\mathbf{w}+\mathbf{c}) = -\frac{\beta N}{2}\|\mathbf{w}+\mathbf{c}\|^2 + N\log\operatorname{ch}(\beta m^* + h) + \beta\Big(m^* \sum_{i\le N} \boldsymbol{\eta}_i\Big) \cdot (\mathbf{w}+\boldsymbol{\theta}) \\ + \sum_{i\le N} \xi(\boldsymbol{\eta}_i \cdot (\mathbf{w}+\boldsymbol{\theta})) \ ,\end{aligned}$$

and since $N\mathbf{c} = m^* \sum_{i\le N} \boldsymbol{\eta}_i$ this means that

$$\begin{aligned}\psi(\mathbf{w}+\mathbf{c}) &= -\frac{\beta N}{2}\|\mathbf{w}\|^2 - \frac{\beta N}{2}\|\mathbf{c}\|^2 + \beta N \mathbf{c}\cdot\boldsymbol{\theta} + \sum_{i\le N} \xi(\boldsymbol{\eta}_i \cdot (\mathbf{w}+\boldsymbol{\theta})) \\ &\quad + N\log\operatorname{ch}(\beta m^* + h) \\ &= -\frac{\beta N}{2}\|\mathbf{w}\|^2 + \frac{\beta N}{2}\|\boldsymbol{\theta}\|^2 + \sum_{i\le N} \xi(\boldsymbol{\eta}_i \cdot (\mathbf{w}+\boldsymbol{\theta})) + Nb^*\end{aligned}$$

because $\|\mathbf{c}\|^2 = m^{*2} + \|\boldsymbol{\theta}\|^2$ and $\mathbf{c}\cdot\boldsymbol{\theta} = \|\boldsymbol{\theta}\|^2$. Thus (10.26) implies that

$$\begin{aligned}\overline{G}(C+\mathbf{c}) &\le W \int_C \exp\Big(-\frac{\beta N}{2}\|\mathbf{w}\|^2 + \sum_{i\le N} \xi(\boldsymbol{\eta}_i \cdot (\mathbf{w}+\boldsymbol{\theta}))\Big) \mathrm{d}\mathbf{w} \\ &= \int_C \exp\Big(\sum_{i\le N} \xi(\boldsymbol{\eta}_i \cdot (\mathbf{w}+\boldsymbol{\theta}))\Big) \mathrm{d}\gamma(\mathbf{w}) \ .\end{aligned}$$

To conclude it suffices to observe that the function ξ of (10.23) is convex, so that

$$\xi(x+y) \le \frac{1}{2}(\xi(2x) + \xi(2y)) \ . \qquad \square$$

To use (10.24) we need to control quantities such as $\sum_{i\le N} \xi(2\boldsymbol{\eta}_i \cdot \mathbf{w})$. It will certainly help to gain some further understanding of the function ξ. We use the notation

$$x^- = \max(0, -x) \ .$$

Lemma 10.3.7. *We have*

$$\xi(x) \le L\exp\left(-\frac{\beta}{L}\right) x^2 + \beta(1+2x)^- \ . \tag{10.27}$$

Thus, (for large β) we may think of $\xi(x)$ as the sum of a very small multiple of x^2, and of a term that is equal to 0 for $x \geq -1/2$. It seems that for our purposes this lemma really captures the behavior of the function ξ.

Proof. We observe that $\xi(0) = 0$, $\xi'(0) = 0$ and

$$\xi''(x) = \frac{\beta^2}{\mathrm{ch}^2(\beta x + \beta m^* + h)} ,$$

we may assume that β_0 is large enough so that $m^* \geq 3/4$ for $\beta \geq \beta_0$. Hence if $x \geq -1/2$ we have

$$\xi''(x) \leq \frac{\beta^2}{\mathrm{ch}^2(\beta/4)} \leq L \exp(-\beta/L) .$$

Moreover, since $|\xi'(x)| \leq 2\beta$, whenever $x \leq -1/2$, we have

$$\xi(x) \leq \xi(-1/2) + 2\beta \left|\frac{1}{2} + x\right| = \xi(-1/2) + 2\beta \left(\frac{1}{2} + x\right)^- . \qquad \square$$

The first term in (10.27) will not be a problem, because the coefficient is very small, but controlling the second term requires a real effort. In the case of the sum $\sum_{i \leq N} \xi(2\boldsymbol{\eta}_i \cdot \boldsymbol{\theta})$ occurring in (10.24) this control is achieved by the following result.

Lemma 10.3.8. *Consider $0 < c \leq 1/8$. Then if $\alpha \leq c$ with overwhelming probability it holds*

$$\sum_{i \leq N} (1 + 4\boldsymbol{\eta}_i \cdot \boldsymbol{\theta})^- \leq LN \exp\left(-\frac{1}{Lc}\right) .$$

We will comment and prove this technical fact last. First, we combine it with (10.24) in the next corollary.

Corollary 10.3.9. *Consider $0 < c \leq 1/8$. Then, if $\alpha \leq c$, with overwhelming probability, for any set $C \subset \{\|\mathbf{w}\| \,;\, \|\mathbf{w}\| \leq 1\}$ we have*

$$\overline{G}(C + \mathbf{c}) \leq \exp LN\left(\exp\left(-\frac{\beta}{L}\right) + \beta \exp\left(-\frac{1}{Lc}\right)\right) \times \int_C \exp\left(\frac{\beta}{2} \sum_{i \leq N} (1 + 4\boldsymbol{\eta}_i \cdot \mathbf{w})^-\right) \mathrm{d}\gamma(\mathbf{w}) . \qquad (10.28)$$

Proof. We recall that from (A.51) we have with overwhelming probability

$$\forall \mathbf{w} , \quad \sum_{i \leq N} (\boldsymbol{\eta}_i \cdot \mathbf{w})^2 \leq LN \|\mathbf{w}\|^2 , \qquad (10.29)$$

so that in particular

$$\sum_{i\leq N}(\boldsymbol{\eta}_i\cdot\boldsymbol{\theta})^2 \leq LN\|\boldsymbol{\theta}\|^2 \ ,$$

and with overwhelming probability $\|\boldsymbol{\theta}\|^2 \leq L$ (as follows from (10.19)), so that combining with (10.27) and Lemma 10.3.8, with overwhelming probability,

$$\sum_{i\leq N}\xi(2\boldsymbol{\eta}_i\cdot\boldsymbol{\theta}) \leq LN\left(\exp\left(-\frac{\beta}{L}\right)+\beta\exp\left(-\frac{1}{Lc}\right)\right) \ .$$

Also, using (10.27), (10.29) and since $\|\mathbf{w}\|\leq 1$ for $\mathbf{w}$ in C, with overwhelming probability

$$\mathbf{w}\in C \quad\Rightarrow\quad \frac{1}{2}\sum_{i\leq N}\xi(2\boldsymbol{\eta}_i\cdot\mathbf{w}) \leq LN\exp\left(-\frac{\beta}{L}\right)+\frac{\beta}{2}\sum_{i\leq N}(1+4\boldsymbol{\eta}_i\cdot\mathbf{w})^- \ ,$$

and the conclusion follows from (10.24). □

To apply Corollary 10.3.9 we must learn how to control the sum

$$\sum_{i\leq N}(1+4\boldsymbol{\eta}\cdot\mathbf{w})^-$$

which occurs in (10.28). This is our next goal.

Lemma 10.3.10. *We have*

$$(1+x)^- \leq |x|\mathbf{1}_{\{|x|\geq 1\}} \leq x^2\mathbf{1}_{\{|x|\geq 1\}} \ . \tag{10.30}$$

For any $x\in\mathbb{R}$ the function $a\mapsto(a+x)^-$ is non-increasing. (10.31)

Proof. Since $(1+x)^-=0$ for $x\geq -1$, to prove (10.30) we may assume that $x\leq -1$. But then $(1+x)^- = -x-1\leq -x=|x|$. This proves (10.30) and (10.31) is obvious since

$$y\mapsto y^- = \max(0,-y)$$

is non-increasing. □

Lemma 10.3.11. *There exists a number L with the following property. Consider numbers $0<d<1/L$, $0<c<1/2$. Then if $\alpha\leq c$, with overwhelming probability*

$$\|\mathbf{w}\|\leq d \quad\Rightarrow\quad \sum_{i\leq N}(1+4\boldsymbol{\eta}_i\cdot\mathbf{w})^- \leq LNd^2\left(c+\exp\left(-\frac{1}{Ld^2}\right)\right) \ . \tag{10.32}$$

Unfortunately it is not easy to provide an intuitive explanation for the somewhat unusual terms that occur in the right-hand side of (10.32). The author feels that these terms are not necessarily an artifact of his approach but might rather capture the true behavior of the left-hand side.

Proof. We observe that for $\|\mathbf{w}\| \le d$ inequality (10.30) yields

$$\sum_{i\le N}(1+4\boldsymbol{\eta}_i\cdot\mathbf{w})^- \le 16\sum_{I(\mathbf{w})}|\boldsymbol{\eta}_i\cdot\mathbf{w}|^2 \ ,$$

where

$$I(\mathbf{w}) = \left\{i\le N\ ;\ |\boldsymbol{\eta}_i\cdot\mathbf{w}| \ge \frac{1}{4}\right\} \subset J(\mathbf{w}) := \left\{i\le N\ ;\ |\boldsymbol{\eta}_i\cdot\mathbf{w}| \ge \frac{1}{4d}\|\mathbf{w}\|\right\} \ .$$

Recalling the constant L_1 of Proposition 4.2.6, consider now $b = 1/4d$ and define a by

$$a^2 = c + 4\exp\left(-\frac{1}{8L_1^2d^2}\right) \ .$$

If $c < 1/2$ and $d \le 1/8L_1$, we have $a^2 < 1$. Moreover, since $2/a \le \exp(1/16L_1^2d^2)$, the inequality

$$L_1\sqrt{\log\frac{2}{a}} \le L_1\sqrt{\frac{1}{16L_1^2d^2}} = \frac{1}{4d} = b$$

holds. We can then use Proposition 4.2.6 for these values of a and b and this concludes the proof. □

Proposition 10.3.12. *If β_0 and L_0 are large enough, then for $\beta \ge \beta_0$ and $\alpha \le c := 1/L_0\log\beta$ the set $C + \mathbf{c}$ is negligible, where*

$$C = \left\{\mathbf{w}\ ;\ \frac{1}{4\sqrt{\beta}} \le \|\mathbf{w}\| \le \frac{1}{L}\right\} \ .$$

Proof. Consider d as in Lemma 10.3.11, and define

$$C_d = \left\{\mathbf{w}\ ;\ \frac{d}{2} \le \|\mathbf{w}\| \le d\right\} \ .$$

Combining (10.32) with Corollary 10.3.9 yields that with overwhelming probability

$$\overline{G}(C_d + \mathbf{c}) \le (\exp LN\delta)\gamma(C_d) \tag{10.33}$$

where

$$\delta = d^2\beta\left(c + \exp\left(-\frac{1}{Ld^2}\right)\right) + \exp\left(-\frac{\beta}{L}\right) + \beta\exp\left(-\frac{1}{cL}\right) \ .$$

On the other hand, it is shown in (4.55) that if $\alpha \le \beta d^2/16$, then

$$\gamma(C_d) \le \gamma\left(\left\{\|\mathbf{w}\|^2 \ge \frac{d^2}{4}\right\}\right) \le \exp\left(-\frac{N\beta d^2}{32}\right) . \tag{10.34}$$

Let us assume that $d \ge 1/(2\sqrt{\beta})$. Then for $\alpha \le 1/(2^6 \log\beta)$ and $\beta_0 \ge e$, it is true that $\alpha \le 2^{-6} \le \beta d^2/16$ so that (10.34) holds. Combining with (10.33) we see that with overwhelming probability

$$\overline{G}(C_d + \mathbf{c}) \le \exp N(L^{\sim}\delta - \beta d^2/L^{\sim}) ,$$

where $L^{\sim}$ is a new constant. Now we observe that if $\beta \ge \beta_0$ large enough, L_0 is large enough and if $d \le 1/L_0$, $c \le 1/L_0 \log\beta$, we have

$$c + \exp\left(-\frac{1}{Ld^2}\right) \le \frac{1}{6L^{\sim 2}}$$

and therefore

$$d^2\beta\left(c + \exp\left(-\frac{1}{Ld^2}\right)\right) \le \frac{\beta d^2}{6L^{\sim 2}} ,$$

and also

$$\exp\left(-\frac{\beta}{L}\right) \le \frac{1}{12L^{\sim 2}} \le \frac{\beta d^2}{6L_0^2} ,$$

$$\beta \exp\left(-\frac{1}{cL}\right) \le \beta^{1-L_0/L} \le \frac{1}{12L^{\sim 2}} \le \frac{\beta d^2}{6L_0^2} .$$

Consequently we have $L^{\sim}\delta - \beta d^2/L^{\sim} < 0$.

We have proved that the set $C_d + \mathbf{c}$ is negligible provided $1/(2\sqrt{\beta}) \le d \le 1/L$. A few sets of this type cover $C + \mathbf{c}$, so the set $C + \mathbf{c}$ is negligible. □

Proposition 10.3.12 shows that from now on we may only concern ourselves with the case $\|\mathbf{w}\| < 1/4\sqrt{\beta}$. In that case, the next argument will give a better control than Lemma 10.3.11.

Lemma 10.3.13. *If* $\|\mathbf{w}\| \le 1/(4\sqrt{\beta})$ *then*

$$\mathsf{E} \exp\frac{\beta}{2}(1 + 4\boldsymbol{\eta}_i \cdot \mathbf{w})^- \le 1 + \exp\left(-\frac{\beta}{2}\right) . \tag{10.35}$$

Proof. Using (A.32) we have

$$\mathsf{E} \exp\frac{\beta}{2}(1 + 4\boldsymbol{\eta}_i \cdot \mathbf{w})^- = 1 + \frac{\beta}{2}\int_0^\infty \exp\frac{\beta t}{2}\mathsf{P}((1 + 4\boldsymbol{\eta}_i \cdot \mathbf{w})^- \ge t)\mathrm{d}t$$

and, by the subgaussian inequality (A.16),

$$\begin{aligned}\mathsf{P}((1+4\boldsymbol{\eta}_i\cdot\mathbf{w})^- \ge t) = \mathsf{P}\left(\boldsymbol{\eta}_i\cdot\mathbf{w} \le -\frac{(1+t)}{4}\right) &\le \exp\left(-\frac{(1+t)^2}{32\|\mathbf{w}\|^2}\right)\\ &\le \exp\left(-\frac{\beta(1+t)^2}{2}\right)\\ &\le \exp\left(-\frac{\beta}{2}-\frac{\beta t}{2}\right)\end{aligned}$$

so that

$$\begin{aligned}\frac{\beta}{2}\int_0^\infty \exp\frac{\beta t}{2}\mathsf{P}((1+4\boldsymbol{\eta}_i\cdot\mathbf{w})^- \ge t)\mathrm{d}t &\le \frac{\beta}{2}\int_0^\infty \exp\left(-\frac{\beta}{2}-\frac{\beta t}{2}\right)\mathrm{d}t\\ &= \exp\left(-\frac{\beta}{2}\right) ,\end{aligned}$$

which finishes the proof. □

Proposition 10.3.14. *If β_0 and L_0 are large enough, for $\beta \ge \beta_0$ and $\alpha \le c := 1/(L_0 \log\beta)$, the set $C + \mathbf{c}$ is negligible where*

$$C = \left\{\mathbf{w}\ ;\ z_1 \le -\beta^{-20}\ ,\ \|\mathbf{w}\| \le \frac{1}{4\sqrt{\beta}}\right\} .$$

Proof. Lemma 10.3.13, the inequality $1+x \le \exp x$ and independence imply

$$\begin{aligned}\mathsf{E}\int_C \exp\frac{\beta}{2}\sum_{i\le N}(1+4\boldsymbol{\eta}_i\cdot\mathbf{w})^-\mathrm{d}\gamma(\mathbf{w}) &= \int_C \mathsf{E}\exp\frac{\beta}{2}\sum_{i\le N}(1+4\boldsymbol{\eta}_i\cdot\mathbf{w})^-\mathrm{d}\gamma(\mathbf{w})\\ &\le \exp\left(N\exp\left(-\frac{\beta}{2}\right)\right)\gamma(C) ,\end{aligned}$$

so that with overwhelming probability

$$\int_C \exp\frac{\beta}{2}\sum_{i\le N}(1+4\boldsymbol{\eta}_i\cdot\mathbf{w})^-\mathrm{d}\gamma(\mathbf{w}) \le \exp\left(2N\exp\left(-\frac{\beta}{2}\right)\right)\gamma(C) ,$$

and by Corollary 10.3.9 we have

$$G(C+\mathbf{c}) \le (\exp LN\delta)\gamma(C)$$

where now

$$\delta = \exp\left(-\frac{\beta}{2}\right) + \beta\exp\left(-\frac{1}{Lc}\right) \le \exp\left(-\frac{\beta}{2}\right) + \beta^{1-L_0/L} .$$

Moreover since the law of z_1 under γ is Gaussian with variance $\mathsf{E}z_1^2 = 1/\beta\sqrt{N}$, using (A.4) we have

$$\gamma(C) \le \gamma(\{\mathbf{z}\ ;\ z_1 \le -\beta^{-20}\}) \le \exp\left(-\frac{N\beta}{2}(\beta^{-20})^2\right) .$$

To conclude the proof, we observe that if $\beta \ge \beta_0$ and $L_0 \ge 41L$, we have $L\delta < \beta^{-39}/2$. □

Proof of Proposition 10.3.5. Combining Propositions 10.3.12 and 10.3.14 we have shown that the set $C + \mathbf{c}$ is negligible for $\overline{G}$, where

$$C = \left\{\mathbf{z}\ ;\ \|\mathbf{z}\| \le \frac{1}{L}\ ,\ z_1 < -\frac{1}{\beta^{20}}\right\} ,$$

so that

$$C + \mathbf{c} = \left\{\mathbf{z}\ ;\ \|\mathbf{z} - \mathbf{c}\| \le \frac{1}{L_1}\ ,\ z_1 < m^* - \frac{1}{\beta^{20}}\right\} .$$

Since $\mathbf{c} = m^*\mathbf{e}_1 + \boldsymbol{\theta}$ and since with overwhelming probability we have $\|\boldsymbol{\theta}\| \le 2\sqrt{\alpha} \le 2/\sqrt{L_0 \log \beta} \le 1/2L_1$ when $\beta \ge \beta_0$ and β_0 is large enough, we have

$$C + \mathbf{c} \supset \left\{\mathbf{z}\ ;\ \|\mathbf{z} - m^*\mathbf{e}_1\| \le \frac{1}{2L_1}\ ,\ z_1 < m^* - \frac{1}{\beta^{20}}\right\}$$

and thus this latter set is negligible for $\overline{G}$. Therefore it suffices to prove that the set

$$D_1 = \left\{\mathbf{z}\ ;\ \|\mathbf{z} - m^*\mathbf{e}_1\| \ge \frac{1}{2L_1}\right\}$$

is negligible for $\overline{G}$.

By Theorem 4.4.4 (used for $\rho_0 = 1/4L_1$, and since $\alpha \le 1/(L_0 \log \beta) \le m^{*2}\rho_0^2/L$ for $\beta \ge 2$ and L_0 large enough) the set

$$D_2 = \left\{\mathbf{z}\ ;\ \|\mathbf{z} - m^*\mathbf{e}_1\| \ge \frac{1}{4L_1}\right\}$$

is negligible for G'. Now, by definition of $\overline{G} = G' * \gamma$ we have

$$\overline{G}(D_1) = G' \otimes \gamma(\{(\mathbf{z}, \mathbf{w})\ ;\ \mathbf{z} + \mathbf{w} \in D_1\}) . \tag{10.36}$$

If $\mathbf{z} + \mathbf{w} \in D_1$, i.e. $\|\mathbf{z} + \mathbf{w} - m^*\mathbf{e}_1\| \ge 1/2L_1$, either $\|\mathbf{w}\| \ge 1/4L_1$ or else $\|\mathbf{z} - m^*\mathbf{e}_1\| \ge 1/4L_1$, so that then $\mathbf{z} \in D_2$ and hence

$$G' \otimes \gamma(\{(\mathbf{z}, \mathbf{w})\ ;\ \mathbf{z} + \mathbf{w} \in D_1\}) \le G'(D_2) + \gamma\left(\left\{\|\mathbf{w}\| \ge \frac{1}{4L_1}\right\}\right) .$$

By (10.36), and since $\gamma(\{\|\mathbf{w}\| \ge 1/4L_1\})$ is very small, D_1 is negligible for $\overline{G}$. □

Finally we can complete the proof of Theorem 10.3.1 by proving Lemma 10.3.8.

Proof of Lemma 10.3.8. Given a vector $\mathbf{x}$ with $\|\mathbf{x}\|^2 \le 2c$, it is not very hard to see that with overwhelming probability we have

$$\sum_{i\le N}(1+4\boldsymbol{\eta}_i\cdot\mathbf{x})^- \le LN\exp\left(-\frac{1}{Lc}\right)\,. \tag{10.37}$$

Therefore, for the typical realization of $(\boldsymbol{\eta}_i)_{i\le N}$, very few such vectors $\mathbf{x}$ will fail to satisfy (10.37). It seems very unlikely that $\boldsymbol{\theta}$ is one of these, and this is what we shall measure. The technical difficulty is that the quantities $(\boldsymbol{\eta}_i\cdot\boldsymbol{\theta})_{i\le N}$ are dependent, and that it is hard to work without independence. This makes the proof rather technical, and needless to say, a special gift for the really motivated reader.

As usual, the basic idea is to bring out some independence. For a subset I of $\{1,\ldots,N\}$ we define

$$\boldsymbol{\theta}_I = \frac{1}{N}\sum_{i\in I}(\boldsymbol{\eta}_i-\mathbf{e}_1) \tag{10.38}$$

so that $\boldsymbol{\theta}=\boldsymbol{\theta}_{\{1,\ldots,N\}}$. We have

$$\|\boldsymbol{\theta}_I\|^2 = \sum_{2\le k\le M}\left(\frac{1}{N}\sum_{i\in I}\eta_{i,k}\right)^2\,,$$

so that by Lemma 10.3.4 (used for $\rho_i=1/N$ if $i\in I$ and $\rho_i=0$ otherwise, so that $\sum_{i\le N}\rho_i^2\le 1/N$) we have, since $\alpha=M/N\le c$,

$$\mathsf{P}(\|\boldsymbol{\theta}_I\|^2 > 8c) \le 2^{M/2}\exp(-2Nc)\le 2^{-cN}\,. \tag{10.39}$$

In words, $\|\boldsymbol{\theta}_I\|^2\le 8c$ with overwhelming probability. We write, using the Cauchy-Schwarz inequality and (10.30),

$$\begin{aligned}\sum_{i\notin I}(1+16\boldsymbol{\eta}_i\cdot\boldsymbol{\theta}_I)^- &\le 16\sum_{i\notin I}|\boldsymbol{\eta}_i\cdot\boldsymbol{\theta}_I|\mathbf{1}_{\{16|\boldsymbol{\eta}_i\cdot\boldsymbol{\theta}_I|\ge1\}}\\ &\le 16\left(\sum_{i\notin I}(\boldsymbol{\eta}_i\cdot\boldsymbol{\theta}_I)^2\right)^{1/2}\left(\sum_{i\notin I}\mathbf{1}_{\{16|\boldsymbol{\eta}_i\cdot\boldsymbol{\theta}_I|\ge1\}}\right)^{1/2}.\end{aligned}$$

Now using (10.29) yields that, with overwhelming probability,

$$\sum_{i\notin I}(1+16\boldsymbol{\eta}_i\cdot\boldsymbol{\theta}_I)^- \le L\sqrt{N}\left(\sum_{i\notin I}\mathbf{1}_{\{16|\boldsymbol{\eta}_i\cdot\boldsymbol{\theta}_I|\ge1\}}\right)^{1/2}. \tag{10.40}$$

Given $\mathbf{x}$ with $\|\mathbf{x}\|^2\le 8c$, the r.v.s

$$\mathbf{1}_{\{16|\boldsymbol{\eta}_i\cdot\mathbf{x}|\ge1\}}$$

for $i \notin I$ are i.i.d. with expectation $\le L\exp(-1/Lc)$ (as follows from the subgaussian inequality). Consequently, using elementary properties about the tail of the binomial law, we see that with probability $\ge 1 - K\exp(-N/K)$ their sum is $\le LN\exp(-1/Lc)$. Using this for $\mathbf{x} = \boldsymbol{\theta}_I$, given $\boldsymbol{\theta}_I$ and combining with (10.39) we have proved that with probability $\ge 1 - K\exp(-N/K)$ we have

$$\sum_{i\notin I}(1 + 16\boldsymbol{\eta}_i \cdot \boldsymbol{\theta}_I)^- \le LN\exp\left(-\frac{1}{Lc}\right) ,$$

i.e. there exists an event Ω_I with $\mathsf{P}(\Omega_I) \ge 1 - K\exp(-N/K)$ such that

$$\mathbf{1}_{\Omega_I}\sum_{i\notin I}(1 + 16\boldsymbol{\eta}_i \cdot \boldsymbol{\theta}_I)^- \le LN\exp\left(-\frac{1}{Lc}\right) . \tag{10.41}$$

Let us denote by Av an average over all possible choices of I. Then

$$\mathrm{Av}\sum_{i\notin I}(1 + 16\boldsymbol{\eta}_i \cdot \boldsymbol{\theta}_I)^- \le \mathrm{Av}\mathbf{1}_{\Omega_I}\sum_{i\notin I}(1 + 16\boldsymbol{\eta}_i \cdot \boldsymbol{\theta}_I)^- + \mathcal{R} \tag{10.42}$$

where

$$\mathcal{R} = \mathrm{Av}\mathbf{1}_{\Omega_I^c}\sum_{i\notin I}(1 + 16\boldsymbol{\eta}_i \cdot \boldsymbol{\theta}_I)^- .$$

Since it holds that $\|\boldsymbol{\eta}_i\| \le \sqrt{M} \le \sqrt{N}$ and $\|\boldsymbol{\theta}_I\| \le \sqrt{N}$, we have

$$\left|\sum_{i\in I}(1 + 16\boldsymbol{\eta}_i \cdot \boldsymbol{\theta}_I)\right| \le \sum_{i\in I}(1 + 16N) \le 17N^2 ,$$

and consequently

$$\begin{aligned}\mathsf{E}\mathcal{R} &= \mathrm{Av}\mathsf{E}\left(\mathbf{1}_{\Omega_I^c}\sum_{i\in I}(1 + 16\boldsymbol{\eta}_i \cdot \boldsymbol{\theta}_I)^-\right)\\ &\le 17N^2\mathrm{Av}\mathsf{E}\mathbf{1}_{\Omega_I^c} \le KN^2\exp\left(-\frac{N}{K}\right) .\end{aligned}$$

In particular with overwhelming probability $\mathcal{R} \le LN\exp(-1/(Lc))$, and (10.41) and (10.42) show that with overwhelming probability

$$\mathrm{Av}\sum_{i\notin I}(1 + 16\boldsymbol{\eta}_i \cdot \boldsymbol{\theta}_I)^- \le LN\exp\left(-\frac{1}{Lc}\right) . \tag{10.43}$$

Now the function $x \to (1 + x)^-$ is convex, so that by Jensen's inequality

$$\begin{aligned}\mathrm{Av}\sum_{i\notin I}(1 + 16\boldsymbol{\eta}_i \cdot \boldsymbol{\theta}_I)^- &= \mathrm{Av}\sum_{i\le N}(\mathbf{1}_{\{i\notin I\}}(1 + 16\boldsymbol{\eta}_i \cdot \boldsymbol{\theta}_I))^-\\ &\ge \sum_{i\le N}(\mathrm{Av}(\mathbf{1}_{\{i\notin I\}}(1 + 16\boldsymbol{\eta}_i \cdot \boldsymbol{\theta}_I))^- .\end{aligned} \tag{10.44}$$

Of course $\mathrm{Av}(\mathbf{1}_{\{i \notin I\}}) = 1/2$, and by definition (10.38)

$$\boldsymbol{\theta}_I = \frac{1}{N} \sum_{j \leq N} \mathbf{1}_{\{j \in I\}} (\boldsymbol{\eta}_j - \mathbf{e}_1) \ ,$$

so that

$$\begin{aligned} \mathrm{Av} \mathbf{1}_{\{i \notin I\}} \boldsymbol{\eta}_i \cdot \boldsymbol{\theta}_I &= \mathrm{Av} \frac{1}{N} \sum_{j \leq N} \mathbf{1}_{\{i \notin I\}} \mathbf{1}_{\{j \in I\}} \boldsymbol{\eta}_i \cdot (\boldsymbol{\eta}_j - \mathbf{e}_1) \\ &= \frac{1}{4N} \sum_{j \neq i} \boldsymbol{\eta}_i \cdot (\boldsymbol{\eta}_j - \mathbf{e}_1) \\ &= \frac{1}{4} \boldsymbol{\eta}_i \cdot \boldsymbol{\theta} - \frac{1}{4N} \boldsymbol{\eta}_i \cdot (\boldsymbol{\eta}_i - \mathbf{e}_1) \ . \end{aligned}$$

Thus from (10.44) we get

$$\mathrm{Av} \sum_{i \in I} (1 + 16 \boldsymbol{\eta}_i \cdot \boldsymbol{\theta}_I)^- \geq \sum_{i \leq N} \left(\frac{1}{2} + 4 \boldsymbol{\eta}_i \cdot \boldsymbol{\theta} - \frac{4}{N} \boldsymbol{\eta}_i \cdot (\boldsymbol{\eta}_i - \mathbf{e}_1) \right)^- \ . \tag{10.45}$$

Now $\boldsymbol{\eta}_i \cdot (\boldsymbol{\eta}_i - \mathbf{e}_1) = \sum_{2 \leq k \leq M} \eta_{i,k}^2 = M - 1$, so that since $\alpha = M/N \leq 1/8$ we obtain $|4 \boldsymbol{\eta}_i \cdot (\boldsymbol{\eta}_i - \mathbf{e}_1)| \leq N/2$. Hence, by (10.31), we have

$$\left(\frac{1}{2} + 4 \boldsymbol{\eta}_i \cdot \boldsymbol{\theta} - \frac{4}{N} \boldsymbol{\eta}_i \cdot (\boldsymbol{\eta}_i - \mathbf{e}_1) \right)^- \geq (1 + 4 \boldsymbol{\eta}_i \cdot \boldsymbol{\theta})^- \ .$$

Combining with (10.43) and (10.45) completes the proof. □

10.4 Controlling $m_k(\boldsymbol{\sigma})$, $k \geq 2$

In this section we prove that if (α, β, h) belongs to the admissible region (10.1) then for $k \geq 2$, $m_k(\boldsymbol{\sigma})$ is typically small, about $1/\sqrt{N}$. This technical fact will be essential in Section 10.6, in order to estimate the error terms created by the "approximate integration by parts" of (4.197). It is here that the condition $h > 0$ is really being used.

Theorem 10.4.1. *If (α, β, h) belongs to the admissible region (10.1) (and if L_0 is large enough) then there exists a constant K depending only on β and h such that, for all $k \geq 2$, we have*

$$\mathsf{E} \left\langle \exp \frac{N m_k^2}{K} \right\rangle \leq K \ . \tag{10.46}$$

In particular for each p we have $\mathsf{E} \langle m_k^{2p} \rangle \leq K N^{-p}$, where K now depends also on p. To prove Theorem 10.4.1 we can assume by symmetry that $k = M$.

The proof uses a "cavity upon M" argument. We denote by $\langle\cdot\rangle_\sim$ the Gibbs measure relative to the Hamiltonian

$$-H_{N,M-1}(\boldsymbol{\sigma}) = \frac{\beta}{2}\sum_{k\le M-1} N m_k^2(\boldsymbol{\sigma}) + Nhm_1(\boldsymbol{\sigma})\ . \tag{10.47}$$

It should be obvious that for a function f on Σ_N we have

$$\langle f\rangle = \frac{\big\langle f\exp\frac{N\beta}{2}m_M^2\big\rangle_\sim}{\big\langle\exp\frac{N\beta}{2}m_M^2\big\rangle_\sim}\ . \tag{10.48}$$

To prove Theorem 10.4.1, we will consider the set

$$U = \{\boldsymbol{\sigma}\ ;\ |m_1(\boldsymbol{\sigma}) - m^*| \le \rho\}\ , \tag{10.49}$$

where ρ will be suitably chosen. To lighten notation we write $\mathbf{1}_U$ for $\mathbf{1}_U(\boldsymbol{\sigma})$ and $\mathbf{1}_{U^c}$ for $\mathbf{1}_{U^c}(\boldsymbol{\sigma})$. The main step of the proof is as follows.

Proposition 10.4.2. *Assume that*

$$\beta(1 - m^{*2} + 2\rho m^*) < 1\ . \tag{10.50}$$

Then we may find $\beta_1 > \beta$ and $s > 0$ such that

$$\mathsf{E}\left(\left\langle \mathbf{1}_U \exp\frac{N\beta_1}{2}m_M^2\right\rangle_\sim^s\right) \le K\ . \tag{10.51}$$

Proof. Writing η_i rather than $\eta_{i,M}$, we obtain

$$m_M = m_M(\boldsymbol{\sigma}) = \frac{1}{N}\left(\sum_{i\le N}\eta_i m^* + \sum_{i\le N}\eta_i(\sigma_i - m^*)\right)\ . \tag{10.52}$$

Given $t > 0$, we have

$$(a+b)^2 \le (1+t)a^2 + \left(1+\frac{1}{t}\right)b^2\ ,$$

and applying this to (10.52) we get

$$Nm_M^2 \le \frac{1+t}{N}\left(\sum_{i\le N}\eta_i m^*\right)^2 + \frac{1+t}{Nt}\left(\sum_{i\le N}\eta_i(\sigma_i - m^*)\right)^2\ , \tag{10.53}$$

and thus

$$\begin{aligned}&\left\langle \mathbf{1}_U\exp\frac{N\beta_1}{2}m_M^2\right\rangle_\sim^s\\ &\le \exp\left(s(1+t)\frac{\beta_1}{2N}m^{*2}\left(\sum_{i\le N}\eta_i\right)^2\right)\\ &\times\left\langle \mathbf{1}_U\exp\left(\frac{(1+t)\beta_1}{2Nt}\left(\sum_{i\le N}\eta_i(\sigma_i - m^*)\right)^2\right)\right\rangle_\sim^s\ .\end{aligned} \tag{10.54}$$

Thus, by the Cauchy-Schwarz inequality applied to (10.54) we have

$$\begin{aligned}
&\left(\mathsf{E}\left\langle \mathbf{1}_U \exp \frac{N\beta_1}{2} m_M^2 \right\rangle_\sim^s\right)^2 \\
&\leq \mathsf{E} \exp\left(s(1+t)\frac{\beta_1}{N} m^{*2} \left(\sum_{i\leq N} \eta_i\right)^2\right) \\
&\times \mathsf{E} \left\langle \mathbf{1}_U \exp\left(\frac{(1+t)\beta_1}{2Nt} \left(\sum_{i\leq N} \eta_i(\sigma_i - m^*)\right)^2\right)\right\rangle_\sim^{2s} . \qquad (10.55)
\end{aligned}$$

There is no loss of generality to assume $2s \leq 1$. For a r.v. $X \geq 0$, Hölder's inequality implies that $\mathsf{E}X^{2s} \leq (\mathsf{E}X)^{2s}$. Thus the last term of (10.55) is at most

$$\left(\mathsf{E} \left\langle \mathbf{1}_U \exp\left(\frac{(1+t)\beta_1}{2Nt} \left(\sum_{i\leq N} \eta_i(\sigma_i - m^*)\right)^2\right)\right\rangle_\sim \right)^{2s} . \qquad (10.56)$$

Now,

$$\begin{aligned}
\sum_{i\leq N} (\sigma_i - m^*)^2 &= N(1 - 2m_1(\boldsymbol{\sigma})m^* + m^{*2}) \\
&= N(1 - m^{*2} + 2(m^* - m_1(\boldsymbol{\sigma}))m^*) \\
&\leq N(1 - m^{*2} + 2\rho m^*) ,
\end{aligned}$$

for $\boldsymbol{\sigma} \in U$. Using (10.50), we choose $\beta_1 > \beta$ and t large enough so that

$$\frac{1+t}{t}\beta_1(1 - m^{*2} + 2\rho m^*) < 1 . \qquad (10.57)$$

Denote by E_0 integration in the r.v.s $(\eta_i)_{i\leq N}$ alone. Then (10.57) and the subgaussian inequality (A.19) show that for $\boldsymbol{\sigma}$ in U, we have

$$\mathsf{E}_0 \exp\left(\frac{(1+t)\beta_1}{2Nt} \left(\sum_{i\leq N} \eta_i(\sigma_i - m^*)\right)^2\right) \leq K ,$$

where K depends neither on N nor $\boldsymbol{\sigma}$. Since the disorder in $\langle\cdot\rangle_\sim$ is independent of the r.v.s η_i, when taking expectation in (10.56), we can first take expectation E_0 inside the bracket $\langle\cdot\rangle_\sim$. Then the quantity (10.56) is bounded independently of N. To finish the proof, we choose s small enough so that

$$s(1+t)\beta_1 m^{*2} < \frac{1}{2} .$$

Then the first term on the right-hand side of (10.55) is bounded independently of N, again by (A.19). □

Lemma 10.4.3. *If the constant L_0 in the definition of the admissible region (10.1) has been chosen large enough, at each point (α, β, h) of this admissible region we can find ρ such that*

$$\beta(1 - m^{*2} + 2\rho m^*) < 1 \tag{10.58}$$

and K depending only on β and h such that

$$\mathsf{E}\, G(\{\boldsymbol{\sigma}\ ;\ |m_1(\boldsymbol{\sigma}) - m^*| > \rho\}) \le K \exp\left(-\frac{N}{K}\right) , \tag{10.59}$$

where as usual G denotes Gibbs' measure.

Proof. We take

$$\rho = \frac{1 - \beta(1 - m^{*2})}{4\beta m^*}.$$

This ensures that (10.58) holds because $1-\beta(1-m^{*2}) > 0$ and hence $2\rho\beta m^* < 1 - \beta(1 - m^{*2})$. Using (4.38), we have

$$\rho \ge \rho' := \frac{m^*}{L\beta} .$$

We recall the vector $\mathbf{c}$ of (10.18). Since its first component is m^*, we have

$$\|\mathbf{m}(\boldsymbol{\sigma}) - \mathbf{c}\| \ge |m_1(\boldsymbol{\sigma}) - m^*|$$

and

$$\{\boldsymbol{\sigma}\ ;\ |m_1(\boldsymbol{\sigma}) - m^*| \ge \rho\} \subset \{\boldsymbol{\sigma}\ ;\ \|\mathbf{m}(\boldsymbol{\sigma}) - \mathbf{c}\| \ge \rho'\} . \tag{10.60}$$

Given $L_1 > 0$, there exists β_0 such that $\rho' \ge L_1/\beta^{10}$ for $\beta \ge \beta_0$, and Theorem 10.3.1 shows that then the sets (10.60) are negligible. For $\beta \le \beta_0$, we have $\rho' \ge \rho_0 := m^*/L_2$, hence

$$\{\boldsymbol{\sigma}\ ;\ |m_1(\boldsymbol{\sigma}) - m^*| \ge \rho_0\} \subset \{\boldsymbol{\sigma}\ ;\ \|\mathbf{m}(\boldsymbol{\sigma}) - m^*\mathbf{e}_1\| \ge \rho_0\} . \tag{10.61}$$

Theorem 4.4.4 shows that this set is negligible if the constant L_0 is large enough. □

Proof of Theorem 10.4.1. If U is given by (10.49), where ρ is as in (10.58), for any number A then

$$\mathsf{E}\left\langle \exp \frac{N}{A} m_M^2 \right\rangle = \mathrm{I} + \mathrm{II}$$

where

$$\mathrm{I} = \mathsf{E}\left\langle \mathbf{1}_{U^c} \exp \frac{N}{A} m_M^2 \right\rangle \quad \text{and} \quad \mathrm{II} = \mathsf{E}\left\langle \mathbf{1}_U \exp \frac{N}{A} m_M^2 \right\rangle .$$

Now $|m_M| \le 1$, so that

$$\mathrm{I} \le \mathsf{E}\langle \mathbf{1}_{U^c} \rangle \exp \frac{N}{A} \le K \exp N\left(\frac{1}{A} - \frac{1}{K}\right) \tag{10.62}$$

by (10.59).

Given a number $u > 1$, by Hölder's inequality used for $\langle\cdot\rangle$ we have

$$\mathrm{II} \le \mathsf{E}\left(\left\langle \mathbf{1}_U \exp \frac{Nu}{A} m_M^2 \right\rangle^{1/u}\right) .$$

By (10.48) we have

$$\left\langle \mathbf{1}_U \exp \frac{Nu}{A} m_M^2 \right\rangle \le \left\langle \mathbf{1}_U \exp N\left(\frac{u}{A} + \beta\right) m_M^2 \right\rangle_\sim . \tag{10.63}$$

Using (10.58), Proposition 10.4.2 provides us with $s > 0$ and $\beta_1 > \beta$ such that (10.51) holds. We take $u = 1/s$, and then A large enough so that $\beta + u/A \le \beta_1$ and $A \ge 2K$, where K occurs in (10.62). From (10.63) and (10.51) we get $\mathrm{II} \le K$ as well as $\mathrm{I} \le K$ by (10.62). The result follows. □

10.5 The Smart Path

In this section we begin the study of the influence of the last spin on Gibbs' measure, by decoupling it from the other spins. This method has already been used in Section 4.5 (but the present section can be read without having seen Section 4.5). Some of the technicalities will now be quite *simpler* because we are interested only in studying $\nu(f)$ for functions f that may be random, but will never depend on the r.v.s $\eta_k = \eta_{N,k}$. We recall some notation from Section 4.5. Given $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_N) \in \Sigma_N$, we write $\boldsymbol{\rho} = (\sigma_1, \ldots, \sigma_{N-1}) \in \Sigma_{N-1}$ and

$$n_k = n_k(\boldsymbol{\sigma}) = n_k(\boldsymbol{\rho}) = \frac{1}{N} \sum_{i \le N-1} \eta_{i,k} \sigma_i , \tag{10.64}$$

so that

$$m_k = m_k(\boldsymbol{\sigma}) = n_k(\boldsymbol{\sigma}) + \frac{\eta_k \sigma_N}{N} . \tag{10.65}$$

We write

$$- H_{N-1,M}(\boldsymbol{\rho}) = \frac{N\beta}{2} \sum_{1 \le k \le M} n_k^2(\boldsymbol{\rho}) + Nhn_1(\boldsymbol{\rho}) , \tag{10.66}$$

and we recall that by simple algebra we have (4.178) i.e. (ignoring an irrelevant constant)

$$- H_{N,M}(\boldsymbol{\sigma}) = -H_{N-1,M}(\boldsymbol{\rho}) + \beta\sigma_N \sum_{1 \le k \le M} \eta_k n_k(\boldsymbol{\rho}) + \sigma_N h . \tag{10.67}$$

Recalling that (q, μ, r) denotes the solution of the replica-symmetric equations (10.3) to (10.5), we consider a standard Gaussian r.v. z and for $0 \le t \le 1$ the Hamiltonian

$$- H_{N,M,t}(\boldsymbol{\sigma}) = -H_{N-1,M}(\boldsymbol{\rho}) + \beta\sigma_N \Bigg(\sqrt{t} \sum_{2 \le k \le M} \eta_k n_k(\boldsymbol{\rho}) + \sqrt{1-t} z \sqrt{r} + t n_1(\boldsymbol{\rho}) + (1-t)\mu \Bigg) + h\sigma_N \ . \tag{10.68}$$

Thus since $\eta_1 = 1$ we have

$$H_{N,M,1} = H_{N,M} \tag{10.69}$$

$$H_{N,M,0}(\boldsymbol{\sigma}) = H_{N-1,M}(\boldsymbol{\rho}) + \sigma_N Y \tag{10.70}$$

where

$$Y = \beta(z\sqrt{r} + \mu) + h \ . \tag{10.71}$$

We denote by $\langle \cdot \rangle_t$ an average for the Gibbs measure (or its tensor products) with Hamiltonian (10.68), and as usual we write $\nu_t(f) = \mathsf{E}\langle f \rangle_t$. Although we will not need this fact, it is reassuring that this definition coincides with definition (4.181) when f does not depend on the r.v.s ξ^ℓ.

We recall the notation $\varepsilon_\ell = \sigma_N^\ell$. As expected, the measure ν_0 indeed decouples the last spin.

Lemma 10.5.1. *Given a function f^- on Σ_{N-1}^n, which does not depend on the r.v.s η_k, and given a subset I of $\{1, \dots, n\}$ we have*

$$\nu_0\Bigg(f^- \prod_{\ell \in I} \varepsilon_\ell\Bigg) = \nu_0\Bigg(\prod_{\ell \in I} \varepsilon_\ell\Bigg)\nu_0(f^-) = \mathsf{E}(\mathrm{th} Y)^{\mathrm{card} I}\nu_0(f^-) \ . \tag{10.72}$$

Proof. As for Lemma 1.6.2. □

We need to compute the derivative $\nu_t'(f) = \mathrm{d}\nu_t(f)/\mathrm{d}t$. The Hamiltonian $H_{N,M,t}$ depends on t in 3 different ways, as shown by its definition (10.68). This is why there are three terms is the forthcoming (straightforward) formula, where we recall the notation $n_k^\ell = n_k(\boldsymbol{\sigma}^\ell)$.

Lemma 10.5.2. *We have*

$$\nu_t'(f) = \mathrm{I} + \mathrm{II} + \mathrm{III}$$

where

$$\mathrm{I} = \beta\Bigg(\sum_{\ell \le n} \nu_t(f\varepsilon_\ell(n_1^\ell - \mu)) - n\nu_t(f\varepsilon_{n+1}(n_1^{n+1} - \mu))\Bigg) \tag{10.73}$$

$$\mathrm{II} = \frac{\beta}{2\sqrt{t}}\Bigg(\sum_{\ell\le n}\nu_t\Big(f\varepsilon_\ell\sum_{2\le k\le M}\eta_k n_k^\ell\Big) - n\nu_t\Big(f\varepsilon_{n+1}\sum_{2\le k\le M}\eta_k n_k^{n+1}\Big)\Bigg) \tag{10.74}$$

$$\mathrm{III} = -\frac{\beta\sqrt{r}}{2\sqrt{1-t}}\Bigg(\sum_{\ell\le n}\nu_t(fz\varepsilon_\ell) - n\nu_t(fz\varepsilon_{n+1})\Bigg)\,. \tag{10.75}$$

To make this decomposition usable, we need to integrate by parts. First we integrate by parts in III. We have already done this many times:

$$\mathrm{III} = -\beta r\Bigg(\sum_{1\le\ell'<\ell\le n}\nu_t(f\varepsilon_\ell\varepsilon_{\ell'}) - n\sum_{\ell\le n}\nu_t(f\varepsilon_\ell\varepsilon_{n+1}) + \frac{n(n+1)}{2}\nu_t(f\varepsilon_{n+1}\varepsilon_{n+2})\Bigg)\,. \tag{10.76}$$

For II, since the r.v.s η_k are Bernoulli rather than Gaussian, we need to use "approximate integration by parts" as in (4.198). (This will be detailed again below). This approximate integration by parts produces "main terms" and "error terms", and the "main terms" are the same as if one would perform Gaussian integration by parts (which we have already done many times). Writing for simplicity

$$S_{\ell,\ell'} = \sum_{2\le k\le M} n_k^\ell n_k^{\ell'}\,, \tag{10.77}$$

these main terms are

$$\frac{\beta^2}{2}\Bigg(\sum_{\ell\le n}\nu_t(fS_{\ell,\ell}) - n\nu_t(fS_{n+1,n+1})\Bigg) + \beta^2\Bigg(\sum_{1\le\ell<\ell'\le n}\nu_t(f\varepsilon_\ell\varepsilon_{\ell'}S_{\ell,\ell'}) - n\sum_{\ell\le n}\nu_t(f\varepsilon_\ell\varepsilon_{n+1}S_{\ell,n+1}) + \frac{n(n+1)}{2}\nu_t(f\varepsilon_{n+1}\varepsilon_{n+2}S_{n+1,n+2})\Bigg)\,. \tag{10.78}$$

Gathering the terms (10.76) and (10.78), we have shown that to complete the proof of the following Proposition, it will suffice to control the "error terms" occurring from approximate integration by parts.

Proposition 10.5.3. *For a function f of Σ_N^n, that does not depend on the r.v.s (η_k), we have*

$$\nu_t'(f) = \mathrm{I} + \mathrm{IV} + \mathrm{V} + \mathcal{R} \tag{10.79}$$

where I *is given by (10.73),*

$$\mathrm{IV} = \frac{\beta^2}{2}\Big(\sum_{\ell\le n}\nu_t(fS_{\ell,\ell}) - n\nu_t(fS_{n+1,n+1})\Big) \tag{10.80}$$

$$\mathrm{V} = \beta^2\Big(\sum_{1\le\ell<\ell'\le n}\nu_t(f(S_{\ell,\ell'}-r)\varepsilon_\ell\varepsilon_{\ell'}) - n\sum_{\ell\le n}\nu_t(f(S_{\ell,n+1}-r)\varepsilon_\ell\varepsilon_{n+1}) + \frac{n(n+1)}{2}\nu_t(f(S_{n+1,n+2}-r)\varepsilon_{n+1}\varepsilon_{n+2})\Big) \tag{10.81}$$

and

$$|\mathcal{R}| \le K\sum_{2\le k\le M}\sum_{\ell\le n+1}\nu_t(|n_k^\ell|^4|f|)\ , \tag{10.82}$$

where the number K depends only on β and n.

One should really think of $\mathcal{R}$ as a lower order error term. That this is indeed the case when $h>0$ and (α,β,h) belongs to the admissible region (10.1) will be proved using the results of Section 10.4. We will soon explain how to use Proposition 10.5.3, but let us prove it first.

Proof. As pointed out, the issue is to control the "error terms" arising from approximate integration by parts, and this is quite easier than in Section 4.5 because we can now use differential inequalities, so let us enjoy the argument. Fixing $\ell\le n+1$ and $2\le k\le M$, let us compute

$$\nu_t(f\eta_k\varepsilon_\ell n_k^\ell) = \mathsf{E}(\eta_k\langle f\varepsilon_\ell n_k^\ell\rangle_t)\ . \tag{10.83}$$

Thinking of t as fixed, let us consider the random function

$$v(x) = \langle f\varepsilon_\ell n_k^\ell\rangle_{t,x}$$

that is obtained by replacing in the explicit expression of $\langle f\varepsilon_\ell n_k^\ell\rangle_t$ every occurrence of η_k by x, so the quantity (10.83) is

$$\mathsf{E}(\eta_k v(\eta_k))\ .$$

When η is a Bernoulli r.v., i.e. $\mathsf{P}(\eta=\pm1)=1/2$, then approximate integration by parts is obtained by (4.198), which states that

$$\mathsf{E}\,\eta v(\eta) = \mathsf{E}\,v'(\eta) + \frac{1}{4}\int_{-1}^{1}(x^2-1)v'''(x)\mathrm{d}x\ .$$

Here we will use simply that, as an obvious consequence, if E_k denotes expectation in η_k only, then

$$\mathsf{E}_k(\eta_k v(\eta_k)) = \mathsf{E}_k v'(\eta_k) + \mathcal{R}$$

where $|\mathcal{R}|\le\sup_{|x|\le1}|v^{(3)}(x)|$. The term $\mathsf{E}_k v'(\eta_k)$ is the main term (which has already been taken into account) so that we turn to the control of $\mathsf{E}\mathcal{R}$. Given a function f on n' replicas, it is straightforward that

$$\frac{\mathrm{d}}{\mathrm{d}x}\langle f\rangle_{t,x} = \beta\sqrt{t}\left(\sum_{\ell'\le n'}\langle\varepsilon_{\ell'} n_k^{\ell'} f\rangle_{t,x} - n'\langle\varepsilon_{n'+1} n_k^{n'+1} f\rangle_{t,x}\right) . \tag{10.84}$$

Iterating this formula twice, using it for $\varepsilon_\ell n_k^\ell f$ rather than f and using the inequality $\left|\prod_{\ell\le 4} a_\ell\right| \le \sum_{\ell\le 4} a_\ell^4$, it should be obvious that

$$|v^{(3)}(x)| \le K(n,\beta)\sum_{\ell\le n+1}\langle |n_k^\ell|^4 |f|\rangle_{t,x} . \tag{10.85}$$

Moreover it follows from (10.84), and since $|n_k| \le 1$, that for a function $f^* \ge 0$ on n' replicas we have

$$\frac{\mathrm{d}}{\mathrm{d}x}\langle f^*\rangle_{t,x} \le K(n',\beta)\langle f^*\rangle_{t,x}$$

so that by integration

$$\begin{aligned}\langle f^*\rangle_{t,x} \le K(n',\beta)\langle f^*\rangle_{t,1} &\le K(n',\beta)\left(\frac{1}{2}\langle f^*\rangle_{t,1} + \frac{1}{2}\langle f^*\rangle_{t,-1}\right)\\ &= K(n',\beta)\mathsf{E}_k\langle f^*\rangle_t ,\end{aligned}$$

and thus, taking the supremum over x and then expectation,

$$\mathsf{E}\sup_{|x|\le 1}\langle f^*\rangle_{t,x} \le K(n',\beta)\nu_t(f^*) .$$

Using this in (10.85) we see that

$$\mathsf{E}\sup_{|x|\le 1}|v^{(3)}(x)| \le K(n,\beta)\sum_{\ell\le n+1}\nu_t(|n_k^\ell|^4|f|) ,$$

and this concludes the proof. □

We plan to use Proposition 10.5.3 to show that in certain circumstances $\nu_t'(f)$ is small. We certainly expect that $S_{\ell,\ell'} \simeq r$, but we do not know this yet. (In the range $\alpha \le 1/L\beta$ we have proved it in Chapter 4, but the results of that chapter do not hold in the entire admissible region (10.1).) Before proving that $\nu_t'(f)$ is very small, we must be able to prove that it is not very large. This is not obvious: the coefficient β might be very large. The best hope to offset this large coefficient is that (say, in the term V), we know a priori that $S_{\ell,\ell'} - r$ is typically small. We will show that this is a simple consequence of the localization theorems of Sections 4.3 and 10.3. This is why in the statement of Proposition 10.5.3 (in contrast with Proposition 2.2.2) we have not (yet) used the symmetry between the values of $2 \le k \le M$, i.e. the fact that

$$\nu_t(S_{\ell,\ell'} f) = (M-1)\nu_t(n_M^\ell n_M^{\ell'} f) .$$

This fact will be used only at a later stage.

A first technical obstacle to use the smallness of $S_{\ell,\ell'} - r$ is that we know it for ν, not for ν_t. The following lemma takes care of this.

Lemma 10.5.4. *Consider a function f on Σ_N^n, with $0 \le f \le 1$. Assume that for some number K' independent of N we have*

$$\nu(f) \le K' \exp\left(-\frac{N}{K'}\right) .$$

Then for some number K'' depending only on K' and β and every $t \le 1$ we have

$$\nu_t(f) \le K'' \exp\left(-\frac{N}{K''}\right) .$$

In words, if a function is negligible for ν, then it is negligible for ν_t uniformly on t.

Proof. We will obtain a differential inequality through Proposition 10.5.3. In order to deduce from Proposition 10.5.3 that $\nu_t'(f) \le Ln^2\beta^2\nu(f)$ we would need to have

$$|S_{\ell,\ell'}| \le L , \qquad |r| \le L .$$

This is not quite true, but (4.150), used for $N-1$ rather than N shows that with overwhelming probability

$$A := \max_{\boldsymbol{\rho}} \sum_{k \le M} n_k^2(\boldsymbol{\rho}) \le L ,$$

and the Cauchy-Schwarz inequality implies $|S_{\ell,\ell'}| \le A$. We observe that r stays bounded in the admissible region (10.1) , and thus

$$|\langle f(S_{\ell,\ell'} - r)\varepsilon_\ell\varepsilon_{\ell'}\rangle_t| \le \langle f|S_{\ell,\ell'} - r|\rangle_t \le L(1+A)\langle f\rangle_t$$

and, since $A \le M$

$$\mathsf{E}A\langle|f|\rangle_t \le L\nu_t(f) + \mathsf{E}(\mathbf{1}_{\{A\ge L\}}M) \le L\nu_t(f) + K\exp\left(-\frac{N}{K}\right) .$$

In this manner we get

$$\nu_t'(f) \le K\nu_t(f) + K\exp\left(-\frac{N}{K}\right)$$

where K depends only on n and β. Integration (Lemma A.11.1) yields the result. □

Now let us examine precisely which kind of information we may deduce from the localization theorems. From this point on, we **always assume that** (α, β, h) **belongs to the admissible region (10.1)** of Section 10.2 **and that the constant L_0 determining this region has been chosen large enough.**

Proposition 10.5.5. *The sets*

$$\{|S_{1,2} - r| \geq s^2\} \; ; \quad \{|S_{1,1} - r| \geq s^2\} \; ; \quad \{|n_1 - \mu| \geq s\} \tag{10.86}$$

are negligible, where

$$s = L \min\left(\frac{m^*}{\beta^5}, \frac{m^*}{\sqrt{L_0}}\right) . \tag{10.87}$$

We should note that as $\beta \to 1$ then s becomes small. This is crucial for the use of the cavity method when β is close to 1. Next we note that we may assume that L_0 is so large that $s \leq m^*/L_0 \leq 1$, so $s^2 \leq s$. The reader certainly noticed that the definition of the third set in (10.86) involved s rather than the smaller number s^2. It does not seem possible to do better. As the reader will gradually come to realize, this is the very point that makes the proofs delicate.

Proof. We first prove that the sets (10.86) are negligible when $s = Lm^*/\sqrt{L_0}$. Using (10.2) and Theorem 4.4.4 with $\rho_0 = Lm^*/\sqrt{L_0}$ (and assuming L_0 large enough to have $\rho_0 \leq m^*/2$), we first observe that the set

$$A = \left\{\boldsymbol{\sigma} \; ; \; \|\mathbf{m}(\boldsymbol{\sigma}) - m^*\mathbf{e}_1\| \geq \frac{Lm^*}{\sqrt{L_0}}\right\}$$

is negligible. Let $\mathbf{n}(\boldsymbol{\sigma}) = (n_k(\boldsymbol{\sigma}))_{k\geq 1}$. Since $|m_k(\boldsymbol{\sigma}) - n_k(\boldsymbol{\sigma})| \leq 1/N$, we have $\|\mathbf{m}(\boldsymbol{\sigma}) - \mathbf{n}(\boldsymbol{\sigma})\| \leq \sqrt{M}/N$, so that the set

$$B = \left\{\boldsymbol{\sigma} \; ; \; \|\mathbf{n}(\boldsymbol{\sigma}) - m^*\mathbf{e}_1\| \geq \frac{2Lm^*}{\sqrt{L_0}}\right\}$$

is negligible. On the complement of B we have

$$|n_1 - m^*| \leq \frac{2Lm^*}{\sqrt{L_0}} \; ; \qquad \sum_{2\leq k\leq M} (n_k)^2 \leq \left(\frac{2Lm^*}{\sqrt{L_0}}\right)^2 = \frac{L'm^{*2}}{L_0} .$$

The second relation above and the Cauchy-Schwarz inequality imply that if $\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2 \notin B$, then

$$|S_{1,2}| = \left|\sum_{2\leq k\leq M} n_k^1 n_k^2\right| \leq \frac{Lm^{*2}}{L_0} .$$

Now, it follows from (10.14) that $r \leq Lm^{*2}/L_0$ and from (10.11) that $|\mu - m^*| \leq Lm^*/L_0$, so, for $\boldsymbol{\sigma} \notin B$ it holds

$$|n_1 - \mu| \leq \frac{Lm^*}{\sqrt{L_0}} , \qquad |S_{1,1} - r| \leq \frac{Lm^{*2}}{L_0}$$

and for $\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2 \notin B$ we have $|S_{1,2} - r| \leq Lm^{*2}/L_0$.

This proves that the sets of (10.86) are negligible when $s = Lm^*/\sqrt{L_0}$, and we now prove that this is also the case when $s = Lm^*/\beta^5$. We only need to consider the case $\beta \geq 2$ because when $\beta \leq 2$ then $Lm^*/\sqrt{L_0} \leq Lm^* \leq L2^5 m^*/\beta^5$. When $\beta \geq 2$, m^* stays bounded below, and it suffices to prove that we can take $s = L/\beta^5$. It follows from Theorem 10.3.1 that (arguing as before to replace $\mathbf{m}(\boldsymbol{\sigma})$ by $\mathbf{n}(\boldsymbol{\sigma})$) the set

$$A = \left\{\boldsymbol{\sigma}\ ;\ \|\mathbf{n}(\boldsymbol{\sigma}) - \mathbf{c}\| \geq \frac{L}{\beta^{10}}\right\}$$

is negligible, where $c_1 = m^*$ and $c_k = m^* N^{-1} \sum_{i \leq N} \eta_{i,k}$ for $2 \leq k \leq M$. On the complement of A we have

$$|n_1 - m^*| \leq \frac{L}{\beta^{10}}$$

so that using (10.11) we get $|n_1 - \mu| \leq L/\beta^{10}$.

On the complement of A, we also have

$$\sum_{2 \leq k \leq M} (n_k - c_k)^2 \leq \frac{L}{\beta^{20}}\ , \tag{10.88}$$

so that for $\boldsymbol{\sigma}^\ell, \boldsymbol{\sigma}^{\ell'} \notin A$ we have

$$\begin{aligned} \left| S_{\ell,\ell'} - \sum_{2 \leq k \leq M} c_k^2 \right| &= \left| \sum_{2 \leq k \leq M} (n_k^\ell n_k^{\ell'} - c_k^2) \right| \\ &= \left| \sum_{2 \leq k \leq M} (n_k^\ell - c_k) n_k^{\ell'} + c_k (n_k^{\ell'} - c_k) \right| \\ &\leq 2 \Big(\sum_{2 \leq k \leq M} (n_k^\ell - c_k)^2 \Big)^{1/2} \Big(\sum_{2 \leq k \leq M} (n_k^\ell)^2 \Big)^{1/2} \\ &\quad + 2 \Big(\sum_{2 \leq k \leq M} c_k^2 \Big)^{1/2} \Big(\sum_{2 \leq k \leq M} (n_k^{\ell'} - c_k)^2 \Big)^{1/2} \\ &\leq \frac{L}{\beta^{10}} \left(\Big(\sum_{2 \leq k \leq M} (n_k^\ell)^2 \Big)^{1/2} + \Big(\sum_{2 \leq k \leq M} c_k^2 \Big)^{1/2} \right) . \end{aligned} \tag{10.89}$$

Recalling that $\|\boldsymbol{\theta}\|^2 = \sum_{2 \leq k \leq M} c_k^2$, we see from (10.19) that with overwhelming probability,

$$\left| \sum_{2 \leq k \leq M} c_k^2 - \alpha m^{*2} \right| \leq \frac{L}{\beta^{10}}\ .$$

When this occurs, we also have $\sum_{2 \leq k \leq M} (n_k^\ell)^2 \leq L$ by (10.88), and (10.89) implies

$$|S_{\ell,\ell'} - \alpha m^{*2}| \leq \frac{L}{\beta^{10}}.$$

Thus by (10.13) it holds that $|S_{\ell,\ell'} - r| \le L/\beta^{10}$. □

We have already gathered the tools to obtain useful bounds on $\nu_t'(f)$, the key to the next result.

Proposition 10.5.6. *Consider a function f on n replicas, $|f| \le N$, that does not depend on the r.v.s η_k. Then*

$$\nu_t(|f|) \le L^{n^2} \nu(|f|) + K \exp\left(-\frac{N}{K}\right) \tag{10.90}$$

$$|\nu_t(f) - \nu_0(f)| \le L^{n^2} \beta^2 s \nu(|f|) + K \exp\left(-\frac{N}{K}\right) , \tag{10.91}$$

where K depends only on β and h and s is as in (10.87).

Inequality (10.90) will be used constantly below. The reason for the condition $|f| \le N$ is that we want to use this result for functions f such as $S_{1,2} - r$ that can in principle be as large as (about) N.

Proof. In this proof, K denotes a number depending only on β and h. It follows from Proposition 10.5.5 and Lemma 10.5.4 that we have

$$\nu_t(\mathbf{1}_{\{|S_{\ell,\ell'} - r| \ge s^2\}}) \le K \exp\left(-\frac{N}{K}\right) \;\; ; \quad \nu_t(\mathbf{1}_{\{|n^\ell - \mu| \ge s\}}) \le K \exp\left(-\frac{N}{K}\right) .$$

Using the bounds $|f| \le N$, $|S_{\ell,\ell'}| \le M$, and writing (10.80) as

$$\mathrm{IV} = \frac{\beta^2}{2} \left(\sum_{\ell \le N} \nu_t(f(S_{\ell,\ell} - r)) - n\nu_t(f(S_{n+1,n+1} - r)) \right) ,$$

we get from Proposition 10.5.3 that

$$|\nu_t'(f)| \le Ln^2 \beta^2 s \nu_t(|f|) + |\mathcal{R}| + K \exp\left(-\frac{N}{K}\right) , \tag{10.92}$$

where $\mathcal{R}$ is as in (10.82).

Next, we will prove that

$$|\mathcal{R}| \le n^2 \beta^2 s \nu_t(|f|) + K \exp\left(-\frac{N}{K}\right) . \tag{10.93}$$

Combining with (10.92) this yields

$$|\nu_t'(f)| \le Ln^2 \beta^2 s \nu_t(|f|) + K \exp\left(-\frac{N}{K}\right) . \tag{10.94}$$

In particular, since $\beta^2 s \le L$, we have

$$|\nu_t'(f)| \le Ln^2\nu_t(|f|) + K\exp\left(-\frac{N}{K}\right) .$$

Integration using Lemma A.11.1 yields (10.90), and combining (10.90) with (10.94) yields (10.91).

We turn to the proof of (10.93). We prove that in fact for any positive constant a, it holds

$$|\mathcal{R}| \le La^2\nu_t(|f|) + K\exp\left(-\frac{N}{K}\right) . \tag{10.95}$$

It follows from Theorem 10.4.1 that $\nu(\mathbf{1}_{\{|n_k|\ge a\}}) \le K\exp(-N/K)$, so by Lemma 10.5.4 we also have $\nu_t(\mathbf{1}_{\{|n_k|\ge a\}}) \le K\exp(-N/K)$. Since $|f|\le N$ we get

$$\nu_t(|n_k^\ell|^4|f|) \le a^2\nu_t(|n_k^\ell|^2|f|) + K\exp\left(-\frac{N}{K}\right) .$$

Finally since $\sum_{k\le M}|n_k^\ell|^2 \le L$, with overwhelming probability,

$$\nu_t(\sum_{k\le M}|n_k^\ell|^2|f|) \le L\nu_t(|f|) + K\exp\left(-\frac{N}{K}\right) . \qquad \square$$

Now we must break the bad news. It will not suffice to use the inequality

$$|\nu(f)-\nu_0(f)| \le \sup_t|\nu_t'(f)| . \tag{10.96}$$

This inequality is an order 1 estimate, and even to study the system of replica-symmetric equations and prove Theorem 10.2.1 we had to make an order 2 expansion. Instead of (10.96) we will need something like

$$|\nu(f)-\nu_0(f)-\nu_0'(f)| \le \sup_t|\nu_t''(f)| . \tag{10.97}$$

Fortunately (Gauss forbids!), we will not really need to care about the exact expression of $\nu_t''(f)$.

We define

$$\rho = \alpha\frac{1-\beta(1-q)^2}{(1-\beta(1-q))^2} \tag{10.98}$$

and we write (10.80) as

$$\mathrm{IV} = \frac{\beta^2}{2}\left(\sum_{\ell\le n}\nu_t(f(S_{\ell,\ell}-\rho)) - n\nu_t(f(S_{n+1,n+1}-\rho))\right) , \tag{10.99}$$

in the hope that $S_{\ell,\ell}\simeq\rho$. Of course it takes some foresight (Theorem 10.7.1 below) to guess the value of ρ.

We now state the precise form we need for (10.97).

Theorem 10.5.7. *Given numbers $\tau_1, \tau_2 \geq 1$ with $1/\tau_1 + 1/\tau_2 = 1$ and a function f on Σ_N^n with $|f| \leq 1$ (that does not depend on the r.v.s η_k) we have*

$$\nu(f) = \nu_0(f) + \mathrm{VI} + \mathrm{VII} + \mathrm{VIII} + \mathcal{R} \tag{10.100}$$

where

$$|\mathcal{R}| \leq B(f) := L^{n^2} \beta^4 \nu(|f|^{\tau_1})^{1/\tau_1} \Big(\nu(|S_{1,1} - \rho|^{2\tau_2})^{1/\tau_2} + \nu(|S_{1,2} - r|^{2\tau_2})^{1/\tau_2}$$
$$+ \nu(|n_1 - \mu|^{2\tau_2})^{1/\tau_2} + \frac{K}{N}\Big) + K \exp\left(-\frac{N}{K}\right) \tag{10.101}$$

and

$$\mathrm{VI} = \beta \Big(\sum_{\ell \leq n} \nu_0(f\varepsilon_\ell(n_1^\ell - \mu)) - n\nu_0(f\varepsilon_{n+1}(n_1^{n+1} - \mu)) \Big) \tag{10.102}$$

$$\mathrm{VII} = \frac{\beta^2}{2} \Big(\sum_{\ell \leq n} \nu_0(f(S_{\ell,\ell} - \rho)) - n\nu_0(f(S_{n+1,n+1} - \rho)) \Big) \tag{10.103}$$

$$\mathrm{VIII} = \beta^2 \Big(\sum_{1 \leq \ell < \ell' \leq n} \nu_0(f\varepsilon_\ell\varepsilon_{\ell'}(S_{\ell,\ell'} - r)) - n \sum_{\ell \leq n} \nu_0(f\varepsilon_\ell\varepsilon_{n+1}(S_{\ell,n+1} - r))$$
$$+ \frac{n(n+1)}{2} \nu_0(f\varepsilon_{n+1}\varepsilon_{n+2}(S_{n+1,n+2} - r)) \Big) . \tag{10.104}$$

Proof. Let us write (10.79) as $\nu_t'(f) = A(t) + \mathcal{R}(t)$, where $A(t)$ is the sum of the terms (10.73), (10.80) and (10.81). Let us observe that $A(0)$ is the sum of the terms (10.102) to (10.104). We have

$$\nu(f) = \nu_0(f) + \int_0^1 \nu_t'(f) \mathrm{d}t$$
$$= \nu_0(f) + A(0) + \int_0^1 (A(t) - A(0) + \mathcal{R}(t)) \mathrm{d}t .$$

We shall prove that $|A'(t)| \leq B(f)$ and $|\mathcal{R}(t)| \leq B(f)$ which will imply (10.100) and (10.101). First we study $\mathcal{R}(t)$. We simply use Hölder's inequality and (10.90) to get

$$\nu_t(|n_k^\ell|^4 |f|) \leq \nu_t(|n_k|^{4\tau_2})^{1/\tau_2} \nu_t(|f|^{\tau_1})^{1/\tau_1}$$
$$\leq K\nu(|n_k|^{4\tau_2})^{1/\tau_2} \nu(|f|^{\tau_1})^{1/\tau_1} + K \exp\left(-\frac{N}{K}\right) .$$

It follows from Theorem 10.4.1 that $\nu(|n_k|^{4\tau_2})^{1/\tau_2} \leq K/N^2$ so the previous inequality yields (since $M \leq N$)

$$|\mathcal{R}(t)| \leq \frac{K}{N} \nu(|f|^{\tau_1})^{1/\tau_1} + K \exp\left(-\frac{N}{K}\right) .$$

We compute $A'(t)$ through Proposition 10.5.3 (writing the term IV as in (10.99)). There is a remainder term which is bounded as previously. The other terms, besides f, feature *two* factors of the type $(S_{\ell,\ell} - \rho)$ or $(S_{\ell,\ell'} - r)$ or $(n_1^\ell - \mu)$; we simply use that $2|U_1 U_2| \le U_1^2 + U_2^2$ and Hölder's inequality to obtain

$$2\nu_t(|fU_1U_2|) \le \nu_t(|f|^{\tau_1})^{1/\tau_1}(\nu_t(|U_1|^{2\tau_2})^{1/\tau_2} + \nu_t(|U_2|^{2\tau_2})^{1/\tau_2}) ,$$

and (10.90) to conclude. □

Sometimes we shall only need the following simpler fact.

Proposition 10.5.8. *If f is as above and s is as in (10.87), we have*

$$\nu(f) = \nu_0(f) + \mathrm{VI} + \mathcal{R}' \tag{10.105}$$

where

$$|\mathcal{R}'| \le L^{n^2}\beta^4 s^2 \nu(|f|) + K\exp\left(-\frac{N}{K}\right) .$$

This should be compared with (10.91). Use of one more step in the expansion provided us with a factor s^2 rather than s in the error term, and this will be essential at some later stage.

Proof. This is a simple variation on the previous argument. We note that, using first (10.90) and then Proposition 10.5.5 we have

$$|\nu_0(f\varepsilon_\ell\varepsilon_{\ell'}(S_{\ell,\ell'} - r))| \le L^{n^2}\nu(|f||S_{\ell,\ell'} - r|) \le L^{n^2}s^2\nu(|f|) + K\exp\left(-\frac{N}{K}\right) ,$$

and proceeding in this manner we get

$$|\mathrm{VII} + \mathrm{VIII}| \le L^{n^2}\beta^2 s^2\nu(|f|) + K\exp\left(-\frac{N}{K}\right) .$$

Next, we bound the terms $\nu_t(fU_1U_2)$ in $A'(t)$ where U_1 and U_2 are both of the type $(S_{\ell,\ell} - \rho)$ or $(S_{\ell,\ell'} - r)$ or $(n_1^\ell - \mu)$, by using again (10.90) and Proposition 10.5.5 to get

$$\nu_t(|fU_1U_2|) \le L^{n^2}s^2\nu(|f|) + K\exp\left(-\frac{N}{K}\right) ,$$

with the factor s^2 occurring since each of the terms U_1 and U_2 gives us a factor s (or $s^2 \le s$). Thus we get

$$|A'(t)| \le L^{n^2}\beta^4 s^2\nu(|f|) + K\exp\left(-\frac{N}{K}\right) .$$

To conclude the proof, we simply use (10.95) to bound $|\mathcal{R}(t)|$. □

10.6 Integration by Parts

In the previous section we have seen how quantities such as $\nu(S_{1,1}f)$ or $\nu(S_{1,2}f)$ occur naturally. We expect that $S_{1,1} \simeq \rho$ and $S_{1,2} \simeq r$. We cannot prove it yet, but we may try to calculate $\nu(S_{1,2}f)$ and $\nu(S_{1,1}f)$. We note that

$$\nu(S_{1,1}f) = \sum_{2\le k\le M} \nu((n_k^1)^2 f) \ ; \qquad \nu(S_{1,2}f) = \sum_{2\le k\le M} \nu(n_k^1 n_k^2 f) \ . \tag{10.106}$$

We will try to compute $\nu((n_M^1)^2 f)$ and $\nu(n_M^1 n_M^2 f)$. Since n_k is of order $1/\sqrt{N}$, it is more natural to consider instead the quantities $N\nu((n_M^1)^2 f)$ and $N\nu(n_M^1 n_M^2 f)$. We define

$$D = 1 - \beta(1-q) \tag{10.107}$$

$$\widehat{r} = \frac{q}{D^2}\ \Big(= \frac{r}{\alpha}\Big) \ \ ; \ \widehat{\rho} = \frac{1-\beta(1-q)^2}{D}\ \Big(= \frac{\rho}{\alpha}\Big) \ . \tag{10.108}$$

We remind the reader that $R_{\ell,\ell'} = N^{-1}\sum_{i\le N} \sigma_i^\ell \sigma_i^{\ell'}$.

Theorem 10.6.1. *Consider a function f on Σ_N^n. This function may be random, but it must not depend on the r.v.s $(\eta_{i,M})_{i\le N}$. Let us define*

$$V_1 = \beta N\Big(\sum_{2\le \ell\le n+1} \nu((R_{1,\ell}-q) n_M^1 n_M^\ell f) - (n+1)\nu((R_{1,n+2}-q)n_M^1 n_M^{n+2} f)\Big) \tag{10.109}$$

and, for $2\le p\le n+1$ let us define

$$V_p = \nu((R_{1,p}-q)f) + \beta N\Big(\sum_{\ell\ne p,\ell\le n+1} \nu((R_{p,\ell}-q) n_M^1 n_M^\ell f) - (n+1)\nu((R_{1,n+2}-q)n_M^1 n_M^{n+2} f)\Big) \ . \tag{10.110}$$

Consider a number $\tau > 0$. Then we have

$$N\nu((n_M^1)^2 f) = \widehat{\rho}\nu(f) + \frac{V_1}{D} + \frac{\beta q}{D^2}\Big(\Big(\sum_{p\le n} V_p\Big) - nV_{n+1}\Big) + \mathcal{R}_1 \tag{10.111}$$

$$N\nu(n_M^1 n_M^2 f) = \widehat{r}\nu(f) + \frac{V_2}{D} + \frac{\beta q}{D^2}\Big(\Big(\sum_{p\le n} V_p\Big) - nV_{n+1}\Big) + \mathcal{R}_2 \tag{10.112}$$

where

$$|\mathcal{R}_1|, |\mathcal{R}_2| \le \frac{K}{N}\nu(|f|^\tau)^{1/\tau} \ . \tag{10.113}$$

It is of real importance to allow f to be random. For example, if we want to compute $\nu(S_{1,2}^2)$, we write

$$\nu(S_{1,2}^2) = (M-1)\nu(n_M^1 n_M^2 S_{1,2}) \simeq (M-1)\nu(n_M^1 n_M^2 S_{1,2}^{\sim})$$

where $S_{1,2}^{\sim} = \sum_{2\le k\le M-1} n_k^1 n_k^2$ is random but does not depend on the r.v.s $\eta_{i,M}$.

We should think of Theorem 10.6.1 as a kind of expansion. Besides $\nu(f)$, the terms in the right-hand sides of (10.111) and (10.112) all have a factor $R_{1,\ell}-q$, that should be small. Some of the terms contain the products $n_M^1 n_M^\ell$, and the process can be iterated. Throughout the argument, the integer p will represent a replica index. We use the notation

$$A_p = N\nu(n_M^1 n_M^p f) \; . \tag{10.114}$$

We consider these quantities for $p = 1, \ldots, n+2$. Since f is a function on Σ_N^n, we have

$$A_{n+2} = N\nu(n_M^1 n_M^{n+2} f) = N\nu(n_M^1 n_M^{n+1} f) = A_{n+1} \; . \tag{10.115}$$

Lemma 10.6.2. *We have, for each $p \le n+1$*

$$A_p = \nu(R_{1,p} f) + N\beta\Big(\sum_{\ell\le n+1} \nu(R_{p,\ell} n_M^1 n_M^\ell f) - (n+1)\nu(R_{p,n+2} n_M^1 n_M^{n+2} f)\Big) + \mathcal{R}_p \; , \tag{10.116}$$

where $|\mathcal{R}_p|$ is bounded as in (10.113).

Proof. The basic idea of the entire approach is to write

$$N n_M^p = \sum_{i\le N-1} \eta_{i,M} \sigma_i^p$$

so that

$$A_p = \sum_{i\le N-1} \mathsf{E}(\eta_{i,M} \langle n_M^1 \sigma_i^p f\rangle) \; , \tag{10.117}$$

and to try to use approximate integration by parts. We have to make explicit the dependence of $\langle\cdot\rangle$ on $\eta_{i,M}$. For this we recall the bracket $\langle\cdot\rangle_\sim$ of Section 10.4 so that, since $p \le n+1$,

$$\mathsf{E}(\eta_{i,M}\langle n_M^1 \sigma_i^p f\rangle) = \mathsf{E}\left(\eta_{i,M} \frac{\langle n_M^1 \sigma_i^p f \exp \frac{\beta N}{2} \sum_{\ell\le n+1} (m_M^\ell)^2\rangle_\sim}{\big\langle \exp \frac{\beta N}{2} (m_M^1)^2 \big\rangle_\sim^{n+1}}\right) \; .$$

Abusing notation in an obvious manner, we have

$$\forall \ell \le n+1\ ,\ \frac{\partial m_M^\ell}{\partial \eta_{i,M}} = \frac{\sigma_i^\ell}{N} \quad ; \quad \frac{\partial n_M^1}{\partial \eta_{i,M}} = \frac{\sigma_i^1}{N}$$

so that the main term obtained from approximate integration by parts is

$$\frac{1}{N}\nu(\sigma_i^1\sigma_i^p f) + \beta \sum_{\ell \le n+1} \nu(n_M^1 m_M^\ell \sigma_i^p \sigma_i^\ell f) \\ - (n+1)\beta\nu(n_M^1 m_M^{n+2}\sigma_i^p\sigma_i^{n+2} f)\ .$$

Here we use in an essential way that f does not depend on the variables $\eta_{i,M}$. Proceeding as in the proof of Proposition 10.5.3, (that is, using differential inequalities) one finds that the error term $\mathcal{R}_i$ in the approximate integration by parts satisfies

$$|\mathcal{R}_i| \le K\nu\Bigg(|n_M^1||f| \sum_{\ell \le n+2} |m_M^\ell|^3\Bigg)\ .$$

Using Theorem 10.4.1, we get

$$|\mathcal{R}_i| \le \frac{K}{N^2}\nu(|f|^\tau)^{1/\tau}\ ,$$

the constant K being independent of i. Thus we have obtained the relation

$$\mathsf{E}(\eta_{i,M}\langle n_M^1 \sigma_i^p f\rangle) = \frac{1}{N}\nu(\sigma_i^1\sigma_i^p f) + \beta \sum_{\ell \le n+1} \nu(n_M^1 m_M^\ell \sigma_i^p \sigma_i^\ell f) \\ - (n+1)\beta\nu(n_M^1 m_M^{n+2}\sigma_i^p\sigma_i^{n+2} f) + \mathcal{R}_i\ . \quad (10.118)$$

We recall the notation $R^-_{\ell,\ell'} = N^{-1}\sum_{i \le N-1}\sigma_i^\ell\sigma_i^{\ell'}$, so summation of (10.118) over $i \le N-1$ yields

$$A_p = \nu(R^-_{1,p}f) + N\beta \sum_{\ell \le n+1} \nu(n_M^1 m_M^\ell R^-_{p,\ell} f) \\ - (n+1)N\beta(n_M^1 m_M^{n+2} R^-_{p,n+2} f) + \mathcal{R}\ , \quad (10.119)$$

where $|\mathcal{R}| \le K\nu(|f|^\tau)^{1/\tau}/N$. This relation is not exactly (10.116) because it involves m_M^ℓ rather than n_M^ℓ and $R^-_{p,\ell}$ rather than $R_{p,\ell}$. It is rather obvious that replacing $R^-_{p,\ell}$ by $R_{p,\ell}$ creates an error term $\le K\nu(|f|)/N \le K\nu(|f|^\tau)^{1/\tau}/N$. Replacing then m_M^ℓ by n_M^ℓ creates an error term

$$\mathcal{R}' = \beta \sum_{\ell \le n+1} \nu(n_M^1 \eta_M \varepsilon_\ell R_{p,\ell} f) - (n+1)\beta\nu(n_M^1\eta_M\varepsilon_{n+2}R_{p,n+2}f)\ .$$

Interestingly, it is not immediately obvious that $|\mathcal{R}'| \le K\nu(|f|^\tau)^{1/\tau}/N$; but this becomes apparent if one performs approximate integration by parts in the r.v. η_M. □

Proof of Theorem 10.6.1. The remainder of this proof is elementary linear algebra. When $p = \ell$, we will use that $R_{p,\ell} = 1$; when $p \neq \ell$ we will bring out the term $R_{p,\ell} - q$. We first consider (10.116) when $p = 1$. It reads

$$A_1 = \nu(R_{1,1} f) + N\beta\nu(R_{1,1}(n_M^1)^2 f) + N\beta\Big(\sum_{2\le\ell\le n+1} \nu(R_{1,\ell} n_M^1 n_M^\ell f) - (n+1)\nu(R_{1,n+2} n_M^1 n_M^{n+2} f)\Big) + \mathcal{R}_1 \ . \tag{10.120}$$

Since $R_{1,1} = 1$ and $N\nu((n_M^1)^2 f) = A_1$, we write (10.120) as

$$A_1 = \nu(f) + \beta A_1 + N\beta\Big(\sum_{2\le\ell\le n+1} \nu(R_{1,\ell} n_M^1 n_M^\ell f) - (n+1)\nu(R_{1,n+2} n_M^1 n_M^{n+2} f)\Big) + \mathcal{R}_1 \ . \tag{10.121}$$

Now we write

$$N\nu(R_{1,\ell} n_M^1 n_M^\ell f) = qA_\ell + N\nu((R_{1,\ell} - q) n_M^1 n_M^\ell f) \ ,$$

and substitution in (10.121) yields

$$A_1 = \nu(f) + \beta A_1 + \beta q\Big(\sum_{2\le\ell\le n+1} A_\ell - (n+1)A_{n+2}\Big) + V_1 + \mathcal{R}_1 \tag{10.122}$$

where V_1 is as in (10.109), i.e.

$$V_1 = \beta N\Big(\sum_{2\le\ell\le n+1} \nu((R_{1,\ell} - q) n_M^1 n_M^\ell f) - (n+1)\nu((R_{1,n+2} - q) n_M^1 n_M^{n+2} f)\Big) \ .$$

Using that $A_{n+2} = A_{n+1}$ by (10.115), we rewrite (10.122) as

$$DA_1 = \nu(f) + \beta q\Big(\sum_{1\le\ell\le n} A_\ell - nA_{n+1}\Big) + V_1 + \mathcal{R}_1 \ . \tag{10.123}$$

Next, we consider the case $1 < p \le n$. Using again (10.116) we have

$$A_p = \nu(fR_{1,p}) + \beta A_p + N\beta\Big(\sum_{1\le\ell\le n+1, \ell\neq p} \nu(R_{p,\ell} n_M^1 n_M^\ell f) - (n+1)\nu(R_{p,n+2} n_M^1 n_M^{n+2} f)\Big) + \mathcal{R}_p$$

and, as before we get

$$DA_p = q\nu(f) + \beta q\Big(\sum_{1\le\ell\le n} A_\ell - nA_{n+1}\Big) + V_p + \mathcal{R}_p \ , \tag{10.124}$$

where now V_p is as in (10.110), i.e.

$$V_p = \nu(f(R_{1,p} - q)) + \beta N\biggl(\sum_{\ell\leq n+1,\ell\neq p} \nu((R_{p,\ell} - q)n_M^p n_M^\ell f) - (n+1)\nu((R_{p,n+2} - q)n_M^p n_M^{n+2} f)\biggr) .$$

Finally, we consider the case $p = n + 1$, where (10.116) now becomes

$$A_{n+1} = \nu(fR_{1,n+1}) + N\beta\nu(n_M^1 n_M^{n+1} f) + N\beta\biggl(\sum_{1\leq\ell\leq n} \nu(R_{1,\ell} n_M^1 n_M^\ell f) - (n+1)\nu(R_{1,n+2} n_M^1 n_M^{n+2} f)\biggr) + \mathcal{R}_{n+1} .$$

Using again that $A_{n+2} = A_{n+1}$by (10.115) we get

$$A_{n+1} = q\nu(f)+\beta A_{n+1}+\beta q\biggl(\sum_{1\leq\ell\leq n} A_\ell-(n+1)A_{n+1}\biggr)+V_{n+1}+\mathcal{R}_{n+1} \quad (10.125)$$

where V_{n+1} is as in (10.110) i.e.

$$V_{n+1} = \nu(f(R_{1,n+1} - q)) + N\beta\biggl(\sum_{\ell\leq n} \nu((R_{\ell,n+1} - q)n_M^\ell n_M^{n+1} f) - (n+1)\nu((R_{n+1,n+2} - q)n_M^{n+1} n_M^{n+2} f)\biggr) ,$$

and this gives again

$$DA_{n+1} = q\nu(f) + \beta q\biggl(\sum_{1\leq\ell\leq n} A_\ell - nA_{n+1}\biggr) + V_{n+1} + \mathcal{R}_{n+1} . \quad (10.126)$$

Thus, for $1 \leq p \leq n + 1$ we have obtained the equations

$$DA_p = \beta q\biggl(\sum_{1\leq\ell\leq n} A_\ell - nA_{n+1}\biggr) + b_p , \quad (10.127)$$

where $b_1 = \nu(f) + V_1 + \mathcal{R}_1$ and $b_p = q\nu(f) + V_p + \mathcal{R}_p$ for $2 \leq p \leq n + 1$. We add the first n equations and subtract n times the $(n + 1)^{\text{th}}$ equation to obtain the relation

$$D\biggl(\sum_{1\leq\ell\leq n} A_\ell - nA_{n+1}\biggr) = \sum_{1\leq\ell\leq n} b_\ell - nb_{n+1}$$

which we substitute in (10.127) to get

$$A_p = \frac{b_p}{D} + \frac{\beta q}{D^2}\Bigl(\sum_{1\le\ell\le n} b_\ell - n b_{n+1}\Bigr) . \tag{10.128}$$

The coefficient of $\nu(f)$ in $\sum_{1\le\ell\le n} b_\ell - n b_{n+1}$ is $1-q$ so the coefficient of $\nu(f)$ in A_1 is

$$\frac{1}{D} + \frac{\beta q(1-q)}{D^2} = 1 - \frac{\beta(1-q)+\beta q(1-q)}{D^2} = \frac{1-\beta(1-q)^2}{D^2} = \widehat{\rho}$$

and the coefficient of $\nu(f)$ in A_p for $p \ge 2$ is

$$\frac{q}{D} + \frac{\beta q(1-q)}{D^2} = \frac{q}{D^2} = \widehat{r} .$$

Therefore (10.128) proves (10.111) and (10.112). □

10.7 The Replica-Symmetric Solution

Throughout this section we assume that (α, β, h) belongs to the admissible region (10.1) (so that $h > 0$), and that the constant L_0 which defines (10.1) has been chosen large enough. We recall that (q, μ, r) is the (unique) solution of the replica-symmetric equations (10.3) to (10.4) that satisfies (10.6), while ρ is given by (10.98). These quantities depend on β, h and on N and M through the ratio $\alpha = M/N$. The goal of this section is to prove the following, which is one of the main results of the chapter.

Theorem 10.7.1. *If (α, β, h) belongs to the admissible region (10.1), and the constant L_0 which defines this region has been chosen large enough, then*

$$Q_N := \nu\bigl((R_{1,2} - q)^2\bigr) \le \frac{K}{N} \tag{10.129}$$

$$R_N := \nu\bigl((n_1 - \mu)^2\bigr) \le \frac{K}{N} \tag{10.130}$$

$$S_N := \nu\bigl((S_{1,2} - r)^2\bigr) \le \frac{K}{N} \tag{10.131}$$

$$T_N := \nu\bigl((S_{1,1} - \rho)^2\bigr) \le \frac{K}{N} , \tag{10.132}$$

where K depends on β and h only.

The quantities $Q_N, \ldots$ depend also on M but we keep this dependence implicit in order to lighten the notation. Of course, it follows from (10.130) that $\nu((m^1 - \mu)^2) \le K/N$.

When $h = 0$, the Bovier-Gayrard localization theorem shows that (typically) exactly one of the quantities m_k, $k \le M$, is close to m^*, but by symmetry between the values of $k \le M$ it cannot always be the case that m_1 is close to m^*, and we cannot expect (10.130) to hold in that case.

Research Problem 10.7.2. (Level 2) Is it true that in the admissible region (10.1) the estimates (10.129), (10.131) and (10.132) hold with a constant K that depends on β only (and not on h)?

Throughout this section we denote by K a number depending only on β and h, that need not be the same at each occurrence.

Proposition 10.7.3. *The following estimates hold:*

$$Q_N \le Lm^* Q_N^{1/2} R_N^{1/2} + \frac{L}{\beta^2} Q_N^{1/2}(S_N^{1/2} + T_N^{1/2}) + \frac{K}{N} \tag{10.133}$$

$$S_N \le \frac{L\beta^4}{L_0^2} Q_N + \frac{K}{N} \tag{10.134}$$

$$T_N \le \frac{L\beta^4}{L_0^2} Q_N + \frac{K}{N} \tag{10.135}$$

$$R_N \le \left(1 - \frac{m^{*2}}{L}\right) R_N + \frac{Lm^*}{\beta^2} R_N^{1/2}(S_N^{1/2} + T_N^{1/2}) + \frac{K}{N} . \tag{10.136}$$

Some of the coefficients above, such as the powers 2 for β in the denominators are somewhat arbitrary. The coefficients β^4 in (10.134) and (10.135) look threatening when β is large, but in fact there is plenty of room then. It is for β close to 1 that one has to struggle.

Proof of Theorem 10.7.1. From (10.136) we get

$$R_N \le \frac{L}{m^*\beta^2} R_N^{1/2}(S_N^{1/2} + T_N^{1/2}) + \frac{K}{N} . \tag{10.137}$$

The reader should observe the nasty small factor m^* in the denominator. (Recall that $m^* \to 0$ as $\beta \to 1$.)

We use the relation

$$ab \le \frac{t}{2}a^2 + \frac{1}{2t}b^2 \tag{10.138}$$

for $a = R_N^{1/2}$, $b = (S_N^{1/2} + T_N^{1/2})$ and $t = m^*\beta^2/L$ to obtain

$$R_N \le \frac{1}{2} R_N + \frac{L}{m^{*2}\beta^4}(S_N + T_N) + \frac{K}{N} ,$$

so that

$$R_N \le \frac{L}{m^{*2}\beta^4}(S_N + T_N) + \frac{K}{N} .$$

Combining with (10.134) and (10.135) yields

$$R_N \le \frac{L}{L_0^2 m^{*2}} Q_N + \frac{K}{N} . \tag{10.139}$$

If we substitute (10.139), (10.134) and (10.135) in (10.133) we get

$$Q_N \le \frac{L}{L_0} Q_N + \frac{K}{N} ,$$

so that for L_0 is large enough we obtain $Q_N \le K/N$; then (10.134), (10.135) and (10.139) conclude the proof. □

We observe that the factor m^* in the first term on the right-hand side of (10.133) is exactly what is required to offset the nasty coefficient m^* in the denominator in (10.137). We also observe that the coefficient β^{-2} in the second term of the right-hand side of (10.133) is exactly what is required to offset the coefficient β^4 in (10.134) and (10.135).

The principle of the proof of Proposition 10.7.3 is simple. To bound Q_N, we write as usual

$$Q_N = \nu\big((R_{1,2} - q)^2\big) = \nu((\varepsilon_1\varepsilon_2 - q)f) \tag{10.140}$$

where $f = R_{1,2} - q$. We then use Theorem 10.5.7 with $\tau_1 = \tau_2 = 2$ (and $(\varepsilon_1\varepsilon_2 - q)f$ rather than f). We have

$$\nu_0((\varepsilon_1\varepsilon_2 - q)f) = \nu_0((\varepsilon_1\varepsilon_2 - q)(R_{1,2} - q)) \tag{10.141}$$

$$= \frac{1}{N}\nu_0((\varepsilon_1\varepsilon_2 - q)\varepsilon_1\varepsilon_2) + \nu_0((\varepsilon_1\varepsilon_2 - q)(R^-_{1,2} - q))$$

$$= \frac{1}{N}\nu_0((\varepsilon_1\varepsilon_2 - q)\varepsilon_1\varepsilon_2) \le \frac{2}{N} , \tag{10.142}$$

using Lemma 10.5.1. Using (10.90) and the Cauchy-Schwarz inequality, we see that all the other terms provided by Theorem 10.5.7 are bounded by

$$C(\beta) Q_N^{1/2}(R_N^{1/2} + S_N^{1/2} + T_N^{1/2}) + K\exp(-N/K) .$$

The difficulty is to obtain small enough coefficients in the various terms involved. The way around this difficulty is no magic: keep expanding until your estimates become good enough. This requires patience only. While performing this computation, we will meet the quantities

$$\widehat{\mu} = \mathsf{E}\mathrm{th}^3 Y = \nu_0(\varepsilon_1\varepsilon_2\varepsilon_3) \; ; \quad \widehat{q} = \mathsf{E}\mathrm{th}^4 Y = \nu_0(\varepsilon_1\varepsilon_2\varepsilon_3\varepsilon_4) ,$$

where $Y = \beta(z\sqrt{r} + \mu) + h$.

Lemma 10.7.4. *We have*

$$\mu \le Lm^* \; , \; q \le Lm^{*2} \le Lm^* \tag{10.143}$$

$$\beta^4(1 - q^2) \le L \; ; \quad \beta^4(q - q^2) \le L \; ; \quad \beta^4(\widehat{q} - q^2) \le L \tag{10.144}$$

$$\beta^4|\mu(1-q)| \le Lm^* , \qquad \beta^4|\widehat{\mu} - \mu q| \le Lm^* \tag{10.145}$$

$$|1 - q| \le L\beta^{-10} \; ; \; |1 - \widehat{q}| \le L\beta^{-10} \; ; \; |1 - \mu| \le L\beta^{-10} \; ; \; |1 - \widehat{\mu}| \le L\beta^{-10} . \tag{10.146}$$

Proof. First we note that (10.6) implies (10.143). Since $1 - m^* \le L\beta^{-10}$ by (4.37), it follows from (10.11) and (10.12) that $|1-\mu| \le L\beta^{-10}$ and $|1-q| \le L\beta^{-10}$. Also, since for $-1 < x < 1$ it holds that $1 - x^3 \le 3(1-x)$, we have

$$0 \le 1 - \widehat{\mu} = \mathsf{E}(1 - \mathrm{th}^3 Y) \le 3\mathsf{E}(1 - \mathrm{th} Y) = 3(1-\mu)$$

so that $|1-\widehat{\mu}| \le L\beta^{-10}$ and similarly $|1-\widehat{q}| \le L\beta^{-10}$. This proves (10.144) and (10.145) for $\beta \ge 2$. All that remains to prove is (10.145) for $\beta \le 2$. But this follows from (10.143) since $|\widehat{\mu}| = |\mathsf{E}\mathrm{th}^3 Y| \le \mathsf{E}\mathrm{th}^2 Y = q$. □

Proof of (10.133). Using (10.140), (10.141) and Theorem 10.5.7 for $n = 2$, $f(\varepsilon_1\varepsilon_2 - q)$ rather than f and $\tau_1 = \tau_2 = 1/2$ we get

$$Q_N \le \frac{4}{N} + \mathrm{I} + \mathrm{II} + \mathrm{III} + \mathcal{R} \tag{10.147}$$

where $\mathcal{R}$ is as in (10.101) and

$$\mathrm{I} = \beta\Bigg(\sum_{\ell\le 2} \nu_0\big(\varepsilon_\ell(\varepsilon_1\varepsilon_2 - q)(n_1^\ell - \mu) f\big) - 2\nu_0\big((\varepsilon_1\varepsilon_2 - q)\varepsilon_3(n_1^3 - \mu) f\big)\Bigg) \tag{10.148}$$

$$\mathrm{II} = \beta^2\Bigg(\sum_{\ell\le 2} \nu_0\big((\varepsilon_1\varepsilon_2 - q)(S_{\ell,\ell} - \rho) f\big) - 2\nu_0\big((\varepsilon_1\varepsilon_2 - q)(S_{3,3} - \rho) f\big)\Bigg) \tag{10.149}$$

$$\begin{aligned}\mathrm{III} = \beta^2\Bigg(&\nu_0\big(\varepsilon_1\varepsilon_2(\varepsilon_1\varepsilon_2 - q)(S_{1,2} - r) f\big) - 2\sum_{\ell\le 2}\nu_0\big(\varepsilon_\ell\varepsilon_3(\varepsilon_1\varepsilon_2 - q)(S_{\ell,3} - r) f\big) \\ &+ 3\nu_0\big(\varepsilon_3\varepsilon_4(\varepsilon_1\varepsilon_2 - q)(S_{3,4} - r) f\big)\Bigg)\,.\end{aligned} \tag{10.150}$$

Since here $f = R_{1,2} - q$, we have $\nu(f^2) = Q_N$. Applying the Cauchy-Schwarz inequality in a straightforward manner we get the inequality

$$Q_N \le L\beta^2 Q_N^{1/2}(R_N^{1/2} + Q_N^{1/2} + S_N^{1/2}).$$

This is quite far from (10.133), because the required small coefficients m^* and β^{-2} are not featured above. So, in each of the terms we will try to bring out a "small factor". By a "small factor" here and in the following few pages, we mean a quantity that will create precisely the required coefficients such as m^* or β^{-2}. Generally speaking, Proposition 10.5.5 is a precious tool for creating such small factors.

Fortunately, all terms can be handled by the same method. First we consider the terms occurring in the summation of (10.148). Given $\ell \le 2$ we have

$$\left|\nu_0\big(\varepsilon_\ell(\varepsilon_1\varepsilon_2-q)(n_1^\ell-\mu)f\big)-\nu_0(\varepsilon_\ell(\varepsilon_1\varepsilon_2-q))\nu_0((n_1^\ell-\mu)f)\right|\le\frac{2}{N}\ . \quad (10.151)$$

This is because by Lemma 10.5.1 and since $f^- = R^-_{1,2} - q$ does not depend on the last spin, it holds that

$$\nu_0\big(\varepsilon_\ell(\varepsilon_1\varepsilon_2-q)(n_1^\ell-\mu)f^-\big)=\nu_0(\varepsilon_\ell(\varepsilon_1\varepsilon_2-q))\nu_0((n_1^\ell-\mu)f^-)\ ,$$

and since $|f-f^-|\le 1/N$ we obtain (10.151). Now, the small factor we are looking for is simply

$$\nu_0(\varepsilon_\ell(\varepsilon_1\varepsilon_2-q))=\nu_0(\varepsilon_1)(1-q)=\mu(1-q)\ .$$

Using (10.151) we have

$$\left|\nu_0\big(\varepsilon_\ell(\varepsilon_1\varepsilon_2-q)(n_1^\ell-\mu)f\big)\right|\le\frac{K}{N}+|\mu(1-q)|\nu_0(|n_1^\ell-\mu||f|)\ .$$

Using (10.90) and the Cauchy-Schwarz inequality we see that

$$\begin{aligned}\nu_0(|n_1^\ell-\mu||f|)&\le L\nu(|n_1^\ell-\mu||f|)+K\exp\left(-\frac{N}{K}\right)\\&\le LR_N^{1/2}Q_N^{1/2}+K\exp\left(-\frac{N}{K}\right)\ ,\end{aligned}\quad (10.152)$$

and thus

$$\begin{aligned}\left|\nu_0\big(\varepsilon_\ell(\varepsilon_1\varepsilon_2-q)(n_1^\ell-\mu)f\big)\right|&\le\frac{K}{N}+|\mu(1-q)|R_N^{1/2}Q_N^{1/2}+K\exp\left(-\frac{N}{K}\right)\\&\le\frac{K}{N}+\frac{Lm^*}{\beta^4}R_N^{1/2}Q_N^{1/2}+K\exp\left(-\frac{N}{K}\right)\end{aligned}$$

because $|\mu|\le Lm^*$ and $\beta^4(1-q)\le L$.

We then proceed exactly in the same way for each of the other terms. For the term

$$\nu_0\big((\varepsilon_1\varepsilon_2-q)\varepsilon_3(n_1^3-\mu)f\big),$$

the "small factor" is $\nu_0((\varepsilon_1\varepsilon_2-q)\varepsilon_3)=\widehat{\mu}-\mu q$, and using (10.145) we see that

$$\left|\nu_0\big((\varepsilon_1\varepsilon_2-q)\varepsilon_3(n_1^3-\mu)f\big)\right|\le\frac{K}{N}+\frac{Lm^*}{\beta^4}R_N^{1/2}Q_N^{1/2}+K\exp\left(-\frac{N}{K}\right)\ .$$

Thus

$$|\mathrm{I}|\le\frac{K}{N}+Lm^*R_N^{1/2}Q_N^{1/2}+K\exp\left(-\frac{N}{K}\right)\ .$$

We could actually get a factor β^2 in the denominator, but it is not useful here. To study the terms II and III, we note that the function $S_{\ell,\ell'}$ does not depend on the last spin (so we may use Lemma 10.5.1) and does not depend either on the r.v.s $\eta_k = \eta_{N,k}$ (so that we may use (10.90)). As for the term

$$\nu_0\big(\varepsilon_1\varepsilon_2(\varepsilon_1\varepsilon_2 - q)(S_{1,2} - r)f\big)$$

the "small factor" is $\nu_0(\varepsilon_1\varepsilon_2(\varepsilon_1\varepsilon_2 - q)) = 1 - q^2$. As in (10.151) we have

$$\left|\nu_0\big(\varepsilon_1\varepsilon_2(\varepsilon_1\varepsilon_2 - q)(S_{1,2} - r)f\big) - \nu_0(\varepsilon_1\varepsilon_2(\varepsilon_1\varepsilon_2 - q))\nu_0(S_{1,2} - r)f)\right| \le \frac{K}{N} ,$$

and as in (10.152) we get

$$|\nu_0((S_{1,2} - r)f)| \le LQ_N^{1/2} S_N^{1/2} + K \exp\left(-\frac{N}{K}\right) .$$

For the other terms of III the small factor is respectively $q - q^2$ or $\widehat{q} - q^2$. In this manner we get, using (10.144),

$$\mathrm{III} \le \frac{K}{N} + \frac{L}{\beta^2} Q_N^{1/2} S_N^{1/2} + K \exp\left(-\frac{N}{K}\right) .$$

For the term II, the estimate is surprisingly straightforward. Indeed as in (10.151) we now have

$$\left|\nu_0\big((\varepsilon_1\varepsilon_2 - q)(S_{\ell,\ell} - \rho)f\big) - \nu_0(\varepsilon_1\varepsilon_2 - q)\nu_0((S_{\ell,\ell} - \rho)f)\right| \le \frac{K}{N} ,$$

and since $\nu_0(\varepsilon_1\varepsilon_2 - q) = 0$ we have $|\mathrm{II}| \le K/N$.

It remains to control the term $\mathcal{R}$ of (10.101) in (10.147). This term obviously satisfies

$$|\mathcal{R}| \le L\beta^4 Q_N^{1/2}\Big(\nu\big(|S_{1,1} - \rho|^4\big)^{1/2} + \nu\big(|S_{1,2} - r|^4\big)^{1/2} + \nu\big(|n_1^1 - \mu|^4\big)^{1/2}\Big) + \frac{K}{N} .$$

To bring out a small factor we use Proposition 10.5.5. Recalling the quantity s of (10.87), the sets

$$\{|S_{1,2} - r| \ge s^2\} \; ; \quad \{|S_{1,1} - \rho| \ge s^2 + |\rho - r|\} \; ; \quad \{|n_1^1 - \mu| \ge s\}$$

are negligible. Now

$$\rho - r = \frac{\alpha(1-q)}{1 - \beta(1-q)} \tag{10.153}$$

and since $1 - \beta(1-q) \ge m^{*2}/L$ by (10.15) and $\alpha \le m^{*4}/L_0$ by (10.2) we have $|\rho - r| \le Lm^{*2}(1-q)/L_0$, so since $1 - q \le L/\beta^{10}$ by (10.146) we obtain $|\rho - r| \le Ls^2$. Therefore the set

$$\{|S_{1,1} - \rho| \ge Ls^2\} \tag{10.154}$$

is negligible. This should make it obvious that

$$|\mathcal{R}| \le L\beta^4 s^2 Q_N^{1/2}(S_N^{1/2} + T_N^{1/2}) + L\beta^4 s Q_N^{1/2} R_N^{1/2} + \frac{K}{N} ,$$

and since $s \le Lm^*/\beta^5$ this finishes the proof. □

Proof of (10.134). We write

$$\begin{aligned} S_N = \nu\big((S_{1,2} - r)^2\big) &= \nu\left(\left(\sum_{2\le k\le M} n_k^1 n_k^2 - r\right)(S_{1,2} - r)\right) \\ &= \nu\big(((M-1)n_M^1 n_M^2 - r)(S_{1,2} - r)\big) \\ &= (M-1)\nu(n_M^1 n_M^2 (S_{1,2} - r)) - r\nu(S_{1,2} - r) . \end{aligned} \tag{10.155}$$

We would like to compute $\nu(n_M^1 n_M^2(S_{1,2} - r))$ using Theorem 10.6.1. Before using this theorem, we must first remove the small dependence of $S_{1,2}$ on the r.v.s $\eta_{i,M}$. For this we define

$$S_{1,2}^{\sim} = \sum_{2\le k\le M-1} n_k^1 n_k^2$$

and we observe that by Theorem 10.4.1 we have

$$(M-1)\nu(n_M^1 n_M^2 (S_{1,2} - r)) \le \frac{K}{N} + (M-1)\nu(n_M^1 n_M^2 (S_{1,2}^{\sim} - r)) . \tag{10.156}$$

Let us consider formula (10.112) for $\tau = 2$ and $f = S_{1,2}^{\sim} - r$. Multiplying this formula by $\alpha' = (M-1)/N = \alpha - 1/N$, we get

$$\begin{aligned} &(M-1)\nu(n_M^1 n_M^2 (S_{1,2}^{\sim} - r)) \\ &= \alpha' \widehat{r} \nu(S_{1,2}^{\sim} - r) + \alpha' \left(\frac{V_2}{D} + \frac{\beta q}{D^2}(V_1 + V_2 - 2V_3)\right) + \mathcal{R} \end{aligned}$$

where

$$|\mathcal{R}| \le \frac{K}{N}\nu\big((S_{1,2}^{\sim} - r)^2\big)^{1/2} \le \frac{K}{N} .$$

Now, since $r = \alpha\widehat{r}$

$$\big| r\nu(S_{1,2} - r) - \alpha'\widehat{r}\nu(S_{1,2}^{\sim} - r)\big| \le \frac{K}{N}$$

and (10.155) yields

$$S_N \le \alpha' \left(\frac{V_2}{D} + \frac{\beta q}{D^2}(V_1 + V_2 - 2V_3)\right) + \frac{K}{N} .$$

In each of the terms of V_1, V_2, V_3 there is a factor $f = S_{1,2}^{\sim} - r$. Theorem 10.4.1 implies that we make an error of order K/N when replacing this factor by $f' = S_{1,2} - r$. We perform this replacement and we call V'_ℓ the corresponding term. Therefore we obtain

$$S_N \le \alpha \left(\frac{V_2'}{D} + \frac{\beta q}{D^2}(V_1' + V_2' - 2V_3') \right) + \frac{K}{N} \,. \tag{10.157}$$

We now examine the respective contributions of the various quantities in (10.157). First V_2' and V_3' involve a term $\nu((R_{1,p} - q)f')$, which satisfies $|\nu((R_{1,p} - q)f')| \le Q_N^{1/2} S_N^{1/2}$. The coefficient of this term is bounded by

$$L\alpha \left(\frac{1}{D} + \frac{\beta q}{D^2} \right) \le L \frac{\alpha\beta}{m^{*2}} \le \frac{L\beta m^{*2}}{L_0} \le \frac{L\beta^2}{L_0} \tag{10.158}$$

because $D \ge m^{*2}/L$ by (10.15), $q \le Lm^{*2}$ by (10.12) and $\alpha \le Lm^{*4}/L_0$ by (10.2). All the other terms in V_1', V_2', V_3' are of the type

$$\beta N \nu\big((R_{1,\ell} - q) n_M^1 n_M^\ell (S_{1,2} - r)\big) \,,$$

and, by symmetry between the values of $k \le M$ we have

$$\alpha N \nu\big((R_{1,\ell} - q) n_M^1 n_M^\ell (S_{1,2} - r)\big) = \nu\big((R_{1,\ell} - q) S_{1,\ell} (S_{1,2} - r)\big) \,.$$

It is the term $S_{1,\ell}$ which provides here the crucial small factor. By Proposition 10.5.5 we know that the set

$$\{|S_{1,\ell}| \ge s^2 + r\}$$

is negligible. Therefore, using the Cauchy-Schwarz inequality we have shown that the contribution of all these other terms in V_1', V_2', V_3' is at most

$$L\beta \left(\frac{1}{D} + \frac{\beta q}{D^2} \right) (s^2 + r) Q_N^{1/2} S_N^{1/2} + K \exp\left(-\frac{N}{K} \right) \,.$$

We use the inequalities $D \ge m^{*2}/L$, $q \le Lm^{*2}$, $s^2 \le Lm^{*2}/L_0$ and (by (10.14)) $r \le L\alpha/m^{*2} \le Lm^{*2}/L_0$ to see that

$$\beta \left(\frac{1}{D} + \frac{\beta q}{D^2} \right) (s^2 + r) \le \frac{\beta^2 L}{L_0} \,.$$

Finally collecting these estimates we have shown that

$$S_N \le \frac{L\beta^2}{L_0} S_N^{1/2} Q_N^{1/2} + \frac{K}{N} \,.$$

Using (10.138) with $t = L_0/\beta^2 L$ we obtain

$$S_N \le \frac{1}{2} S_N + \frac{L\beta^4}{L_0^2} Q_N + \frac{K}{N} \,,$$

hence (10.134).

Proof of (10.135). It is nearly identical to the proof of (10.134), once one observes from (10.153) that $\rho - r \le L\alpha/m^{*2}$, so that $\rho \le Lm^{*2}/L_0$. □

Proof of (10.136). This is the most delicate inequality, because for certain estimates we are not permitted to make errors bigger than $m^{*2}R_N^2/L$ and obtaining this factor m^{*2} (rather than the much easier factor m^*) requires caution. As may be expected we write

$$
\begin{aligned}
R_N &= \nu\left(\left(\frac{1}{N}\sum_{i\le N-1}\sigma_i - \mu\right)^2\right) = \nu\left(\left(\frac{1}{N}\sum_{2\le i\le N}\sigma_i - \mu\right)^2\right)\\
&= \nu\left(\left(\frac{N-1}{N}\varepsilon_1 - \mu\right)\left(\frac{1}{N}\sum_{2\le i\le N}\sigma_i - \mu\right)\right)\\
&\le \nu((\varepsilon_1-\mu)f) + \frac{K}{N}
\end{aligned}
$$

where $f = n_1^1 - \mu$. We use Theorem 10.5.7 with $n = 1$, $\tau_1 = \tau_2 = 2$ and we get

$$R_N \le \nu_0((\varepsilon_1-\mu)f) + \mathrm{IV} + \mathrm{V} + \mathrm{VI} + \mathcal{R} + \frac{K}{N}$$

where $\mathcal{R}$ is as in (10.101) and where

$$\mathrm{IV} = \beta\big(\nu_0(\varepsilon_1(\varepsilon_1-\mu)(n_1^1-\mu)f) - \nu_0(\varepsilon_2(\varepsilon_1-\mu)(n_1^2-\mu)f)\big) \tag{10.159}$$

$$\mathrm{V} = \frac{\beta^2}{2}\big(\nu_0((\varepsilon_1-\mu)(S_{1,1}-\rho)f) - \nu_0((\varepsilon_1-\mu)(S_{2,2}-\rho)f)\big) \tag{10.160}$$

$$
\begin{aligned}
\mathrm{VI} = \beta^2\big(&-\nu_0(\varepsilon_1\varepsilon_2(\varepsilon_1-\mu)(S_{1,2}-r)f)\\
&+ \nu_0(\varepsilon_2\varepsilon_3(\varepsilon_1-\mu)(S_{2,3}-r)f)\big)\ .
\end{aligned}
\tag{10.161}
$$

Thus use of Lemma 10.5.1 is facilitated by the fact that none of the functions f, $S_{\ell,\ell'}$ or n^ℓ depends on the last spin. Thus $\mathrm{V} = 0$, and

$$\mathrm{IV} = \beta\big((1-\mu^2)\nu_0((n_1^1-\mu)f) - (q-\mu^2)\nu_0((n_1^2-\mu)f)\big) \tag{10.162}$$

$$\mathrm{VI} = \beta^2\big(-\mu(1-q)\nu_0((S_{1,2}-r)f) + (\widehat{\mu}-\mu q)\nu_0((S_{2,3}-r)f)\big)\ . \tag{10.163}$$

The easiest term to dispose of is VI. We simply use (10.90) and the Cauchy-Schwarz inequality to write

$$|\nu_0((S_{\ell,\ell'}-r)f)| \le LS_N^{1/2}R_N^{1/2} + K\exp\left(-\frac{N}{K}\right)$$

and we use (10.145) to obtain

$$|\mathrm{VI}| \le \frac{Lm^*}{\beta^2}S_N^{1/2}R_N^{1/2} + K\exp\left(-\frac{N}{K}\right)\ .$$

Next we study the term IV. The difficulty there is that the coefficient $1-\mu^2$ is *not* small. Observing that $q = \mathsf{E}\mathrm{th}^2 Y \geq (\mathsf{E}\mathrm{th} Y)^2 = \mu^2$ we write

$$|\mathrm{IV}| \leq \beta(1-\mu^2)|\nu_0((n_1^1-\mu)f)| + \beta(q-\mu^2)|\nu_0((n_1^2-\mu)f)| \;.$$

Recalling that $f = n_1^1 - \mu$ and using the Cauchy-Schwarz inequality we get $|\nu_0((n_1^2-\mu)f)| \leq \nu_0((n_1^1-\mu)^2)$ and thus

$$|\mathrm{IV}| \leq \beta(1-2\mu^2+q)\nu_0((n_1^1-\mu)^2) \;. \tag{10.164}$$

The strategy now is to use (10.17) and to relate $\nu_0((n_1^1-\mu)^2)$ and $\nu((n_1^1-\mu)^2) = R_N$ rather accurately. It would not suffice to use (10.91). Rather, we use (10.105) for $f = (n_1^1-\mu)^2$ to get

$$\nu((n_1^1-\mu)^2) = R_N = \nu_0((n_1^1-\mu)^2) + \mathrm{VII} + \mathcal{R}' \tag{10.165}$$

where, recalling yet another time the quantity s of (10.87), we get

$$|\mathcal{R}'| \leq L\beta^4 s^2 \nu((n_1^1-\mu)^2) + K\exp\left(-\frac{N}{K}\right) = L\beta^4 s^2 R_N + K\exp\left(-\frac{N}{K}\right) \;,$$

and

$$\begin{aligned}\mathrm{VII} &= \beta\biggl(\sum_{\ell\leq 2}\nu_0(\varepsilon_\ell(n_1^1-\mu)^2(n_1^\ell-\mu)) - 2\nu_0(\varepsilon_3(n_1^1-\mu)^2(n_1^3-\mu))\biggr)\\ &= \beta\mu\biggl(\sum_{\ell\leq 2}\nu_0((n_1^1-\mu)^2(n_1^\ell-\mu)) - 2\nu_0((n_1^1-\mu)^2(n_1^3-\mu))\biggr) \;.\end{aligned}$$

Therefore, using (10.90) and Hölder's inequality we get

$$|\mathrm{VII}| \leq L\beta\mu\nu(|n_1^1-\mu|^3) + K\exp\left(-\frac{N}{K}\right)$$

and since by Proposition 10.5.5 the set $\{|n_1^1-\mu| \geq s\}$ is negligible we get

$$|\mathrm{VII}| \leq L\beta\mu s R_N + K\exp\left(-\frac{N}{K}\right) \;.$$

In this manner we have shown that

$$|\nu_0((n_1^1-\mu)^2) - R_N| \leq L\beta^4(s^2+s\mu)R_N + K\exp\left(-\frac{N}{K}\right) \;, \tag{10.166}$$

and thus, using $s \leq Lm^*$ and $\mu \leq Lm^*$ in the second line:

$$\begin{aligned}\nu_0((n_1^1-\mu)^2) &\leq R_N + L\beta^4(s^2+s\mu)R_N + K\exp\left(-\frac{N}{K}\right)\\ &\leq R_N + L\beta^4 m^{*2} + K\exp\left(-\frac{N}{K}\right) \;.\end{aligned}$$

Combining with (10.164) and (10.17) we see that

$$|\mathrm{IV}| \leq \left(1 - \frac{m^{*2}}{L_2} + L_2\beta^4 s m^*\right) R_N + K \exp\left(-\frac{N}{K}\right) .$$

Since $s \leq Lm^*/\beta^5$, there exists β_0 such that $L_2\beta^4 s m^* \leq m^{*2}/2L_2$ for $\beta \geq \beta_0$. Since $s \leq Lm^*/\sqrt{L_0}$, if L_0 is large enough, we have $L_2\beta^4 s m^* \leq m^{*2}/2L_2$ for $\beta \leq \beta_0$. Thus, if L_0 is large enough we have

$$|\mathrm{IV}| \leq \left(1 - \frac{m^{*2}}{L}\right) R_N + K \exp\left(-\frac{N}{K}\right) . \tag{10.167}$$

Let us now turn to the control of the innocent looking remainder (10.101), which satisfies

$$|\mathcal{R}| \leq L\beta^4 R_N^{1/2}\big(\nu\big(|S_{1,1} - \rho|^4\big)^{1/2} + \nu\big(|S_{1,2} - r|^4\big)^{1/2} + \nu\big(|n_1^1 - \mu|^4\big)^{1/2}\big) + \frac{K}{N} . \tag{10.168}$$

The best we can do about the term $\nu(|n_1 - \mu|^4)$ is to use the negligible property of the set $\{|n_1 - \mu| \geq s\}$ to get

$$\nu\big(|n_1^1 - \mu|^4\big)^{1/2} \leq s R_N^{1/2} + K \exp\left(-\frac{N}{K}\right) .$$

In this manner, and as we have already noticed that the sets $\{|S_{1,2} - r| \geq s^2\}$ and $\{|S_{1,1} - \rho| \geq s^2\}$ are negligible we get

$$|\mathcal{R}| \leq L\beta^4\big(sR_N + s^2 R_N^{1/2}(S_N^{1/2} + T_N^{1/2})\big) + \frac{K}{N} . \tag{10.169}$$

Unfortunately the coefficient of R_N is s, and s is bounded by m^*, and not by m^{*2}. We definitely need an extra factor m^* in the bound. The only way to improve the estimate (10.169) is to start by replacing (10.97) by

$$\Big|\nu(f) - \nu_0(f) - \nu_0'(f) - \frac{1}{2}\nu_0''(f)\Big| \leq \sup_t \big|\nu_t^{(3)}(f)\big| .$$

Wait, why are you leaving? Didn't I promise earlier that we wouldn't even need to calculate $\nu''(f)$? It is not going to be bad at all. We follow the method of proof of Theorem 10.5.7 to get a higher order expansion

$$\nu(f) = \nu_0(f) + \mathrm{IV} + \mathrm{V} + \mathrm{VI} + \mathrm{VII} + \mathcal{R}' ,$$

where the terms IV, V, VI are as in (10.159), (10.160), (10.161), where the term VII will soon be discussed , and where the remainder $\mathcal{R}'$ now satisfies

$$|\mathcal{R}'| \leq L\beta^6 R_N^{1/2}\big(\nu\big(|S_{1,1} - \rho|^6\big)^{1/2} + \nu\big(|S_{1,2} - r|^6\big)^{1/2} + \nu\big(|n_1 - \mu|^6\big)^{1/2}\big) + \frac{K}{N} , \tag{10.170}$$

and instead of (10.169) we now get

$$|\mathcal{R}'| \le L\beta^6 s^2 R_N^{1/2}(S_N^{1/2} + T_N^{1/2} + R_N^{1/2}) + \frac{K}{N}\ , \tag{10.171}$$

where the factor s^2 provides the desired power m^{*2}. The term VII (which I promised not to write explicitly) is a sum of terms of the form

$$\beta^a \nu_0(fUV)\ , \tag{10.172}$$

where $a \le 4$, $f = (\varepsilon_1 - \mu)(n_1^1 - \mu)$, and both U and V are of the type either $\varepsilon_\ell \varepsilon_{\ell'}(S_{\ell,\ell'} - r)$ or $S_{\ell,\ell} - \rho$ or $\varepsilon_\ell(n_1^\ell - \mu)$. To control such a term we first consider the case where either U or V is not of the type $\varepsilon_\ell(n_1^\ell - \mu)$. Assuming e.g. that V is not of this type, we write

$$\begin{aligned}|\beta^a \nu_0(fUV)| &\le L\beta^a \nu(|f||U||V|) + K\exp\left(-\frac{N}{K}\right)\\ &\le L\beta^a s^2 \nu(|f||U|) + K\exp\left(-\frac{N}{K}\right)\ ,\end{aligned}$$

because the set $\{|V| \ge Ls^2\}$ is negligible, and we use the Cauchy-Schwarz inequality to see that the sum of these terms is bounded as in (10.171) and in particular (since $s \le Lm^*/\beta^5$) by

$$L\beta^4 s m^* R_N + \frac{Lm^*}{\beta^2} R_N^{-1/2}(S_N^{1/2} + T_N^{1/2})\ .$$

Next let us consider the case where both U and V are of the type $\varepsilon_\ell(n_1^\ell - \mu)$, so that the term (10.172) is

$$\begin{aligned}&\beta^a \nu_0\big((\varepsilon_1 - \mu)(n_1^1 - \mu)\varepsilon_{\ell_1}\varepsilon_{\ell_2}(n_1^{\ell_1} - \mu)(n_1^{\ell_2} - \mu)\big)\\ &= \beta^a \nu_0\big((\varepsilon_1 - \mu)\varepsilon_{\ell_1}\varepsilon_{\ell_2}\big)\nu_0\big((n_1^1 - \mu)(n_1^{\ell_1} - \mu)(n_1^{\ell_2} - \mu)\big)\ .\end{aligned} \tag{10.173}$$

Now $\nu_0(\varepsilon_1\varepsilon_{\ell_1}\varepsilon_{\ell_2}) = \widehat{\mu}$ when the indices $\{1, \ell_1, \ell_2\}$ are all distinct and $= \mu$ otherwise. In any case we have $|\nu_0((\varepsilon_1 - \mu)\varepsilon_{\ell_1}\varepsilon_{\ell_2})| \le Lm^*$ and we can bound the term (10.173) by $L\beta^4 m^* \nu(|n_1^\ell - \mu|^3)$ and then by

$$L\beta^4 m^* s R_N + K\exp\left(-\frac{N}{K}\right)\ .$$

Finally we have reached the bound

$$R_N \le \left(1 - \frac{m^{*2}}{L} + L\beta^4 s m^*\right) R_N + \frac{Lm^*}{\beta^2} R_N^{1/2}(S_N^{1/2} + T_N^{1/2}) + \frac{K}{N}\ ,$$

which as explained in (10.167) is the required bound if L_0 is large enough. □

10.8 Computing $p_{N,M}$

Throughout this section we still assume that (α, β, h) belongs to the admissible region (10.1) (so that $h > 0$), and that the constant L_0 which defines (10.1) has been chosen large enough. As usual (see e.g. (2.11)), we define

$$p_{N,M}(\beta, h) = \frac{1}{N} \mathsf{E} \log \sum_{\boldsymbol{\sigma}} \exp(-H_{N,M}(\boldsymbol{\sigma})) \ .$$

Let us consider the quantity

$$\begin{aligned} F(\alpha, \beta, h, r, \mu, q) = \log 2 - \frac{\mu^2 \beta}{2} + \frac{\alpha}{2}\left(\frac{\beta q}{1-\beta(1-q)} - \log(1-\beta(1-q))\right) \\ - \frac{\beta^2 r}{2}(1-q) + \mathsf{E} \log \mathrm{ch}(\beta(z\sqrt{r} + \mu) + h) \ , \end{aligned} \tag{10.174}$$

where (q, μ, r) is the unique solution of the replica-symmetric equations (10.3) to (10.5) that satisfies (10.6). The aim of this section is to prove the following.

Theorem 10.8.1. *If (α, β, h) belongs to the admissible region (10.1) we have*

$$\left| p_{N,M}(\beta, h) - F(\alpha, \beta, h, r, \mu, q) \right| \leq \frac{K}{N} \ , \tag{10.175}$$

where $\alpha = M/N$ and K depends only on β and h.

Research Problem 10.8.2. Prove that if (α, β, h_0) belongs to the admissible region, then (10.175) holds uniformly over $h \leq h_0$ (that is, the constant K depends only on β and h_0).

This problem is of course closely related to the Research Problem 10.7.2.

Lemma 10.8.3. *Thinking of $\alpha, \beta, h, r, \mu, q$ as independent variables, conditions (10.3) to (10.5) are equivalent to $\partial F/\partial \mu = \partial F/\partial r = \partial F/\partial q = 0$.*

Proof. Writing $Y = \beta(z\sqrt{q} + \mu) + h$ and $D = 1 - \beta(1-q)$, the relations

$$\begin{aligned} \frac{\partial F}{\partial \mu} &= -\beta\mu + \beta \mathsf{E} \mathrm{th} Y \\ \frac{\partial F}{\partial r} &= -\frac{\beta}{2}(1-q) + \frac{\beta^2}{2} \mathsf{E} \frac{1}{\mathrm{ch}^2 Y} \\ \frac{\partial F}{\partial q} &= \frac{\alpha}{2}\left(\frac{\beta}{D} - \frac{\beta^2 q}{D^2} - \frac{\beta}{D}\right) + \frac{\beta^2}{2} r = \frac{\beta^2}{2}\left(r - \frac{\alpha q}{D^2}\right) \end{aligned}$$

hold. □

As a consequence, considering $F(\alpha,\beta,h,r,\mu,q)$ as a function of α,β and h, we have

$$\frac{\partial F}{\partial \alpha} = f(\alpha,\beta,h) := \frac{1}{2}\left(\frac{\beta q}{1-\beta(1-q)} - \log(1-\beta(1-q))\right) .$$

Let us recall that we denote by $\langle\cdot\rangle_\sim$ an average for the Hamiltonian $H_{N,M-1}$ of (10.47), so the identity

$$N(p_{N,M}(\beta,h) - p_{N,M-1}(\beta,h)) = \mathsf{E}\log\left\langle\exp\frac{N\beta m_M^2}{2}\right\rangle_\sim \tag{10.176}$$

holds. As in the second proof of Theorem 2.4.2 the method of proof consists in computing $p_{N,M}(\beta,h)-p_{N,0}(\beta,h)$ by summation of the relations (10.176). It is for this purpose that we have always insisted that our constants K do not depend on either M or N (provided $M/N \le \alpha_0$) as e.g. in the definition of negligible sets (see Definition 4.2.3 and the discussion afterwards).

So, proceeding as in the second proof of Theorem 2.4.2 it suffices to prove the following.

Proposition 10.8.4. *If (α,β,h) belongs to the admissible region (10.1) we have*

$$\left|\mathsf{E}\log\left\langle\exp\frac{\beta N m_M^2}{2}\right\rangle_\sim - f(\alpha,\beta,h)\right| \le \frac{K}{N} , \tag{10.177}$$

where K depends on β and h only.

The idea is simply to consider the function

$$\varphi(u) = \mathsf{E}\log\left\langle\exp\frac{uNm_M^2}{2}\right\rangle_\sim$$

and to show that the quantity

$$\varphi'(u) = \frac{N}{2}\mathsf{E}\frac{\langle m_M^2\exp\frac{uNm_M^2}{2}\rangle_\sim}{\langle\exp\frac{uNm_M^2}{2}\rangle_\sim} \tag{10.178}$$

is nearly equal to

$$\begin{aligned}\frac{\partial f}{\partial u}(\alpha,u,h) &= \frac{1}{2}\left(\frac{q}{1-u(1-q)} + \frac{uq}{(1-u(1-q))^2} + \frac{1-q}{1-u(1-q)}\right)\\ &= \frac{1}{2}\frac{1-u(1-q)^2}{(1-(1-q))^2} .\end{aligned} \tag{10.179}$$

The formula (10.178) motivates the introduction of the Hamiltonian

$$-H_{M,N,u} = \frac{\beta N}{2} \sum_{k<M} m_k^2 + \frac{uN}{2} m_M^2 + Nhm_1 \,. \tag{10.180}$$

Indeed, if $\langle \cdot \rangle_u$ denotes an average for this Hamiltonian, and $\nu_u(f) = \mathsf{E}\langle f \rangle_u$, then (10.178) reads

$$\varphi'(u) = \frac{N}{2} \nu_u(m_M^2) \,. \tag{10.181}$$

Proposition 10.8.5. *If (α_0, β, h) belongs to the admissible region and $\alpha \le \alpha_0$, $0 \le u \le \beta$ we have*

$$\left| N\nu_u(m_M^2) - \frac{1 - u(1-q)^2}{(1 - u(1-q))^2} \right| \le \frac{K}{N} \,, \tag{10.182}$$

where K depends on α_0, β and h only.

Of course, since (10.182) means that $|\varphi'(u) - \partial f / \partial u| \le K/N$, (10.177) follows from (10.182) by integration.

To compute $N\nu_u(m_M^2)$ we would like to integrate by parts as in Theorem 10.6.1. To control the error terms we need to control the size of m_M.

Lemma 10.8.6. *For any number a the function*

$$u \mapsto \nu_u(\exp Nam_M^2) \tag{10.183}$$

is non-decreasing.

Consequently Theorem 10.4.1 implies that for $\alpha \le \alpha_0$ and $u \le 1$

$$\nu_u \left(\exp \frac{Nm_M^2}{K} \right) \le K \,, \tag{10.184}$$

where K depends only on h, β and α_0.

Proof. Let

$$\psi(u) = \nu_u(\exp Nam_M^2) = \mathsf{E} \frac{\left\langle \exp N \left(a + \frac{u}{2}\right) m_M^2 \right\rangle_\sim}{\left\langle \exp \frac{Nu}{2} m_M^2 \right\rangle_\sim}$$

so

$$\psi'(u) = \frac{N}{2} \left(\nu_u(m_M^2 \exp Nam_M^2) - \nu_u(m_M(\boldsymbol{\sigma}^2)^2 \exp Nam_M(\boldsymbol{\sigma}^1)^2) \right) \ge 0 \,,$$

as is seen by expansion of the exponential as a power series and use of Hölder's inequality $\langle m_M(\boldsymbol{\sigma}^2)^2 m_M(\boldsymbol{\sigma}^1)^{2k} \rangle_u \le \langle m_M^{2k+2} \rangle_u$. □

Proposition 10.8.7. *We have*

$$\nu_u\big((R_{1,2}-q)^4\big) \le \frac{K}{N^2} \tag{10.185}$$

$$|\nu_u(R_{1,2}-q)| \le \frac{K}{N} \ . \tag{10.186}$$

Here q is as in (10.3) and as usual K depends on (β, h, α_0) only. The proof will be given after that of Proposition 10.8.5.

Proof of Proposition 10.8.5. First, we reproduce the proof of Theorem 10.6.1, replacing β by u and ν by ν_u, so that (10.111) and (10.112) hold for ν_u rather than ν and u rather than β, and the values

$$\widehat{r} = \frac{q}{(1-u(1-q))^2} \ ; \ \widehat{\rho} = \frac{1-u(1-q)^2}{(1-u(1-q))^2} \ .$$

we may also replace everywhere n_M by m_M without changing the nature of the error terms. We use (10.111) for $n=1$ and $f=1$. It suffices to prove that $|V_n| \le K/N$ for $n=1,2$. Looking at the definition (10.109) and (10.110) of V_n, and using (10.186) it suffices to prove that for any two replicas ℓ_1 and ℓ_2 we have $N|\nu_u((R_{1,2}-q)m_M^{\ell_1}m_M^{\ell_2})| \le K/N$. To prove this we use (10.111) and (10.112) with $f=R_{1,2}-q$. Using that $\nu_u((R_{1,2}-q)^2) \le K/N$ by (10.185) it suffices to prove that

$$N\big|\nu_u((R_{1,2}-q)(R_{\ell,\ell'}-q)m_M^{\ell_1}m_M^{\ell_2})\big| \le \frac{K}{N} \ ,$$

which is the case because each of the four factors "counts as $N^{-1/2}$". This follows from (10.184), (10.185) and Hölder's inequality. □

Proof of Proposition 10.8.7. We start by explaining how to prove the inequality

$$\nu_u\big((R_{1,2}-q)^2\big) \le K/N \ . \tag{10.187}$$

We were not able to deduce this result from the case $u=\beta$. To prove it, it seems necessary to rewrite the entire proof of Theorems 10.7.1 and 10.6.1. It is hardly possible to actually do this here. Completing this project in full detail should be a straightforward, but nonetheless useful exercise for the reader wishing to really understand these arguments. There are many details to be changed, but each of them is trivial. There is no longer symmetry between the values of $k=2,\dots,M$ but only between the values $k=2,\dots,M-1$. So, for example, rather than (10.155) we now write

$$\begin{aligned} S_N = \nu_u\big((S_{1,2}-r)^2\big) &= \nu_u\Bigg(\bigg(\sum_{2\le k\le M} n_k^1 n_k^2 - r\bigg)(S_{1,2}-r)\Bigg) \\ &= \nu_u\big(((M-2)n_{M-1}^1 n_{M-1}^2 - r)(S_{1,2}-r)\big) \\ &\quad + \nu_u(n_M^1 n_M^2(S_{1,2}-r)) \ . \end{aligned}$$

The last term is $\le K/N$ by (10.183), and one would like to compute $\nu_u(n^1_{M-1}n^2_{M-1}f)$ by Theorem 10.6.1, used for $M-1$ rather than M. Theorem 10.6.1 will hold with the same proof provided we can control the size of n_{M-1}, namely we can prove that $\nu_u(\exp Nm^2_{M-1}/K) \le K$. For $u = \beta$ this follows from Theorem 10.4.1. Despite the intuitive feeling that decreasing u can only improve matters, we do not see how to deduce simply the case $u < \beta$ from the case $u = \beta$ (see Problem 10.8.8 below). We can however repeat verbatim the proof of Theorem 10.4.1, provided that the set $\{|m_1(\boldsymbol{\sigma}) - m^*| \ge \rho\}$ is negligible for ν_u, where ρ is as in (10.58). This fact is deduced from the case $u = \beta$ in Lemma 10.8.9 below.

Research Problem 10.8.8. Is it true that the function

$$u \mapsto \nu_u\left(\exp\left(\frac{Nm^2_{M-1}}{K}\right)\right)$$

is increasing?

We now discuss (10.185) and (10.186) (which we have not even yet proved for $u = \beta$). In the case of (10.185) we will claim in Section 10.9 that this is done simply by "iterating" the proof of (10.129), an idea that has already been used e.g. in Proposition 1.6.7, and that works similarly here. As for (10.186), it will be proved in the case $u = \beta$ in Section 10.10, as an introduction to central limit theorems, and these arguments are straightforwardly adapted to the case of ν_u. □

Lemma 10.8.9. *There exists a constant K such that for any subset A of Σ_N and any $u \le \beta$ we have*

$$\nu_u(\mathbf{1}_A) \le 4\nu(\mathbf{1}_A)^{1/K} \,. \tag{10.188}$$

Proof. Let $\varphi(u) = \nu_u(\mathbf{1}_A)$ so that

$$\varphi'(u) = \nu_u\left(\frac{Nm^2_M}{2}\mathbf{1}_A\right) - \nu_u\left(\frac{Nm_M(\boldsymbol{\sigma}^1)^2}{2}\mathbf{1}_A(\boldsymbol{\sigma}^2)\right) \,. \tag{10.189}$$

Consider a number $\tau > 1$ and τ' such that $1/\tau + 1/\tau' = 1$. Then use of (10.189) and Hölder's inequality yields

$$\varphi'(u) \ge -\varphi(u)^{1/\tau'}\nu_u(|Nm^2_M|^\tau)^{1/\tau} \,. \tag{10.190}$$

Combining (10.184) and the fact that $e^x \ge (x/\tau)^\tau$ implies that $\nu_u(|Nm^2_M|^\tau) \le (K\tau)^\tau$, so (10.190) yields

$$\varphi'(u) \ge -K\tau\varphi(u)^{1/\tau'} = -K\tau\varphi(u)^{1-1/\tau}$$

and thus

$$\frac{1}{\tau}\varphi'(u)\varphi(u)^{1/\tau-1} \geq -K \ .$$

By integration, for $u_0 \leq u_1$ this gives

$$\varphi(u_1)^{1/\tau} - \varphi(u_0)^{1/\tau} \geq -K(u_1 - u_0)$$

i.e.

$$\varphi(u_0)^{1/\tau} \leq K(u_1 - u_0) + \varphi(u_1)^{1/\tau} \ . \tag{10.191}$$

Assuming $\varphi(u_1) < 1/4$, let $\tau > 1$ with $\varphi(u_1) = 4^{-\tau}$. Then, if $K(u_1 - u_0) \leq 1/4$, we obtain

$$\varphi(u_0)^{1/\tau} \leq K(u_1 - u_0) + \varphi(u_1)^{1/\tau} \leq \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

so that $\varphi(u_0) \leq 2^{-\tau} = \sqrt{\varphi(u_1)}$. In particular taking $u_0 = \max(0, u_1 - 1/4K)$ we see that

$$\varphi\Big(\max\Big(0, u_1 - \frac{1}{4K}\Big)\Big) \leq \sqrt{\varphi(u_1)} \ .$$

Iteration of this inequality yields, for any integer k,

$$\varphi\Big(\max\Big(0, u_1 - \frac{k}{4K}\Big)\Big) \leq (\varphi(u_1))^{2^{-k}} \ .$$

If $\varphi(1) \geq 1/4$ then (10.188) holds. If $\varphi(1) < 1/4$ we take $u_1 = 1$ and for k the smallest integer for which $k \geq 4K$ to prove the result. □

Research Problem 10.8.10. Find a shorter proof of Proposition 10.8.4.

10.9 Higher Moments, the TAP Equations

Theorem 10.7.1 generalizes to higher moments.

Theorem 10.9.1. *If (α, β, h) belongs to the admissible region (10.1), and the constant L_0 that defines this region has been chosen large enough, then for each p we have*

$$\nu\big((R_{1,2} - q)^{2p}\big) \leq \left(\frac{Kp}{N}\right)^p \tag{10.192}$$

$$\nu\big((m_1 - \mu)^{2p}\big) \leq \left(\frac{Kp}{N}\right)^p \tag{10.193}$$

$$\nu\big((S_{1,2} - r)^{2p}\big) \leq \left(\frac{Kp}{N}\right)^p \tag{10.194}$$

$$\nu\big((S_{1,1} - \rho)^{2p}\big) \leq \left(\frac{Kp}{N}\right)^p \ , \tag{10.195}$$

where K depends only on b, h and p.

We have explained twice (in Proposition 1.6.7 and Theorem 2.5.1) how to control higher moments in the same manner as one controls second moments. The reader who has been energetic enough to follow the proof of Theorem 10.7.1 shall produce by herself a similar proof of Theorem 10.9.1. Other readers, of course, would hardly even glance at the argument. Therefore the best option seems to leave that proof to the reader.

We recall that $\langle\cdot\rangle_-$ denotes an average for the Hamiltonian (10.67), and that $\eta_k = \eta_{N,k}$.

Proposition 10.9.2. *Under the conditions of Theorem 10.9.1, for each p there is a number K independent of N such that*

$$\mathsf{E}\left(\langle\sigma_N\rangle - \operatorname{th}\left(\beta \sum_{2\le k\le M} \eta_k \langle n_k\rangle_- + \beta\mu + h\right)\right)^{2p} \le \frac{K}{N^p} . \tag{10.196}$$

Research Problem 10.9.3. (Level 2) Under the conditions of Theorem (10.9.1), given an integer n, prove that there exist independent standard normal r.v.s $(z_i)_{i\le n}$ such that

$$\forall i \le n\ , \quad \mathsf{E}(\langle\sigma_i\rangle - \operatorname{th}(\beta(z_i\sqrt{r} + \mu) + h))^2 \le \frac{K}{N} , \tag{10.197}$$

where K is independent of N.

Most likely, given any integer p one can even require

$$\forall i \le n\ , \quad \mathsf{E}(\langle\sigma_i\rangle - \operatorname{th}(\beta(z_i\sqrt{r} + \mu) + h))^{2p} \le \frac{K}{N^p} . \tag{10.198}$$

Of course one expects Proposition 10.9.2 to be the starting point of a solution to Problem 10.9.3, but even when $n = 1$, replacing the nearly Gaussian r.v. $\sum_{2\le k\le M} \eta_k \langle n_k\rangle_-$ by an actual Gaussian r.v. while preserving the rate K/N is not obvious. The following weaker statement is easier.

Exercise 10.9.4. Consider a point (α_0, β, h) in the admissible region (10.1). Prove that given n, the joint law of $(\langle\sigma_1\rangle, \dots, \langle\sigma_n\rangle)$ converges as $N \to \infty$ and $M/N \to \alpha_0$ to the joint law of the sequence $(\operatorname{th}(\beta(z_i\sqrt{r} + \mu) + h))_{i\le n}$, where (q, r, μ) is the solution of the replica-symmetric equations when the parameter α there is equal to α_0.

Our next result complements (10.196) by the T.A.P. equations. We use the notation $\xi_{i,j} = N^{-1} \sum_{2\le k\le M} \eta_{i,k}\eta_{j,k}$.

Theorem 10.9.5. *If (α, β, h) belongs to the admissible region (10.1) then for each p there is a constant K depending only on β, h and p such that*

$$\mathsf{E}\bigg(\langle\sigma_N\rangle - \operatorname{th}\Big(\beta \sum_{i\le N-1} \xi_{i,N}\langle\sigma_i\rangle - \beta(\rho - q)\langle\sigma_N\rangle + \beta\mu + h\Big)\bigg)^{2p} \le \frac{K}{N^p}\ .$$

As in Chapter 1 this implies that the equations

$$\langle\sigma_i\rangle \simeq \operatorname{th}\Big(\beta\sum_{j\ne i}\xi_{j,i}\langle\sigma_j\rangle - \beta(\rho - q)\langle\sigma_i\rangle + \beta\mu + h\Big)$$

are simultaneously nearly satisfied. (In physics, the exact equations are called the T.A.P. equations.)

As in Section 1.7 both Proposition 10.9.2 and Theorem 10.9.5 are consequences of a general principle, which is the main result of this section, and to which we turn now.

As usual we assume that (α, β, h) belongs to the admissible region (10.1) and that the constant L_0 defining (10.1) has been chosen large enough.

Theorem 10.9.6. *Consider a function U, that is infinitely differentiable. Assume that for each integer ℓ there is a constant $C(= C_\ell)$ such that*

$$|U^{(\ell)}(x)| \le C \exp C|x|\ . \tag{10.199}$$

Consider independent Bernoulli r.v.s $(\zeta_k)_{k\le M}$ and a standard Gaussian r.v. ξ which are all independent of the randomness of $H_{N,M}$. Then, using the notation $\dot m_k = m_k - \langle m_k\rangle$, we have for each p,

$$\mathsf{E}\bigg(\Big\langle U\Big(\sum_{2\le k\le M}\zeta_k\dot m_k\Big)\Big\rangle - \mathsf{E}_\xi U(\xi\sqrt{\rho - q})\bigg)^{2p} \le \frac{K}{N^p}\ , \tag{10.200}$$

where q is as in (10.3), ρ as in (10.98) and K does not depend on N.

In the application of Theorem 10.9.6 just below one needs a little more than just the fact that K does not depend on N. One needs K to stay bounded when the parameters (α, β, h) stay inside a compact set entirely contained in the admissible region (10.1) and when U might depend on N, but with a fixed constant C in (10.199); this is shown by the same arguments.

Before we get into the main argument, let us establish a simple technical fact.

Lemma 10.9.7. *Consider numbers a and $C > 0$ independent of N. Consider also a standard Gaussian r.v. ξ independent of all the other randomnesses, and let $Y = \sum_{2\le k\le M}\zeta_k m_k + a\xi$. Then we have*

$$\mathsf{E}\langle\exp C|Y|\rangle \le K\ . \tag{10.201}$$

The same relation holds if one replaces m_k by $\dot m_k$ in the definition of Y.

Proof. We observe that for any numbers b, a_k, and since $\exp|x| \le \exp x + \exp(-x)$ and $\log \mathrm{ch} x \le x^2/2$ we have

$$\begin{aligned}\mathsf{E}\exp\Big|b+\sum_k a_k\zeta_k\Big| &\le \mathsf{E}\exp\Big(b+\sum_k a_k\zeta_k\Big)+\mathsf{E}\exp\Big(-b-\sum_k a_k\zeta_k\Big)\\ &= 2\mathrm{ch} b\exp\sum_k \log\mathrm{ch} a_k \le 2\mathrm{ch} b\exp\sum_k \frac{a_k^2}{2}\,.\end{aligned}$$

Using this for $b = c\xi$, and taking expectation also in ξ we get

$$\mathsf{E}\exp\Big|c\xi+\sum_k a_k\zeta_k\Big| \le 2\exp\Big(\frac{c^2}{2}+\sum_k\frac{a_k^2}{2}\Big)\,. \tag{10.202}$$

Therefore if E_0 denotes expectation in the r.v.s ζ_k and ξ only, and since these r.v.s are independent of the randomness of the bracket, we then have

$$\mathsf{E}\langle\exp C|Y|\rangle = \mathsf{E}\langle\mathsf{E}_0\exp C|Y|\rangle \le K\mathsf{E}\Big\langle\exp\Big(\frac{C^2}{2}\sum_{2\le k\le M} m_k^2\Big)\Big\rangle\,.$$

Let $T = \sup_{\boldsymbol{\sigma}}\sum_{2\le k\le M} m_k^2$. It then suffices to prove that given a number C independent of N we have $\mathsf{E}\exp CT \le K$, which is proved in Corollary 4.5.8. When we replace m_k by $\dot{m}_k$ the argument is identical. □

Proof of Proposition 10.9.2. The idea is to apply (10.200) to the $(N-1)$-spin system with Hamiltonian (10.66). Writing this Hamiltonian as

$$\begin{aligned}-H_{N-1,M}(\boldsymbol{\rho}) = \frac{N-1}{2}\Big(\frac{\beta(N-1)}{N}\Big)\sum_{1\le k\le M}\Big(\sum_{i\le N-1}\frac{\eta_{i,k}\sigma_i}{N-1}\Big)^2\\ +(N-1)h\Big(\frac{1}{N-1}\sum_{i\le N-1}\eta_{i,1}\sigma_i\Big)\,,\end{aligned}$$

shows that this is the Hamiltonian of an $(N-1)$-spin Hopfield model where β has been changed in $\beta_- = \beta(N-1)/N$ and α in $\alpha_- = \alpha N/(N-1)$. Let us denote by (q_-,ρ_-,μ_-) the values of (q,ρ,μ) after this change of parameters. We leave it to the reader to check that

$$|q-q_-| \le \frac{K}{N}\,,\quad |\rho-\rho_-| \le \frac{K}{N}\,,\quad |\mu-\mu_-| \le \frac{K}{N}\,. \tag{10.203}$$

We shall now use (10.200) for the $(N-1)$-spin system with parameters α_- and β_-, and for $\zeta_k = \eta_{N,k}$. The quantity that corresponds to m_k for this system is $N/(N-1)n_k$, which motivates the use of the function

$$U(x) = \exp\Big(\varepsilon\beta x\frac{N-1}{N}\Big)\,.$$

Setting

$$A = \left\langle \exp \varepsilon\beta \sum_{2\le k\le M} \zeta_k \dot{n}_k \right\rangle_- - \exp\left(\frac{\beta^2}{2}(\rho_- - q_-)\right) ,$$

we then obtain from (10.200) that

$$\mathsf{E}A^{2p} \le \frac{K}{N^p} .$$

Let us define

$$B = \exp \varepsilon \left(\beta \sum_{2\le k\le M} \zeta_k \langle n_k\rangle_- + \beta\mu + h\right) ,$$

so that using Lemma 10.9.7 for the $(N-1)$ spin system we have $\mathsf{E}B^{4p} \le K$ and therefore

$$\mathsf{E}(AB)^{2p} \le (\mathsf{E}A^{4p}\mathsf{E}B^{4p})^{1/2} \le \frac{K}{N^p} ,$$

i.e.

$$\mathsf{E}\Bigg(\left\langle \exp \varepsilon \left(\beta \sum_{2\le k\le M} \zeta_k n_k + \beta\mu + h\right)\right\rangle_- \tag{10.204}$$
$$- \exp \frac{\beta^2}{2}(\rho_- - q_-) \exp \varepsilon \left(\beta \sum_{2\le k\le M} \zeta_k \langle n_k\rangle_- + \beta\mu + h\right)\Bigg)^{2p} \le \frac{K}{N^p} .$$

From (10.67) we have

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

where $\mathcal{E} = \exp \varepsilon \left(\beta \sum_{1\le k\le M} \zeta_k n_k + h\right)$. Recalling that $\zeta_1 = \eta_{N,1} = 1$, only trivial bounds are required to prove from (10.193) that "we may replace $\zeta_1 n_1 = n_1$ by μ" to get

$$\mathsf{E}\left(\langle \sigma_N\rangle - \frac{\langle \mathrm{Av}\varepsilon\mathcal{E}'\rangle_-}{\langle \mathrm{Av}\mathcal{E}'\rangle_-}\right)^{2p} \le \frac{K}{N^p}$$

where

$$\mathcal{E}' = \exp \varepsilon \left(\beta \sum_{2\le k\le M} \zeta_k n_k + \beta\mu + h\right) .$$

Use of (10.204) and of Lemma 1.7.14 concludes the proof of Proposition 10.9.2. □

Exercise 10.9.8. Complete the proof of Theorem 10.9.5.

Hint: Proceed as in (1.208) to show that

$$\mathsf{E}\Big(\sum_{2\le k\le M}\zeta_k\langle n_k\rangle-(\rho-q)\langle\sigma_N\rangle-\sum_{2\le k\le M}\zeta_k\langle n_k\rangle_-\Big)^{2p}\le\frac{K}{N^p}\,.$$

Combine with (10.196) to obtain that

$$\mathsf{E}\bigg(\langle\sigma_N\rangle-\mathrm{th}\Big(\beta\sum_{2\le k\le M}\zeta_k\langle n_k\rangle-\beta(\rho-q)\langle\sigma_N\rangle+\beta\mu+h\Big)\bigg)^{2p}\le\frac{K}{N^p}\,,$$

and substitute the value $n_k=N^{-1}\sum_{1\le i<N}\eta_{i,k}\sigma_i$.

We turn to the proof of Theorem 10.9.6, which occupies the remainder of this section. The proof is an elaboration on the proof of Theorem 1.7.11. The main difficulty is that we have to be more sophisticated about approximate integration by parts. In the remainder of this section a and b denote integers.

Lemma 10.9.9. *There exist numbers $(c_a)_{a\ge1}$ and polynomials $(P_a(x))_{a\ge1}$ such that for any infinitely differentiable function v and any integer n,*

$$\begin{aligned}v(1)-v(-1)=&\sum_{1\le a\le n}c_a(v^{(2a-1)}(1)+v^{(2a-1)}(-1))\\&+\int_{-1}^{1}P_n(x)v^{(2n+1)}(x)\mathrm{d}x\,.\end{aligned}\tag{10.205}$$

Proof. We have proved this formula in (4.197) for $n=1$, $c_1=1$, $P_1(x)=(x^2-1)/2$. The general case follows by induction over n. To perform the induction, together with (10.205), we assume that $P_n(x)=P_n(-x)$. We find a polynomial $Q_n(x)$ such that $Q_n'(x)=P_n(x)$ and $Q_n(0)=0$, so that $Q_n(x)=-Q_n(-x)$, and by integration by parts

$$\begin{aligned}\int_{-1}^{1}P_n(x)v^{(2n+1)}(x)\mathrm{d}x=&\,Q_n(1)(v^{(2n+1)}(1)+v^{(2n+1)}(-1))\\&-\int_{-1}^{1}Q_n(x)v^{(2n+2)}(x)\mathrm{d}x\,.\end{aligned}$$

We set $c_{n+1}=Q_n(1)$. Next, we consider a polynomial $P_{n+1}(x)$ such that $P'_{n+1}(x)=Q_n(x)$, so that $P_{n+1}(x)=P_{n+1}(-x)$. we may and do assume that $P_{n+1}(1)=P_{n+1}(-1)=0$. Integration by parts yields

$$\int_{-1}^{1}Q_n(x)v^{(2n+2)}(x)\mathrm{d}x=-\int_{-1}^{1}P_{n+1}(x)v^{(2n+3)}(x)\mathrm{d}x\,.$$

This completes the induction. □

Consider the function $V(x) = U(x) - \mathsf{E}_\xi U(\xi\sqrt{\rho - q})$ so that

$$\mathsf{E}_\xi V(\xi\sqrt{\rho - q}) = 0 \,. \tag{10.206}$$

Using replicas, the left-hand side of (10.200) is

$$\mathsf{E}\left\langle \prod_{\ell\le 2p} V\Big(\sum_{2\le k\le M} \zeta_k \dot{m}_k^\ell\Big)\right\rangle \,.$$

We define the function F by setting $F(\mathbf{x}) = \prod_{\ell\le 2p} V(x_\ell)$ for $\mathbf{x} = (x_\ell)_{\ell\le 2p}$. Let us consider i.i.d. standard Gaussian r.v.s $(\xi^\ell)_{\ell\le 2p}$ and let us define $\mathbf{X}_t = (X_\ell)_{\ell\le 2p}$ for

$$X_\ell = \sqrt{t}\sum_{2\le k\le M} \zeta_k \dot{m}_k^\ell + \sqrt{1-t}\,\xi^\ell\sqrt{\rho - q} \,. \tag{10.207}$$

We define

$$\varphi(t) = \mathsf{E}\langle F(\mathbf{X}_t)\rangle$$

so that (10.200) means that $|\varphi(1)| \le KN^{-p}$. To prove this we will show that

$$\forall n \le 2p\,, \quad |\varphi^{(n)}(0)| \le \frac{K}{N^p}\,; \quad \forall t\,,\ 0\le t\le 1\,, \quad |\varphi^{(2p)}(t)| \le \frac{K}{N^p}\,.$$

Of course, the first task is to understand the structure of these derivatives. Given numbers $\ell(1), \ldots, \ell(b) \le 2p$ we define $T(\ell(1), \ldots, \ell(b))$ as follows:
If $b = 2$ and $\ell(1) = \ell(2)$ we define

$$T(\ell(1), \ell(2)) = \sum_{2\le k\le M} (\dot{m}_k^{\ell(1)})^2 - (\rho - q)$$

If $b = 2$ and $\ell(1) \ne \ell(2)$ we define

$$T(\ell(1), \ell(2)) = \sum_{2\le k\le M} \dot{m}_k^{\ell(1)} \dot{m}_k^{\ell(2)}$$

If $b > 2$ we define

$$T(\ell(1), \ldots, \ell(b)) = \sum_{2\le k\le M} \prod_{a\le b} \dot{m}_k^{\ell(a)} \,.$$

An important fact is that "a factor $T(\ell(1), \ldots, \ell(b))$ counts as $N^{-1/2}$ for $b = 2$ and as $N^{-(b-2)/2}$ for $b \ge 3$". More precisely, we have the following.

Lemma 10.9.10. *For any n it holds*

$$\mathsf{E}\langle T(\ell(1), \ell(2))^{2n}\rangle \le \frac{K}{N^n} \tag{10.208}$$

$$\mathsf{E}\langle T(\ell(1), \ell(2), \ldots, \ell(b))^{2n}\rangle \le \frac{K}{N^{n(b-2)+1}} \,. \tag{10.209}$$

Proof. To prove (10.208) we will use (10.194), (10.195) and Jensen's inequality as in the proof of (1.200) but the complete details are tedious. First, let us define

$$S'_{\ell,\ell'} = \sum_{2\le k\le M} m_k^\ell m_k^{\ell'} \ .$$

We show that

$$\nu\big((S'_{1,2}-r)^{2n}\big)\le \frac{K}{N^n} \tag{10.210}$$

and

$$\nu\big((S'_{1,1}-\rho)^{2n}\big)\le \frac{K}{N^n} \ . \tag{10.211}$$

To prove (10.211) we use that $m_k = n_k + \eta_{N,k}\sigma_N/N$ to get

$$S'_{1,1} = S_{1,1} + \mathrm{I} + \mathrm{II} \ ,$$

where

$$\mathrm{I} = \frac{2}{N}\sigma_N \sum_{2\le k\le M} \eta_{N,k} n_k \ ,$$

$$\mathrm{II} = \frac{M-1}{N^2} \ .$$

We simply use

$$|\mathrm{I}| \le \frac{2\sqrt{M-1}}{N}\sqrt{\sum_{2\le k\le M}(n_k)^2} \ ,$$

and Lemma 4.5.7 to obtain

$$\mathsf{E}\nu(\mathrm{I}^{2n}) \le \frac{K}{N^n} \ .$$

Of course the control of the term II is trivial, and this completes the proof of (10.211). The proof of (10.210) is entirely similar. Let us now prove (10.208) in the case $\ell(1)\ne\ell(2)$, and without loss of generality we assume $\ell(1)=1$ and $\ell(2)=2$. Then we have, using Jensen's inequal

$$\begin{aligned}
&\Big\langle\Big(\sum_{2\le k\le M}\dot m_k^1 \dot m_k^2\Big)^{2n}\Big\rangle \\
&= \Big\langle\Big(\sum_{2\le k\le M}(m_k^1 m_k^2 - m_k^1\langle m_k\rangle - m_k^2\langle m_k\rangle - \langle m_k\rangle^2)\Big)^{2n}\Big\rangle \\
&\le \big\langle (S'_{1,2}-S'_{1,3}-S'_{2,4}+S'_{3,4})^{2n}\big\rangle \ ,
\end{aligned}$$

and the conclusion follows from (10.210). Next, let us prove (10.208) in the case $\ell(1)=\ell(2)=1$. Then we write

$$T_{1,1} = \sum_{2\le k\le M}(\dot m_k)^2 - (\rho - r) = S'_{1,1} - \rho + \mathrm{III} + \mathrm{IV} \ ,$$

where

$$\mathrm{III} = -2\biggl(\sum_{2\le k\le M} m_k\langle m_k\rangle - r\biggr) \;,$$

and

$$\mathrm{IV} = \sum_{2\le k\le M} (\langle m_k\rangle^2 - r) \;.$$

As before (10.210) and Jensen's inequality imply that $\mathsf{E}\langle \mathrm{III}^{2n}\rangle \le K/N^n$ and similarly for the term IV.

To prove (10.209) we use Theorem 10.4.1 to deduce from Hölder's inequality that for $k \ge 2$,

$$\mathsf{E}\left\langle \left(\prod_{a\le b} \dot{m}_k^{\ell(a)}\right)^{2n}\right\rangle \le \frac{K}{N^{bn}} \;,$$

where K does not depend on k. We then use the inequality $\left(\sum_{2\le k\le M} x_k\right)^{2n} \le M^{2n-1}\sum_{2\le k\le M} x_k^{2n}$ to obtain (10.209). □

Corollary 10.9.11. *If b is even, for every n we have*

$$\mathsf{E}\langle T(\ell(1),\ldots,\ell(b))^{2n}\rangle \le \frac{K}{N^{nb/2}} \;. \tag{10.212}$$

Proof. For $b = 2$ this is (10.208) and for $b \ge 4$, this follows from (10.209) since $b - 2 \ge b/2$. □

Given a sequence $\mathbf{s} = (s_\ell)_{\ell\le 2p}$, we denote by $F^{(\mathbf{s})}$ the corresponding partial derivative of F.

Proposition 10.9.12. *For $n \le 2p$, the derivative $\varphi^{(n)}(t)$ is a sum of "main terms" and of a "remainder" $\mathcal{R}_n(t)$ with the following properties:*

$$\forall \ell\;;\; 0\le \ell\le 2p-n\;, \qquad |\mathcal{R}_n^{(\ell)}(t)| \le \frac{K}{N^p} \;, \tag{10.213}$$

where K is independent of N. Each main term is of the type

$$ct^d\mathsf{E}\bigl\langle T(\ell_1(1),\ldots,\ell_1(b_1))\cdots T(\ell_m(1),\ldots,\ell_m(b_m))F^{(\mathbf{s})}(\mathbf{X}_t)\bigr\rangle \tag{10.214}$$

where c is a number, $m \le n$, $b_1,\ldots,b_m$ are even *integers, and*

$$d := \frac{b_1+\cdots+b_m}{2} - n \ge 0 \;. \tag{10.215}$$

Moreover the sequence $\mathbf{s} = (s_\ell)_{\ell\le 2n}$ is obtained as follows. For each $\ell \le 2p$, s_ℓ counts the number of times the index ℓ occurs in the list of $b_1+b_2+\cdots+b_m$ integers:

$$\ell_1(1),\ldots,\ell_1(b_1),\ldots,\ell_m(1),\ldots,\ell_m(b_m) \;. \tag{10.216}$$

Of course, the number of main terms is independent of N.

Lemma 10.9.13. *For each* $\mathbf{s}$ *there is a number* K *such that for each* t *we have* $\mathsf{E}\langle(F^{(\mathbf{s})}(\mathbf{X}_t))^2\rangle \le K$.

Proof. We recall that $F(\mathbf{x}) = \prod_{l\le 2p} V(x_\ell)$, so using Hölder's inequality it suffices to prove that for each ℓ and p there exists a number K independent of N and t such that $\mathsf{E}\langle|U^{(\ell)}(X)|^p\rangle \le K$. Using (10.199) it suffices to show that for each number C' we have $\mathsf{E}\langle\exp C'|X|\rangle < K$ where $X = \sqrt{t}\sum_{2\le k\le M}\zeta_k\dot{m}_k + \sqrt{1-t}\xi\sqrt{\rho-q}$. That this is the case was proved in Lemma 10.9.7 (the proof of which shows that K does not depend on t). □

Proof of Theorem 10.9.6. First we prove that $\sup_t|\varphi^{(2p)}(t)| \le KN^{-p}$. For this we use Proposition 10.9.12 for $n = 2p$, so that $|\mathcal{R}_n(t)| \le K/N^p$ by (10.213). To control the main term (10.214) we use Hölder's inequality and that $m \le n$ to get

$$\begin{aligned}&\big|\mathsf{E}\langle T(\ell_1(1),\ldots,\ell_1(b_1))\cdots T(\ell_m(1),\ldots,\ell_m(b_m))F^{(\mathbf{s})}(\mathbf{X}_t)\rangle\big|\\ &\le \mathsf{E}\langle T(\ell_1(1),\ldots,\ell_1(b_1))^{2n}\rangle^{1/2n}\cdots\mathsf{E}\langle T(\ell_m(1),\ldots,\ell_m(b_m))^{2n}\rangle^{1/2n}\\ &\quad\times\mathsf{E}\langle F^{(\mathbf{s})}(\mathbf{X}_t)^2\rangle^{1/2}\,.\end{aligned}$$

The last term is $\le K$ by Lemma 10.9.13. Corollary 10.9.11 implies that for $r \le m$

$$\mathsf{E}\langle T(\ell_r(1),\ldots,\ell_r(b_r))^{2n}\rangle^{1/2n} \le \frac{K}{N^{b_r/4}}\,.$$

Now (10.215) yields

$$b_1+\cdots+b_m \ge 2n = 4p \tag{10.217}$$

and this proves that $\sup_t|\varphi^{(2p)}(t)| \le KN^{-p}$.

Now we prove that $|\varphi^{(n)}(0)| \le KN^{-p}$ for $n \le 2p$. Using (10.213), it suffices to prove that if for $t=0$ a term (10.214) is not 0, then it is $\le KN^{-p}$. So, let us assume that the term (10.214) is not zero for $t = 0$. Denoting by E_0 expectation in $\xi^1,\ldots,\xi^{2p}$ only, then, for $t=0$ this term is

$$c\mathsf{E}\langle T(\ell_1(1),\ldots,\ell_1(b_1))\cdots T(\ell_m(1),\ldots,\ell_m(b_m))\rangle\mathsf{E}_0F^{(\mathbf{s})}(\mathbf{X}_0)\,. \tag{10.218}$$

Thus for each $\ell \le 2n$ we have $s_\ell \ge 1$, for otherwise, by (10.206) and since $\mathbf{X}_0 = (\xi^\ell\sqrt{\rho-q})_{\ell\le 2p}$ we have $\mathsf{E}_0F^{(\mathbf{s})}(\mathbf{X}_0) = 0$. But then we must have $s_\ell \ge 2$ for each $\ell \le 2n$. This is because if ℓ occurs exactly once in the list (10.216), taking average on the ℓ^{th}-replica shows that the bracket in (10.218) is zero. Therefore the list (10.216) has at least $4p$ elements, i.e. (10.217) holds and we conclude by Corollary 10.9.11 as before. □

We turn to the proof of Proposition 10.9.12. The basic idea is to compute recursively $\varphi^{(n)}(t)$, integrating by parts in the r.v.s ξ^ℓ and using (10.205) as "approximate integration by parts" in the r.v.s ζ_k. It is this recursion which necessitates condition (10.213). First we consider the case $n = 1$ and we detail the computation of $\varphi'(t)$. This quantity is a sum of $2p$ terms of the type

$$\mathsf{E}\left\langle\left(\frac{1}{2\sqrt{t}}\sum_{2\le k\le M}\zeta_k\dot{m}_k^\ell - \frac{1}{2\sqrt{1-t}}\xi^\ell\sqrt{\rho-q}\right)F^{(\mathbf{s})}(\mathbf{X}_t)\right\rangle , \qquad (10.219)$$

where $\mathbf{s} = (s_\ell)$ is such that $s_{\ell'} = 0$ if $\ell \neq \ell'$ and $s_\ell = 1$. For each k, we consider the term

$$\frac{1}{2\sqrt{t}}\mathsf{E}\zeta_k\langle\dot{m}_k^\ell F^{(\mathbf{s})}(\mathbf{X}_t)\rangle , \qquad (10.220)$$

for which we are going to perform approximate integration by parts in the random sign ζ_k. Given t and k, let us use the notation $\mathbf{X}_{t,x}$ to mean the quantity obtained by replacing in each component X_ℓ of $\mathbf{X}_t$ every occurrence of ζ_k by x. There is one such occurrence in each term X_ℓ of (10.207), and this occurrence is $\zeta_k\sqrt{t}\dot{m}_k^\ell$. Let us define the function $v_k(x) = \langle\dot{m}_k^\ell F^{(\mathbf{s})}(\mathbf{X}_{t,x})\rangle$, so we may rewrite (10.220) as

$$\frac{1}{2\sqrt{t}}\mathsf{E}\zeta_k v_k(\zeta_k) . \qquad (10.221)$$

To estimate this quantity, we think of (10.205) as

$$\mathsf{E}\zeta v(\zeta) = \sum_{1\le a\le 2p} c_a\mathsf{E}v^{(2a-1)}(\zeta) + \frac{1}{2}\int_{-1}^{1} P_{2p}(x)v^{(4p+1)}(x)\mathrm{d}x . \qquad (10.222)$$

The main terms in Proposition 10.9.12 will be produced by the terms in the summation of (10.222), while the remainder term of this proposition will be produced by the integral term.

Keeping in mind that the occurrence of ζ_k in X_ℓ has a factor $\sqrt{t}\dot{m}_k^\ell$ we see that $\mathsf{E}v_k^{(2a-1)}(\zeta)$ is a sum of terms of the type

$$t^{a-1/2}\mathsf{E}\langle\dot{m}_k^\ell\dot{m}_k^{\ell(1)}\cdots\dot{m}_k^{\ell(2a-1)}F^{(\mathbf{s}')}(\mathbf{X}_t)\rangle ,$$

where $\mathbf{s}' = (s_{\ell'})_{\ell'\le 2p}$ and $s_{\ell'}$ counts the number of occurrences of ℓ' in the list $\ell, \ell(1), \ldots, \ell(2a-1)$. Thus $t^{-1/2}\mathsf{E}v_k^{(2a-1)}(\zeta)$ is a sum of terms of the type

$$t^{a-1}\mathsf{E}\langle\dot{m}_k^\ell\dot{m}_k^{\ell(1)}\cdots\dot{m}_k^{\ell(2a-1)}F^{(\mathbf{s}')}(\mathbf{X}_t)\rangle . \qquad (10.223)$$

When $a \ge 2$ we gather the corresponding terms over k to obtain a term as in (10.214), with $m = 1$, $b_1 = 2a$ and $d = a - 1 = b_1/2 - 1$.

When $a = 1$ the terms we obtain when computing $c_1\mathsf{E}v^{(2a-1)}\zeta = \mathsf{E}v'(\zeta)$ are the same as those we would get by integration by parts if the r.v.s ζ_k were standard Gaussian. When $\ell(1) \neq \ell$ we gather the terms over k to get a term as in (10.214) for $m = 1$, $b_1 = 2$, $d = 0 = b_1/2 - 1$. When $\ell(1) = \ell$ we gather the terms over k with the corresponding term obtained by integration by parts of (10.219) with respect to ξ^ℓ. This creates the magic quantity $\sum_{2\le k\le M}(\dot{m}_k^\ell)^2-(\rho-q)=T(\ell,\ell)$, and therefore we still obtain a term as in (10.214), again with $m=1$, $b_1=2$ and $d=0=b_1/2-1$.

The remainder term in $\varphi'(t)$ is obtained by collecting the integral terms of (10.222) for the different values of k, so it is

$$\mathcal{R}_1(t)=\frac{1}{4\sqrt{t}}\sum_{2\le k\le M}\int_{-1}^{1}P_{2p}(x)v_k^{(4p+1)}(x)\mathrm{d}x\ ,$$

and to conclude the proof of the case $n=1$ it remains only to show that all derivatives of order $s\le 2p-1$ of $\mathcal{R}_1(t)$ in t are bounded by KN^{-p}. To prove this we show that for given k all derivatives of order $s\le 2p$ of

$$\frac{1}{4\sqrt{t}}\int_{-1}^{1}P_{2p}(x)v_k^{(4p+1)}(x)\mathrm{d}x \tag{10.224}$$

are bounded by KN^{-p-1} (and in fact even by KN^{-2p-1}) where K does not depend on k. The claim will then follow by summation over k. We recall that the notation $\mathbf{X}_{t,x}$ means that every occurrence of ζ_k in every component of $\mathbf{X}_t$ has been replaced by x and that $v_k(x)=\mathsf{E}\langle\dot{m}_k^\ell F^{(\mathbf{s})}(\mathbf{X}_{t,x})\rangle$. Computing the derivatives of v_k we see that the quantity (10.224) itself is a sum of terms

$$\frac{1}{2}t^{2p}\mathsf{E}\left\langle\dot{m}_k^\ell\dot{m}_k^{\ell(1)}\cdots\dot{m}_k^{\ell(4p+1)}\int_{-1}^{1}P_{2p}(x)F^{(\mathbf{s})}(\mathbf{X}_{t,x})\mathrm{d}x\right\rangle\ . \tag{10.225}$$

By Theorem 10.4.1, the quantity $\dot{m}_k^\ell\cdots\dot{m}_k^{\ell(4p+1)}$ "counts as N^{-2p-1}", and by Lemma 10.9.13 (or more exactly by the version of this lemma, with identical proof, where ζ_k has been replaced by x) the term $F^{(\mathbf{s})}(\mathbf{X}_{t,x})$ "counts as 1". This proves that $|\mathcal{R}_1(t)|\le KN^{-2p-1}$. To control the derivatives of $\mathcal{R}_1(t)$ one first computes them from (10.225) and (10.207). We then immediately see that a potential problem to bound these derivatives is the presence of the factors $\sqrt{t}$ and $\sqrt{1-t}$ in (10.207), which, when differentiated, create the large factors $1/\sqrt{t}$ and $1/\sqrt{1-t}$.

When computing the derivatives of the function

$$t\mapsto\mathsf{E}\left\langle\dot{m}_k^\ell\dot{m}_k^{\ell(1)}\cdots\dot{m}_k^{\ell(4p+1)}\int_{-1}^{1}P_{2p}(x)F^{(\mathbf{s})}(\mathbf{X}_{t,x})\mathrm{d}x\right\rangle\ ,$$

a good strategy is after each differentiation to integrate by parts in the Gaussian r.v.s ξ^ℓ, but NOT to perform approximate integration by parts in the Bernoulli r.v.s $\zeta_{k'}$. The Gaussian integration by parts removes the factor $1/\sqrt{1-t}$ that was created by the differentiation (as we have seen so many times).

In this manner, we may express a derivative in t of order s of the quantity (10.225) as a linear combination of terms of the form

$$t^{2p-s_1-s_2/2}\mathsf{E}\left\langle \dot{m}_k^\ell \dot{m}_k^{\ell(1)} \cdots \dot{m}_k^{\ell(4p+1)} A_{\ell_1} \cdots A_{\ell_{s_2}} \int_{-1}^{1} P_{2p}(x) F^{(\mathbf{s}')}(\mathbf{X}_{t,x}) \mathrm{d}x \right\rangle ,$$

where $s_1, s_2 \geq 0$, $s = s_1 + s_2$ and

$$A_\ell = \sum_{2 \leq k' \leq M} \zeta_{k'} \dot{m}_{k'}^\ell .$$

We now observe that each quantity A_ℓ "counts as one", i.e. that for each integer n' we have

$$\mathsf{E}\left\langle \left(\sum_{2 \leq k' \leq M} \zeta_{k'} \dot{m}_{k'}^\ell \right)^{2n'} \right\rangle \leq K .$$

This is simply a consequence of Lemma 10.9.7 since actually

$$\mathsf{E}\left\langle \exp \left| \sum_{2 \leq k' \leq M} \zeta_{k'} \dot{m}_{k'} \right| \right\rangle \leq K .$$

In this manner we complete the proof of (10.196) when $n = 1$.

To complete the proof of Proposition 10.9.12 we proceed by induction over n, with exactly the same type of arguments as above. The reason why we have to consider terms with $m \leq n$ (rather than only with $m = n$) is that terms are produced by the factor t^d when differentiating (10.214) . Differentiating this factor does not change m, and d decreases by one unit while n has increased by one unit, preserving the relation (10.215). □

10.10 Central Limit Theorems

Using Theorems 10.5.7 and 10.6.1 one can recursively compute higher moments. The principle of the computations is the same as in Section 1.10 and 9.7, but the algebra is a bit more complicated. We start with a simple result.

Theorem 10.10.1. *If (α, β, h) belongs to the admissible region (10.1), then*

$$\nu(m_1 - \mu) = O(2) \tag{10.226}$$

$$\nu(S_{1,2} - r) = O(2) \tag{10.227}$$

$$\nu(R_{1,2} - q) = O(2) . \tag{10.228}$$

Here and everywhere, $O(k)$ denotes a quantity U such that $|U| \leq K/N^{k/2}$ where K depends only on β and h.

Theorem 10.10.1 proves the claim made in Section 4.5 that the quantities $\mu_{N,M} = \nu(m_1)$, $r_{N,M} = \nu\left(\sum_{2 \leq k \leq M} m_k^1 m_k^2\right)$ and $q_{N,M} = \nu(R_{1,2})$ "satisfy the replica-symmetric equations (10.3) to (10.5) with accuracy of order $1/N$". These quantities are within $O(2)$ of the true solution of these equations.

When using Theorem 10.5.7 we will take $\tau_1 = \tau_2 = 2$. Using Theorem 10.7.1 the remainder $\mathcal{R}$ of (10.101) then satisfies

$$|\mathcal{R}| \le K(n)\frac{\nu(f^2)^{1/2}}{N} + K \sup|f| \exp\left(-\frac{N}{K}\right) . \tag{10.229}$$

This will be used only in the case where $|f| \le N^K$ (e.g. $f = S_{1,2} - r$), so that for any k the last term is $O(k)$. We also note that Proposition 10.5.3 together with (10.90) show that

$$|\nu(f) - \nu_0(f)| \le K(n)\frac{\nu(f^2)^{1/2}}{\sqrt{N}} + K \sup|f| \exp\left(-\frac{N}{K}\right) . \tag{10.230}$$

When using Theorem 10.6.1, it will be convenient to use the following form, together with the notation

$$D = 1 - \beta(1-q) . \tag{10.231}$$

Proposition 10.10.2. *Consider a function f on Σ_N^n, that does not depend on the r.v.s $(\eta_{i,k})_{i\le N,k\le M}$. Let us define*

$$U_1 = \beta\Bigg(\sum_{2\le\ell\le n+1} \nu((R_{1,\ell}-q)S_{1,\ell}f) - (n+1)\nu((R_{1,n+2}-q)S_{1,n+2}f)\Bigg) \tag{10.232}$$

and, for $2 \le p \le n+1$ let us define

$$\begin{aligned} U_p = \alpha\nu((R_{1,p}-q)f) + \beta\Bigg(&\sum_{\ell\ne p,\ell\le n+1} \nu((R_{p,\ell}-q)S_{1,\ell}f) \\ &-(n+1)\nu((R_{p,n+2}-q)S_{1,n+2}f)\Bigg) . \end{aligned} \tag{10.233}$$

Then we have

$$\nu(S_{1,1}f) = \rho\nu(f) + \frac{U_1}{D} + \frac{\beta q}{D^2}\left(\sum_{p\le n} U_p - nU_{n+1}\right) + \mathcal{R}_1 \tag{10.234}$$

$$\nu(S_{1,2}f) = r\nu(f) + \frac{U_2}{D} + \frac{\beta q}{D^2}\left(\sum_{p\le n} U_p - nU_{n+1}\right) + \mathcal{R}_2 \tag{10.235}$$

where

$$|\mathcal{R}_1|, |\mathcal{R}_2| \le K\frac{\nu(f^2)^{1/2}}{N} .$$

Proof. For every $2 \leq k \leq M$ we write the version of Theorem 10.6.1 for k rather than M, and we sum over k. □

Proof of Theorem 10.10.1. Let us define

$$A = \nu(m_1 - \mu)\ ; \quad B = \nu(S_{1,2} - r)\ ; \quad C = \nu(R_{1,2} - q)\ . \tag{10.236}$$

The principle of the proof is to establish 3 linear relations between A, B and C with accuracy $O(2)$. The first step is to evaluate B by using (10.235) for $n = 2$ and $f = 1$. We observe that (10.129), (10.131) and the Cauchy-Schwarz inequality imply that for $\ell \neq 1$ we have $\nu((R_{1,\ell}-q)(S_{1,\ell}-r)) = O(2)$. Therefore for $\ell \neq 1$ we have

$$\begin{aligned}\nu((R_{1,\ell} - q)S_{1,\ell}) &= r\nu(R_{1,\ell} - q) + \nu((R_{1,\ell} - q)(S_{1,\ell} - r)) \\ &= rC + O(2)\ .\end{aligned}$$

Similarly for $p \neq 1$ we get

$$\nu((R_{p,1} - q)S_{1,1}) = \rho C + O(2)\ .$$

Therefore by (10.232) we have

$$U_1 = -\beta r C + O(2)$$

and by (10.233)

$$U_2 = U_3 = (\alpha + \beta(\rho - 2r))C + O(2)$$

and thus

$$U_1 + U_2 - 2U_3 = -(\alpha + \beta(\rho - r))C + O(2)\ .$$

Consequently (10.235) gives

$$B = \nu(S_{1,2} - r) = C\left(\frac{\alpha + \beta(\rho - 2r)}{D} - \frac{\beta q}{D^2}(\alpha + \beta(\rho - r))\right) + O(2)\ . \tag{10.237}$$

This is our first relation. The other two required relations will follow from Theorem 10.5.7. We observe that

$$A = \nu(m_1 - \mu) = \nu(\varepsilon_1 - \mu)$$

so we may use Theorem 10.5.7 with $n = 1$, $f = \varepsilon_1 - \mu$ and (10.229) to get

$$A = \nu_0(f) + \mathrm{I} + \mathrm{II} + \mathrm{III} + O(2)\ , \tag{10.238}$$

where

$$\begin{aligned}\mathrm{I} &= \beta\big(\nu_0(\varepsilon_1 - \mu)\varepsilon_1(n_1^1 - \mu)) - \nu_0((\varepsilon_1 - \mu)\varepsilon_2(n_1^2 - \mu))\big) \\ \mathrm{II} &= \frac{\beta^2}{2}\big(\nu_0((\varepsilon_1 - \mu)(S_{1,1} - \rho)) - \nu_0((\varepsilon_1 - \mu)(S_{2,2} - \rho))\big) \\ \mathrm{III} &= \beta^2\big(-\nu_0(\varepsilon_1\varepsilon_2(\varepsilon_1 - \mu)(S_{1,2} - r)) + \nu_0(\varepsilon_2\varepsilon_3(\varepsilon_1 - \mu)(S_{2,3} - r))\big)\ .\end{aligned}$$

We compute these terms using Lemma 10.5.1 e.g.

$$\nu_0((\varepsilon_1 - \mu)\varepsilon_1(n_1^1 - \mu)) = (1 - \mu^2)\nu_0(n_1 - \mu) \ .$$

We then replace ν_0 by ν in the last term with error $O(2)$ using (10.230). In this manner we get

$$\begin{aligned} \mathrm{I} &= \beta(1-q)A + O(2) \\ \mathrm{II} &= O(2) \\ \mathrm{III} &= \beta^2(\widehat{\mu} - \mu)B + O(2) \ , \end{aligned}$$

where $\widehat{\mu} = \mathsf{E}\mathrm{th}^3(\beta(z\sqrt{r} + \mu) + h)$. Therefore, since $\nu_0(f) = 0$, (10.238) yields the relation

$$DA = \beta^2(\widehat{\mu} - \mu)B + O(2) \ . \tag{10.239}$$

To get the third and last required relation, we use again Theorem 10.5.7, this time with $n = 2$, $f = \varepsilon_1\varepsilon_2 - q$, so that

$$C = \nu(\varepsilon_1\varepsilon_2 - q) = \nu_0(\varepsilon_1\varepsilon_2 - q) + \mathrm{IV} + \mathrm{V} + \mathrm{VI} + O(2) \tag{10.240}$$

where

$$\begin{aligned} \mathrm{IV} &= \beta\Big(\sum_{\ell\le 2}\nu_0\big((\varepsilon_1\varepsilon_2 - q)\varepsilon_1(n_1^\ell - \mu)\big) - 2\nu_0\big((\varepsilon_1\varepsilon_2 - q)\varepsilon_3(n_1^3 - \mu)\big)\Big) \\ \mathrm{V} &= \frac{\beta^2}{2}\Big(\sum_{\ell\le 2}\nu_0\big((\varepsilon_1\varepsilon_2 - q)(S_{\ell,\ell} - \rho)\big) - 2\nu_0\big((\varepsilon_1\varepsilon_2 - q)(S_{3,3} - \rho)\big)\Big) \\ \mathrm{VI} &= \beta^2\Big(\nu_0\big((\varepsilon_1\varepsilon_2 - q)\varepsilon_1\varepsilon_2(S_{1,2} - r)\big) - 2\sum_{\ell\le 2}\nu_0\big((\varepsilon_1\varepsilon_2 - q)\varepsilon_\ell\varepsilon_3(S_{\ell,3} - r)\big) \\ &\quad + 3\nu_0\big((\varepsilon_1\varepsilon_2 - q)\varepsilon_3\varepsilon_4(S_{3,4} - r)\big)\Big) \ . \end{aligned}$$

Proceeding as in the proof of (10.239), and recalling the notation $\widehat{q} = \mathsf{E}\mathrm{th}^4(\beta(z\sqrt{r} + \mu) + h)$ we find

$$\begin{aligned} \mathrm{IV} &= 2\beta(\mu - \widehat{\mu})A + O(2) \\ \mathrm{V} &= O(2) \\ \mathrm{VI} &= \beta^2(1 - 4q + 3\widehat{q})B + O(2) \end{aligned}$$

so that (10.240) implies

$$C = 2\beta(\mu - \widehat{\mu})A + \beta^2(1 - 4q + 3\widehat{q})B + O(2) \ . \tag{10.241}$$

It remains to prove that the relations (10.237), (10.239) and (10.241) imply that A, B and C are $O(2)$. Substitution of (10.239) in (10.241) yields

$$C = W_1 B + O(2) \tag{10.242}$$

for
$$W_1 = \frac{2\beta^3(\widehat{\mu}-\mu)^2}{D} + \beta^2(1-4q+3\widehat{q}) \ .$$
Letting
$$W_2 = \frac{\alpha+\beta(\rho-2r)}{D} - \frac{\beta q}{D^2}(\alpha+\beta(\rho-r)) \ ,$$
it follows from (10.237) that $B = W_2 C + O(2)$. Combining with (10.242) yields
$$C(1-W_1W_2) = O(2) \ .$$

Therefore it suffices to show that in the admissible region (10.1) the coefficient of C is not zero, provided that the parameter L_0 which defines this region has been taken large enough. This follows from the estimates we have proved (and similar ones). First, we recall that by (10.15) we have $D \geq m^{*2}/L$. Moreover by (10.14) we have $r \leq L\alpha/m^{*2}$, and combining with (10.153) we have also $\rho \leq L\alpha/m^{*2}$. Finally $q \leq Lm^{*2}$ by (10.12). Combining these and using (10.2) it is straightforward to see that
$$|W_2| \leq L\frac{\alpha}{m^{*4}} \leq \frac{L}{L_0} \ . \tag{10.243}$$

Next, we claim that
$$|W_1| \leq L \ . \tag{10.244}$$
These two relations complete the proof.

To prove that W_1 stays bounded we first observe that by (10.146) this is the case for $\beta \geq 2$ since then $D \geq 1/L$. So it suffices to show that W_1 stays bounded for $\beta \leq 2$. For this we use that (10.11) implies $|\mu| \leq Lm^*$, that (10.12) implies $q \leq Lm^{*2}$, and thus
$$|\widehat{\mu}| = |\mathsf{E}\mathrm{th}^3 Y| \leq \mathsf{E}|\mathrm{th}^3 Y| \leq q = \mathsf{E}\mathrm{th}^2 Y \leq Lm^{*2} \ .$$
Finally we use that $D \geq m^{*2}/L$. □

The proof of Proposition 10.10.2 clearly brings forward the fact that the underlying algebra is non trivial. Compared with Chapter 8, the new feature here is created by the term IV in Theorem 10.5.7. To understand the situation, it would most likely be useful to consider first the case of the SK model with ferromagnetic interaction given by (4.22), where the same phenomenon occurs. We recall that for this model the "replica-symmetric" equations write
$$\mu = \mathsf{E}\mathrm{th}(\beta_2 z\sqrt{q} + \beta_1\mu + h)$$
$$q = \mathsf{E}\mathrm{th}^2(\beta_2 z\sqrt{q} + \beta_1\mu + h) \ .$$

Research Problem 10.10.3. (Level 1) For the SK model with ferromagnetic interaction, compute the joint asymptotic behavior of the quantities $m_1^\ell - \mu$ and $R_{\ell,\ell'} - q$. That is, given integers k_1 and k_2, integers $\ell_1, \ell'_1, \ldots, \ell_{k_1}, \ell'_{k_1}$ and $\ell''_1 \ldots, \ell''_{k_2}$ compute

$$\lim_{N\to\infty} N^{(k_1+k_2)/2} \nu\biggl(\prod_{k\le k_1} (R_{\ell_k,\ell'_k} - q) \prod_{k\le k_2} (m_1^{\ell''_k} - \mu)\biggr) .$$

The main difficulty is that the quantities $R_{\ell,\ell'} - q$ on the one hand and the quantities $m_1^\ell - \mu$ on the other hand are correlated.

Research Problem 10.10.4. (Level 1) Same as Research Problem 10.10.3, but now for the Hopfield model.

Of course one might be more ambitious, and want to include the quantities $S_{\ell,\ell'} - r$, $S_{\ell,\ell} - \rho$, etc. We will prove one single result in this direction, which is simple because symmetry allows for reasonable computations.

We use the notation $\mathbf{b} = \langle \boldsymbol{\sigma} \rangle$,

$$T_{\ell,\ell'} = \frac{(\boldsymbol{\sigma}^\ell - \mathbf{b}) \cdot (\boldsymbol{\sigma}^{\ell'} - \mathbf{b})}{N}$$

of (1.244) and $\widehat{q} = \mathsf{E}\mathrm{th}^4(\beta(z\sqrt{r} + \mu) + h)$. We recall that $D = 1 - \beta(1-q)$.

Theorem 10.10.5. *For (α, β, h) in the admissible region (10.1), consider*

$$A = \frac{1 - 2q + \widehat{q}}{N(1 - \alpha\beta^2(1 - 2q + \widehat{q})D^{-2})} . \tag{10.245}$$

Consider an integer n and for $1 \le \ell < \ell' \le n$ consider integers $k(\ell,\ell')$. Let $k = \sum_{1\le\ell<\ell'\le n} k(\ell,\ell')$. Then

$$\nu\biggl(\prod_{1\le\ell<\ell'\le n} T_{\ell,\ell'}^{k(\ell,\ell')}\biggr) = \prod_{1\le\ell<\ell'\le n} a(k(\ell,\ell'))A^k + O(k+1)$$

where $a(k) = \mathsf{E}g^k$ for g standard Gaussian.

If we examine the proof of Proposition 1.10.4 in the case $k_1 = k_2 = 0$, we see that the proof of Theorem 10.10.5 follows by the same argument once we know the following, where $R^-_{\ell,\ell'} = N^{-1}\sum_{i<N} \sigma_i^1 \sigma_i^2$.

Proposition 10.10.6. *Consider a function f on Σ_{N-1}^n. Assume that f is the product of $k-1$ terms of the type $R^-_{\ell,\ell'}$. Then we have*

$$\begin{aligned}\nu((\varepsilon_1 - \varepsilon_2)(\varepsilon_3 - \varepsilon_4)f) = \frac{\alpha\beta^2(1-2q+\widehat{q})}{D^2}&\nu((T_{1,3} - T_{1,4} - T_{2,3} + T_{2,4})f)\\ &+ O(k+1) .\end{aligned}$$

This is shown by combination of the next two results.

Lemma 10.10.7. *If f is as above, then*

$$\nu((\varepsilon_1-\varepsilon_2)(\varepsilon_3-\varepsilon_4)f)=\beta^2(1-2q+\widehat{q})\nu((S_{1,3}-S_{1,4}-S_{2,3}+S_{2,4})f) + O(k+1)\ .$$

Proof. We use Theorem 10.5.7, and the argument of Lemma 1.8.4 to see that the terms above are the only ones which do not cancel. □

The second part of the proof of Proposition 10.10.6 is as follows.

Lemma 10.10.8. *If f is as above, then*

$$\begin{aligned}\nu((S_{1,3}-S_{1,4}-S_{2,3}+S_{2,4})f) &= \frac{\alpha}{D^2}\nu((R_{1,3}-R_{1,4}-R_{2,3}+R_{2,4})f)\\ &\quad + O(k+1)\\ &= \frac{\alpha}{D^2}\nu((T_{1,3}-T_{1,4}-T_{2,3}+T_{2,4})f)\\ &\quad + O(k+1)\ .\end{aligned}$$

Proof. We apply Proposition 10.10.2. Spectacular simplifications occur immediately from (10.235) since

$$\nu((S_{1,3}-S_{1,4})f)=\frac{U_3-U_4}{D}+O(k+1), \tag{10.246}$$

where U_3 and U_4 are given by (10.233). Now since $\nu((R_{p,1}-q)(S_{1,1}-\rho)f)=O(k+1)$ we have

$$\nu((R_{p,1}-q)S_{1,1}f)=\rho\nu((R_{p,1}-q)f)+O(k+1)\ ,$$

and, similarly,

$$\ell>1\Rightarrow\nu((R_{p,\ell}-q)S_{1,\ell}f)=r\nu((R_{p,\ell}-q)f)+O(k+1)\ .$$

Defining

$$B_p=\beta r\Bigg(\sum_{5\le\ell\le n+1}\nu((R_{p,\ell}-q)f)-(n+1)\nu((R_{p,n+2}-q)f)\Bigg)\ ,$$

we then get, writing explicitly the terms in the summation of (10.233) for $\ell=1,2,3,4,\ \ell\neq p$

$$\begin{aligned}U_3 &= \alpha\nu((R_{1,3}-q)f)+\beta\rho\nu((R_{1,3}-q)f)\\ &\quad +\beta r\nu((R_{2,3}-q)f)+\beta r\nu((R_{3,4}-q)f)+B_3+O(k+1)\\ U_4 &= \alpha\nu((R_{1,4}-q)f)+\beta\rho\nu((R_{1,4}-q)f)\\ &\quad +\beta r\nu((R_{2,4}-q)f)+\beta r\nu((R_{3,4}-q)f)+B_4+O(k+1)\ ,\end{aligned}$$

so that

$$U_3 - U_4 = \alpha\nu((R_{1,3} - R_{1,4})f) + \beta\rho\nu((R_{1,3} - R_{1,4})f) \\ + \beta r\nu((R_{2,3} - R_{2,4})f) + B_3 - B_4 + O(k+1) \ .$$

We substitute this expression in (10.246) to compute $\nu((S_{1,3} - S_{1,4})f)$. In the resulting formula we exchange the indices 2 and 1 to compute $\nu((S_{2,3} - S_{2,4})f)$. We then subtract these two formulas and we conclude the proof with the help of the relation $\alpha + \beta(\rho - r) = \alpha/D$, which itself is a consequence of (10.153). □

10.11 The *p*-Spin Hopfield Model

Consider an integer $p \geq 3$. Keeping the notation $m_k(\boldsymbol{\sigma}) = N^{-1}\sum_{i\leq N}\eta_{i,k}\sigma_i$, we consider the Hamiltonian

$$-H_{M,N}(\boldsymbol{\sigma}) = \frac{N\beta}{p}\sum_{k\leq M} m_k(\boldsymbol{\sigma})^p + Nhm_1(\boldsymbol{\sigma}) \ . \tag{10.247}$$

For definiteness, we will consider only the case where p is even and is fixed once and for all. As in the Hopfield model, we assume that $\eta_{k,1} = 1$. As we shall explain soon, the case of interest is now when M is of order N^{p-1}.

Let us first point out the really new feature of this model. It turns out that in a certain domain of parameters, $m_k(\boldsymbol{\sigma})$ is typically of order $N^{-1/2}$ for $k \geq 2$. Writing as usual $n_k(\boldsymbol{\sigma}) = N^{-1}\sum_{i<N}\eta_{i,k}\sigma_i$ and $\eta_k = \eta_{N,k}$ let us bring out the influence of the last spin in the Hamiltonian:

$$\begin{aligned} -H_{N,M} &= \frac{N\beta}{p}\sum_{k\leq M}\left(n_k(\boldsymbol{\sigma}) + \frac{\eta_k\sigma_N}{N}\right)^p + Nhn_1(\boldsymbol{\sigma}) + h\sigma_N \\ &= \frac{N\beta}{p}\sum_{k\leq M} n_k(\boldsymbol{\sigma})^p + Nhn_1(\boldsymbol{\sigma}) + \sigma_N\Bigg(\beta\sum_{2\leq k\leq M}\eta_k n_k(\boldsymbol{\sigma})^{p-1} \\ &\qquad + \beta n_1(\boldsymbol{\sigma})^{p-1} + h\Bigg) + \cdots \end{aligned} \tag{10.248}$$

where the other terms $\cdots$ are, hopefully, of lower order. Writing as usual $n_k^\ell = n_k(\boldsymbol{\sigma}^\ell)$, the quantity $\sum_{2\leq k\leq M}(n_k^1 n_k^2)^{p-1}$ measures some kind of overlap, and we may expect that it will play an important role. Since each n_k^ℓ should be of order $N^{-1/2}$, $(n_k^1 n_k^2)^{p-1}$ should be of order N^{-p+1}, so that the case where M is a proportion of N^{p-1}, that is $M = \alpha N^{p-1}$ should be of special interest. we may then hope that for a certain number r we will have $\sum_{2\leq k\leq M}(n_k^1 n_k^2)^{p-1} \simeq r$. Looking at (10.248) we expect that m_1 will typically be nearly equal to the number μ such that

$$\mu = \mathsf{E}\mathrm{th}(\beta(z\sqrt{r} + \mu^{p-1}) + h) \tag{10.249}$$

where z is standard Gaussian, and that $R_{1,2}$ will be typically nearly equal to the number q given by

$$q = \mathsf{E}\mathrm{th}^2(\beta(z\sqrt{r} + \mu^{p-1}) + h) . \tag{10.250}$$

Finding the value of r is the real fun. We should calculate

$$\nu\bigg(\sum_{2\le k\le M}(n_k^1 n_k^2)^{p-1}\bigg) = (M-1)\nu((n_M^1 n_M^2)^{p-1}) . \tag{10.251}$$

Let us then write

$$(M-1)\nu((n_M^1 n_M^2)^{p-1}) = \frac{M-1}{N}\mathsf{E}\bigg(\sum_{i\le N}\eta_{i,M}\langle\sigma_i^1(n_M^1)^{p-2}(n_M^2)^{p-1}\rangle\bigg) .$$

To make sense of this expression we have to perform approximate integration by parts. Some of the "main terms" occur because n_M^1 and n_M^2 depend on $\eta_{i,M}$. These are

$$\frac{M-1}{N}(p-2)\nu((n_M^1)^{p-3}(n_M^2)^{p-1}) + \frac{M-1}{N}(p-1)\nu(R_{1,2}(n_M^1)^{p-2}(n_M^2)^{p-2}) . \tag{10.252}$$

There are also "main terms" occurring because the Hamiltonian (and hence the corresponding the bracket) depend on $\eta_{i,M}$. These terms are

$$\begin{aligned}(M-1)\beta\bigg(&\sum_{\ell=1,2}\nu(R_{1,\ell}(n_M^1)^{p-2}(n_M^2)^{p-1}(n_M^\ell)^{p-1}\\ &- 2\nu(R_{1,3}(n_M^1)^{p-2}(n_M^2)^{p-1}(n_M^3)^{p-1})\bigg) .\end{aligned} \tag{10.253}$$

Since each factor n_M^ℓ counts for $N^{-1/2}$, these terms are of lower order. Indeed, they are of order $(M-1)N^{-(3p-4)/2}$, while the terms (10.252) are of order $(M-1)N^{-1}N^{-(2p-4)/2} = (M-1)N^{-p+1}$, and for $p > 2$ we have $3p-4 > 2(p-1)$. Therefore since $R_{1,2} \simeq q$ we should expect that

$$\begin{aligned}(M-1)\nu((n_M^1 n_M^2)^{p-1}) \simeq &\frac{M-1}{N}(p-2)\nu((n_M^1)^{p-3}(n_M^2)^{p-1})\\ &+ q\frac{M-1}{N}(p-1)\nu((n_M^1 n_M^2)^{p-2}) .\end{aligned}$$

To pursue the computation we now replace in each term one power of n_M^2 by $\sum_{i<N}\eta_{i,M}\sigma_i^2$ and we integrate by parts again. As before, the dependence of the Hamiltonian on $\eta_{i,M}$ plays essentially no role. In this manner we find recursion relations between the quantities

$$A_s = (M-1)\nu((n_M^1 n_M^2)^s) \; ; \; A_s^- = (M-1)\nu((n_M^1)^{s-2}(n_M^2)^s) ,$$

namely

$$A_s \simeq \frac{s-1}{N} A_s^- + \frac{sq}{N} A_{s-1} \tag{10.254}$$

$$A_s^- \simeq \frac{(s-2)q}{N} A_{s-1}^- + \frac{s-1}{N} A_{s-1} , \tag{10.255}$$

and we can compute these quantities by induction over s.

To describe the result of the computation, for two standard Gaussian r.v.s g_1 and g_2 with $q = \mathsf{E} g_1 g_2$, let us define the quantities

$$Q_s(q) = \mathsf{E}(g_1 g_2)^s \; ; \qquad Q_s^-(q) = \mathsf{E} g_1^s g_2^{s-2} = \mathsf{E} g_1^{s-2} g_2^s \; .$$

Writing $(g_1 g_2)^s = g_1(g_1^{s-1} g_2^s)$ and using integration by parts, we get the relation

$$Q_s(q) = (s-1) Q_s^-(q) + sq Q_{s-1}(q) \; . \tag{10.256}$$

Writing $g_1^s g_2^{s-2} = g_1(g_1^{s-1} g_2^{s-2})$ we get (for $s \geq 2$)

$$Q_s^-(q) = (s-2) q Q_{s-1}^-(q) + (s-1) Q_{s-2}(q) \; . \tag{10.257}$$

We can then compute these polynomials by induction:

$$Q_0(q) = 1 \; , \quad Q_1(q) = q \; , \quad Q_2^-(q) = 1 \; , \quad Q_2(q) = 1 + 2q^2$$

$$Q_3^-(q) = 2Q_1(q) + q Q_2^-(q) = 3q$$

$$Q_3(q) = 2Q_3^-(q) + 3q Q_2(q) = 6q + 3q + 6q^2 = 9q + 6q^2 \; .$$

Comparing the relations (10.254) and (10.255) with the relations (10.256) and (10.257) we see that

$$A_{p-1} \simeq \frac{M-1}{N^{p-1}} Q_{p-1}(q) \; .$$

Of course this simply formalizes the fact that in the computation of A_s we may pretend that the r.v.s $N^{1/2} n_M^1$ and $N^{1/2} n_M^2$ are standard Gaussian r.v.s with correlation q and that ν is expectation.

Since

$$r \simeq \nu \Bigg(\sum_{2 \leq k \leq M} (n_k^1 n_k^2)^{p-1} \Bigg) = (M-1) \nu ((n_M^1 n_M^2)^{p-1}) = A_{p-1} \; ,$$

the third of the "replica-symmetric" equations is

$$r = \alpha Q_{p-1}(q) \; ,$$

setting $\alpha = M/N^{p-1}$.

Research Problem 10.11.1. (Level 2) Study the p-spin Hopfield model for $p = 4$ with the same accuracy as for $p = 2$. In particular, find the shape of the "admissible region".

Research Problem 10.11.2. (Level 1) Are there fruitful variations on the formula (10.247)? For example, if γ is a given number, is the Hamiltonian

$$-H_{M,N}(\boldsymbol{\sigma}) = \frac{N^\gamma \beta}{p} \sum_{k \le M} m_k(\boldsymbol{\sigma})^p + Nhm_1(\boldsymbol{\sigma})$$

of interest? (The most interesting case should be when M is of order $N^{p-\gamma}$.)

Despite some similarities with the case of the usual Hopfield model the p-spin Hopfield model is rather different, and we will perform only some of the first steps necessary towards its understanding.

Let us first consider the case $M = 1$. We recall the function $\mathcal{I}$ of (A.22). As in Section 4.1 we see that

$$\lim_{N\to\infty} \mathsf{E} \log \sum_{\boldsymbol{\sigma}} \exp(-H_{N,1}(\boldsymbol{\sigma})) = \log 2 + \sup_t \left(\frac{\beta t^p}{t} + th - \mathcal{I}(t) \right) . \quad (10.258)$$

The supremum is obtained for a value t such that $\beta t^{p-1} + h = \mathcal{I}'(t)$, i.e.

$$t = \mathrm{th}(\beta t^{p-1} + h) .$$

It seems that for h small (the only case we will consider) and β large this equation has 3 solutions, the largest of which achieves the supremum in (10.258). We denote by m^* this solution, so that

$$m^* = \mathrm{th}(\beta (m^*)^{p-1} + h) .$$

We will now focus on the most interesting case, where M is a proportion of N^{p-1}. A first observation is that the quantity

$$d(p) := \mathsf{E} m_k(\boldsymbol{\sigma})^p \quad (10.259)$$

(that does not depend on the choice of $\boldsymbol{\sigma}$) is of order $N^{-p/2}$ (remember that we assume that p is even), so that $\sum_{1\le k\le M} \mathsf{E} m_k(\boldsymbol{\sigma})^p$ is about $\alpha N^{p/2-1}$ (where $\alpha = M/N^{p-1}$) and $N^{-1} \sum_{1\le k \le N} \mathsf{E} m_k(\boldsymbol{\sigma})^p$ is not bounded as $N \to \infty$. In the same line of thought, since $\mathsf{E}|m_k(\boldsymbol{\sigma})|^a$ is of order $N^{-a/2}$, if we want $\sum_{k\le M} |m_k(\boldsymbol{\sigma})|^a$ to have a chance to be of order 1, the natural choice of a is $a = 2p - 2$. For $\mathbf{x} = (x_k)_{k\le M}$ we consider the norm $|||\mathbf{x}||| = \left(\sum_{k\le M} x_k^{2(p-1)}\right)^{1/2(p-1)}$, and we recall that $\mathbf{m}(\boldsymbol{\sigma}) = (m_k(\boldsymbol{\sigma}))_{k\le M}$.

Theorem 10.11.3. *Let us fix p and $h \ge 0$, and β large enough. Then, given a number $s > 0$ we may find $c > 0$ such that if $M \le cN^{p-1}$ the set*

$$\big\{\boldsymbol{\sigma}\ ;\ \forall k \le M\ ,\ |||\mathbf{m}(\boldsymbol{\sigma}) \pm m^* \mathbf{e}_k||| \ge s\big\}$$

is negligible.

What this means is that the relevant configurations are such that for some $k \le M$ and some $\eta = \pm 1$ we have $|m_k(\boldsymbol{\sigma}) - \eta m^*| \le s$, and that all the other $m_{k'}(\boldsymbol{\sigma})$, $k' \ne k$, are small in the sense that $\sum_{k' \ne k} m_{k'}(\boldsymbol{\sigma})^{2(p-1)} \le s^{2(p-1)}$. The question as to whether for a given k $m_k(\boldsymbol{\sigma})$ is actually close to 0 is addressed at the end of this section.

Research Problem 10.11.4. In Theorem 10.11.3, what is the correct (order of) dependence of c as a function of s when $s \to 0$?

The strategy of proof of Theorem10.11.3 is rather interesting, and we will try to outline it before getting into the details. The issue is to prove that, typically, among the numbers $m_k(\boldsymbol{\sigma})$ only one is not small. Consider the quantities

$$u_p(\boldsymbol{\sigma}) = \sup_I \bigg(\sum_{k \in I} m_k(\boldsymbol{\sigma})^p\bigg)^{1/p}\ ,\ u_2(\boldsymbol{\sigma}) = \sup_I \bigg(\sum_{k \in I} m_k(\boldsymbol{\sigma})^2\bigg)^{1/2}\ . \tag{10.260}$$

There, the supremum is over all subsets I of $1, \dots M$ of cardinality n, a suitably chosen integer. We will show that typically $u_p(\boldsymbol{\sigma})$ and $u_2(\boldsymbol{\sigma})$ are very close to each other, and that this implies that only one of the numbers $m_k(\boldsymbol{\sigma})$ is not small. It is a general fact that $u_p(\boldsymbol{\sigma}) \le u_2(\boldsymbol{\sigma})$, and to prove that these quantities are close to each other we will bound $u_p(\boldsymbol{\sigma})$ from below and $u_2(\boldsymbol{\sigma})$ from above. The lower bound on $u_p(\boldsymbol{\sigma})$ is obtained by combining Lemma 10.11.5 and Proposition 10.11.9. The upper bound for $u_2(\boldsymbol{\sigma})$ is obtained through Lemma 10.11.5 and Proposition 10.11.10. The results are summarized in Proposition 10.11.11 below. The choice of the cardinality n of the sets I in (10.260) is guided by a simple idea. It helps us to bound $u_2(\boldsymbol{\sigma})$ from above if n is small. Fortunately, direct arguments show that typically there can be at most n of the quantities $m_k(\boldsymbol{\sigma})$ that are not small, where n is not too large. This is obtained in Lemmas 10.11.6 and 10.11.7.

A first step in the proof of Theorem 10.11.3 is to find a lower bound for the Hamiltonian. We can see a contribution to the Hamiltonian coming from the fact that it contains so many terms ($M = \alpha N^{p-1}$ terms) that are not very small (about $N^{-p/2}$). Another obvious contribution is from the fact that we might have $m_1 \simeq m^*$. To lighten notation we often write $\exp Na$ rather than $\exp(Na)$ when there is no ambiguity. We recall the notation $d(p)$ of (10.259).

Lemma 10.11.5. *If $M \le N^{p-1}$, then, with overwhelming probability,*

$$Z_{N,M} = \sum_{\boldsymbol{\sigma}} \exp(-H_{N,M}(\boldsymbol{\sigma}))$$
$$\geq \frac{2^N}{K\sqrt{N}} \exp N\left(\frac{\beta}{p} m^{*p} + hm^* - \mathcal{I}(m^*) + \frac{\beta}{p} M d(p)\right). \quad (10.261)$$

Proof. Since $m_1(\boldsymbol{\sigma}) = N^{-1}\sum_{i\leq N}\sigma_i$, given an integer b the trivial bound

$$Z_{N,M} \geq \operatorname{card}\left\{\boldsymbol{\sigma} \; ; \sum_{i\leq N}\sigma_i = b\right\} \exp N\left(\frac{\beta}{p}\left(\frac{b}{N}\right)^p + h\frac{b}{N}\right.$$
$$\left. + \frac{\beta}{p}(M-1)d(p) + \frac{\beta}{p}\inf_{\boldsymbol{\sigma}} \sum_{2\leq k\leq M}(m_k(\boldsymbol{\sigma})^p - d(p))\right) \quad (10.262)$$

holds. The key point of the proof is to show that with overwhelming probability

$$\forall \boldsymbol{\sigma} \in \Sigma_N \;, \qquad \sum_{2\leq k\leq M}(m_k(\boldsymbol{\sigma})^p - d(p)) \geq -K(p) \;. \quad (10.263)$$

Once this is proved (10.262) implies that with overwhelming probability:

$$Z_{N,M} \geq \frac{1}{K}\operatorname{card}\left\{\boldsymbol{\sigma} \; ; \sum_{i\leq N}\sigma_i = b\right\} \exp N\left(\frac{\beta}{p}\left(\frac{b}{N}\right)^p + h\frac{b}{N} + \frac{\beta}{p}Md(p)\right)$$
$$\geq \frac{2^N}{K\sqrt{N}} \exp N\left(\frac{\beta}{p}\left(\frac{b}{N}\right)^p + h\frac{b}{N} - \mathcal{I}\left(\frac{b}{N}\right) + \frac{\beta}{p}Md(p)\right) ,$$

using (A.25) in the last line. Finally we choose b so that $|b/N - m^*| \leq 1/N$ to complete the proof of (10.261).

We turn to the proof of (10.263). Let us fix $\boldsymbol{\sigma}$ and define

$$Y_k = \min\left(m_k(\boldsymbol{\sigma})^p, \frac{1}{N}\right)$$

and $X_k = Y_k - \mathsf{E}Y_k$. Thus

$$|X_k| \leq \frac{1}{N} \;;\; \mathsf{E}X_k^2 \leq \mathsf{E}Y_k^2 \leq \frac{K(p)}{N^p} \;.$$

Then Bernstein's inequality (A.34) implies

$$\mathsf{P}\left(\left|\sum_{k\leq M} X_k\right| \geq t\right) \leq 2\exp\left(-\frac{1}{L}\min\left(\frac{t^2}{M\mathsf{E}X_k^2}, Nt\right)\right) .$$

Since $M \leq N^{p-1}$ and $\mathsf{E}X_k^2 \leq K(p)N^{-p}$ we get

$$M\mathsf{E}(X_k^2) \leq N^{p-1}K(p)N^{-p} \leq \frac{K(p)}{N} ,$$

hence

$$\mathsf{P}\left(\left|\sum_{k\le M} X_k\right| \ge t\right) \le 2\exp\left(-\frac{1}{K(p)}\min(Nt^2, Nt)\right) ,$$

and therefore there exists $K(p)$ such that, with overwhelming probability, for each $\boldsymbol{\sigma}$,

$$\left|\sum_{k\le M} X_k\right| \le K(p) .$$

Setting $d = d(N,p) = \mathsf{E}Y_k$ (which does not depend on $\boldsymbol{\sigma}$ or k) then with overwhelming probability, for each $\boldsymbol{\sigma}$,

$$\sum_{k\le M}(Y_k - d) \ge -K(p) ,$$

and therefore

$$\begin{aligned}\sum_{k\le M}(m_k(\boldsymbol{\sigma})^p - d(p)) &\ge \sum_{k\le M}(Y_k - d(p)) \\ &= \sum_{k\le M}(Y_k - d) + M(d - d(p)) \\ &\ge -K(p) + M(d - d(p)) ,\end{aligned}$$

and to conclude the proof it suffices to show by using (A.32) (and that $p \ge 3$) that $|d - d(p)| \le K\exp(-N/K)$. □

Let us define the integer j_0 as the largest one for which

$$2^{-j_0} \ge \left(\frac{8\log(2N^{p-1})}{N}\right)^{1/2} ,$$

so 2^{-j_0} is somewhat larger than $N^{-1/2}$, the typical size of $m_k(\boldsymbol{\sigma})$. A first step towards Theorem 10.11.3 is the proof of the following: if we consider only those terms $m_k(\boldsymbol{\sigma})$ that are not larger than 2^{-j_0}, the contribution of these terms to the sum $\sum_{k\le M} m_k(\boldsymbol{\sigma})^p$ is not much more than $Md(p)$.

Lemma 10.11.6. *Consider a number $c > 0$. Then if*

$$\alpha = \frac{M}{N^{p-1}} \le c ,$$

with overwhelming probability it is true that for each $\boldsymbol{\sigma} \in \Sigma_N$

$$\sum_{2\le k\le M} m_k(\boldsymbol{\sigma})^p \mathbf{1}_{\{|m_k(\boldsymbol{\sigma})|\le 2^{-j_0}\}} \le Md(p) + K(p)\sqrt{c} \tag{10.264}$$

and

$$\sum_{2\le k\le M} m_k(\boldsymbol{\sigma})^{2(p-1)} \mathbf{1}_{\{|m_k(\boldsymbol{\sigma})|\le 2^{-j_0}\}} \le K(p)c . \tag{10.265}$$

Proof. We first prove (10.264). Let us fix $\boldsymbol{\sigma}$ and for $2 \le k \le M$, let us consider the r.v.s

$$Y_k = m_k(\boldsymbol{\sigma})^p \mathbf{1}_{\{|m_k(\boldsymbol{\sigma})| \le 2^{-j_0}\}} \; ; \qquad X_k = Y_k - \mathsf{E}Y_k \; .$$

The r.v.s X_k are independent, centered, and satisfy $|X_k| \le 2^{-pj_0}$,

$$\mathsf{E}X_k^2 \le \mathsf{E}m_k(\boldsymbol{\sigma})^{2p} \le \frac{K(p)}{N^p} \; .$$

Thus, Bernstein's inequality (A.34) implies that for $t > 0$,

$$\begin{aligned}\mathsf{P}\left(\left|\sum_{2 \le k \le M} X_k\right| \ge t\right) &\le 2\exp\left(-\frac{1}{L}\min\left(\frac{N^p t^2}{K(p)M}, \frac{t}{2^{-pj_0}}\right)\right)\\ &\le 2\exp\left(-\frac{1}{L}\min\left(\frac{N t^2}{K(p)c}, \frac{t}{2^{-pj_0}}\right)\right)\end{aligned}$$

since $M \le cN^{p-1}$. Now $2^{-j_0} \le K(p)N^{-1/3}$, so, since $p \ge 3$ we have $2^{-pj_0} \le K(p)/N$, and taking $t = K(p)\sqrt{c}$ we obtain

$$\mathsf{P}\left(\left|\sum_{2 \le k \le M} X_k\right| \ge K(p)\sqrt{c}\right) \le 2\exp(-N)$$

and hence

$$\mathsf{P}\left(\sum_{2 \le k \le M} Y_k \ge M\mathsf{E}Y_1 + K(p)\sqrt{c}\right) \le 2\exp(-N) \; .$$

It follows that with overwhelming probability

$$\forall \boldsymbol{\sigma} \; , \qquad \sum_{2 \le k \le M} Y_k \le M\mathsf{E}Y_1 + K(p)\sqrt{c}$$

and since $\mathsf{E}Y_1 \le d(p)$ this proves (10.264). To prove (10.265) we proceed similarly, with now

$$Y_k = m_k(\boldsymbol{\sigma})^{2(p-1)} \mathbf{1}_{\{|m_k(\boldsymbol{\sigma})| \le 2^{-j_0}\}} \; ,$$

so that

$$\mathsf{E}Y_k \le \frac{K(p)}{N^{p-1}} \; ; \qquad \mathsf{E}X_k^2 \le \mathsf{E}Y_k^2 \le \frac{K(p)}{N^{2(p-1)}} \; .$$

Bernstein's inequality implies

$$\mathsf{P}\left(\left|\sum_{2 \le k \le M} X_k\right| \ge t\right) \le \exp\left(-\frac{1}{L}\min\left(\frac{t^2 N^{2(p-1)}}{K(p)M}, t2^{2j_0(p-1)}\right)\right) \; ,$$

and since $N^{2(p-1)}/M \geq N^{2(p-1)}$ and

$$2^{2j_0(p-1)} \geq N^{2(p-1)/3}/K(p) \geq N^{4/3}/K(p)$$

it follows (with plenty of room) that with overwhelming probability we have $\left|\sum_{2\leq k\leq M} X_k\right| \leq c$ for each $\boldsymbol{\sigma}$ and therefore

$$\sum_{2\leq k\leq M} Y_k \leq M\mathsf{E}Y_1 + c \leq cK(p) + c\,. \qquad \square$$

Now we study the contribution to the Hamiltonian of the quantities $m_k(\boldsymbol{\sigma})$ that are quite larger than $N^{-1/2}$. There are surprisingly few of then, as the next lemma shows.

Lemma 10.11.7. *Assume $j \leq j_0$. Then with overwhelming probability, for each $\boldsymbol{\sigma}$ in Σ_N,*

$$\operatorname{card}\{2 \leq k \leq M\ ;\ |m_k(\boldsymbol{\sigma})| \geq 2^{-j}\} < 2^{2j+2}\,. \tag{10.266}$$

Proof. Given $\boldsymbol{\sigma}$ and k, the subgaussian inequality (A.16) implies

$$\mathsf{P}\left(|m_k(\boldsymbol{\sigma})| \geq 2^{-j}\right) \leq 2\exp\left(-\frac{N2^{-2j}}{2}\right)\,.$$

Consider a subset I of $\{1,\dots,M\}$, with $n = \operatorname{card} I$. By independence, given $\boldsymbol{\sigma}$, we have

$$\mathsf{P}\left(\forall\, k\in I,\ |m_k(\boldsymbol{\sigma})| \geq 2^{-j}\right) \leq 2^n \exp\left(-\frac{nN2^{-2j}}{2}\right)\,.$$

Moreover,

$$\operatorname{card}\{k\,;\ |m_k(\boldsymbol{\sigma})| \geq 2^{-j}\} \geq n \Rightarrow \exists I\ ,\ \operatorname{card} I = n\,,\ \forall k\in I\,,\ |m_k(\boldsymbol{\sigma})| \geq 2^{-j}\,.$$

Thus, since there are at most M^n choices for I, we get

$$\mathsf{P}\left(\exists\boldsymbol{\sigma}\in\Sigma_N,\ \operatorname{card}\{k\,;\ |m_k(\boldsymbol{\sigma})| \geq 2^{-j}\} \geq n\right) \leq 2^N M^n\, 2^n \exp\left(-\frac{nN2^{-2j}}{2}\right)$$
$$\leq \exp\left(N\log 2 + n\log(2M) - \frac{nN2^{-2j}}{2}\right)\,.$$

If $n = 2^{2j+2}$, and if $j \leq j_0$, then $\log(2M) \leq \log(2N^{p-1}) \leq 2^{-2j}N/8$, and the exponent in the last term is at most

$$N\log 2 + n\left(\log(2M) - \frac{N2^{-2j}}{2}\right) \leq N - \frac{3nN2^{-2j}}{8} \leq N - \frac{3N}{2} = -\frac{N}{2}\,.$$

This finishes the proof. $\square$

Corollary 10.11.8. *With overwhelming probability, given $0 \le j_1 < j_0$, for each $\boldsymbol{\sigma}$, we have*

$$\sum_{2\le k\le M} m_k(\boldsymbol{\sigma})^p \mathbf{1}_{\{2^{-j_0}\le |m_k(\boldsymbol{\sigma})|\le 2^{-j_1}\}} \le K(p) 2^{-(p-2)j_1} \tag{10.267}$$

$$\sum_{2\le k\le M} m_k(\boldsymbol{\sigma})^{2(p-1}\mathbf{1}_{\{2^{-j_0}\le |m_k(\boldsymbol{\sigma})|\le 2^{-j_1}\}} \le K(p) 2^{-2(p-2)j_1}. \tag{10.268}$$

Proof. With overwhelming probability (10.266) holds for all $0 \le j \le j_0$. To prove (10.267) we simply write that then

$$\sum_{2\le k\le M} m_k(\boldsymbol{\sigma})^p \mathbf{1}_{\{2^{-j}\le |m_k(\boldsymbol{\sigma})|\le 2^{-j+1}\}} \le 2^{2j+2}(2^{-j+1})^p ,$$

and that, since $p \ge 3$, the sum $\sum_{j_1\le j<j_0} 2^{-j(p-2)}$ behaves like the first term. We proceed in a similar manner for (10.268). □

We are now ready to show that the contribution to the Hamiltonian (besides the "bulk contribution" of those many $m_k(\boldsymbol{\sigma})$'s that are not large) comes entirely from very few of the largest terms of the sequence $(m_k(\boldsymbol{\sigma}))$.

Proposition 10.11.9. *Consider a number $0 < c < 1$ and the smallest integer j_1 such that*

$$2^{-(p-2)j_1} \le \sqrt{c} .$$

Let us define $u_p(\boldsymbol{\sigma})^p$ as the sum of the 2^{2j_1+2} largest terms of the sequence $m_k(\boldsymbol{\sigma})^p$. Equivalently, let us

$$\textit{define } \mathcal{C}(j) \textit{ as the collection of subsets } I \subset \{1,\cdots,M\} \textit{ with } \mathrm{card} I = 2^{2j+2} , \tag{10.269}$$

and

$$u_p(\boldsymbol{\sigma})^p = \sup_{I\in\mathcal{C}(j_1)} \sum_{k\in I} m_k(\boldsymbol{\sigma})^p . \tag{10.270}$$

Then, with overwhelming probability,

$$\forall\boldsymbol{\sigma} , -H_{M,N}(\boldsymbol{\sigma}) \le \frac{\beta N}{p} Md(p) + \frac{\beta N}{p} u_p(\boldsymbol{\sigma})^p + Nhu_p(\boldsymbol{\sigma}) + N\beta K(p)\sqrt{c} . \tag{10.271}$$

Proof. For large N we have $j_1 \le j_0$. Combining Lemma 10.11.6 and Corollary 10.11.8 implies that with overwhelming probability

$$\begin{aligned}\sum_{2\le k\le M} m_k(\boldsymbol{\sigma})^p \mathbf{1}_{\{|m_k(\boldsymbol{\sigma})|\le 2^{-j_1}\}} &\le Md(p) + K(p)2^{-(p-2)j_1} \\ &\le Md(p) + K(p)\sqrt{c} ,\end{aligned}$$

by the choice of j_1. Therefore

$$\sum_{2\le k\le M} m_k(\boldsymbol{\sigma})^p \le Md(p) + K(p)\sqrt{c} + \sum_{k\in J} m_k(\boldsymbol{\sigma})^p \ , \tag{10.272}$$

where $J = \{2 \le k \le M \ ; \ |m_k(\boldsymbol{\sigma})| > 2^{-j_1}\}$. By Lemma 10.11.7 with overwhelming probability we have $\mathrm{card} J < 2^{2j_1+2}$, so that (10.272) implies that if we set $I = J \cup \{1\}$ then, since $\mathrm{card} I \le 2^{2j_1+2}$,

$$\begin{aligned}\sum_{1\le k\le M} m_k(\boldsymbol{\sigma})^p &\le Md(p) + K(p)\sqrt{c} + \sum_{k\in I} m_k(\boldsymbol{\sigma})^p \\ &\le Md(p) + K(p)\sqrt{c} + u_p(\boldsymbol{\sigma})^p \ ,\end{aligned}$$

and (10.271) follows because $Nhm_1(\boldsymbol{\sigma}) \le Nhu_p(\boldsymbol{\sigma})$ since $m_1(\boldsymbol{\sigma}) \le u_p(\boldsymbol{\sigma})$. □

Defining $u_p(\boldsymbol{\sigma})$ as in Proposition 10.11.9, it will obviously help to control for how many values of $\boldsymbol{\sigma}$ we can have $u_p(\boldsymbol{\sigma}) \ge t$. It turns out that this is not so easy, so we will replace $u_p(\boldsymbol{\sigma})$ by the larger quantity (recalling (10.269))

$$u_2(\boldsymbol{\sigma}) = \sup_{I\in\mathcal{C}(j_1)} \left(\sum_{k\in I} m_k(\boldsymbol{\sigma})^2\right)^{1/2} . \tag{10.273}$$

To see that $u_p(\boldsymbol{\sigma}) \le u_2(\boldsymbol{\sigma})$ we simply notice that for numbers $b_k \ge 0$, since $p \ge 2$, we have

$$\sum_{k\in I} b_k^{p/2} \le \left(\sum_{k\in I} b_k\right)^{p/2} \tag{10.274}$$

so that taking $b_k = m_k^2(\boldsymbol{\sigma})$ proves the inequality. Of course the reader might wonder how using something as crude as (10.274) can be useful; but (10.274) is a near equality in the case where all the numbers b_k but one are very small, which turn out to be the case here (as shown by Theorem 10.11.3).

Proposition 10.11.10. *There exists a constant L with the following property. Consider $\varepsilon > 0$ and an integer n. Assume*

$$Ln \log\left(\frac{LM}{n}\right) \le N\varepsilon^2 \ , \tag{10.275}$$

and define $u_2(\boldsymbol{\sigma})$ by (10.273), where the supremum is taken over all choices of I with $\mathrm{card} I \le n$. Then with overwhelming probability we have

$$\forall \boldsymbol{\sigma} \ , u_2(\boldsymbol{\sigma}) \le \sqrt{1+\varepsilon} \tag{10.276}$$

and for each $0 \le t \le \sqrt{1+\varepsilon}$ it holds

$$\mathrm{card}\{\boldsymbol{\sigma} \ ; \ u_2(\boldsymbol{\sigma}) \ge t\} \le 2^N N^n \left(1 + \frac{1}{\varepsilon}\right)^n \exp(-N\mathcal{I}((1-3\varepsilon)t)) \ . \tag{10.277}$$

Proof. We have shown in Proposition A.8.3 that with overwhelming probability, for every set $I \subset \{1, \dots, M\}$ with $n = \mathrm{card} I$ and any numbers $(x_k)_{k \in I}$, then

$$\sum_{i \le N} \Big(\sum_{k \in I} x_k \eta_{i,k} \Big)^2 \le N(1+\varepsilon) \sum_{k \in I} x_k^2 \; . \tag{10.278}$$

In the remainder of the proof we show that (10.278) implies (10.276) and (10.277). For this we will prove that for any subset I of $\{1, \dots, M\}$ with $\mathrm{card} I = n$, we have

$$\forall \boldsymbol{\sigma} \; , \qquad \sum_{k \in I} m_k^2(\boldsymbol{\sigma}) \le 1 + \varepsilon \tag{10.279}$$

and

$$\begin{gathered} \forall t \; , \quad 0 \le t \le \sqrt{1+\varepsilon} \quad , \\ \mathrm{card}\Big\{ \boldsymbol{\sigma} \; ; \; \sum_{k \in I} m_k^2(\boldsymbol{\sigma}) \ge t^2 \Big\} \le 2^N \Big(1 + \frac{1}{\varepsilon} \Big)^n \exp(-N\mathcal{I}((1-3\varepsilon)t) \; . \end{gathered} \tag{10.280}$$

To see that this suffices, we first note that (10.279) obviously implies (10.276). Moreover

$$\begin{aligned} \mathrm{card}\{ \boldsymbol{\sigma} \; ; \; u_2(\boldsymbol{\sigma}) \ge t \} &= \mathrm{card}\Big\{ \boldsymbol{\sigma} \; ; \; \sup_{\mathrm{card} I \le n} \sum_{k \in I} m_k^2(\boldsymbol{\sigma}) \ge t^2 \Big\} \\ &\le \sum_{\mathrm{card} I \le n} \mathrm{card}\Big\{ \boldsymbol{\sigma} \; ; \; \sum_{k \in I} m_k^2(\boldsymbol{\sigma}) \ge t^2 \Big\} \; , \end{aligned}$$

and since there are at most N^n choices for I, (10.280) implies (10.277).

We turn to the proof of (10.279) and (10.280). Consider numbers $(\alpha_k)_{k \in I}$, and note the identity

$$\sum_{k \in I} \alpha_k m_k(\boldsymbol{\sigma}) = \sum_{i \le N} a_i \sigma_i \; , \tag{10.281}$$

where $a_i = N^{-1} \sum_{k \in I} \alpha_k \eta_{i,k}$, so by (10.278) this entails

$$\sum_{i \le N} a_i^2 \le \frac{1}{N}(1+\varepsilon) \sum_{k \in I} \alpha_k^2 \; . \tag{10.282}$$

When $\sum_{k \in I} \alpha_k^2 \le 1$, (10.281), the Cauchy-Schwarz inequality and (10.282) imply

$$\sum_{k \in I} \alpha_k m_k(\boldsymbol{\sigma}) = \sum_{i \le N} a_i \sigma_i \le \sqrt{N} \Big(\sum_{i \le N} a_i^2 \Big)^{1/2} \le \sqrt{1+\varepsilon} \; . \tag{10.283}$$

The choice

$$\alpha_k = \frac{m_k(\boldsymbol{\sigma})}{\sum_{k' \le M} m_{k'}(\boldsymbol{\sigma})^2}$$

then proves (10.279). Also, it follows from Proposition A.7.1 that there exists a set $\mathcal{A}$ of sequences $\boldsymbol{\alpha} = (\alpha_k)_{k\in I}$ such that $\sum_{k\in I} \alpha_k^2 \le 1$, with card$\mathcal{A} \le (1+1/\varepsilon)^n$ and

$$\sup_{\boldsymbol{\alpha}\in\mathcal{A}} \sum_{k\in I} \alpha_k m_k(\boldsymbol{\sigma}) \ge (1-2\varepsilon)\sqrt{\sum_{k\in I} m_k(\boldsymbol{\sigma})^2} \; .$$

Thus

$$\mathrm{card}\Big\{\boldsymbol{\sigma} \; ; \; \sum_{k\in I} m_k^2(\boldsymbol{\sigma}) \ge t^2\Big\} \le \sum_{\boldsymbol{\alpha}\in\mathcal{A}} \mathrm{card}\Big\{\boldsymbol{\sigma} \; ; \; \sum_{k\in I} \alpha_k m_k(\boldsymbol{\sigma}) \ge (1-2\varepsilon)t\Big\} \; . \tag{10.284}$$

Now (10.281) and (A.26) imply

$$\begin{aligned} &\mathrm{card}\Big\{\boldsymbol{\sigma} \; ; \; \sum_{k\in I} \alpha_k m_k(\boldsymbol{\sigma}) \ge (1-2\varepsilon)t\Big\} \\ &= \mathrm{card}\Big\{\boldsymbol{\sigma} \; ; \; \sum_{i\le N} \sqrt{N} a_i \sigma_i \ge (1-2\varepsilon)t\sqrt{N}\Big\} \\ &\le 2^N \exp\left(-N\mathcal{I}\left(\frac{(1-2\varepsilon)t}{N\sum_{i\le N} a_i^2}\right)\right) \\ &\le 2^N \exp(-N\mathcal{I}(1-3\varepsilon)t) \end{aligned} \tag{10.285}$$

since $N\sum_{i\le N} a_i^2 \le 1+\varepsilon$ and $(1-2\varepsilon)/(1+\varepsilon) \ge 1-3\varepsilon$. Combining (10.285) with (10.284) proves (10.280). □

Proposition 10.11.11. *Consider $w > 0$. we may find $c > 0$ small enough such that if j_1 is the smallest integer with $2^{-(p-2)j_1} \le \sqrt{c}$, and $u_p(\boldsymbol{\sigma})$ and $u_2(\boldsymbol{\sigma})$ are given respectively by (10.270) and (10.273), then for $M \le cN^{p-1}$ the sets*

$$B_1 = \{\boldsymbol{\sigma} \; ; \; u_p(\boldsymbol{\sigma}) \le m^* - w\} \quad \text{and} \quad B_2 = \{\boldsymbol{\sigma} \; ; \; u_2(\boldsymbol{\sigma}) \ge m^* + w\}$$

are negligible.

Proof. For any set $B \subset \mathbb{R}^N$ we have

$$G_{N,M}(B) = \frac{S(B)}{Z_{N,M}} \; ,$$

where

$$S(B) = \sum_{\boldsymbol{\sigma}\in B} \exp(-H_{N,M}(\boldsymbol{\sigma})) \; .$$

A lower bound for $Z_{N,M}$ is provided by Lemma 10.11.5, so we try to find an upper bound for $S(B)$. As a start we observe that the key property of m^* is that the function $t \mapsto f(t) - \mathcal{I}(t)$ attains its maximum at $t = m^*$, where we define

$$f(t) = \frac{\beta}{p} t^p + th \ . \tag{10.286}$$

Therefore we may find $\delta > 0$ and $\varepsilon > 0$ such that

$$0 \le t \le m^* - w \Rightarrow f(t) - \mathcal{I}((1-3\varepsilon)t) \le f(m^*) - \mathcal{I}(m^*) - \delta \tag{10.287}$$

$$m^* + w \le t \le \sqrt{1+\varepsilon} \Rightarrow f(t) - \mathcal{I}((1-3\varepsilon)t) \le f(m^*) - \mathcal{I}(m^*) - \delta \ . \tag{10.288}$$

To bound $S(B)$ we use that (10.271) holds with overwhelming probability. Thus, with overwhelming probability, using again the notation (10.286),

$$S(B) \le \exp\left(\frac{\beta N}{p} M d(p) + N\beta K(p)\sqrt{c}\right) \sum_{\boldsymbol{\sigma} \in B} \exp N f(u_p(\boldsymbol{\sigma})) \ . \tag{10.289}$$

By (A.32) we have (since $u_p(\boldsymbol{\sigma}) \le m^* - w$ for $\boldsymbol{\sigma} \in B_1$)

$$\begin{aligned} &\sum_{\boldsymbol{\sigma} \in B_1} \exp N f(u_p(\boldsymbol{\sigma})) \qquad (10.290) \\ &\le \mathrm{card} B_1 + \int_0^{m^*-w} N f'(t) \exp(N f(t)) \mathrm{card}\{\boldsymbol{\sigma} \ ; \ u_p(\boldsymbol{\sigma}) \ge t\} \mathrm{d}t \\ &\le \mathrm{card} B_1 + \int_0^{m^*-w} N f'(t) \exp(N f(t)) \mathrm{card}\{\boldsymbol{\sigma} \ ; \ u_2(\boldsymbol{\sigma}) \ge t\} \mathrm{d}t \ , \end{aligned}$$

because $u_2(\boldsymbol{\sigma}) \ge u_p(\boldsymbol{\sigma})$. Now we observe that (10.275) holds for N large enough if $n = 2^{2j_1+2}$, and in that case the quantity $u_2(\boldsymbol{\sigma})$ of Proposition 10.11.10 is the same as the quantity (10.273), so that (10.277) holds with overwhelming probability. Using (10.277), we get, for a certain number K independent of N,

$$\begin{aligned} \exp N f(t) \mathrm{card}\{\boldsymbol{\sigma} \ ; \ u_2(\boldsymbol{\sigma}) \ge t\} &\le 2^N K N^n \exp N(f(t) - \mathcal{I}((1-3\varepsilon)t)) \\ &\le 2^N K N^n \exp N(f(m^*) - \mathcal{I}(m^*) - \delta) \ , \end{aligned}$$

using (10.287) in the second line. Since $f(m^*) - \mathcal{I}(m^*) > f(0) - \mathcal{I}(0) = 0$, we may assume without loss of generality that $f(m^*) - \mathcal{I}(m^*) - \delta > 0$. Since $\mathrm{card} B_1 \le 2^N$, (10.290) implies,

$$\sum_{\boldsymbol{\sigma} \in B_1} \exp N f(u_p(\boldsymbol{\sigma})) \le K 2^N N^{n+1} \exp N(f(m^*) - \mathcal{I}(m^*) - \delta) \ .$$

Combining with (10.289) we have with overwhelming probability

$$S(B_1) \le K 2^N N^{n+1} \exp N\left(\frac{\beta}{p} M d(p) + \beta K(p)\sqrt{c} + f(m^*) - \mathcal{I}(m^*) - \delta\right) \ .$$

Combining with Lemma 10.11.5 yields that with overwhelming probability

$$G_{N,M}(B_1) \le N^K \exp N(\beta K(p)\sqrt{c} - \delta) ,$$

so that if we have chosen c small enough, B_1 is negligible. In the case of B_2 we conclude exactly as before. □

Now the idea is to prove that the only way we may have $u_p(\boldsymbol{\sigma}) \simeq u_2(\boldsymbol{\sigma}) \simeq m^*$ is that there exists k with $m_k(\boldsymbol{\sigma}) \simeq \pm m^*$ and all the other values of $m_k(\boldsymbol{\sigma})$ are small.

Proof of Theorem 10.11.3. Consider $w > 0$, to be chosen later. Let c be as in Proposition 10.11.11. Consider $\boldsymbol{\sigma}$ in Σ_N and $u_p(\boldsymbol{\sigma})$, $u_2(\boldsymbol{\sigma})$ as in Proposition 10.11.11. Assuming that

$$m^* - w \le u_p(\boldsymbol{\sigma}) \le u_2(\boldsymbol{\sigma}) \le m^* + w , \tag{10.291}$$

we will show that $\mathbf{m}(\boldsymbol{\sigma})$ is close to $\pm m^* \mathbf{e}_k$ for some k. Consider a set I with $\mathrm{card} I \le 2^{2j_1+2}$ such that

$$u_p(\boldsymbol{\sigma})^p = \sum_{k \in I} m_k(\boldsymbol{\sigma})^p .$$

Let $S = \max_{k\in I} |m_k(\boldsymbol{\sigma})|$. Then

$$u_p(\boldsymbol{\sigma})^p \le S^{p-2} \sum_{k\in I} m_k(\boldsymbol{\sigma})^2 \le S^{p-2} u_2(\boldsymbol{\sigma})^2$$

and thus (10.291) entails

$$S \ge \left(\frac{(m^* - w)^p}{(m^* + w)^2}\right)^{\frac{1}{p-2}} \ge m^* - wK(p,\beta,h) \tag{10.292}$$

for w small enough, with $K(p,\beta,h)$ depending only on p, β and h. Let $k_0 \in I$ with $S = |m_{k_0}(\boldsymbol{\sigma})|$. Then

$$S^2 + \sum_{k\in I, k\ne k_0} m_k(\boldsymbol{\sigma})^2 \le u_2(\boldsymbol{\sigma})^2 \le (m^* + w)^2$$

and therefore using (10.292) we get

$$\sum_{k\in I, k\ne k_0} m_k(\boldsymbol{\sigma})^2 \le K(p,\beta,h) w .$$

Since $p - 1 \ge 1$ and $|m_k(\boldsymbol{\sigma})| \le 1$ it follows that

$$\sum_{k\in I, k\ne k_0} m_k(\boldsymbol{\sigma})^{2(p-1)} \le K(p,\beta,h) w . \tag{10.293}$$

We recall that $2^{-2(p-2)j_1} \le c$, so that combining (10.265) and (10.268) proves that with overwhelming probability

$$\sum_{2 \le k \le M} m_k(\boldsymbol{\sigma})^{2(p-1)} \mathbf{1}_{\{|m_k(\boldsymbol{\sigma})| \le 2^{-j_1}\}} \le K(p)c \,. \tag{10.294}$$

Now, it follows from (10.266) that $k \in I$ whenever $|m_k(\boldsymbol{\sigma})| > 2^{-j_1}$, so that combining (10.293) and (10.294), we get

$$\sum_{k \ne k_0} m_k(\boldsymbol{\sigma})^{2(p-1)} \le K(p, \beta, h)(w + c) \,.$$

Moreover, using (10.292), and since $S = |m_{k_0}(\boldsymbol{\sigma})| \le u_2(\boldsymbol{\sigma}) \le m^* + w$ we have $||m_{k_0}(\boldsymbol{\sigma})| - m^*| \le K(p, \beta, h)w$. It should then be obvious that if w and c have been chosen small enough, for some $\eta = \pm 1$ we have $|||\mathbf{m}(\boldsymbol{\sigma}) - \eta \mathbf{e}_{k_0}||| \le s$. □

We will leave it to the reader to check as in Section 4.4 that when $h > 0$ among all the possibilities left open by Theorem 10.11.3, the one that always occurs is that $m_1(\boldsymbol{\sigma})$ is close to m^*. As a consequence, the following holds:

Proposition 10.11.12. *Assume $h > 0$. Then given $s > 0$ we may find $c > 0$ such that if $M \le cN^{p-1}$ the set $\{\boldsymbol{\sigma} \,;\, |m_1(\boldsymbol{\sigma}) - m^*| > s\}$ is negligible.*

The last result we will prove concerns the control of the quantities $\nu(m_k^{2n})$ for $k \ge 2$ and n not too large. This control is essential to perform approximate integration by parts.

When β is small we have

$$0 < t < 1 \quad \Rightarrow \quad \frac{\beta t^p}{p} - \mathcal{I}(t) \le 0 \,. \tag{10.295}$$

Define β_0 as the largest value of β for which (10.295) occurs, so that

$$0 < t < 1 \quad \Rightarrow \quad \frac{\beta_0 t^p}{p} - \mathcal{I}(t) \le 0 \,. \tag{10.296}$$

We expect that for $\beta > \beta_0$ the maximum of $\beta t^p/p - \mathcal{I}(t)$ is attained at the largest root of the equation $t = \operatorname{th}(\beta t^{p-1})$ (although we did not muster the energy to prove it).

Proposition 10.11.13. *Assume that*

$$\beta(1 - m^*)^{p/2} < \beta_0 \,. \tag{10.297}$$

Then there is $c > 0$ such that if $\alpha \le c$, we have

$$\forall\, n \le \frac{N^{1-2/p}}{2 \log N} \,, \; \mathsf{E}\, \langle m_M^{2n} \rangle \le \left(\frac{Kn}{N} \right)^n$$

where K depends only on β, h, p.

This provides a very good control for many moments of m_M. We do not know however if condition (10.297) always holds when $\beta > \beta_0$. It certainly holds when β is large enough because, as in the case $p = 2$, m^* approaches 1 exponentially fast as $\beta \to \infty$.

Lemma 10.11.14. *Consider numbers* $(a_i)_{i \le N}$ *with*

$$\left(\frac{1}{N}\sum_{i\le N} a_i^2\right)^{p/2} < \beta_0 \ .$$

Then we have

$$\mathsf{E}\exp\frac{N}{p}\left(\frac{1}{N}\sum_{i\le N}\eta_{i,M}\, a_i\right)^p \le K$$

where K *depends only on* $N^{-1}\sum_{i\le N} a_i^2$.

Proof. Define $a = (N^{-1}\sum_{i\le N} a_i^2)^{1/2}$. We have, by (A.31), and using (A.26) in the last line,

$$\begin{aligned}
&\mathsf{E}\exp\frac{N}{p}\left(\frac{1}{N}\sum_{i\le N}\eta_{i,M}\, a_i\right)^p \\
&= 1 + \int_0^\infty N t^{p-1}\exp\frac{N t^p}{p}\ \mathsf{P}\left(\sum_{i\le N}\eta_{i,M}\, a_i \ge tN\right)\mathrm{d}t \\
&\le 1 + N\int_0^a t^{p-1}\exp N\left(\frac{t^p}{p} - \mathcal{I}\left(\frac{t}{a}\right)\right)\mathrm{d}t \ .
\end{aligned}$$

By change of variables, the integral is

$$N\, a^p\int_0^1 t^{p-1}\exp N\left(\frac{a^p\, t^p}{p} - \mathcal{I}(t)\right)\mathrm{d}t \ . \tag{10.298}$$

Now (10.296) implies $a^p t^p/p - \mathcal{I}(t) \le -(\beta_0 - a^p)t^p/p$, and since $a^p < \beta_0$, change of variable in (10.298) proves that this integral is bounded independently of N. □

Proof of Proposition 10.11.13. We define $\Omega_0 = \left\{\left|\sum_{i\le N}\eta_{i,M}\right| \le N^{1-1/p}\right\}$ and we write

$$\mathsf{E}\,\langle m_M^{2n}\rangle \le \mathrm{I} + \mathrm{II} \tag{10.299}$$

where

$$\begin{aligned}
\mathrm{I} &= \mathsf{E}\,(\mathbf{1}_{\Omega_0}\,\langle m_M^{2n}\rangle) \\
\mathrm{II} &= \mathsf{E}\,(\mathbf{1}_{\Omega_0^c}\,\langle m_M^{2n}\rangle) \ .
\end{aligned}$$

Thus the subgaussian inequality (A.16) implies

$$\mathrm{II} \le \mathsf{P}\left(\Omega_0^c\right) \le 2\exp\left(-\frac{N^{1-2/p}}{2}\right) \le K\left(\frac{n}{N}\right)^n \tag{10.300}$$

provided $n \le N^{1-2/p}/2\log N$.

Now we choose $b > 1$, $\tau_2 > 1$ and $\rho > 0$ such that

$$\beta b^{p-1}\,\tau_2(1 - m^{*2} + 2m^*\rho)^{p/2} < \beta_0 \,. \tag{10.301}$$

This is possible by (10.297). We consider the number a with $1/a + 1/b = 1$, and the number τ_1 with $1/\tau_1 + 1/\tau_2 = 1$. Given x, y, the convexity of the function $t \mapsto t^p$ shows that

$$\left(\frac{x}{a} + \frac{y}{b}\right)^p \le \frac{x^p}{a} + \frac{y^p}{b} \,.$$

Hence

$$(x+y)^p \le a^{p-1}\,x^p + b^{p-1}\,y^p \,,$$

and thus

$$\exp\left(\frac{N\beta}{p} m_M^p\right) \le \exp\left(\frac{N\beta\, a^{p-1}}{p}\left(\frac{1}{N}\sum_{i\le N}\eta_{i,M}\right)^p m^{*p}\right) \times \exp\left(\frac{N\beta\, b^{p-1}}{p}\left(\frac{1}{N}\sum_{i\le N}\eta_{i,M}(\sigma_i - m^*)\right)^p\right) \tag{10.302}$$

Consider the set

$$A = \{\boldsymbol{\sigma}\,;\ |m_1(\boldsymbol{\sigma}) - m^*| \le \rho\} \,.$$

We write

$$\langle m_M^{2n}\rangle \le \langle m_M^{2n}\,\mathbf{1}_A\rangle + \langle m_M^{2n}\,\mathbf{1}_{A^c}\rangle \tag{10.303}$$

and thus

$$\mathsf{E}\,\mathbf{1}_{\Omega_0}\langle m_M^{2n}\rangle \le \mathsf{E}\,\mathbf{1}_{\Omega_0}\langle m_M^{2n}\,\mathbf{1}_A\rangle + \mathsf{E}\,\langle \mathbf{1}_{A^c}\rangle \,. \tag{10.304}$$

Now, using the familiar notation $\langle\cdot\rangle_\sim$ for "cavity in M" we have

$$\langle m_M^{2n}\,\mathbf{1}_A\rangle = \frac{\langle m_M^{2n}\,\mathbf{1}_A \exp(N\beta\, m_M^p/p)\rangle_\sim}{\langle \exp(N\beta m_M^p/p)\rangle_\sim} \le \langle m_M^{2n}\,\mathbf{1}_A \exp(N\beta\, m_M^p/p)\rangle_\sim \,, \tag{10.305}$$

because the denominator is at least 1. On Ω_0 it holds that

$$N\left(\frac{1}{N}\sum_{i\le N}\eta_{i,M}\right)^p \le 1 \,,$$

and combining with (10.302) and (10.305) we get

$$\mathsf{E}\left(\mathbf{1}_{\Omega_0}\langle m_M^{2n}\,\mathbf{1}_A\rangle\right) \le K\mathsf{E}\left\langle \mathbf{1}_A\, m_M^{2n} \exp\left(\frac{N\beta\, b^{p-1}}{p}\left(\frac{1}{N}\sum_{i\le N}\eta_{i,M}(\sigma_i - m^*)\right)^p\right)\right\rangle_\sim \,. \tag{10.306}$$

If $\boldsymbol{\sigma} \in A$, we have

$$\sum_{i \le N} (\sigma_i - m^*)^2 = N - 2m^* \left(\sum_{i \le N} \sigma_i \right) + N m^{*2} \le N(1 - m^{*2} + 2m^* \rho) \ . \quad (10.307)$$

It then follows from (10.301) and Lemma 10.11.14 that if we denote by E_M expectation in the r.v.s $(\eta_{i,M})_{i \le M}$ only, and if we set

$$B = \mathsf{E}_M \exp\left(\frac{N\beta\, b^{p-1} \tau_2}{p} \left(\frac{1}{N} \sum_{i \le N} \eta_{i,M}(\sigma_i - m^*) \right)^p \right) ,$$

then $B \le K$. By Hölder's inequality we get

$$\mathsf{E}\, (\mathbf{1}_{\Omega_0} \langle m_M^{2n}\, \mathbf{1}_A \rangle) \le K \mathsf{E}\, \langle \mathbf{1}_A\, (\mathsf{E}_M m_M^{2n\tau_1})^{1/\tau_1}\, B^{1/\tau_2} \rangle \ . \quad (10.308)$$

Now Khinchin's inequality (A.20) implies

$$\mathsf{E}_M (m_M^{2n\tau_1})^{1/\tau_1} \le \left(\frac{K(\tau_1) n}{N} \right)^n \ ,$$

so we get from (10.308) that

$$\mathsf{E}\, \mathbf{1}_{\Omega_0} \langle m_M^{2n} \rangle \le \left(\frac{K(\tau_1) n}{N} \right)^n + \mathsf{E}\, \langle \mathbf{1}_{A^c} \rangle \ .$$

Proposition 10.11.12 shows that if c is small enough, and $\alpha \le c$, the set A^c is negligible, i.e. $\mathsf{E}\, \langle \mathbf{1}_{A^c} \rangle \le K \exp(-N/K)$. Since if K' is large enough we have $\exp(-N/K) \le (K' n/N)^n$ for each n, this finishes the proof. □

10.12 Proof of Theorem 10.2.1

The proof of Theorem 10.2.1 requires some caution when β is close to 1, but there is a lot of room for larger β. In particular, many of the powers of β found in our subsequent estimates are pretty much arbitrary and only convenient choices.

We first study $r(q)$ given by (10.9), i.e. $r(q) = \alpha q/(1 - \beta(1-q))^2$. We recall that we have proved in Lemma 4.2.5 that

$$1 - \beta(1 - m^{*2}) \ge \frac{m^{*2}}{L_3} \quad (10.309)$$

and we choose $L_1 = 2L_3$.

Lemma 10.12.1. *If*

$$|q - m^{*2}| \le \frac{m^{*2}}{2L_3\beta^2} = \frac{m^{*2}}{L_1\beta^2} \tag{10.310}$$

we have

$$r(q) \le \frac{L_4\alpha}{m^{*2}} \tag{10.311}$$

$$\left|\frac{\mathrm{d}r}{\mathrm{d}q}(q)\right| \le \frac{L\alpha\beta}{m^{*4}} \,. \tag{10.312}$$

Proof. We observe that

$$1 - \beta(1-q) = 1 - \beta(1 - m^{*2}) + \beta(q - m^{*2}) \ge \frac{m^{*2}}{2L_3} \,, \tag{10.313}$$

using (10.309) and (10.310). Also, (assuming without loss of generality that $L_1 \ge 1$) (10.310) implies that $q \le 2m^{*2}$, and (10.311) follows. Straightforward computation shows that

$$\frac{\mathrm{d}r}{\mathrm{d}q} = \alpha \frac{1 - \beta(1+q)}{(1 - \beta(1-q))^3} \,,$$

so that from (10.313) we get

$$\left|\frac{\mathrm{d}r}{\mathrm{d}q}\right| \le \frac{L\alpha}{m^{*6}}((\beta - 1) + q\beta) \,.$$

Now $q \le 2m^{*2}$, and from (10.309) we have $\beta - 1 \le \beta m^{*2}$, so (10.312) follows. □

We turn to the study of the functions Φ and Ψ of (10.7) and (10.8). Recalling the integration by parts formula

$$\frac{\mathrm{d}}{\mathrm{d}x}\mathsf{E} f(z\sqrt{x}) = \frac{1}{2}\mathsf{E} f''(z\sqrt{x})$$

we obtain the relations

$$\frac{\partial \Phi}{\partial r} = \beta^2 \mathsf{E}\, \frac{1 - 2\mathrm{sh}^2 Y}{\mathrm{ch}^4 Y} \tag{10.314}$$

$$\frac{\partial \Phi}{\partial \mu} = \beta \mathsf{E}\, \frac{2\mathrm{th} Y}{\mathrm{ch}^2 Y} \tag{10.315}$$

$$\frac{\partial \Psi}{\partial r} = -\beta^2 \mathsf{E}\, \frac{\mathrm{th} Y}{\mathrm{ch}^2 Y} \tag{10.316}$$

$$\frac{\partial \Psi}{\partial \mu} = \beta \mathsf{E}\, \frac{1}{\mathrm{ch}^2 Y} \,. \tag{10.317}$$

Lemma 10.12.2. *There exists a constant L with the following property. If we have chosen L_0 large enough, if (α, β, h) belongs to the admissible region (10.1), if*

$$|\mu - m^*| \le \frac{m^*}{2}\,, \qquad r \le \frac{L_4 \alpha}{m^{*2}}\,, \tag{10.318}$$

and if Ξ denotes any partial derivative of order ≤ 3 of either Φ or Ψ, we have

$$|\Xi| \le \frac{L}{\beta^{10}}\,. \tag{10.319}$$

There is nothing specific about the power 10 in (10.319).

Recall that we think of L_0 as a parameter, which may be adjusted as large as we wish, that in our computations all dependences on L_0 are explicit, and that when we write "L" for a universal constant, the value of L has been determined through calculations that do not depend on the value of L_0.

Proof. It should be obvious that for $\beta \le L$ we have $\Xi \le L'$ so that it suffices to prove (10.319) for $\beta \ge 2$. We recall that by Lemma 4.2.5, for $\beta \ge 2$, m^* stays bounded below i.e. $m^* \ge 1/L$, so, since we assume $|\mu - m^*| \le m^*/2$, we have

$$\mu \ge \frac{m^*}{2} \ge \frac{1}{L}\,,$$

and

$$\beta\mu + h \ge \frac{\beta}{L}$$

so that

$$Y \ge \beta z \sqrt{r} + \frac{\beta}{L}\,.$$

Therefore

$$z \ge -\frac{1}{2L\sqrt{r}} \quad \Rightarrow \quad Y \ge \frac{\beta}{2L}\,. \tag{10.320}$$

This means that Y is of order at least β unless the rare event

$$\Omega = \left\{ z < -\frac{1}{2L\sqrt{r}} \right\}$$

occurs. Consider a function f with $|f| \le 1$. Then by (10.320) we have

$$\mathsf{E} f(Y) \le \mathsf{P}(\Omega) + \sup \left\{ f(x)\,;\, x \ge \frac{\beta}{2L} \right\}\,. \tag{10.321}$$

To bound $\mathsf{P}(\Omega)$, we observe that since we assume $r \le L_4\alpha/m^{*2} \le L\alpha$, and since $\alpha \le 1/(L_0 \log \beta)$ in the admissible region, using (A.4) in the second inequality,

$$\mathsf{P}(\Omega) = \mathsf{P}\left(z < -\frac{1}{2L\sqrt{r}} \right) \le \exp\left(-\frac{1}{Lr} \right) \le \exp\left(-\frac{L_0 \log \beta}{L} \right) = \beta^{-L_0/L}\,.$$

Thus (10.321) yields

$$\mathsf{E}f(Y) \le \beta^{-L_0/L} + \sup\left\{ f(x)\ ;\ x \ge \frac{\beta}{2L} \right\} .$$

Using this for the function $f(x) = (1 - 2\mathrm{sh}^2 x)/\mathrm{ch}^4 x$, we get from (10.314) that

$$\left|\frac{\partial \Phi}{\partial r}\right| \le \beta^2 \left(\beta^{-L_0/L} + \exp\left(-\frac{\beta}{L}\right)\right) \le \frac{L}{\beta^{10}}$$

if L_0 has been chosen large enough. The other cases are similar. For example, using now the function $f(x) = \mathrm{th} x/\mathrm{ch}^2 x$, (10.315) yields

$$\left|\frac{\partial \Phi}{\partial \mu}\right| \le \beta \left(\beta^{-L_0/L} + \exp\left(-\frac{\beta}{L}\right)\right) \le \frac{L}{\beta^{10}} . \qquad \square$$

We will also use the following simple fact.

Lemma 10.12.3. *If the segment between the points (x_0, y_0) and (x_1, y_1) is entirely contained in a domain W,*

$$|F(x_1, y_1) - F(x_0, y_0)| \le |x_1 - x_0| \sup_W \left|\frac{\partial F}{\partial x}\right| + |y_1 - y_0| \sup_W \left|\frac{\partial F}{\partial y}\right| . \tag{10.322}$$

Proof. Combine trivial bounds with identity

$$|F(x_1, y_1) - F(x_0, y_0)| = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t} F(x_0 + t(x_1 - x_0), y_0 + t(y_1 - y_0))\mathrm{d}t . \quad \square$$

We consider a new parameter $A > 1$. Of course when we write "L" for a universal constant, this means that the value of this constant has been determined independently of A (and of L_0).

Lemma 10.12.4. *Assume that (α, β, h) belongs to the admissible region (10.1), and that*

$$r \le \frac{L_4 \alpha}{m^{*4}} , \qquad |\mu - m^*| \le \frac{m^*}{A} . \tag{10.323}$$

Then if A is large enough, and if $L_0 \ge A$, we have

$$\left|\frac{\partial \Phi}{\partial r}(r, \mu)\right| \le \frac{L}{\beta^{10}} \tag{10.324}$$

$$\left|\frac{\partial \Phi}{\partial \mu}(r, \mu)\right| \le \frac{L m^*}{\beta^{10}} \tag{10.325}$$

$$\left|\frac{\partial \Psi}{\partial r}(r, \mu)\right| \le \frac{L m^*}{\beta^{10}} \tag{10.326}$$

$$\left|\frac{\partial \Psi}{\partial \mu}(r, \mu)\right| \le 1 - \frac{m^{*2}}{L} . \tag{10.327}$$

Proof. We observe that (10.319) implies (10.324). To prove (10.326) we observe from (10.316) that

$$\frac{\partial \Psi}{\partial r}(0, m^*) = -\beta^2 \frac{\mathrm{th}(\beta m^* + h)}{\mathrm{ch}^2(\beta m^* + h)} = \beta^2 m^*(1 - m^{*2}) \ .$$

By (4.37) we have $\beta^{20}(1 - m^{*2}) \le L$, and hence

$$\left| \frac{\partial \Psi}{\partial r}(0, m^*) \right| \le \frac{L m^*}{\beta^{10}} \ . \tag{10.328}$$

We use (10.322) for the function $F = \partial \Psi / \partial r$, and thus, using Lemma 10.12.2, we get

$$\left| \frac{\partial \Psi}{\partial r}(0, m^*) - \frac{\partial \Psi}{\partial r}(r, \mu) \right| \le \frac{L}{\beta^{10}}(r + |\mu - m^*|) \ . \tag{10.329}$$

Since we assume $r \le L_4 \alpha / m^{*2}$, (10.2) yields

$$r \le \frac{L m^{*2}}{L_0} \ . \tag{10.330}$$

Combining (10.328) and (10.329) and using that $|\mu - m^*| \le m^*/A \le m^*$ yields (10.326). The proof of (10.325) is similar.

We turn to the proof of (10.327). This proof is more delicate because it requires a kind of order 2 expansion. We start by showing that under (10.323) we have

$$\left| \frac{\partial^2 \Psi}{\partial \mu^2}(r, m^*) \right| \le \frac{L m^*}{\beta^{10}} \le L m^* \ . \tag{10.331}$$

The proof is nearly identical to that of (10.325) and (10.326). We start by observing from (10.317) that

$$\frac{\partial^2 \Psi}{\partial \mu^2} = -2\beta^2 \mathsf{E} \frac{\mathrm{th} Y}{\mathrm{ch}^2 Y}$$

so that, as in (10.328)

$$\left| \frac{\partial^2 \Psi}{\partial \mu^2}(0, m^*) \right| = |-2\beta^2 m^*(1 - m^{*2})| \le L m^* / \beta^{10} \ ,$$

and (10.331) follows as before. Next, since $|\partial^2 \Psi / \partial \mu \partial r| \le L$ by Lemma 10.12.2, applying (10.331) and (10.322) to the function $F = \partial \Psi / \partial \mu$ yields

$$\begin{aligned} \left| \frac{\partial \Psi}{\partial \mu}(r, \mu) - \frac{\partial \Psi}{\partial \mu}(0, m^*) \right| &\le L(r + m^*|\mu - m^*|) \\ &\le L m^{*2} \left(\frac{1}{L_0} + \frac{1}{A} \right) \ , \end{aligned} \tag{10.332}$$

using (10.330) and (10.323). Finally (10.317) entails

$$\frac{\partial \Psi}{\partial \mu}(0, m^*) = \beta(1 - \text{th}^2(\beta m^* + h)) = \beta(1 - m^{*2}) \leq 1 - \frac{m^{*2}}{L},$$

using (10.309) in the last inequality. Combining with (10.332) finishes the proof. □

Corollary 10.12.5. *Assume that A has been chosen as in Lemma 10.12.4 and that $L_0 \geq A$. Consider (r, μ) and (r', μ') as in (10.323). Then*

$$|\Phi(r, \mu) - \Phi(r', \mu')| \leq \frac{L}{\beta^{10}}|r - r'| + \frac{Lm^*}{\beta^{10}}|\mu - \mu'| \tag{10.333}$$

$$|\Psi(r, \mu) - \Psi(r', \mu')| \leq \frac{Lm^*}{\beta^{10}}|r - r'| + \left(1 - \frac{m^{*2}}{L}\right)|\mu - \mu'| \,. \tag{10.334}$$

Moreover we have

$$|\Phi(r, \mu) - m^{*2}| \leq \frac{Lm^{*2}}{A\beta^{10}} \tag{10.335}$$

$$|\Psi(r, \mu) - m^*| \leq \left(1 - m^{*2}\left(\frac{1}{L} - \frac{LA}{L_0}\right)\right)\frac{m^*}{A} \,. \tag{10.336}$$

Proof. To prove (10.333) and (10.334) we combine Lemmas 10.12.3 and 10.12.4. Taking $r' = 0$, $\mu' = m^*$, observing that $\Phi(0, m^*) = \text{th}^2(\beta m^* + h) = m^{*2}$ and $\Psi(0, m^*) = \text{th}(\beta m^* + h) = m^*$ and recalling (10.330) and that $L_0 \geq A$ we deduce (10.335) and

$$|\Psi(r, \mu) - m^*| \leq \frac{Lm^*}{\beta^{10}}\frac{m^{*2}}{L_0} + \left(1 - \frac{m^{*2}}{L}\right)\frac{m^*}{A} \,,$$

which implies (10.336). □

Proof of Theorem 10.2.1. We fix once and for all A large enough so that Lemma 10.12.4 holds, as well as the following

$$\frac{L}{A} \leq \frac{1}{L_1} \,, \tag{10.337}$$

where L is as in (10.335) and where L_1 is as in (10.310). We will prove Theorem 10.2.1 with the value $L_2 = A$. Assume that (r, μ) is as in (10.323). We see from (10.335) and (10.337) that

$$|\Phi(r, \mu) - m^{*2}| \leq \frac{m^{*2}}{L_1\beta^2} \,. \tag{10.338}$$

Moreover, if L_0 is large enough we deduce from (10.336) that

$$|\Psi(r, \mu) - m^*| \leq \frac{m^*}{A} = \frac{m^*}{L_2} \,. \tag{10.339}$$

Let us consider the domain

$$W=\left\{(q,\mu)\;;\;|q-m^{*2}|\le\frac{m^{*2}}{L_1\beta^2}\;,\;|\mu-m^*|\le\frac{m^*}{A}\right\}\;,$$

and for $(q,\mu)\in W$ set $T(q,\mu)=(\Phi(r(q),\mu),\Psi(r(q),\mu))$. It follows from (10.311) that for $(q,\mu)\in W$, the pair $(r(q),\mu)$ satisfies (10.323), so (10.338) and (10.339) show that $T(W)\subset W$. Also, for $(q,\mu)\in W$ and $(q',\mu')\in W$ we deduce from (10.312) and (10.2) that

$$|r(q)-r(q')|\le\frac{L\alpha\beta}{m^{*4}}|q-q'|\le\frac{L}{L_0}\beta|q-q'|\;,$$

and using (10.333) and (10.334) for $r=r(q)$ and $r'=r(q')$ we get respectively

$$|\Phi(r(q),\mu)-\Phi(r(q'),\mu')|\le\frac{L}{L_0}|q-q'|+Lm^*|\mu-\mu'| \tag{10.340}$$

and

$$|\Psi(r(q),\mu)-\Psi(r(q'),\mu')|\le\frac{L}{L_0}m^*|q-q'|+\left(1-\frac{m^{*2}}{L}\right)|\mu-\mu'|\;. \tag{10.341}$$

Consider a number $a>0$, to be determined later, and the distance d on W given by

$$d((q,\mu),(q',\mu'))=a|q-q'|+|\mu-\mu'|\;.$$

It follows from (10.340) and (10.341) that

$$d(T(q,\mu),T(q',\mu'))\le Bd((q,\mu),(q',\mu'))$$

where

$$B=\max\left(\frac{L}{L_0}\left(1+\frac{m^*}{a}\right),1-\frac{m^{*2}}{L}+Lam^*\right)\;.$$

Taking $a=m^*/L$ where L is large, and then L_0 large yields

$$B\le\max\left(\frac{1}{2},1-\frac{m^{*2}}{L}\right)<1\;.$$

Therefore, T is a contraction on W for the distance d, and admits a unique fixed point on W. □

Proof of Lemma 10.2.2. We use (10.334) for $r'=0$ and $\mu'=m^*$ to get

$$|\mu-m^*|\le\frac{Lm^*r}{\beta^{10}}+\left(1-\frac{m^{*2}}{L}\right)|\mu-m^*|$$

so that

$$|\mu-m^*|\le\frac{Lr}{m^*\beta^{10}}$$

and combining with (10.330) this proves (10.11). To prove (10.12) we use (10.333) for $r' = 0$, $\mu' = m^*$ to get

$$|q - m^{*2}| \le \frac{L}{\beta^{10}} r + \frac{Lm^*}{\beta^{10}} |\mu - m^*|$$

and we use (10.11) and (10.330).

To prove (10.13), recalling (10.330) we may assume that $\beta \ge 2$. It suffices to prove the first part since $|1 - m^{*2}| \le L\beta^{-10}$ by (4.37). Now

$$\begin{aligned} r - \alpha &= \alpha \left(\frac{q}{(1 - \beta(1-q))^2} - 1 \right) \\ &= \alpha \left(\frac{q - 1}{(1 - \beta(1-q))^2} + \frac{1 - (1 - \beta(1-q))^2}{(1 - \beta(1-q))^2} \right) . \end{aligned}$$

Since $|1 - (1-x)^2| \le 2x$ for $0 \le x \le 1$ we obtain that $|1 - (1 - \beta(1-q))^2| \le 2\beta(1-q)$, so for $\beta \ge 2$, using (10.313) and $m^* \ge 1/L$, we get

$$|r - \alpha| \le \frac{L\beta(1-q)}{(1 - \beta(1-q))^2} \le L\beta(1-q) .$$

Now, (10.11) and (4.37) yield

$$|1 - q| \le |q - m^{*2}| + |1 - m^{*2}| \le \frac{L}{\beta^{10}} ,$$

and this implies (10.13). To prove (10.16) we use (10.313) and that $q \le Lm^{*2}$ by (10.12). Finally (10.17) follows from (10.309), (10.11) and (10.12). □

11. The SK Model Without External Field

11.1 Overview

In this chapter we study the SK model without external field, with Hamiltonian

$$-H_N(\boldsymbol{\sigma}) = \frac{\beta}{\sqrt{N}} \sum_{i<j} g_{ij}\sigma_i\sigma_j , \tag{11.1}$$

for $\beta \leq 1$. This is the Hamiltonian (1.12) when $h = 0$. This model (at least for $\beta < 1$) is in some sense the simplest non trivial spin glass model, and not surprisingly more detailed results are available than for the more complicated cases. It enjoys some truly special features, one of which is that if $\beta < 1$ we have

$$\mathsf{E}Z_N^2 \leq \frac{1}{\sqrt{1-\beta^2}}(\mathsf{E}Z_N)^2 , \tag{11.2}$$

where Z_N is of course the partition function $\sum_{\boldsymbol{\sigma}} \exp(-H_N(\boldsymbol{\sigma}))$, and where the value of β is kept implicit. A main difference between the cases $h \neq 0$ and $h = 0$ of the SK model (at high temperature) is that if $h \neq 0$ the fluctuations of $\log Z_N$ are typically of order $\sqrt{N}$, while if $h = 0$ they are typically of order 1. Consider the random variable

$$X = \log Z_N - N\left(\log 2 + \frac{\beta^2}{4}\right) .$$

In Section 11.2, we prove exponential bounds for $\mathsf{P}(X \leq -t)$; and in Section 11.3 we prove exponential bounds for $\mathsf{P}(X \geq t)$. Not surprisingly, these bounds are obtained through specialized methods. In Section 11.4, we compute for each k the limit $\lim_{N\to\infty} \mathsf{E}X^k$, establishing a quantitative version of a central limit theorem of Aizenman, Lebowitz and Ruelle.

In Section 11.5 we examine the (random) matrix of the spin correlation $\langle\sigma_i\sigma_j\rangle$. We conjecture that this matrix shares some properties with the random matrix $(g_{ij}/\sqrt{N})$, and in particular that its operator norm remains bounded independently of N, a result that we prove within a logarithmic factor.

In Section 11.6 we examine some natural d-dimensional generalizations of the SK model without external field, for which we show that the high-

M. Talagrand, *Mean Field Models for Spin Glasses*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys in Mathematics 55, DOI 10.1007/978-3-642-22253-5_4,

temperature phase extends much beyond that of the typical situation of Section 1.13.

The final Section 11.7 examines the case $\beta = 1$, and the case $\beta = \beta_N \to 1$ as $N \to \infty$. This situation is replete with exciting problems, and our understanding is still very limited. Even such a basic question as determining the exact order of $\nu(R_{1,2}^2)$ when $\beta = 1$ is wide open (and looks very difficult).

This results of this chapter are largely independent from those of Chapter 1, but the reader should be at least be familiar with Section 1.6.

11.2 Lower Deviations for Z_N

The goal of the section is to prove the following:

Theorem 11.2.1. *Given $\beta < 1$, there exists K depending on β only such that for any N and any $t > 0$:*

$$\mathsf{P}\left(\log Z_N \le N\left(\frac{\beta^2}{4} + \log 2\right) - t\right) \le K \exp\left(-\frac{t^2}{K}\right) .$$

This deviation inequality raises the following "large deviation" problem.

Research Problem 11.2.2. (Level 2) Given $\beta < 1$ and $t > 0$, prove the existence of the limit

$$\lim_{N\to\infty} \frac{1}{N^2} \log \mathsf{P}\left(\frac{1}{N}\log Z_N \le \frac{\beta^2}{4} + \log 2 - t\right)$$

and compute it. More generally, given $0 \le \alpha < 1$, compute the limit

$$\lim_{N\to\infty} \frac{1}{N^{2(1-\alpha)}} \log \mathsf{P}\left(\frac{1}{N}\log Z_N \le \frac{\beta^2}{4} + \log 2 - \frac{t}{N^\alpha}\right) .$$

We now prepare for the proof of Theorem 11.2.1. The fundamental relation (11.2) is the special case $\gamma = 0$ of the following (which will also be useful in its own right):

Lemma 11.2.3. *If $\gamma + \beta^2 < 1$ we have*

$$\mathsf{E} \sum_{\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2} \exp\left(-H_N(\boldsymbol{\sigma}^1) - H_N(\boldsymbol{\sigma}^2) + \frac{\gamma N}{2} R_{1,2}^2\right) \le \frac{1}{\sqrt{1-\beta^2-\gamma}} (\mathsf{E} Z_N)^2 . \tag{11.3}$$

Proof. We recall that

$$\mathsf{E}H_N(\boldsymbol{\sigma}^1)H_N(\boldsymbol{\sigma}^2) = \frac{\beta^2}{2}(NR_{1,2}^2 - 1)\ ;\quad \mathsf{E}H_N^2(\sigma) = \frac{\beta^2}{2}(N-1)\ , \tag{11.4}$$

so that

$$\mathsf{E}(-H_N(\boldsymbol{\sigma}^1) - H_N(\boldsymbol{\sigma}^2))^2 = \beta^2(N-2) + \beta^2 NR_{1,2}^2\ .$$

Using (A.1) we have

$$\mathsf{E}Z_N = 2^N \exp\frac{\beta^2}{4}(N-1)\ , \tag{11.5}$$

and the left-hand side of (11.3) is then

$$\exp\frac{\beta^2}{2}(N-2)\sum_{\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2}\exp\frac{\delta N}{2}R_{1,2}^2 \le (\mathsf{E}Z_N)^2 2^{-2N}\sum_{\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2}\exp\frac{\delta N}{2}R_{1,2}^2\ , \tag{11.6}$$

where $\delta = \beta^2 + \gamma$. To conclude the proof, we note that for any value of $\boldsymbol{\sigma}^2$, (A.19) implies

$$\sum_{\boldsymbol{\sigma}^1}\exp\frac{\delta N}{2}R_{1,2}^2 = \sum_{\boldsymbol{\sigma}}\exp\left(\frac{\delta N}{2}\left(\frac{1}{N}\sum_{i\le N}\sigma_i\right)^2\right) \le 2^N\frac{1}{\sqrt{1-\delta}}\ . \qquad \square$$

By definition of the Gibbs measure we have

$$Z_N^2\left\langle\exp\frac{\gamma N}{2}R_{1,2}^2\right\rangle = \sum_{\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2}\exp\left(-H_N(\boldsymbol{\sigma}^1) - H_N(\boldsymbol{\sigma}^2) + \frac{\gamma N}{2}R_{1,2}^2\right)\ ,$$

so that (11.3) implies

$$\mathsf{E}\left(Z_N^2\left\langle\exp\frac{\gamma N}{2}R_{1,2}^2\right\rangle\right) \le \frac{1}{\sqrt{1-\beta^2-\gamma}}(\mathsf{E}Z_N)^2\ . \tag{11.7}$$

In particular, taking e.g. $\gamma = (1-\beta^2)/2$ in (11.7) we obtain

$$\mathsf{E}\left(Z_N^2\left\langle\exp\frac{1-\beta^2}{4}NR_{1,2}^2\right\rangle\right) \le \sqrt{\frac{2}{1-\beta^2}}(\mathsf{E}Z_N)^2\ , \tag{11.8}$$

whereas taking $\gamma = 0$ yields (11.2).

Lemma 11.2.4. *For some number L, and all $\beta < 1$, we have*

$$\mathsf{P}\left(Z_N \ge \frac{1}{2}\mathsf{E}Z_N\ ,\ N\langle R_{1,2}^2\rangle \le \frac{4}{1-\beta^2}\log\frac{L}{1-\beta^2}\right) \ge \frac{1}{8}\sqrt{1-\beta^2}\ . \tag{11.9}$$

For the time being, one should read this formula as

$$\mathsf{P}\left(Z_N \geq \frac{1}{2}\mathsf{E}Z_N \;,\; N\langle R_{1,2}^2\rangle \leq K(\beta)\right) \geq \frac{1}{K(\beta)} \;,$$

where $K(\beta)$ depends on β only. The actual dependence on β of $K(\beta)$ will not be relevant in this section, and will be used only in Section 11.7.

Proof. Consider the event $A = \{Z_N \geq \mathsf{E}Z_N/2\}$. Combining (11.2) and the Paley-Zygmund inequality (A.61) we get

$$\mathsf{P}(A) \geq \frac{1}{4}\sqrt{1-\beta^2} \,.$$

On the other hand, (11.8) and Markov's inequality imply that for $t > 0$ we have

$$\mathsf{P}\left(\left\{Z_N^2 \left\langle \exp \frac{1-\beta^2}{4} N R_{1,2}^2 \right\rangle \geq t(\mathsf{E}Z_N)^2\right\}\right) \leq \frac{1}{t}\sqrt{\frac{2}{1-\beta^2}} \;,$$

so that if

$$B = \left\{Z_N^2 \left\langle \exp \frac{1-\beta^2}{4} N R_{1,2}^2 \right\rangle \leq t(\mathsf{E}Z_N)^2\right\} \;,$$

then

$$\mathsf{P}(B) \geq 1 - \frac{1}{t}\sqrt{\frac{2}{1-\beta^2}} \,.$$

Since $\mathsf{P}(A\cap B) \geq \mathsf{P}(A) + \mathsf{P}(B) - 1$, for $t = 32/(1-\beta^2)$ it follows that $\mathsf{P}(A\cap B) \geq \sqrt{1-\beta^2}/8$. And on $A \cap B$ we have

$$Z_N^2 \left\langle \exp \frac{1-\beta^2}{4} N R_{1,2}^2 \right\rangle \leq \frac{32}{1-\beta^2}(\mathsf{E}Z_N)^2 \leq \frac{128}{1-\beta^2} Z_N^2$$

so that, using Jensen's inequality,

$$\exp \frac{1-\beta^2}{4} N\langle R_{1,2}^2\rangle \leq \left\langle \exp \frac{1-\beta^2}{4} N R_{1,2}^2 \right\rangle \leq \frac{128}{1-\beta^2} \,. \qquad \square$$

We now think of the quantities $(g_{ij})_{i<j}$ used to define H_N in (11.2) as the coordinates of a point $\mathbf{g}$ of $\mathbb{R}^M$, for $M = N(N-1)/2$. We denote by $\langle\cdot\rangle$ and $\langle\cdot\rangle'$ respectively the brackets corresponding to the realizations $\mathbf{g}$ and $\mathbf{g}'$ of the disorder and Z_N and Z'_N the corresponding partition functions.

Lemma 11.2.5. *We have*

$$\log Z_N - \log Z'_N \geq -\left(\sum_{i<j}(g_{ij}-g'_{ij})^2\right)^{1/2} (N\langle R_{1,2}^2\rangle')^{1/2} \,. \tag{11.10}$$

Proof. The definition of $\langle\cdot\rangle'$ implies the obvious identity

$$Z_N = Z'_N \Big\langle \exp \frac{\beta}{\sqrt{N}} \sum_{i<j} (g_{ij} - g'_{ij}) \sigma_i \sigma_j \Big\rangle' .$$

Using Jensen's inequality for $\langle\cdot\rangle'$ yields

$$\Big\langle \exp \frac{\beta}{\sqrt{N}} \sum_{i<j} (g_{ij} - g'_{ij}) \sigma_i \sigma_j \Big\rangle' \geq \exp \frac{\beta}{\sqrt{N}} \sum_{i<j} (g_{ij} - g'_{ij}) \langle \sigma_i \sigma_j \rangle' ,$$

and the Cauchy-Schwarz inequality implies

$$\sum_{i<j} (g_{ij} - g'_{ij}) \langle \sigma_i \sigma_j \rangle' \geq -\Big(\sum_{i<j} (g_{ij} - g'_{ij})^2 \Big)^{1/2} \Big(\sum_{i<j} \langle \sigma_i \sigma_j \rangle'^2 \Big)^{1/2} .$$

Finally we observe the identity

$$\langle R_{1,2}^2 \rangle' = \frac{1}{N^2} \sum_{i,j} \langle \sigma_i \sigma_j \rangle'^2 \geq \frac{1}{N^2} \sum_{i<j} \langle \sigma_i \sigma_j \rangle'^2 ,$$

which is obvious by replacing $R_{1,2}$ by its value and expanding the square. Combining these relations yields

$$\log Z_N - \log Z'_N \geq -\frac{\beta}{\sqrt{N}} \Big(\sum_{i<j} (g_{ij} - g'_{ij})^2 \Big)^{1/2} (N^2 \langle R_{1,2}^2 \rangle')^{1/2} ,$$

and we use that $\beta \leq 1$ to conclude. □

Keeping (11.5) in mind, Theorem 11.2.1 is a consequence of the following more precise result.

Proposition 11.2.6. *For some number L and all $\beta < 1$ we have for all $t > 0$*

$$\mathsf{P}(Z_N \leq e^{-t} \mathsf{E} Z_N) \leq \frac{L}{1-\beta^2} \exp \left(-\frac{(1-\beta^2)t^2}{L \log \frac{2}{1-\beta^2}} \right) . \tag{11.11}$$

Proof. Consider the subset C of $\mathbb{R}^M$ given by

$$C = \left\{ Z_N \geq \frac{1}{2} \mathsf{E} Z_N \ , \ N \langle R_{1,2}^2 \rangle \leq \frac{4}{1-\beta^2} \log \frac{L}{1-\beta^2} \right\}$$

so that by (11.9) it holds

$$\mathsf{P}(\mathbf{g} \in C) \geq \frac{1}{8} \sqrt{1-\beta^2} . \tag{11.12}$$

From Lemma 11.2.5, and using the notation here from, we have, for $\mathbf{g}' \in C$:

$$\log Z'_N \geq \log \frac{1}{2}\mathsf{E}Z_N = \log \mathsf{E}Z_N - \log 2$$

and

$$N\langle R_{1,2}\rangle'^2 \leq \frac{4}{1-\beta^2}\log\frac{L}{1-\beta^2} := K_0 \ ,$$

so that by (11.10) we get

$$\log Z_N \geq \log \mathsf{E}Z_N - \log 2 - \sqrt{K_0}d(\mathbf{g},\mathbf{g}') \ ,$$

where of course $d(\mathbf{g},\mathbf{g}') = (\sum_{i<j}(g_{ij}-g'_{ij})^2)^{1/2}$ is the Euclidean distance between $\mathbf{g}$ and $\mathbf{g}'$ in $\mathbb{R}^M$. Since this holds for any $\mathbf{g}' \in C$ we have

$$\log Z_N \geq \log \mathsf{E}Z_N - \log 2 - \sqrt{K_0}d(\mathbf{g},C)$$

where $d(\mathbf{g},C)$ is the distance from $\mathbf{g}$ to C, and therefore

$$Z_N \leq e^{-t}\mathsf{E}Z_N \quad \Rightarrow \quad \sqrt{K_0}d(\mathbf{g},C) \geq t - \log 2 \ .$$

Now, recalling (11.12), we have

$$t \geq 2\log 2 + 4\sqrt{K_0}\sqrt{\log\frac{16}{\sqrt{1-\beta^2}}} \quad \Rightarrow \quad t - \log 2 \geq \frac{t}{2} + 2\sqrt{K_0}\sqrt{\log\frac{2}{\mathsf{P}(\mathbf{g}\in C)}} \tag{11.13}$$

and thus, for these values of t we get

$$Z_N \leq e^{-t}\mathsf{E}Z_N \quad \Rightarrow \quad d(\mathbf{g},C) \geq \frac{t}{2\sqrt{K_0}} + 2\sqrt{\log\frac{2}{\mathsf{P}(\mathbf{g}\in C)}} \ .$$

Hence Lemma 9.3.3 implies

$$\mathsf{P}(Z_N \leq e^{-t}\mathsf{E}Z_N) \leq 2\exp\left(-\frac{t^2}{8K_0}\right) \ .$$

This in turn implies (11.11) because one can arrange by a suitably large choice of the constant L there that the right-hand side of (11.11) is ≥ 1 for the values of t that do not satisfy the left-hand inequality of (11.13). □

We denote by K a number depending on β only, that need not be the same at each occurrence.

Theorem 11.2.7. *If $\beta < 1$ we have*

$$\nu\left(\exp\frac{1-\beta^2}{8}NR_{1,2}^2\right) \leq K \ . \tag{11.14}$$

The difference with Theorem 1.4.1 is that (11.14) holds for $\beta < 1$ rather than for $\beta < 1/2$. Later on, in Section 13.7, we will be able to obtain an exponential control as in Theorem 1.4.1 in the entire high-temperature region of the SK model (for all values of h), but this is much more difficult than in the present case.

Proof. We use the Cauchy-Schwarz inequality for $\langle\cdot\rangle$ to get

$$\left\langle \exp \frac{1-\beta^2}{8} NR_{1,2}^2 \right\rangle \le \left\langle \exp \frac{1-\beta^2}{4} NR_{1,2}^2 \right\rangle^{1/2} = \frac{1}{Z_N}\left(Z_N^2 \left\langle \exp \frac{1-\beta^2}{4} NR_{1,2}^2 \right\rangle\right)^{1/2} .$$

Using the Cauchy-Schwarz inequality for E and (11.3) with $\gamma = (1-\beta^2)/2$ we get

$$\begin{aligned}\mathsf{E}\left\langle \exp \frac{1-\beta^2}{8} NR_{1,2}^2 \right\rangle &\le \left(\mathsf{E}\frac{1}{Z_N^2}\right)^{1/2} \left(\mathsf{E}\left(Z_N^2 \left\langle \exp \frac{1-\beta^2}{4} NR_{1,2}^2 \right\rangle\right)\right)^{1/2} \\ &\le \left(\mathsf{E}\frac{1}{Z_N^2}\right)^{1/2} (K(\mathsf{E}Z_N)^2)^{1/2} \\ &= K\left(\mathsf{E}\left(\frac{\mathsf{E}Z_N}{Z_N}\right)^2\right)^{1/2} .\end{aligned}$$

It follows from (11.11) that, for $t > 1$,

$$\mathsf{P}\left(\frac{\mathsf{E}Z_N}{Z_N} > t\right) \le K \exp\left(-\frac{(\log t)^2}{K}\right)$$

and this implies (using (A.27)) that

$$\mathsf{E}\left(\frac{\mathsf{E}Z_N}{Z_N}\right)^2 < K . \qquad \square$$

Theorem 11.2.7 provides a good control of the situation. Here is an example of application of this result. Given a subset I of $\{1,\dots,N\}$ let us denote by $G_I = G_{N,I}$ the average of the Gibbs' measure under the map $\boldsymbol{\sigma} \mapsto (\sigma_i)_{i\in I}$ and by μ_I the uniform measure on $\{0,1\}^I$. The following should be compared to Theorem 1.4.15. (We recall that $\|\cdot\|$ denotes the total variation distance.)

Proposition 11.2.8. *We have*

$$\mathsf{E}\|G_I - \mu_I\| \le K\frac{\mathrm{card} I}{\sqrt{N}} .$$

We refer to [103] Theorem 1.8 to see how this can be deduced from (11.14). Despite the relative simplicity of the situation, many questions remain unanswered.

Research Problem 11.2.9. (Level 1?) Find (optimal, please) pairs of sequences $a_N \to \infty$, $b_N \to 0$ such that

$$\mathsf{P}\left(\sup_{\mathrm{card} I \le a_N} \|G_I - \mu_I\| \ge b_N\right) \to 0 \, .$$

It seems almost certain that this is connected to the following

Research Problem 11.2.10. (Level 1) For a typical realization of Gibbs' measure, study the family of numbers $(\langle \sigma_i \sigma_j \rangle)_{i<j}$. For example, what is the order of $\max_{i<j} \langle \sigma_i \sigma_j \rangle^2$? Does the family of numbers $(\langle \sigma_i \sigma_j \rangle)_{i<j}$ look (after proper rescaling) like the realization of an i.i.d sequence? (The limiting law of $\langle \sigma_i \sigma_j \rangle$ is computed as a very special case of Theorem 1.11.1.)

Research Problem 11.2.11. If $I = \{1, \ldots, n\}$, find competent bounds for

$$\mathsf{P}\left(\|G_I - \mu_I\| \ge \frac{t}{\sqrt{N}}\right) \, .$$

It is likely that the complication inherent to the SK model (a complication that remains hidden at high temperature) is related to the fact that the "energy landscape", i.e. the function $\boldsymbol{\sigma} \mapsto -H_N(\boldsymbol{\sigma})$ has a complex geometry. The physicists seem to say that the collection of configurations $\boldsymbol{\sigma}$ for which $-H_N(\boldsymbol{\sigma})$ is near maximum is made up of rather separated pieces (the so-called multiple valley picture). Does this show up on the structure of the Gibbs measure, even at $\beta < 1$?

Conjecture 11.2.12. (Level 3) If $\beta > 0$, there exists $\theta > 0$ such that with overwhelming probability we can find subsets A, B of Σ_N with $G_N(A) \ge 1/4$, $G_N(B) \ge 1/4$, $d(A, B) \ge \theta$ (where $d(A, B)$ is the distance of A and B, for $d(\boldsymbol{\sigma}, \boldsymbol{\sigma}') = N^{-1} \mathrm{card}\{i \le N \, ; \, \sigma_i \ne \sigma'_i\}$.)

Maybe this formulation that the Gibbs' measure "is made up of different pieces" is too naive. Finding the correct formulation is of course part of the problem.

11.3 Upper Deviations for Z_N

The goal of this section is to prove the following:

Theorem 11.3.1. *If $\beta < 1$ we have*

$$0 < t < \sqrt{N} \quad \Rightarrow \quad \mathsf{P}\left(\log Z_N \geq N\left(\frac{\beta^2}{4} + \log 2\right) + t\right) \leq K \exp\left(-\frac{t^2}{K}\right) . \tag{11.15}$$

The restriction $t < \sqrt{N}$ is essentially necessary. This is because if $\boldsymbol{\sigma}$ is an arbitrary configuration we have $\log Z_N \geq -H_N(\boldsymbol{\sigma})$, and $H_N(\boldsymbol{\sigma})$ is a Gaussian r.v. with $\mathsf{E}H_N(\boldsymbol{\sigma})^2 = \beta^2(N-1)/2$, so that we have

$$\begin{aligned}\mathsf{P}\left(\log Z_N \geq N\left(\frac{\beta^2}{4} + \log 2\right) + N\right) &\geq \mathsf{P}\left(-H_N(\boldsymbol{\sigma}) \geq N\left(\frac{\beta^2}{4} + \log 2 + 1\right)\right) \\ &\geq \exp(-KN) ,\end{aligned}$$

where the last inequality follows from the fact that for a standard Gaussian r.v. z we crudely have $\mathsf{P}(z \geq t) \geq \exp(-t^2)$ for large values of t, see (A.5). Hence (11.15) cannot hold for $t/\sqrt{N}$ much larger than 1.

The meaning of (11.15) is essentially that the r.v.

$$V = \frac{Z_N}{\mathsf{E}Z_N}$$

satisfies

$$0 < t < \sqrt{N} \quad \Rightarrow \quad \mathsf{P}(\log V \geq t) \leq K \exp\left(-\frac{t^2}{K}\right) .$$

The most obvious method to prove such an inequality using moments is to prove that for $n \geq 1$ we have

$$\mathsf{E}V^n \leq K_0^{n^2} .$$

We then deduce from the Markov inequality that for any integer n it holds

$$\mathsf{P}(\log V \geq t) = \mathsf{P}(V \geq e^t) \leq \frac{\mathsf{E}V^n}{e^{tn}} \leq \left(\frac{K_0^n}{e^t}\right)^n . \tag{11.16}$$

If n is such that $K_0^n \leq e^{t/2}$, this implies

$$\mathsf{P}(\log V \geq t) \leq \exp\left(-\frac{nt}{2}\right) ,$$

and taking the largest possible value of n (which is $\geq t/K$ for $t \geq K$) gives $\mathsf{P}(\log V \geq t) \leq \exp(-t^2/K)$ for $t \geq K$.

So, let us study $\mathsf{E}Z_N^n$. We set

$$e_n(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n) = \exp \frac{\beta}{\sqrt{N}} \sum_{i<j} g_{ij} \sum_{\ell \le n} \sigma_i^\ell \sigma_j^\ell \,, \tag{11.17}$$

so that

$$Z_N^n = \sum_{\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n} e_n(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n) \,. \tag{11.18}$$

Recalling (1.396) we have

$$\mathsf{E} e_n(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n) = \exp\biggl(\frac{\beta^2}{4} n(N-n) + \frac{\beta^2}{2} N \sum_{1 \le \ell < \ell' \le n} R_{\ell,\ell'}^2\biggr). \tag{11.19}$$

Therefore,

$$\mathsf{E} Z_N^n = \exp \frac{\beta^2}{4} n(N-n) \sum_{\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n} \exp \frac{\beta^2}{2} N \sum_{1 \le \ell < \ell' \le n} R_{\ell,\ell'}^2 \,.$$

Restricting the summation to the case $\boldsymbol{\sigma}^1 = \cdots = \boldsymbol{\sigma}^n$, implies that, using (11.5),

$$\begin{aligned} \mathsf{E} Z_N^n &\ge 2^N \exp \frac{\beta^2}{4} n(N-n) \exp \frac{\beta^2}{4} Nn(n-1) \\ &\ge (\mathsf{E} Z_N)^n 2^{-N(n-1)} \exp \frac{\beta^2}{4} n(n-1)(N-1) \\ &= (\mathsf{E} Z_N)^n \exp(n-1) \left(n \frac{\beta^2}{4} (N-1) - N \log 2 \right) \,. \end{aligned}$$

Thus if $n\beta^2 > 4 \log 2$, it is not true that $\mathsf{E} Z_N^n \le K(\beta, n)(\mathsf{E} Z_N)^n$. There is a "moment explosion".

Our proof of Theorem 11.3.1 is based on the fact that this moment explosion is created by a small set of configurations. We will prove that there is a decomposition $Z_N^n = U_1 + U_2$ where

$$\mathsf{E} U_1 \le K^{n^2} (\mathsf{E} Z_N)^n$$

and where U_2 is "typically very small", despite the fact that $\mathsf{E} U_2$ is very large. The argument is pretty, but it is unrelated to any other material of this work.

Proposition 11.3.2. *If $\beta < 1$, there exists $c > 0$ and $K > 0$ such that, for each n, if we define*

$$B_n = \left\{ (\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n) \in \Sigma_N^n \,;\, \sum_{1 \le \ell < \ell' \le n} R_{\ell,\ell'}^2 \le c \right\} ,$$

if we denote by B_n^c the complement of B_n, and if we define

$$U_1 = \sum_{B_n} e_n(\boldsymbol{\sigma}^1, \dots, \boldsymbol{\sigma}^n) \; ; \quad U_2 = \sum_{B_n^c} e_n(\boldsymbol{\sigma}^1, \dots, \boldsymbol{\sigma}^n) \; ,$$

then for $n \le \sqrt{N}/K$ we have

$$\mathsf{E} U_1 \le K^{n^2} (\mathsf{E} Z_N)^n \tag{11.20}$$

and

$$\mathsf{P}\left(U_2 \ge \exp\left(-\frac{N}{K}\right)(\max(Z_n, \mathsf{E} Z_N))^n\right) \le K \exp\left(-\frac{N}{K}\right) \; . \tag{11.21}$$

Proof of Theorem 11.3.1. Since $\log \mathsf{E} Z_N \le N(\beta^2/4 + \log 2)$, it suffices to prove that

$$0 < t \le \frac{\sqrt{N}}{K} \quad \Rightarrow \quad \mathsf{P}(Z_N \ge e^t \mathsf{E} Z_N) \le K \exp\left(-\frac{t^2}{K}\right) \; . \tag{11.22}$$

When $Z_N \ge e^t \mathsf{E} Z_N$, we have $Z_N^n \ge e^{tn} (\mathsf{E} Z_N)^n$ and $Z_N = \max(Z_N, \mathsf{E} Z_N)$. Since $Z_N^n = U_1 + U_2$, when $Z_N \ge e^t \mathsf{E} Z_N$, we have either

$$U_2 \ge \frac{1}{2} Z_N^n = \frac{1}{2} (\max(Z_N, \mathsf{E} Z_N))^n$$

or else

$$U_1 \ge \frac{1}{2} Z_N^n \ge \frac{e^{tn}}{2} (\mathsf{E} Z_N)^n \; .$$

Thus (11.21) yields

$$\mathsf{P}(Z_N \ge e^t \mathsf{E} Z_N) \le \mathsf{P}\left(U_1 \ge \frac{e^{tn}}{2} (\mathsf{E} Z_N)^n\right) + K \exp\left(-\frac{N}{K}\right) \; . \tag{11.23}$$

Now, by (11.20) and Markov inequality, whenever $n \le \sqrt{N}/K$ we have

$$\mathsf{P}\left(U_1 \ge \frac{e^{tn}}{2} (\mathsf{E} Z_N)^n\right) \le \frac{2 \mathsf{E} U_1}{e^{tn}} \le \frac{2 K^{n^2}}{e^{tn}} \le \left(\frac{K_1^n}{e^t}\right)^n \; . \tag{11.24}$$

The rest of the proof consists in checking that if n is the largest integer for which $K_1^n \le e^{t/2}$, the bounds (11.23) and (11.24) imply (11.22). First we may assume $K_1 \ge e$, so that $n \le t$ and therefore $n \le \sqrt{N}/K$ as soon as $t \le \sqrt{N}/K$. Also, by definition of n,

$$K_1^{n+1} \ge e^{t/2}$$

so that $n + 1 \ge t/K_2$, and therefore

$$-\frac{nt}{2} \le -\frac{t^2}{2K_2} + \frac{t}{2} \, .$$

Thus, using that $K_1^n \le e^{t/2}$ in the first inequality, we get

$$\left(\frac{K_1^n}{e^t}\right)^2 \le \exp\left(-\frac{tn}{2}\right) \le \exp\left(-\frac{t^2}{K_2} + \frac{t}{2}\right)$$
$$\le K \exp\left(-\frac{t^2}{K}\right) .$$

Combining with (11.23) and (11.24) proves (11.22). □

Lemma 11.3.3. *There exists a universal constant $c > 0$ for which (11.20) holds.*

Proof. Using (11.19) we have

$$\mathsf{E}U_1 \le 2^{-nN}(\mathsf{E}Z_N)^n \exp\left(\frac{\beta^2}{2}N \sum_{1\le\ell<\ell'\le n} R^2_{\ell,\ell'}\right) .$$

Thus it suffices to prove that

$$2^{-nN}\sum_{B_n}\exp\left(\frac{N\beta^2}{2}\sum_{1\le\ell<\ell'\le n} R^2_{\ell,\ell'}\right) \le K^{n^2} . \tag{11.25}$$

Let us denote by P_0 the uniform probability on Σ_N and by E_0 the corresponding expectation, so that (11.25) simply means

$$\mathsf{E}_0\left(\exp\left(\frac{N\beta^2}{2}\sum_{1\le\ell<\ell'\le n} R^2_{\ell,\ell'}\right)\mathbf{1}_{\{\sum_{1\le\ell<\ell'\le n} R^2_{\ell,\ell'}\le c\}}\right) \le K^{n^2}$$

or, equivalently,

$$\mathsf{E}_0(F(X)\mathbf{1}_{\{X\le Nc\}}) \le K^{n^2} \tag{11.26}$$

for $X = N\sum_{\ell<\ell'} R^2_{\ell,\ell'}$ and $F(x) = \exp \beta^2 x/2$. Using the formula (A.31),

$$\mathsf{E}_0 F(X)\mathbf{1}_{\{X\le Nc\}} = \int_{-\infty}^{Nc} F'(t)\mathsf{P}_0(t\le X<Nc)\mathrm{d}t$$
$$\le 1 + \frac{\beta^2}{2}\int_0^{Nc} f(t)\exp\frac{\beta^2 t}{2}\mathrm{d}t , \tag{11.27}$$

where $f(t) = \mathsf{P}_0(X \ge t)$. Using Corollary A.8.7 with $1-2\varepsilon = \sqrt{\beta}$ and $u = \beta t$ yields

$$f(t) \le K^{n^2}\exp\left(-\frac{\beta t}{2}\left(1 - L\sqrt{\frac{t}{N}}\right)\right) ,$$

so that

$$f(t)\exp\frac{\beta^2 t}{2} \le K^{n^2}\exp\left(-\frac{\beta t}{2}\left(1-\beta-L\sqrt{\frac{t}{N}}\right)\right)$$
$$\le K^{n^2}\exp\left(-\frac{t}{K}\right)$$

provided $t \le cN$ for $c = c(\beta)$ small enough. Thus (11.26) follows from (11.27). □

We now fix c as in Lemma 11.3.3.

Lemma 11.3.4. *Let*

$$D_n = \left\{\boldsymbol{\sigma}^1,\dots,\boldsymbol{\sigma}^n\ ;\ \frac{c}{3} \le \sum_{1\le\ell<\ell'\le n} R^2_{\ell,\ell'} \le c\right\}.$$

Then

$$\mathsf{E}\sum_{D_n} e_n(\boldsymbol{\sigma}^1,\dots,\boldsymbol{\sigma}^n) \le (\mathsf{E}Z_N)^n K^{n^2}\exp\left(-\frac{N}{K}\right). \tag{11.28}$$

Proof. This is a variation of the proof of Lemma 11.3.3. Keeping the notation of that lemma it suffices to prove that

$$\mathsf{E}_0(F(X)\mathbf{1}_{\{Nc/3\le X\le Nc\}}) \le K^{n^2}\exp\left(-\frac{N}{K}\right).$$

Using (A.32) we have

$$\begin{aligned}\mathsf{E}_0(F(X)\mathbf{1}_{\{Nc/3\le X\le Nc\}}) &= \int_{-\infty}^{\infty} F'(t)\mathsf{P}(\min(t,c/3)\le X\le c)\mathrm{d}t\\ &\le F\left(\frac{Nc}{3}\right)f\left(\frac{Nc}{3}\right)+\frac{\beta^2}{2}\int_{Nc/3}^{Nc} f(t)\exp\frac{\beta^2 t}{2}\mathrm{d}t\end{aligned}$$

and we have seen that $f(t)\exp\beta^2 t/2 \le K^{n^2}\exp(-t/K)$ for $t \le cN$. □

Consider any subset I of $\{1,\dots,N\}$, with $\mathrm{card} I = n' \ge 3$. Set

$$D_I = \left\{(\boldsymbol{\sigma}^1,\dots,\boldsymbol{\sigma}^n)\ ;\ \frac{c}{3} \le \sum_{\ell,\ell'\in I,\ell<\ell'} R^2_{\ell,\ell'} \le c\right\}.$$

Lemma 11.3.5. *If* $n \le \sqrt{N}/K$ *then*

$$\mathsf{P}\left(\sum_{D_I} e_n(\boldsymbol{\sigma}^1,\dots,\boldsymbol{\sigma}^n) \ge \exp\left(-\frac{N}{K}\right)(\mathsf{E}Z_N)^{n'}Z_N^{n-n'}\right) \le \exp\left(-\frac{N}{K}\right). \tag{11.29}$$

Proof. Without loss of generality we may assume that $I = \{1, \ldots, n'\}$. Then

$$\sum_{D_I} e_n(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n) = Z_N^{n-n'} \sum_{D'} e_n(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^{n'})$$

for

$$D' = \left\{ (\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^{n'}) \; ; \; \frac{c}{3} \le \sum_{1 \le \ell < \ell' \le n'} R_{\ell,\ell'}^2 \le c \right\} .$$

Using Markov's inequality and (11.28) for n' rather than n, we get

$$\mathsf{P}\left(\sum_{D'} e_n(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^{n'}) \ge t(\mathsf{E}Z_N)^{n'} \right) \le \frac{K_5^{n'^2}}{t} \exp\left(-\frac{N}{K_5} \right) .$$

Now, if $n \le \sqrt{N}/K_6$ we have $n'^2 \le N/K_6^2$ and $K_5^{n'^2} \exp(-N/K_5) \le \exp(-N/2K_5)$, and taking $t = \exp(-N/4K_5)$ proves (11.29). □

Consider a subset I of $\{1, \ldots, n\}$ with $\operatorname{card} I = 2$. Writing $I = \{\ell, \ell'\}$ we set

$$D_I = \left\{ (\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n) \; ; \; R_{\ell,\ell'}^2 \ge \frac{c}{3} \right\} .$$

Lemma 11.3.6. *We have*

$$\mathsf{P}\left(\sum_{D_I} e_n(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n) \ge \exp\left(-\frac{N}{K} \right) (\mathsf{E}Z_N)^2 Z_N^{n-2} \right) \le \exp\left(-\frac{N}{K} \right) . \tag{11.30}$$

Proof. Identical to that of Lemma 11.3.5, using instead of (11.28) that for $\gamma = (1 - \beta^2)/2$ we have

$$\exp \frac{cN\gamma}{2} \, \mathsf{E} \sum_{R_{1,2} \ge c} e_2(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) \le K(\mathsf{E}Z_N)^2$$

by (11.3). □

Corollary 11.3.7. *For any subset I of $\{1, \ldots, n\}$ with at least two elements we have*

$$\mathsf{P}\left(\sum_{D_I} e(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n) \ge \exp\left(-\frac{N}{K} \right) (\max(Z_N, \mathsf{E}Z_N))^n \right) \le \exp\left(-\frac{N}{K} \right) .$$

Proof. Use Lemma 11.3.5 if $\operatorname{card} I \ge 3$ and Lemma 11.3.6 if $\operatorname{card} I = 2$. □

To conclude the proof of (11.21) and of Theorem 11.3.1 it suffices to show the following.

Lemma 11.3.8. *We have $B_n^c \subset \bigcup_I D_I$, where the union is over all subsets I of $\{1, \dots, n\}$ with at least two elements.*

Proof. Given $(\boldsymbol{\sigma}^1, \dots, \boldsymbol{\sigma}^n)$ in B_n^c, we consider the smallest integer $n' \geq 2$ such that there exists a subset I of $\{1, \dots, n\}$ with $\operatorname{card} I = n'$ for which

$$\sum_{\ell,\ell' \in I, \ell < \ell'} R_{\ell,\ell'}^2 \geq \frac{c}{3} \,. \tag{11.31}$$

The fact that $(\boldsymbol{\sigma}^1, \dots, \boldsymbol{\sigma}^n)$ belongs to B_n^c is used to show that there exists at least one set I as above, namely $I = \{1, \dots, n\}$.

If $n' = \operatorname{card} I = 2$, then $(\boldsymbol{\sigma}^1, \dots, \boldsymbol{\sigma}^n) \in D_I$ and we are done. If $n' = \operatorname{card} I \geq 3$, the minimality of n' implies that given a subset I' of I with $\operatorname{card} I' = n' - 1$ we have

$$\sum_{\ell,\ell' \in I', \ell < \ell'} R_{\ell,\ell'}^2 < \frac{c}{3} \,,$$

and averaging over all the n' possible choices of I' we get

$$\frac{n'-2}{n'} \sum_{\ell,\ell' \in I, \ell < \ell'} R_{\ell,\ell'}^2 < \frac{c}{3}$$

and since $(n'-2)/n' \geq 1/3$ we get $\sum_{\ell,\ell' \in I, \ell < \ell'} R_{\ell,\ell'}^2 < c$, and combining with (11.31) shows that $(\boldsymbol{\sigma}^1, \dots, \boldsymbol{\sigma}^n) \in D_I$. □

11.4 The Aizenman-Lebowitz-Ruelle Central Limit Theorem

For $\beta < 1$ we define

$$c(\beta) = N\left(\log 2 + \frac{\beta^2}{4}\right) + \frac{1}{4}\log(1-\beta^2) \tag{11.32}$$

$$b(\beta) = \frac{1}{2}\left(\log\frac{1}{1-\beta^2} - \beta^2\right) \tag{11.33}$$

and we recall the notation $a(k) = \mathsf{E} g^k$ where g is standard Gaussian.

Theorem 11.4.1. *([9]). Consider $\beta < 1$ and $k \geq 1$. Then*

$$\left|\mathsf{E}(\log Z_N(\beta) - c(\beta))^k - a(k) b(\beta)^{k/2}\right| \leq \frac{K}{\sqrt{N}} \tag{11.34}$$

where K does not depend on N.

This is a refined version of a result of [9]. It should be compared with Theorem 1.4.11, which in the present case (since $q = 0$ and $b = 0$) yields only

$$\mathsf{E}\left|\log Z_N - N\left(\log 2 + \frac{\beta^2}{2}\right)\right|^k \le K N^{(k-1)/2} .$$

The normalization of Theorem 1.4.11 is not appropriate here because the quantity $Z_N(\beta) - N(\log 2 + \beta^2/4)$ is typically of order 1, not $\sqrt{N}$, as we demonstrated in the previous section. The content of Theorem 11.4.1 is that the k-th moment of $Z_N(\beta) - c(\beta)$ is nearly the same as the k-th moment of $b(\beta)^{1/2} g$ where g is standard Gaussian, and this with accuracy of order $N^{-1/2}$. The reader must wonder how the strange formulas (11.32) and (11.33) arise; let us first explain this.

We define $X(\beta) = \log Z_N(\beta) - c(\beta)$, and, keeping the dependence on k implicit we write

$$f(\beta) = \mathsf{E} X(\beta)^k . \tag{11.35}$$

Lemma 11.4.2. *We have*

$$\begin{aligned} f'(\beta) &= k\frac{\beta}{2}\mathsf{E}\big(N(1 - \langle R_{1,2}^2\rangle) X(\beta)^{k-1}\big) \qquad (11.36)\\ &\quad + k(k-1)\frac{\beta}{2}\mathsf{E}\big((N\langle R_{1,2}^2\rangle - 1) X(\beta)^{k-2}\big) - k c'(\beta) \mathsf{E} X(\beta)^{k-1} . \end{aligned}$$

Proof. We observe first that

$$\begin{aligned} X(\beta)' &= \left\langle \frac{1}{\sqrt{N}} \sum_{i<j} g_{ij}\sigma_i\sigma_j \right\rangle - c'(\beta) \\ &= \left\langle -\frac{H_N}{\beta} \right\rangle - c'(\beta) . \end{aligned}$$

Therefore differentiation of (11.35) yields

$$f'(\beta) = k\mathsf{E}\left(\left(\left\langle -\frac{H_N}{\beta}\right\rangle - c'(\beta)\right) X(\beta)^{k-1}\right) .$$

Integration by parts using (11.4) then implies

$$\begin{aligned} \mathsf{E}\left(\left\langle -\frac{H_N}{\beta}\right\rangle X(\beta)^{k-1}\right) &= \frac{\beta}{2}\mathsf{E}\big(N(1 - R_{1,2})^2 X(\beta)^{k-1}\big) \\ &\quad + (k-1)\frac{\beta}{2}\mathsf{E}\big((N\langle R_{1,2}^2\rangle - 1) X(\beta)^{k-2}\big) . \end{aligned}$$

In this formula, the first term comes as usual from the dependence of the bracket $\langle\cdot\rangle$ on the r.v.s g_{ij} and the second term comes from the dependence of $X(\beta)$ on these variables. □

The central fact is as follows.

Lemma 11.4.3. *We have*

$$\mathsf{E}\big(N\langle R_{1,2}^2\rangle X(\beta)^k\big) = \frac{1}{1-\beta^2}\mathsf{E}X(\beta)^k + \mathcal{R} \tag{11.37}$$

where $\mathcal{R}$ denotes a quantity such that $\sqrt{N}|\mathcal{R}|$ remains bounded independently of N and of β whenever $\beta \le \beta_0 < 1$.

Throughout this section $\mathcal{R}$ denotes such a quantity, that need not be the same at each occurrence.

Proof of Theorem 11.4.1. Using (11.37) for $k-1$ or $k-2$ rather than k in (11.36) we get

$$\begin{aligned} f'(\beta) = k\left(\frac{\beta}{2}N - \frac{\beta}{2(1-\beta^2)} - c'(\beta)\right)\mathsf{E}X(\beta)^{k-1} \\ + k(k-1)\frac{\beta^3}{2(1-\beta^2)}\mathsf{E}X(\beta)^{k-2} + \mathcal{R}\,. \end{aligned} \tag{11.38}$$

With the choices (11.32) and (11.33) we have

$$\begin{aligned} c'(\beta) &= \frac{\beta}{2}N - \frac{\beta}{2(1-\beta^2)} \\ b'(\beta) &= \frac{\beta^3}{2(1-\beta^2)} \end{aligned} \tag{11.39}$$

and (11.38) becomes

$$f'(\beta) = \frac{k(k-1)}{2}b'(\beta)\mathsf{E}X(\beta)^{k-2} + \mathcal{R}\,,$$

from which the formula

$$\mathsf{E}X(\beta)^k = a(k)b(\beta)^{k/2} + \mathcal{R}$$

follows by induction over k and integration (since $X(0) = 0$ and $a(k) = (k-1)a(k-2)$). □

Lemma 11.4.4. *For each k we have $\mathsf{E}X(\beta)^k \le K$ where K does not depend on N and stays bounded as $\beta \le \beta_0 < 1$.*

Proof. Using Hölder's inequality we may assume that k is even. Using (11.36) and (11.39) we obtain

$$\begin{aligned} f'(\beta) = -\frac{k\beta}{2}\mathsf{E}\big(N\langle R_{1,2}^2\rangle X(\beta)^{k-1}\big) + \frac{k\beta}{2(1-\beta^2)}\mathsf{E}X(\beta)^{k-1} \\ + k(k-1)\frac{\beta}{2}\mathsf{E}\big((N\langle R_{1,2}^2\rangle - 1)X(\beta)^{k-2}\big)\,. \end{aligned}$$

It follows from (11.14) that for any k, $\mathsf{E}\langle (NR_{1,2}^2)^k\rangle$ remains bounded for $\beta \le \beta_0$. Using Hölder's inequality implies

$$f'(\beta) \le C(k,\beta)(f(\beta)^{\frac{k-1}{k}} + f(\beta)^{\frac{k-2}{k}}) ,$$

where $C(k,\beta)$ remains bounded as $\beta \le \beta_0 \le 1$. Since $x^\tau \le 1 + x$ for $\tau < 1$, this gives

$$f'(\beta) \le 2C(k,\beta)(f(\beta)+1)$$

and the result follows by integration. □

Exercise 11.4.5. Deduce Lemma 11.4.4 from Theorems 11.2.1 and 11.3.1.

It is obvious that something like (11.37) should be true. Indeed, the Cauchy-Schwarz inequality implies

$$\begin{aligned}&\mathsf{E}\left(\left(N\langle R_{1,2}^2\rangle - \frac{1}{1-\beta^2}\right) X(\beta)^k\right)\\ &\le \mathsf{E}\left(\left(N\langle R_{1,2}^2\rangle - \frac{1}{1-\beta^2}\right)^2\right)^{1/2} (\mathsf{E} X(\beta)^{2k})^{1/2} .\end{aligned}$$

From Lemma 11.4.4 the last term is bounded, and using replicas and (11.48) below, we obtain

$$\mathsf{E}\left(N\langle R_{1,2}^2\rangle - \frac{1}{1-\beta^2}\right)^2 = \mathcal{R} .$$

This argument gives only a rate $N^{-1/4}$ in (11.37). To obtain the true rate $N^{-1/2}$ one has to work harder. We sketch the argument, which uses the cavity method to decouple the last spin. First, we observe that symmetry among sites implies

$$N\mathsf{E}\big(\langle R_{1,2}^2\rangle X(\beta)^k\big) = N\mathsf{E}\big(\langle \varepsilon_1\varepsilon_2 R_{1,2}^-\rangle X(\beta)^k\big) + \mathsf{E} X(\beta)^k , \tag{11.40}$$

where $\varepsilon_\ell = \sigma_N^\ell$ and $R_{1,2}^- = N^{-1}\sum_{i<N}\sigma_i^1\sigma_i^2$. Fixing $\beta < 1$, consider the usual interpolating Hamiltonian

$$-H_t(\boldsymbol{\sigma}) = \frac{\beta}{\sqrt{N}}\sum_{i<j<N} g_{ij}\sigma_i\sigma_j + \frac{\beta\sqrt{t}}{\sqrt{N}}\sigma_N\sum_{i<N} g_{iN}\sigma_i , \tag{11.41}$$

and denote by $\langle\cdot\rangle_t$ an average for the corresponding Gibbs measure. Let

$$Y(t) = \log\sum_{\boldsymbol{\sigma}}\exp(-H_t(\boldsymbol{\sigma})) - c(\beta) ,$$

so $Y(1) = X(\beta)$. We consider the function

$$\varphi(t) = N\mathsf{E}\big(\langle\varepsilon_1\varepsilon_2 R_{1,2}^-\rangle_t Y(t)^k\big) ,$$

so that $\varphi(0) = 0$. Using that $|\varphi(1) - \varphi(0) - \varphi'(0)| \le \sup_{0\le t\le 1} |\varphi''(t)|$, the objective is to show that

$$\varphi'(0) = \beta^2 N\mathsf{E}\big(\langle R_{1,2}^2\rangle X(\beta)^k\big) + \mathcal{R} \tag{11.42}$$

and

$$\sup_t |\varphi''(t)| = \mathcal{R}\,. \tag{11.43}$$

Indeed, since $\varphi(1) = N\mathsf{E}(\langle \varepsilon_1\varepsilon_2 R_{1,2}^-\rangle X(\beta)^k)$, it then follows from (11.40) that

$$N\mathsf{E}\big(\langle R_{1,2}^2\rangle X(\beta)^k\big) = \beta^2 N\mathsf{E}\big(\langle R_{1,2}^2\rangle X(\beta)^k\big) + \mathsf{E}X(\beta)^k + \mathcal{R}\,,$$

which is (11.37). The reader having reached this point should find the proof of (11.42) and (11.43) to be a mere exercise.

Exercise 11.4.6. Complete the proof of (11.42) and (11.43). Hint: A preliminary step is to use a differential inequality and Lemma 11.4.4 in order to prove that for each n the quantity $\mathsf{E}Y(t)^{2n}$ remains bounded independently of t. Moreover, observe that $\varphi'(0) = \beta^2 N\mathsf{E}(\langle (R_{1,2}^-)^2\rangle_0 Y(0)^k)$, and prove (11.42) by showing that $\psi(t) = \mathsf{E}(\langle (R_{1,2}^-)^2\rangle_t Y(t)^k)$ satisfies $|\psi'(t)| \le \mathcal{R}$.

11.5 The Matrix of Spin Correlations

In this section we consider the matrix $\mathcal{M} = (\langle \sigma_i\sigma_j\rangle)_{i,j}$ of spin correlations, and its operator norm

$$\|\mathcal{M}\| = \sup\left\{\sum_{i,j\le N} x_i x_j \langle\sigma_i\sigma_j\rangle\ ;\ \sum_{i\le N} x_i^2 \le 1\right\}.$$

Conjecture 11.5.1. For $\beta < 1$ we have $\mathsf{E}\|\mathcal{M}\| \le K$.

The size of $\|\mathcal{M}\|$ seems to be related to the Research Problem 11.2.11. The intuition behind Conjecture 11.5.1 is that the main contribution to $\langle\sigma_i\sigma_j\rangle$ is created by the interaction term $\beta g_{ij}\sigma_i\sigma_j/\sqrt{N}$; and that the matrix $\mathcal{M}' = (g_{ij}/\sqrt{N})_{i,j}$ satisfies $\mathsf{E}\|\mathcal{M}'\| \le K$. However the main appeal of Conjecture 11.5.1 is that its study leads to rather pretty mathematics. A standard way to control the operator norm $\|\mathcal{M}\|$ of $\mathcal{M}$ is through the traces of the powers of $\mathcal{M}$. Since $\mathcal{M}$ is symmetric, its norm is the absolute value of its largest eigenvalue. The sum of the k^{th} powers of the eigenvalues is the trace of $\mathcal{M}^k$, so that, for k even

$$\|\mathcal{M}\| \le (\text{trace } \mathcal{M}^k)^{1/k}\,. \tag{11.44}$$

The trace of $\mathcal{M}^k$ has a nice expression, as the next lemma shows.

Lemma 11.5.2. *We have, for all positive integers k,*

$$\text{trace}\ \mathcal{M}^k = N^k \langle R_{1,2}\, R_{2,3} \cdots R_{k,1} \rangle\ . \tag{11.45}$$

Proof. It is immediate by induction over k that the entries of $\mathcal{M}^k$ are

$$N^{k-1} \langle \sigma_i^1\, R_{1,2}\, R_{2,3} \cdots R_{k-1,k}\, \sigma_j^k \rangle\ .$$ □

Corollary 11.5.3. *For all even integers k, we have*

$$\mathsf{E}\, \|\mathcal{M}\| \le (\mathsf{E}\, N^k \langle R_{1,2}\, R_{2,3} \cdots R_{k,1} \rangle)^{1/k}\ . \tag{11.46}$$

Proof. Combine (11.44) and (11.45). □

Thus, this approach leads to the question of evaluating

$$\nu(N^k R_{1,2} R_{2,3} \cdots R_{k,1})\ . \tag{11.47}$$

Some information on this quantity is provided by Theorem 1.10.1. For convenience of the reader we state this theorem in the present case.

Theorem 11.5.4. *Consider numbers $k(\ell,\ell')$ for $1 \le \ell < \ell' \le n$ and $k = \sum_{1\le\ell<\ell'\le n} k(\ell,\ell')$. Then for $\beta < 1$ we have*

$$\nu\left(\prod_{1\le\ell<\ell'\le n} R_{\ell,\ell'}^{k(\ell,\ell')}\right) = \frac{1}{(N(1-\beta^2))^{k/2}} \prod_{1\le\ell<\ell'\le n} a(k(\ell,\ell')) + O(k+1) \tag{11.48}$$

where $a(k) = \mathsf{E} g^k$, and where $O(k)$ denotes a quantity A such that $|A| \le K N^{-k/2}$ for a number K that does not depend on N.

Theorem 1.10.1 was proved only for $\beta < 1/2$. In the present situation $\beta < 1$ we know (11.14) and, consequently

$$\nu(R_{1,2}^{2k}) \le \left(\frac{Kk}{N}\right)^k \tag{11.49}$$

where K depends on β and k only. It should be obvious that this suffices to have the proof of Theorem 1.10.1 go through. (Of course, this proof is very much simpler in the present case than in the general case.)

In the case (11.47), it follows from (11.48) that (for $k \ge 3$)

$$|\nu(N^k R_{1,2} R_{2,3} \cdots R_{k,1})| \le K N^{(k-1)/2}\ ,$$

which is not very useful for our purpose. In the present section we will improve the previous inequality into

$$|\nu(N^k R_{1,2} \cdots R_{k,1})| \le K(\beta, k)N \tag{11.50}$$

which combined with (11.46) yields

$$\mathsf{E}\|\mathcal{M}\| \le K(\beta, k)N^{1/k} .$$

In a further effort we will show that in fact

$$\mathsf{E}\|\mathcal{M}\| \le (K(\beta)k)^8 N^{1/k} , \tag{11.51}$$

so that taking k about $\log N$ yields

$$\mathsf{E}\|\mathcal{M}\| \le K(\log N)^8 .$$

It is probably not difficult to improve on the exponent 8, which simply occurs as consequence of several convenient choices, but there is little point in doing this. It is a entirely different matter to remove the factor $\log N$ altogether. One should of course expect that the solution of Conjecture 11.5.1 will be obtained through an improvement of the estimate (11.51). Considering the case $k = 2$ in (11.50) we see that the dependence on N seems optimal. The dependence on k is a much more delicate matter. The author does not see any place where his estimates are obviously suboptimal, but of course there could be much more cancellation than is readily apparent.

The method of the proof is somewhat similar to that of Theorem 11.5.4, and relies on the cavity method. We recall the interpolating Hamiltonian (11.41) and that we denote as usual $\langle\cdot\rangle_t$ an average for the corresponding Gibbs measure and $\nu_t(\cdot) = \mathsf{E}\langle\cdot\rangle_t$.

It will not suffice to make order 2 expansions, and we will have to make higher order expansions. The proofs will ultimately rely on the formula

$$\left|\nu(f) - \sum_{k=0}^{s} \frac{1}{k!}\nu_0^{(k)}(f)\right| \le \sup_{0\le t\le 1} |\nu_t^{(s+1)}(f)| . \tag{11.52}$$

The remainder term will be bounded somewhat crudely, and the terms $\nu_0^{(k)}(f)$ will be bounded using a suitable induction hypothesis and some combinatorics.

Let us turn to the description of the induction hypothesis. We consider a fixed family $\mathcal{I}$ of pairs (ℓ, ℓ'). Each pair might be repeated several times. Setting

$$T_{\mathcal{I}} = \prod_{(\ell,\ell')\in\mathcal{I}} NR_{\ell,\ell'} ,$$

we aim at computing the order of magnitude of $\nu(T_{\mathcal{I}})$ as $N \to \infty$. A first observation is that $\nu(T_{\mathcal{I}}) = 0$ unless

$$\text{Each integer } \ell_0 \text{ appears an even number of times as an element of a pair } (\ell, \ell') \in \mathcal{I} . \tag{11.53}$$

This is because if ℓ_0 appears an odd number of times, then the transformation $\boldsymbol{\sigma}^{\ell_0} \mapsto -\boldsymbol{\sigma}^{\ell_0}$ changes the sign of $T_{\mathcal{I}}$, while leaving $\langle\cdot\rangle$ invariant, so that $\nu(T_{\mathcal{I}}) = 0$.

Thus we assume that (11.53) is satisfied. Let us say that $\mathcal{I}$ is a *cycle* if it consists of the pairs $(\ell_1, \ell_2), (\ell_2, \ell_3), \ldots, (\ell_p, \ell_1)$, where the elements $\ell_1, \ldots, \ell_p$ are all different. A cycle satisfies condition (11.53), and so does a union of cycles. Conversely if $\mathcal{I}$ satisfies condition (11.53), it is a union of cycles. To see this, it suffices to show that if $\mathcal{I}$ satisfies (11.53), it contains a cycle $\mathcal{I}_1$, because then $\mathcal{I}\backslash\mathcal{I}_1$ will still satisfy (11.53). To construct a cycle in $\mathcal{I}$, we start with a pair (ℓ_1, ℓ_2) in $\mathcal{I}$ and we recursively choose integers ℓ_s such that $(\ell_{s-1}, \ell_s) \in \mathcal{I}$ and $\ell_s \notin (\ell_1, \ell_2, \ldots, \ell_{s-1})$, as long as we can. When the construction stops, we have constructed ℓ_s and we pick $r \in \{1, \ldots, s-1\}$ such that $(\ell_s, \ell_r) \in \mathcal{I}$. The existence of r follows from the fact that ℓ_s must occur twice as an element of a pair of $\mathcal{I}$. Then $(\ell_r, \ell_{r+1}), \ldots, (\ell_{s-1}, \ell_s), (\ell_s, \ell_r)$ is the required cycle in $\mathcal{I}$.

In general, a family $\mathcal{I}$ satisfying (11.53) can be decomposed as a union of cycles in several different manners. For example,

$$\mathcal{I} = \{(1,2), (1,2), (2,3), (2,3), (3,1), (3,1)\}$$

is the union of two cycles of length 3, and is also the union of three cycles of length 2. We denote by $C(\mathcal{I})$ the *maximum* number of cycles in which $\mathcal{I}$ can be decomposed.

Theorem 11.5.5. *We have*

$$|\nu(T_{\mathcal{I}})| \leq K(\beta, \mathcal{I}) N^{C(\mathcal{I})} . \tag{11.54}$$

Research Problem 11.5.6. (Level 2) Prove that the limit

$$\lim_{N\to\infty} N^{-C(\mathcal{I})} \nu(T_{\mathcal{I}})$$

exists, *and* find a manageable expression for it.

Proving existence of the limit should be a mere exercise; there is no guarantee however that the limit is a simple function of the "geometry" of $\mathcal{I}$.

In order to establish Theorem 11.5.5, we will prove a more general statement.

Proposition 11.5.7. *Given an integer* s*, given* $\beta_0 < 1$*, given a family* $\mathcal{I}$*, we have*

$$2C(\mathcal{I}) \leq \mathrm{card}\mathcal{I} - s \Rightarrow \forall\, \beta \leq \beta_0\, , |\nu(T_{\mathcal{I}})| \leq K(\beta_0, \mathcal{I}, s) N^{(\mathrm{card}\mathcal{I} - s)/2} . \tag{11.55}$$

The case $s < 0$ is allowed. This case does not make much sense since $2C(\mathcal{I}) \leq \mathrm{card}\mathcal{I}$, but will be useful in the proof.

The choice $s = \mathrm{card}\mathcal{I} - 2C(\mathcal{I})$ shows that this is more general than Theorem 11.5.5. The point of the formulation is to allow induction on s. The case $s \le 0$ follows from (11.49) and Hölder's inequality.

Proof. The proof is by induction over s. For the induction step, we assume that (11.55) holds for each value $s' < s$, and we prove by induction over $\mathrm{card}\,\mathcal{I}$ that it holds for s. Thus, in the course of the proof, we may use (11.55) for a family $\mathcal{I}_0$, and a number s_0 if *either* $s_0 < s$ (in which case $\mathrm{card}\,\mathcal{I}_0$ can be whatever we want) *or* $s_0 = s$ and $\mathrm{card}\,\mathcal{I}_0 < \mathrm{card}\,\mathcal{I}$. we may also use (11.55) for s_0 and $\mathcal{I}_0$ as before, but for $N-1$ rather than N.

We fix $\beta_0 < 1$ and we denote throughout the proof a number K depending only on $s, \mathcal{I}, \beta_0$ and which may vary between occurrences. We assume $\beta \le \beta_0$.

Given a family $\mathcal{I}_1$ of pairs (ℓ, ℓ'), we set

$$T^-_{\mathcal{I}_1} = \prod_{(\ell,\ell')\in\mathcal{I}_1} (NR^-_{\ell,\ell'}) \,.$$

Keeping in mind that

$$NR_{\ell,\ell'} = \sum_{i\le N} \sigma_i^\ell \sigma_i^{\ell'} \;;\; NR^-_{\ell,\ell'} = \sum_{i\le N-1} \sigma_i^\ell \sigma_i^{\ell'} \,,$$

we see that $T^-_{\mathcal{I}_1}$ is obtained from $T_{\mathcal{I}_1}$ by replacing N by $N-1$. The principle of the proof is to establish the inequality

$$|\nu(T_\mathcal{I}) - \beta^2 \nu_0(T^-_\mathcal{I})| \le K N^{(\mathrm{card}\mathcal{I}-s)/2} \,. \tag{11.56}$$

The quantity $\nu_0(T^-_\mathcal{I})$ is of the same nature as $\nu(T_\mathcal{I})$, except that one has replaced N by $N-1$ and β by β^-, where $\beta^- = \beta\sqrt{1-1/N} \le \beta_0$. We consider the quantity

$$A(N_0, \beta_0, \mathcal{I}, s) = \sup\{N^{-(\mathrm{card}\mathcal{I}-s)/2} |\nu(T_\mathcal{I})| \;;\; \beta \le \beta_0 \,,\; N \le N_0\} \,,$$

and we observe (since $\beta^- \le \beta$) that (11.56) implies

$$A(N_0, \beta_0, \mathcal{I}, s) \le \beta_0^2 A(N_0, \beta_0, \mathcal{I}, s) + K \,,$$

so that $A(N_0, \beta_0, \mathcal{I}, s) \le K$, which is the required inequality. So all we have to do is to prove (11.56).

Without loss of generality we assume $(1,2) \in \mathcal{I}$, and we set $\mathcal{I}' = \mathcal{I}\backslash\{(1,2)\}$. (If the pair $(1,2)$ occurs several times in $\mathcal{I}$, we remove it just once.) As always, we use symmetry between sites to write

$$\nu(T_\mathcal{I}) = \nu(f) \,, \tag{11.57}$$

where

$$f = N\varepsilon_1\varepsilon_2 T_{\mathcal{I}'} \,. \tag{11.58}$$

We now appeal to formula (11.52). Since f contains $\operatorname{card}\mathcal{I}$ factors N and $\operatorname{card}\mathcal{I}-1$ factors $R_{\ell,\ell'}$, it follows from (11.49) and Hölder's inequality that

$$\nu(f^2)^{1/2} \le KN^{(\operatorname{card}\mathcal{I}+1)/2} \, .$$

Since "taking a derivative" brings out a factor $N^{-1/2}$ we expect that

$$|\nu_t^{(s+1)}(f)| \le KN^{(\operatorname{card}\mathcal{I}-s)/2} \, . \tag{11.59}$$

To prove this formally, we compute $\nu^{(k)}(f)$ by iteration of (1.151). (It is slightly more convenient here to use (1.151) than (1.150)). We find that $\nu^{(k)}(f)$ is a sum of terms of the type

$$\pm\beta^{2k}\nu_t(f\varepsilon_{\ell_1}\varepsilon_{\ell_2}R_{\ell_1,\ell_2}\varepsilon_{\ell_3}\varepsilon_{\ell_4}R_{\ell_3,\ell_4}\cdots\varepsilon_{\ell_{2k-1}}\varepsilon_{\ell_{2k}}R_{\ell_{2k-1},\ell_{2k}}) \, ,$$

which we bound by $\nu_t(|f||R_{\ell_1,\ell_2}|\cdots|R_{\ell_{2k-1},\ell_{2k}}|)$; we then use (11.49) and Hölder's inequality.

To prove (11.56) we prove that

$$\left|\sum_{k=0}^{s}\frac{1}{k!}\nu_0^{(k)}(f)-\beta^2\nu_0(T^-_{\mathcal{I}_1})\right| \le KN^{(\operatorname{card}\mathcal{I}-s)/2} \, . \tag{11.60}$$

We compute $\nu_0^{(k)}(f)$ when $k\le s$ by iteration of (1.150) (which is now more convenient than (1.151)). We find that it is a sum of terms

$$\pm\,\beta^{2k}\nu_0(\varepsilon_1\varepsilon_2\varepsilon_{\mathcal{I}_1}T_{\mathcal{I}'}T^-_{\mathcal{I}_1}N^{1-k}) \, , \tag{11.61}$$

where $\mathcal{I}_1$ is a family of k pairs (ℓ,ℓ'), and where

$$\varepsilon_{\mathcal{I}_1} = \prod_{(\ell,\ell')\in\mathcal{I}_1}\varepsilon_\ell\varepsilon_{\ell'} \, .$$

To see this, we iterate the formula (1.150) and we observe that

$$\prod_{(\ell,\ell')\in\mathcal{I}_1}R^-_{\ell,\ell'} = N^{-\operatorname{card}\mathcal{I}_1}T^-_{\mathcal{I}_1} \, .$$

Before we may use Lemma 1.6.2 to compute $\nu_0^{(k)}(f)$, we must bring out the dependence of $T_{\mathcal{I}'}$ on the variables ε_ℓ. For this we write each factor $NR_{\ell,\ell'}$ of $T_{\mathcal{I}'}$ as $\varepsilon_\ell\varepsilon_{\ell'}+NR^-_{\ell,\ell'}$, and we expand the resulting product. We get

$$T_{\mathcal{I}'} = \sum_{\mathcal{I}_2}\varepsilon_{\mathcal{I}'\setminus\mathcal{I}_2}T^-_{\mathcal{I}_2} \, ,$$

where the sum is over all subsets $\mathcal{I}_2$ of $\mathcal{I}'$. Thus, we are now concerned with terms of the type

$$\pm\beta^{2k}N^{1-k}\nu_0(\varepsilon_1\varepsilon_2\varepsilon_{\mathcal{I}_1}\varepsilon_{\mathcal{I}'\setminus\mathcal{I}_2}T^-_{\mathcal{I}_2}T^-_{\mathcal{I}_1})$$
$$= \pm\beta^{2k}N^{1-k}\nu_0(\varepsilon_{\mathcal{I}_1}\varepsilon_{\mathcal{I}\setminus\mathcal{I}_2}T^-_{\mathcal{I}_1\cup\mathcal{I}_2}) , \tag{11.62}$$

because $\varepsilon_1\varepsilon_2\varepsilon_{\mathcal{I}'\setminus\mathcal{I}_2} = \varepsilon_{\mathcal{I}\setminus\mathcal{I}_2}$ and $T^-_{\mathcal{I}_1}T^-_{\mathcal{I}_2} = T^-_{\mathcal{I}_1\cup\mathcal{I}_2}$. Now Lemma 1.6.2 shows that the quantity (11.62) is zero unless $\varepsilon_{\mathcal{I}_1}\varepsilon_{\mathcal{I}\setminus\mathcal{I}_2}$ is identically 1. Since $(1,2)\in\mathcal{I}\setminus\mathcal{I}_2$, we must have

$$\mathrm{card}\mathcal{I}_1 + \mathrm{card}(\mathcal{I}\setminus\mathcal{I}_2) = k + \mathrm{card}(\mathcal{I}\setminus\mathcal{I}_2) \geq 2 . \tag{11.63}$$

We first examine the case where there is equality. A first possibility is that $\mathrm{card}(\mathcal{I}\setminus\mathcal{I}_2) = 1 = k$. Then, keeping in mind that $(1,2)\in\mathcal{I}\setminus\mathcal{I}_2$, for $\varepsilon_{\mathcal{I}_1}\varepsilon_{\mathcal{I}\setminus\mathcal{I}_2}$ to be identically 1, we must have $\mathcal{I}_1 = \mathcal{I}\setminus\mathcal{I}_2 = \{(1,2)\}$ so that $\mathcal{I}_1\cup\mathcal{I}_2 = \mathcal{I}$, and the term is

$$\beta^2\nu_0(T^-_{\mathcal{I}}) .$$

We will show that all the other terms are $\leq KN^{(\mathrm{card}\mathcal{I}-s)/s}$, concluding the proof of (11.60) and of the proposition. A second possibility for equality in (11.63) is that $k=0$ and $\mathrm{card}(\mathcal{I}\setminus\mathcal{I}_2) = 2$. For $\varepsilon_{\mathcal{I}\setminus\mathcal{I}_2}$ to be identically 1, we must have $\mathcal{I}\setminus\mathcal{I}_2 = \{(1,2),(1,2)\}$ so that $\mathcal{I}_2 = \mathcal{I}\setminus\{(1,2),(1,2)\}$. Since $\mathcal{I}$ is the union of $\mathcal{I}_2$ and a cycle, we have $C(\mathcal{I})\geq C(\mathcal{I}_2)+1$, i.e. $C(\mathcal{I}_2)\leq C(\mathcal{I})-1$. Also $\mathrm{card}\,\mathcal{I}_2 = \mathrm{card}\,\mathcal{I}-2$, so that the condition $2C(\mathcal{I})\leq \mathrm{card}\,\mathcal{I}-s$ implies

$$2C(\mathcal{I}_2) \leq \mathrm{card}\,\mathcal{I}_2 - s .$$

Since $\mathrm{card}\,\mathcal{I}_2 < \mathrm{card}\,\mathcal{I}$, we can use (11.55) for the same value of s, for β^- rather than β and $N-1$ rather than N to see that the absolute value of the corresponding term in (11.62) is bounded by

$$|N\nu_0(T^-_{\mathcal{I}_2})| \leq KN^{1+(\mathrm{card}\mathcal{I}_2-s)/2} = KN^{(\mathrm{card}\mathcal{I}-s)/2} .$$

Now we examine the case where

$$k + \mathrm{card}(\mathcal{I}\setminus\mathcal{I}_2) \geq 3 , \tag{11.64}$$

and we want to study

$$N^{1-k}\nu_0(T^-_{\mathcal{I}_1\cup\mathcal{I}_2}) . \tag{11.65}$$

We only have to consider the case where the family $\mathcal{I}_1\cup\mathcal{I}_2$ satisfies (11.53), for otherwise, as already pointed out, the quantity (11.65) is zero. In that case, $\mathcal{I}_1\cup\mathcal{I}_2$ is a union of cycles. We claim that

$$C(\mathcal{I}_1\cup\mathcal{I}_2) \leq C(\mathcal{I}) + k - 1 . \tag{11.66}$$

To see this, we consider a decomposition of $\mathcal{I}_1\cup\mathcal{I}_2$ into cycles. Let n_1 (resp. n_2) be the number of cycles that contain at least one pair of $\mathcal{I}_1$ (resp. that consist entirely of pairs of $\mathcal{I}_2$). Then $n_1\leq \mathrm{card}\,\mathcal{I}_1 = k$. Also, $C(\mathcal{I})\geq n_2+1$. This is because the set obtained by removing from $\mathcal{I}$ all the elements of $\mathcal{I}_2$

making up the n_2 cycles consisting only of elements of $\mathcal{I}_2$ is not empty, since it contains $(1,2)$ because $\mathcal{I}_2 \subset \mathcal{I}' = \mathcal{I}\backslash\{(1,2)\}$. Thus

$$n_1 + n_2 \le C(\mathcal{I}) + k - 1 \ ,$$

and this proves (11.66). It follows from (11.66) that

$$\begin{aligned} 2C(\mathcal{I}_1 \cup \mathcal{I}_2) &\le 2C(\mathcal{I}) + 2k - 2 \\ &\le \operatorname{card}\mathcal{I} + 2k - s - 2 \\ &= \operatorname{card}(\mathcal{I}_1 \cup \mathcal{I}_2) - s_1 \ , \end{aligned} \tag{11.67}$$

where, since $\operatorname{card}\mathcal{I}_1 = k$,

$$\begin{aligned} s_1 &= s + 2 + \operatorname{card}(\mathcal{I}_1 \cup \mathcal{I}_2) - \operatorname{card}\mathcal{I} - 2k \\ &= s + 2 + \operatorname{card}\mathcal{I}_2 - \operatorname{card}\mathcal{I} - k \\ &= s + 2 - \operatorname{card}(\mathcal{I}\backslash\mathcal{I}_2) - k \ . \end{aligned} \tag{11.68}$$

Thus, under (11.64) we have $s_1 \le s - 1$, and by (11.67) we can use (11.55) for $\mathcal{I}_1 \cup \mathcal{I}_2$ rather that $\mathcal{I}$, for $N-1$ rather than N and s_1 rather than s to obtain

$$|\nu_0(T^-_{\mathcal{I}_1 \cup \mathcal{I}_2})| \le K N^{(\operatorname{card}(\mathcal{I}_1 \cup \mathcal{I}_2) - s_1)/2} \ ,$$

and therefore

$$N^{1-k} |\nu_0(T^-_{\mathcal{I}_1 \cup \mathcal{I}_2})| \le K N^{1-k+(\operatorname{card}(\mathcal{I}_1 \cup \mathcal{I}_2) - s_1)/2} \ .$$

Now, the first line of (11.68) implies

$$1 - k + \frac{1}{2}(\operatorname{card}(\mathcal{I}_1 \cup \mathcal{I}_2) - s_1) = \frac{1}{2}(\operatorname{card}\mathcal{I} - s) \ , \tag{11.69}$$

and this finishes the proof of (11.56) and of Proposition 11.5.7. □

For the truly energetic reader, we now explain how to carry out an explicit value for the constant $K(\beta_0, \mathcal{I}, s)$ in (11.55).

Proposition 11.5.8. *Given $\beta_0 < 1$, an integer $s \ge 0$, a family $\mathcal{I}$, with $m = \operatorname{card}\mathcal{I}$, we have*

$$\begin{aligned} &2C(\mathcal{I}) \le m - s \Rightarrow \forall\, \beta \le \beta_0, \\ &|\nu(T_{\mathcal{I}})| \le K(\beta_0)^{m+7s}(m+s)^{m+7s} N^{(m-s)/2} \ . \end{aligned} \tag{11.70}$$

Strange quantities such as $m + 7s$ are no magic. They merely correspond to a convenient choice.

We first observe that when $C(\mathcal{I}) = 1$, taking $s = m - 2$ and recalling (11.45) proves (11.51).

The proof of (11.70) consists simply in repeating the proof of (11.55), while attempting to estimate more precisely the various contributions. Through the proof, $K(\beta_0)$ will denote a number depending on β_0 only, that need not be the same at each occurrence.

First, we note that

$$\nu_t(T_{\mathcal{I}}) \le N^m \prod_{(\ell,\ell')\in\mathcal{I}} \nu_t(R_{\ell,\ell'}^{2m})^{1/2m} = N^m \nu_t(R_{1,2}^{2m})^{1/2} \le K(\beta_0)^m m^{m/2} N^{m/2} \tag{11.71}$$

as is seen from (1.24) and (11.49). Therefore Proposition 11.5.8 holds when $s=0$, and we may assume $s \ge 1$. Next, when f is a function on n replicas, by iteration of (1.151) we may describe $\nu_t'(f)$ as the sum of $2n^2$ terms of the type

$$\pm\, \beta^2 \nu_t(f\varepsilon_\ell\varepsilon_{\ell'}R_{\ell,\ell'})\ , \tag{11.72}$$

if we count the terms in (1.150) "with their order of multiplicity"; e.g. $-\beta^2\nu_t(f\varepsilon_\ell\varepsilon_{n+1}R_{\ell,n+1})$ occurs n times if $\ell \le n$. The terms (11.72) involve the function $f\varepsilon_\ell\varepsilon_{\ell'}R_{\ell,\ell'}$, that might depend on $n+2$ replicas. By iteration, $\nu_t^{(k)}(f)$ is the sum of at most

$$2^k(n(n+2)\cdots(n+2k-2))^2 \le 2^k(n+2k)^{2k} \tag{11.73}$$

terms of the type

$$\pm\beta^{2k}\nu_t(f\varepsilon_{\mathcal{I}_1}T_{\mathcal{I}_1}N^{-k})\ ,$$

where $\mathcal{I}_1$ is a family of k pairs (ℓ,ℓ') and where $\varepsilon_{\mathcal{I}_1}$ is as in (11.61). We apply this to $f = N\varepsilon_1\varepsilon_2T_{\mathcal{I}'}$, $k=s+1$. Obviously f depends on $n \le 2m$ replicas, so the total number of terms is at most

$$2^{s+1}(2(m+s+1))^{2(s+1)} \le (4(m+s))^{2(s+1)}\ ,$$

the inequality using that $m \ge 2$ and $s \ge 1$ so that $m+s+1 \le \sqrt{2}(m+s)$. The term $f\varepsilon_{\mathcal{I}_1}T_{\mathcal{I}_1}N^{-k} = N^{1-k}\varepsilon_1\varepsilon_2\varepsilon_{\mathcal{I}_1}T_{\mathcal{I}'\cup\mathcal{I}_1}$ contains $k-1=s$ factors N^{-1}, and since $\operatorname{card}\mathcal{I}' = m-1$ and $\operatorname{card}\mathcal{I}_1 = k$, it contains $m-1+k = m-1+s+1 = m+s$ factors $NR_{\ell,\ell'}$, so that, using (11.49) for ν_t, it is at most

$$K(\beta_0)^{m+s}(m+s)^{(m+s)/2}N^{(m-s)/2}.$$

Since we have at most $(4(m+s))^{2(s+1)}$ such terms, we have improved (11.59) into

$$|\nu_t^{(s+1)}(f)| \le K(\beta_0)^{m+s}(m+s)^{m+7s}N^{(m-s)/2} \tag{11.74}$$

since, s being at least 1, the following holds true

$$\frac{m+s}{2} + 2(s+1) \le m+7s\ .$$

Now we examine the contribution of the terms (11.62). We keep in mind that we attempt to prove by induction that (11.70) holds for a certain constant $K_1(\beta_0)$ that we will determine later. When there is equality in (11.63), we have seen that, besides the term $\beta^2\nu_0(T^-_{\mathcal{I}})$, the only possible contribution is $N\nu^-_0(T^-_{\mathcal{I}_2})$ where $\mathcal{I}_2 = \mathcal{I}\backslash\{(1,2),(1,2)\}$, and that $2C(\mathcal{I}_2) \le \operatorname{card}\mathcal{I}_2 - s$. Using the induction hypothesis (11.70) with $m-2$ rather than m and the same value of s we get

$$N|\nu_0(T^-_{\mathcal{I}_2})| \le K_1(\beta_0)^{m+7s-2}(m+s)^{m+7s}N^{(m-s)/2} . \tag{11.75}$$

Now we examine the contribution of the terms (11.62) (or, equivalently, $\pm\beta^{2k}N^{1-k}\nu_0(T^-_{\mathcal{I}_1\cup\mathcal{I}_2})$) when

$$r := k + \operatorname{card}(\mathcal{I}\backslash\mathcal{I}_2) \ge 3 . \tag{11.76}$$

Let us consider

$$m' = \operatorname{card}(\mathcal{I}_1 \cup \mathcal{I}_2) \le k + m - 1 \tag{11.77}$$

and

$$s_1 = s + 2 - r \; ; \; s' = \max(0, s_1) .$$

Using (11.68) we see that the value of s_1 here is the same as that of (11.68), and (11.67) then shows that

$$2C(\mathcal{I}_1 \cup \mathcal{I}_2) \le \operatorname{card}(\mathcal{I}_1 \cup \mathcal{I}_2) - s_1 ,$$

and since we have anyway $2C(\mathcal{I}) \le \operatorname{card}\mathcal{I}$ we have in fact

$$2C(\mathcal{I}_1 \cup \mathcal{I}_2) \le \operatorname{card}(\mathcal{I}_1 \cup \mathcal{I}_2) - s' . \tag{11.78}$$

We now collect some simple inequalities. Since $s' < s$, and, if $s_1 > 0$,

$$m' + s' \le m + k - 1 + s + 2 - r \le m + s$$

since $r \ge k+1$, while, if $s_1 \le 0$, and since $k \le s$,

$$m' + s' = m' \le k + m - 1 \le m + s . \tag{11.79}$$

Since $r \ge 3$ and $r \ge k+1$ we have $6r \ge 2k + 14$, so that, if $s_1 > 0$,

$$m' + 7s' \le m + k - 1 + 7s - 7r + 14 \le m + 7s - (k + r + 1) .$$

If $s_1 \le 0$, we observe that, since $\operatorname{card}\mathcal{I}_1 \le k$ and $\operatorname{card}\mathcal{I} = m$ we have

$$m' + r \le \operatorname{card}\mathcal{I}_1 + \operatorname{card}\mathcal{I}_2 + k + \operatorname{card}(\mathcal{I}\backslash\mathcal{I}_2) \le 2k + m$$

and thus

$$m' + 7s' = m' \le 2k + m - r \le m + 7s - (k + r + 1) \tag{11.80}$$

since $s \ge k$ and $s \ge 1$.

We now use the induction hypothesis for s' and m'. Using (11.78), it follows from (11.70), (11.79) and (11.80) that

$$N^{1-k}|\nu_0(T^-_{\mathcal{I}_1\cup\mathcal{I}_2})| \le K_1(\beta_0)^{m+7s-(k+r+1)}(m+s)^{m+7s-(k+r+1)}N^{(m-s)/2} \ .$$

Finally we have to count how many such terms occur. Counting terms with their multiplicity, by (11.73), and since $n \le 2m$, there are at most (since $k \le r$)

$$2^k(2(m+k))^{2k} \le 8^r(m+k)^{2k}$$

choices for $\mathcal{I}_1$.

Since $\mathcal{I}_2$ is determined by $\mathcal{I}'\backslash\mathcal{I}_2$, and since $\mathrm{card}(\mathcal{I}'\backslash\mathcal{I}_2) = r-k-1$, there are at most m^{r-k-1} choices for $\mathcal{I}_2$.

Thus there are at most $8^r(m+k)^{r+k-1}$ choices for the pair $(\mathcal{I}_1, \mathcal{I}_2)$. Since we may assume $K_1(\beta_0) \ge 8$, the total contribution of the terms (11.65) for given values of k, r is at most

$$K_1(\beta_0)^{m+7s-1}(m+s)^{m+7s-2}N^{(m-s)/2}$$

and since $k \le s$, $r \le k+m \le s+m$, the total contribution over all values of k, $r \le s+m$ is at most

$$K_1(\beta_0)^{m+7s-1}(m+s)^{m+7s}N^{(m-s)/2} \ .$$

Recalling the estimates (11.74) and (11.75), we have found that for a certain quantity $K_0(\beta_0)$ we have

$$\nu(T_{\mathcal{I}}) \le \beta^2\nu_0(T^-_{\mathcal{I}}) + (K_0(\beta_0)^{m+7s-1} + 2K_1(\beta_0)^{m+7s-1})(m+s)^{m+7s}N^{(m-s)/2} \ ,$$

and if we choose $K_1(\beta_0) = 3K_0(\beta_0)/(1-\beta_0^2)$, we may then complete the proof as in Proposition 11.5.7. □

11.6 The Model with d-Component Spins

We return to the SK model with d-component spins, which was considered in Section 1.12. We recall that each individual spin σ_i is a vector $(\sigma_{i,1}, \dots, \sigma_{i,d})$ of $\mathbb{R}^d$, and that the Hamiltonian is given by

$$-H_N = \frac{\beta}{\sqrt{N}}\sum_{u\le d}\sum_{i<j} g_{ij}\sigma_{i,u}\sigma_{j,u} \ .$$

We will focus on the case where the base measure μ is the uniform measure on the sphere $\mathbb{S}_d$ of center 0 and radius $\sqrt{d}$ of $\mathbb{R}^d$. Therefore,

$$\sigma_i \in \mathbb{S}_d \Rightarrow \sum_{u\le d}\sigma_{i,u}^2 = d \ .$$

The case $d = 1$ is the case of the usual SK model with no external field. As usual we define

$$Z_N = \int \exp(-H_N)\mathrm{d}\mu(\sigma_1)\cdots\mathrm{d}\mu(\sigma_N) = \int \exp(-H_N)\mathrm{d}\mu^{\otimes N} ;$$
$$p_N = \frac{1}{N}\mathsf{E}\log Z_N .$$

The purpose of this section is to prove the following:

Theorem 11.6.1. *If $\beta \le 1/L$ we have*

$$\lim_{N\to\infty} p_N = \frac{\beta^2 d}{4} . \tag{11.81}$$

The difficult part in the proof of this theorem is to reach values of β that do not depend on d. The striking feature is that (as we will prove next) given $\beta > 0$, for d and N large enough, we have

$$\frac{1}{N}\log \mathsf{E}Z_N \ge \frac{\beta^2 d^2}{L} . \tag{11.82}$$

The reader observe the factor d^2 in the right-hand side of (11.82), while there is a factor d in the right-hand side of (11.82). In particular for d and N large enough we have

$$\log \mathsf{E}Z_N \ge \frac{d}{L}\mathsf{E}\log Z_N .$$

Of course, Theorem 11.6.1 is a very special result, begging for a sweeping generalization.

Research Problem 11.6.2. (Level 2) Is it possible to compute the supremum β_d of the values of β for which (11.81) holds? Or at least to compute $\lim_{d\to\infty}\beta_d$?

To explain (11.82) we compute

$$\mathsf{E}Z_N = \int \exp\Bigl(\frac{\beta^2}{2N}\sum_{u,v\le d}\sum_{i<j}\sigma_{i,u}\sigma_{j,u}\sigma_{i,v}\sigma_{j,v}\Bigr)\mathrm{d}\mu(\sigma_1)\cdots\mathrm{d}\mu(\sigma_N) ,$$

and, recalling the notation

$$R^{u,v} = \frac{1}{N}\sum_{i\le N}\sigma_{i,u}\sigma_{i,v} ,$$

we observe the identity

$$\sum_{u,v\le d}\sum_{i<j}\sigma_{i,u}\sigma_{j,u}\sigma_{i,v}\sigma_{j,v} = \frac{1}{2}\sum_{u,v\le d}(NR^{u,v})^2 - \frac{1}{2}\sum_{u,v\le d}\sum_{i\le N}\sigma_{i,u}^2\sigma_{i,v}^2 .$$

Since we integrate on a domain where $\sum_{u\le d}\sigma_{i,u}^2 = d$ for each i, it holds that

$$\sum_{u,v\le d}\sum_{i\le N}\sigma_{i,u}^2\sigma_{i,v}^2 = \sum_{i\le N}\Big(\sum_{u\le d}\sigma_{i,u}^2\Big)\Big(\sum_{v\le d}\sigma_{i,v}^2\Big) = d^2N\ ,$$

and we obtain

$$\mathsf{E}Z_N = \exp\left(-\frac{\beta^2 d^2}{4}\right)\int \exp\Big(\frac{\beta^2 N}{4}\sum_{u,v\le d}(R^{u,v})^2\Big)\mathrm{d}\mu^{\otimes N}\ . \tag{11.83}$$

It is simple to see that for $\mu^{\otimes N}$ we typically have $R^{u,v}\simeq 0$ when $u\ne v$ and $R^{u,u}\simeq 1$ for each $u\le d$. It then follows that the previous integrand is typically about $\exp(\beta^2Nd/4)$. However, as we shall see soon, some very small sets of configurations create a large contribution to the integral. It is a simple property of the sphere $\mathbb{S}_d$ that

$$\mu\bigg(\Big\{(\sigma_1,\ldots,\sigma_d)\in\mathbb{S}_d\ ;\ \sigma_1\ge\frac{\sqrt{d}}{2}\Big\}\bigg)\ge\exp(-Ld)\ .$$

Bounding from below $\sum_{u,v\le d}(R^{u,v})^2$ by $(R^{1,1})^2 = \Big(N^{-1}\sum_{i\le N}\sigma_{i,1}^2\Big)^2$ and integrating only on the domain where $\sigma_{i,1}\ge\sqrt{d}/2$ for all $i\le N$, we get that (since $R^{1,1}\ge d/4$ on this domain)

$$\mathsf{E}Z_N \ge \exp\Big(-\frac{\beta^2 d^2}{4}+\frac{\beta^2 N}{4}\Big(\frac{d}{4}\Big)^2 - NLd\Big)\ . \tag{11.84}$$

Hence

$$\frac{1}{N}\log\mathsf{E}Z_N \ge \frac{\beta^2 d^2}{2^6} - Ld - \frac{\beta^2 d^2}{4N}\ ,$$

which proves (11.82). Let us also note that

$$(R^{u,v})^2 = \Big(\frac{1}{N}\sum_{i\le N}\sigma_{i,u}\sigma_{i,v}\Big)^2 \le \frac{1}{N}\sum_{i\le N}\sigma_{i,u}^2\sigma_{i,v}^2\ ,$$

so that

$$\sum_{u,v}(R^{u,v})^2 \le \frac{1}{N}\sum_{i\le N}\Big(\sum_u \sigma_{i,u}^2\Big)^2 \le d^2$$

on the domain of integration, and thus from (11.83) we have

$$\mathsf{E}Z_N \le \exp\frac{N\beta^2 d^2}{4}\ . \tag{11.85}$$

It is a simple and well known fact about random Gaussian matrices (that is proved in Lemma A.8.1) that for a certain number L, the event Ω defined as

$$\Omega = \left\{ \forall (y_i)_{i \le N} \; ; \quad \left| \sum_{i<j} g_{ij} y_i y_j \right| \le L\sqrt{N} \sum_{i \le N} y_i^2 \right\} \tag{11.86}$$

satisfies

$$\mathsf{P}(\Omega) \ge 1 - L \exp\left(-\frac{N}{L}\right) . \tag{11.87}$$

The main ingredient of Theorem 11.6.1 is the following.

Theorem 11.6.3. *If $L\beta \le 1$ we have*

$$\mathsf{E}(\mathbf{1}_\Omega Z_N) \le K \exp \frac{N\beta^2 d}{4} . \tag{11.88}$$

Here, as well as in the rest of this section, K denotes a number depending on d only. The improvement over (11.85) is that we have a factor d (rather than d^2) in the exponent. It must be stressed that this theorem is a rather precise result. It identifies a set of very small probability that makes $\mathsf{E}Z_N$ large.

Corollary 11.6.4. *If $L\beta \le 1$ we have*

$$\frac{1}{N}\mathsf{E} \log Z_N \le \frac{\beta^2 d}{4} + \frac{K}{N} . \tag{11.89}$$

This is the "hard part" of Theorem 11.6.1.

Proof. Let

$$\varphi(t) = \mathsf{P}\left(\frac{1}{N}\log Z_N \ge \frac{\beta^2 d}{4} + t\right) = \mathsf{P}\left(Z_N \ge \exp\frac{N\beta^2 d}{4} \exp Nt\right) .$$

Using (11.88) and Markov's inequality, we get

$$\varphi(t) \le \mathsf{P}(\Omega^c) + K\exp(-Nt) \le L\exp\left(-\frac{N}{L}\right) + K\exp(-Nt) . \tag{11.90}$$

Using (11.85) and Markov's inequality, we also get

$$\varphi(t) \le \exp\left(\frac{N\beta^2}{4}(d^2 - d) - Nt\right) . \tag{11.91}$$

Using (11.90) for $t \le \beta^2 d^2$ and (11.91) for $t \ge \beta^2 d^2$, it is straightforward to obtain

$$\int_0^\infty \varphi(t)\mathrm{d}t \le \frac{K}{N} ,$$

and hence from (A.32) we get $\mathsf{E}(\max(0, N^{-1}\log Z_N - \beta^2 d/4)) \le K/N$ from which (11.89) follows. □

We define the set

$$C = \left\{ (\sigma_1, \ldots, \sigma_N) \in \mathbb{S}_d^N \; ; \; \sum_{u \neq v} (R^{u,v})^2 \leq 1 \right\} , \tag{11.92}$$

and

$$Z_{N,C} = \int_C \exp(-H_N) \mathrm{d}\mu(\sigma_1) \cdots \mathrm{d}\mu(\sigma_N) \; . \tag{11.93}$$

The number 1 in (11.92) plays no special role, and can be replaced by any other small number. We first reduce the proof of Theorem 11.6.3 to the following simple statement.

Proposition 11.6.5. *If* $L\beta \leq 1$ *we have*

$$\mathsf{E}(\mathbf{1}_\Omega Z_{N,C}) \leq K \exp \frac{N\beta^2 d}{4} \; . \tag{11.94}$$

Proof of Theorem 11.6.3. Consider an orthogonal basis $W = (e_1, \ldots, e_d)$ of $\mathbb{R}^d$, and define

$$R_W^{u,v} = \frac{1}{N} \sum_{i \leq N} (\sigma_i, e_u)(\sigma_i, e_v) \; ,$$

where (x, y) is the Euclidean scalar product of $\mathbb{R}^d$. If

$$C_W = \left\{ (\sigma_1, \ldots, \sigma_N) \in \mathbb{S}_d^N \; ; \; \sum_{u \neq v} (R_W^{u,v})^2 \leq 1 \right\}$$

then we claim that (11.94) holds when C is replaced by C_W. This should be obvious if we recall the formula (1.368) for the Hamiltonian, which shows that the choice of the basis of $\mathbb{R}^d$ is irrelevant.

Every symmetric quadratic form Q on $\mathbb{R}^d$ can be diagonalized in an orthonormal basis, that is, we may find such a basis $e_1, \ldots, e_n$ such that $\sum_{u \neq v} Q(e_u, e_v)^2 = 0$. The set of quadratic forms on $\mathbb{R}^d$ identifies naturally with $\mathbb{R}^{d^2}$ and has therefore a natural topology. Given the orthonormal basis $e_1, \ldots, e_n$, the set of quadratic forms that satisfy

$$\sum_{u \neq v} Q(e_u, e_v)^2 < 1$$

is an open set for this topology. As the orthonormal basis varies, these sets cover the compact set of quadratic forms that satisfy

$$\forall x \; , \; y \; , \; Q(x, y) \leq d \|x\| \|y\| \; , \tag{11.95}$$

so a finite sub-family still covers this set. Therefore we have found a finite family $\mathcal{W}$ of orthonormal bases of $\mathbb{R}^d$ such that, given a quadratic form Q on $\mathbb{R}^d$ as in (11.95), we may find $(e_1, \ldots, e_d) \in \mathcal{W}$ with

$$\sum_{u \neq v} Q(e_u, e_v)^2 \leq 1 \ . \tag{11.96}$$

(Of course the number 1 plays no special role here.) In words, $(e_1, \ldots, e_d)$ nearly diagonalizes Q.

Given any configuration $(\sigma_1, \ldots, \sigma_N)$ (with $\|\sigma_i\| = \sqrt{d}$), the quadratic form

$$Q(x, y) = \frac{1}{N} \sum_{i \leq N} (\sigma_i, x)(\sigma_i, y)$$

satisfies (11.95), so that we may find $(e_1, \ldots, e_d) \in \mathcal{W}$ that satisfies (11.96). Since $Q(e_u, e_v) = R_W^{u,v}$, this means that $(\sigma_1, \ldots, \sigma_N) \in C_W$. Thus

$$Z_N \leq \sum_{W \in \mathcal{W}} Z_{N, C_W} \ ,$$

and taking expectation on Ω proves (11.88). □

Given a subset U of $\{1, \ldots, d\}$, we define

$$C(U) = \{(\sigma_1, \ldots, \sigma_N) \in C \ ; \ \forall u \in U \ , \ R^{u,u} \geq 2 \ ; \ \forall u \notin U \ , \ R^{u,u} < 2\} \ .$$

The idea behind this definition is to identify the set of "dangerous values" of u, those for which $R^{u,u} \geq 2$. We have $C \subset \bigcup_U C(U)$, where the union is over all possible choices of U. We define $Z_{N,C(U)}$ as in (11.93) so that $Z_{N,C} \leq \sum_U Z_{N,C(U)}$ and Proposition 11.6.5 will follow from our next result.

Proposition 11.6.6. *If $L\beta \leq 1$, for any subset U of $\{1, \ldots, d\}$ we have*

$$\mathsf{E}(\mathbf{1}_\Omega Z_{N,C(U)}) \leq K \exp \frac{N\beta^2 d}{4} \ . \tag{11.97}$$

Proof. On the set Ω defined in (11.86), it holds

$$\frac{1}{\sqrt{N}} \sum_{i<j} g_{ij} \sigma_{i,u} \sigma_{j,u} \leq L \sum_{i \leq N} \sigma_{i,u}^2 = LNR^{u,u} \ .$$

We use this bound for all the "dangerous values" $u \in U$, so that

$$\begin{aligned} \mathsf{E}(\mathbf{1}_\Omega Z_{N,C(U)}) \leq \mathsf{E} \int_{C(U)} \exp \Bigg(& LN\beta \sum_{u \in U} R^{u,u} \\ & + \frac{\beta}{\sqrt{N}} \sum_{i<j} \sum_{u \notin U} g_{ij} \sigma_{i,u} \sigma_{j,u} \Bigg) \mathrm{d}\mu(\sigma_1) \cdots \mathrm{d}\mu(\sigma_N) \end{aligned}$$

and, reproducing the computation of (11.83),

$$\mathsf{E}(\mathbf{1}_\Omega Z_{N,C(U)}) \le \int_{C(U)} \exp\Bigl(LN\beta \sum_{u\in U} R^{u,u} + \frac{\beta^2 N}{4} \sum_{u,v\notin U} (R^{u,v})^2\Bigr) \mathrm{d}\mu^{\otimes N} . \tag{11.98}$$

Now we observe that when $\|\sigma_i\|^2 = d$ for all $i \le N$, we have

$$\sum_{u\le d} R^{u,u} = \frac{1}{N} \sum_{i\le N} \sum_{u\le d} \sigma_{i,u}^2 = d ,$$

so that

$$\sum_{u\le d} (R^{u,u})^2 = \sum_{u\le d} (R^{u,u} - 1)^2 + d .$$

Since for $u \in U$ we have $(R^{u,u})^2 \ge (R^{u,u}-1)^2$, this yields

$$\sum_{u\notin U} (R^{u,u})^2 \le \sum_{u\notin U} (R^{u,u} - 1)^2 + d ,$$

and thus, since $\sum_{u,v\notin U} (R^{u,v})^2 \le \sum_{u\notin U} (R^{u,u})^2 + \sum_{u\ne v} (R^{u,v})^2$, we obtain

$$\mathsf{E}(\mathbf{1}_\Omega Z_{N,C(U)}) \le \exp \frac{N\beta^2 d}{4} \int_{C(U)} \exp\Bigl(LN\beta \sum_{u\in U} R^{u,u} + \frac{\beta^2 N}{4} \sum_{u\notin U} (R^{u,u} - 1)^2 + \frac{\beta^2 N}{4} \sum_{u\ne v} (R^{u,v})^2\Bigr) \mathrm{d}\mu^{\otimes N} .$$

Using Hölder's inequality, (11.97) then follows from Lemmas 11.6.7 to 11.6.9 below. □

Lemma 11.6.7. *There exists a number $c > 0$ such that*

$$\int_{C(U)} \exp\Bigl(cN \sum_{u\in U} R^{u,u}\Bigr) \mathrm{d}\mu^{\otimes N} \le 1 . \tag{11.99}$$

Lemma 11.6.8. *There exists a number $c > 0$ such that*

$$\int_{C(U)} \exp\Bigl(cN \sum_{u\ne v} (R^{u,v})^2\Bigr) \mathrm{d}\mu^{\otimes N} \le K . \tag{11.100}$$

Lemma 11.6.9. *There exists a number $c > 0$ such that*

$$\int_{C(U)} \exp\Bigl(cN \sum_{u\notin U} (R^{u,u} - 1)^2\Bigr) \mathrm{d}\mu^{\otimes N} \le K . \tag{11.101}$$

At some point we must do some actual work to bring out special properties of the measure μ. The following describes what we need.

Proposition 11.6.10. a) *If the numbers* $(a_u)_{u\le d}$ *satisfy* $|a_u|\le 1/4$, *then*

$$\int \exp\Big(\sum_{u\le d} a_u(x_u^2-1)\Big)\mathrm{d}\mu(x) \le \exp\Big(L\sum_{u\le d} a_u^2\Big)\,. \qquad (11.102)$$

b) *If the numbers* $(a_{u,v})_{u,v\le d}$ *satisfy* $a_{u,v}=a_{v,u}$, $a_{u,u}=0$, $\sum_{u,v\le d} a_{u,v}^2 \le 1/16$ *and if* $x=(x_u)_{u\le d}$, *then*

$$\int \exp\Big(\sum_{u,v\le d} a_{u,v}x_u x_v\Big)\mathrm{d}\mu(x) \le \exp\Big(L\sum_{u,v\le d} a_{u,v}^2\Big)\,. \qquad (11.103)$$

Since $NR^{u,v}=\sum_{i\le N}\sigma_{i,u}\sigma_{i,v}$, the following is a consequence of Proposition 11.6.10.

Corollary 11.6.11. a) *Consider numbers* $(a_u)_{u\le d}$ *with* $|a_u|\le 1/4$. *Then*

$$\int \exp\Big(N\sum_{u\le d} a_u(R^{u,u}-1)\Big)\mathrm{d}\mu^{\otimes N} \le \exp\Big(LN\sum_{u\le d} a_u^2\Big)\,. \qquad (11.104)$$

b) *Consider numbers* $(a_{u,v})$ *with*

$$a_{u,u}=0\,;\quad a_{u,v}=a_{v,u}\,;\quad \sum_{u,v\le d} a_{u,v}^2 \le \frac{1}{16}\,. \qquad (11.105)$$

Then

$$\int \exp\Big(N\sum_{u,v\le d} a_{u,v}R^{u,v}\Big)\mathrm{d}\mu^{\otimes N} \le \exp\Big(LN\sum_{u,v\le d} a_{u,v}^2\Big)\,. \qquad (11.106)$$

Proof of Lemma 11.6.7. Since $\sum_{u\in U} R^{u,u}\ge 2\mathrm{card}U$ on $C(U)$, for $0\le a\le 1/4$ we get

$$\begin{aligned}
&\int_{C(U)} \exp\Big(aN\sum_{u\in U} R^{u,u}\Big)\mathrm{d}\mu^{\otimes N}\\
&\le \int_{C(U)} \exp\Big(aN\sum_{u\in U} R^{u,u}\Big)\exp\Big(2aN\sum_{u\in U} R^{u,u}-4aN\mathrm{card}U\Big)\mathrm{d}\mu^{\otimes N}\\
&\le \exp(-4aN\mathrm{card}U)\int \exp\Big(3aN\sum_{u\in U} R^{u,u}\Big)\mathrm{d}\mu^{\otimes N}\\
&\le \exp(-aN\mathrm{card}U)\int \exp\Big(3aN\sum_{u\in U} (R^{u,u}-1)\Big)\mathrm{d}\mu^{\otimes N}\\
&\le \exp(N\mathrm{card}U(-a+La^2))
\end{aligned}$$

using (11.104) for $a_u = a$ if $u \in U$ and $a_u = 0$ otherwise. The result follows by choosing $a = 1/L$. □

Proof of Lemma 11.6.8. First, we note that if A denotes a family $(a_{u,v})$ of sequences satisfying (11.105) we have

$$\int \exp\Big(\max_A\Big(N\sum_{u,v\le d} a_{u,v}R^{u,v}\Big)\Big)\mathrm{d}\mu^{\otimes N} \tag{11.107}$$
$$\le \sum_A \int \exp\Big(N\sum_{u,v\le d} a_{u,v}R^{u,v}\Big)\mathrm{d}\mu^{\otimes N} \le \mathrm{card} A\max_A \exp LN\sum_{u\ne v} a_{u,v}^2 \; .$$

It should be obvious that there exists a finite set B of sequences satisfying (11.105) such that, given numbers $(x_{u,v})_{u\ne v}$ with $x_{u,u} = 0$, $x_{u,v} = x_{v,u}$ we have

$$\max_B \sum_{u,v\le d} a_{u,v}x_{u,v} \ge \frac{1}{8}\Big(\sum_{u,v\le d} x_{u,v}^2\Big)^{1/2} .$$

This holds in particular for $x_{u,v} = R^{u,v}$. Given $0 \le t < 1$ we then use (11.107) for the family A of sequences $(ta_{u,v})$ for $(a_{u,v}) \in B$, and we obtain

$$\int_{C(U)} \exp\frac{tN}{8}\sqrt{\sum_{u\ne v} R_{u,v}^2}\,\mathrm{d}\mu^{\otimes N} \le \mathrm{card} B\exp L_0 N t^2 \; . \tag{11.108}$$

Since $\sum_{u\ne v} R_{u,v}^2 \le 1$ in $C(U)$, this also holds for all t. Using this for $t = g/2\sqrt{L_0 N}$ where g is standard Gaussian and taking expectation yields the result. □

Proof of Lemma 11.6.9. Let us define

$$C_1(U) = \Big\{\sum_{u\notin U}(R^{u,u}-1)^2 \le 1\Big\}$$

and

$$C_2(U) = \Big\{\sum_{u\notin U}(R^{u,u}-1)^2 > 1\Big\} .$$

The proof that there exists a number $c > 0$ such that

$$\int_{C_1(U)} \exp\Big(cN\sum_{u\notin U}(R^{u,u}-1)^2\Big)\mathrm{d}\mu^{\otimes N} \le K \; ,$$

is identical to the proof of Lemma 11.6.8, using that whenever $\sum a_u^2 \le 1/16$, we have $|a_u| \le 1/4$ for each u. So we turn to the proof of a similar result for $C_2(U)$.

Using (11.102) we see that if the numbers a_u satisfy $|a_u| \le 1/4$, then

$$\int_{C_2(U)} \exp\Big(\max_A N \sum_{u \notin U} (a_u(R^{u,u}-1) - La_u^2)\Big) d\mu^{\otimes N} \tag{11.109}$$
$$\le \sum_A \int_{C_2(U)} \exp\Big(N \sum_{u \notin U} (a_u(R^{u,u}-1) - La_u^2)\Big) d\mu^{\otimes N} \le \mathrm{card} A\,.$$

Consider a number k_0 to be determined later. We choose for A the family $(a_u)_{u \le d}$ where each a_u is of the type $\pm 2^{-k}$ for $2 \le k \le k_0$. Next, we prove the elementary fact that for any number x with $|x| \le 1$ we have

$$\max_{2 \le k \le k_0} \big(\pm 2^{-k} x - L2^{-2k}\big) \ge \frac{1}{L} x^2 - L2^{-2k_0}\,. \tag{11.110}$$

For clarity let us denote by L_1 the constant in the left-hand side of (11.110). Without loss of generality we assume that $x \ge 0$ and we consider the unique integer $k_1 \in \mathbb{Z}$ for which

$$L_1 2^{-2k_1+1} \le 2^{-k_1} x < L_1 2^{-2k_1+3}\,, \tag{11.111}$$

so that

$$L_1 2^{-k_1+1} \le x < L_1 2^{-k_1+2}\,. \tag{11.112}$$

Using (11.111) in the first inequality and (11.112) in the second we obtain

$$2^{-k_1} x - L_1 2^{-2k_1} \ge 2^{-k_1-1} x \ge \frac{1}{2^3 L_1} x^2\,,$$

and (11.110) is proved whenever $2 \le k_1 \le k_0$. If $k_1 \ge k_0$ then by (11.112) we have $x < L_1 2^{-k_0+2}$ and then

$$2^{-k_0} x - L_1 2^{-2k_0} \ge \frac{x^2}{4L_1} - L_1 2^{-2k_0}\,,$$

and (11.110) is again proved in this case. Finally, since without loss of generality we may assume that $L_1 \ge 1$, and since $x \le 1$, (11.112) shows that we cannot have $k_1 < 2$, and (11.110) is proved.

We recall that for $u \notin U$ we have $|R^{u,u} - 1| \le 1$. Using (11.110) for the family A in (11.109) we get

$$\int_{C_2(U)} \exp N\Big(\frac{1}{L} \sum_{u \notin U} (R^{u,u}-1)^2 - Ld2^{-2k_0}\Big) d\mu^{\otimes N} \le \mathrm{card} A\,.$$

Now on $C_2(U)$ we have $\sum_{u \notin U} (R^{u,u}-1)^2 \ge 1$, so the required inequality follows from (11.110) when $Ld2^{-2k_0} \le 1/2$ and in particular if $k \ge K$. □

Proof of Proposition 11.6.10. We will proceed by comparison with Gaussian r.v.s Consider independent standard Gaussian r.v.s $(h_u)_{u \le d}$ and a r.v. x that is uniform over $\mathbb{S}_d$ and independent of $h_1, \ldots, h_d$. Then the two $\mathbb{R}^d$-valued random vectors

$$(h_1, \dots, h_d) \quad \text{and} \quad x\left(\frac{1}{d}\sum_{u\le d} h_u^2\right)^{1/2} \tag{11.113}$$

have the same distribution. This is because the density of the distribution of $(h_1, \dots, h_d)$ depends only on the distance to the origin, so that it suffices to observe that the distribution of the lengths of the vectors in (11.113) coincide.

Next, we observe that to prove (11.102) we may assume that $\sum_{u\le d} a_u = 0$. This is because if we set $a'_u = a_u - d^{-1}\sum_{v\le d} a_v$, we have $\sum_{u\le d} a'_u = 0$, $\sum_{u\le d} a'^2_u \le \sum_{u\le d} a_u^2$, and, using that $\sum_{u\le d} x_u^2 = d$ on $\mathbb{S}_d$, we also have $\sum_{u\le d} a_u(x_u^2-1) = \sum_{u\le d} a'_u(x_u^2-1)$. Assuming that $\sum_{u\le d} a_u = 0$, and since $\int x_u^2 \mathrm{d}\mu(x) = 1$ for each u, we have $\int \sum_{u\le d} a_d x_u^2 \mathrm{d}\mu(x) = 0$, and Jensen's inequality shows that

$$\int \exp\left(\sum_{u\le d} a_u x_u^2\right)\mathrm{d}\mu(x) \ge 1\,.$$

Hölder's inequality implies that

$$\int \exp\left(r\sum_{u\le d} a_u x_u^2\right)\mathrm{d}\mu(x)$$

is an increasing function of r. Thus

$$\begin{aligned}
&\mathbf{1}_{\{\sum_{u\le d} h_u^2 \ge d\}} \int \exp\left(\sum_{u\le d} a_u(x_u^2-1)\right)\mathrm{d}\mu(x)\\
&\le \int \exp\left(\left(\frac{1}{d}\sum_{v\le d} h_v^2\right)\sum_{u\le d} a_u x_u^2 - \sum_{u\le d} a_u\right)\mathrm{d}\mu(x)\,.
\end{aligned} \tag{11.114}$$

As explained in (11.113), the vector

$$\left(x_u^2\left(\frac{1}{d}\sum_{v\le d} h_v^2\right)\right)_{u\le d}\,,$$

has the same distribution as $(h_u^2)_{u\le d}$ when x varies uniformly over $\mathbb{S}_d$. Thus, performing expectation in (11.114), we obtain

$$\begin{aligned}
&\mathsf{P}\left(\sum_{u\le d} h_u^2 \ge d\right)\int \exp\left(\sum_{u\le d} a_u(x_u^2-1)\right)\mathrm{d}\mu(x)\\
&\le \mathsf{E}\exp\left(\sum_{u\le d} a_u(h_u^2-1)\right) = \prod_{u\le d}\frac{1}{\sqrt{1-2a_u}}\exp(-a_u)\\
&\le \exp L\sum_{u\le d} a_u^2
\end{aligned} \tag{11.115}$$

because $(1-2t)^{-1/2}\exp(-t) \le \exp Lt^2$ for $t \le 1/4$. Now (using the one dimensional central limit theorem), $\mathsf{P}\big(\sum_{u\le d} h_u^2 \ge d\big)$ is bounded below independently of d, and (11.115) gives

$$\int \exp\Big(\sum_{u\le d} a_u(x_u^2-1)\Big)\mathrm{d}\mu(x) \le L\exp\Big(L\sum_{u\le d} a_u^2\Big)\,.$$

To finish the proof of (11.102) we appeal to a general fact (proved implicitly during the proof of Theorem A.6.1): if a r.v. X satisfies $\mathsf{E}X=0$, $\mathsf{E}\exp tX \le L\exp t^2A^2$ for $|t|\le 1$, then for $|t|\le 1$, $\mathsf{E}\exp tX \le \exp Lt^2A^2$ (and we recall that $\int x_u^2 \mathrm{d}\mu_d(x) = 1$).

To prove (11.103) we use the fact that a symmetric matrix $(a_{u,v})$ can be diagonalized in an orthogonal basis, and that the corresponding diagonal matrix $(b_{u,v})$ has the same trace and the same Hilbert-Schmidt norm, i.e.

$$\sum_{u,v\le d} b_{u,v}^2 = \sum_{u\le d} b_{u,u}^2 = \sum_{u,v\le d} a_{u,v}^2 \le \frac{1}{16}$$

$$\sum_{u\le d} b_{u,u} = \sum_{u\le d} a_{u,u} = 0\,.$$

Writing $b_u = b_{u,u}$, we then see that $|b_u| \le 1/4$ for each u, and that

$$\sum_{u,v\le d} b_{u,v}x_u x_v = \sum_{u\le d} b_u x_u^2 = \sum_{u\le d} b_u(x_u^2-1)\,.$$

Thus (11.103) follows from (11.102). □

To complete the proof of Theorem 11.6.1, we will show the following.

Proposition 11.6.12. *For β small enough we have*

$$\lim_{N\to\infty} p_N \ge \frac{\beta^2 d}{4}\,.$$

The idea is to use the symmetries of $\mathbb{S}_d$ to reduce to the case of a model rather similar to the SK model. For $\sigma_1,\ldots,\sigma_N \in \mathbb{S}_d$ and $1\le u\le d$ we write

$$e_u(\sigma_1,\ldots,\sigma_N) = 2^{-N}\sum \exp\Big(\frac{\beta}{\sqrt{N}}\sum_{i<j} g_{ij}\sigma_{i,u}\sigma_{j,u}\eta_i\eta_j\Big)\,,$$

where the summation is over $\eta_i = \pm 1$ for $i \le N$.

Lemma 11.6.13. *We have*

$$\mathsf{E}\log Z_N \ge d\int \mathsf{E}\log e_1(\sigma_1,\ldots,\sigma_N)\mathrm{d}\mu(\sigma_1)\cdots\mathrm{d}\mu(\sigma_N)\,. \tag{11.116}$$

Proof. Let

$$
\begin{aligned}
e(\sigma_1,\ldots,\sigma_N) &= \prod_{u\le d} e_u(\sigma_1,\ldots,\sigma_N) \\
&= 2^{-Nd}\sum \exp\left(\frac{\beta}{\sqrt{N}}\sum_{q\le i<j\le N} g_{ij}\sum_{u\le d}\sigma_{i,u}\sigma_{j,u}\eta_i^u\eta_j^u\right) ,
\end{aligned}
$$

where the sum is over $\eta_i^u = \pm 1$ for $u \le d$, $i \le N$. The symmetries of $\mathbb{S}_d$ imply

$$
Z_N = \int e(\sigma_1,\ldots,\sigma_N)\mathrm{d}\mu(\sigma_1)\cdots\mathrm{d}\mu(\sigma_N) ,
$$

and Jensen's inequality finally shows that

$$
\begin{aligned}
\log Z_N &\ge \int \log e(\sigma_1,\ldots,\sigma_N)\mathrm{d}\mu(\sigma_1)\cdots\mathrm{d}\mu(\sigma_N) \\
&= \sum_{u\le d}\int \log e_u(\sigma_1,\ldots,\sigma_N)\mathrm{d}\mu(\sigma_1)\cdots\mathrm{d}\mu(\sigma_N) .
\end{aligned}
$$

The result follows by taking expectations. □

To study the right-hand side of (11.116) we will use the following.

Lemma 11.6.14. *Consider numbers* $(r_i)_{i\le N}$. *If* $\sum_{i\le N} r_i^4 \le N/2$, *then*

$$
\frac{1}{N}\mathsf{E}\log 2^{-N}\sum\exp\left(\frac{1}{\sqrt{N}}\sum_{i<j} g_{ij}\eta_i\eta_j r_i r_j\right) \ge \frac{1}{4N^2}\left(\sum_{i\le N} r_i^2\right)^2 - \frac{L}{\sqrt{N}} . \tag{11.117}
$$

Proof. Consider

$$
V_N = 2^{-N}\sum\exp\left(\frac{1}{\sqrt{N}}\sum_{i<j} g_{ij}\eta_i\eta_j r_i r_j\right)
$$

so that

$$
\mathsf{E}V_N = \exp\frac{1}{2N}\sum_{i<j} r_i^2 r_j^2
$$

and, writing

$$
V_N^2 = 2^{-2N}\sum\exp\left(\frac{1}{\sqrt{N}}\sum_{i<j} g_{ij}(\eta_i^1\eta_j^2 + \eta_i^2\eta_j^2) r_i r_j\right) ,
$$

where the sum is over $\eta_i^1, \eta_i^2 = \pm 1$, we get

$$
\mathsf{E}V_N^2 = (\mathsf{E}V_N)^2 2^{-2N}\sum\exp\left(\frac{1}{N}\sum_{i<j}\eta_i^1\eta_j^1\eta_i^2\eta_j^2 r_i^2 r_j^2\right) ,
$$

where the sum is again over $\eta_i^1, \eta_i^2 = \pm 1$. Now

$$\frac{1}{N}\sum_{i<j}\eta_i^1\eta_j^1\eta_i^2\eta_j^2 r_i^2 r_j^2 \le \frac{1}{2N}\Big(\sum_{i\le N}\eta_i^1\eta_i^2 r_i^2\Big)^2 ,$$

and it follows from (A.19) that

$$2^{-2N}\sum \exp\frac{1}{2N}\Big(\sum_{i\le N}\eta_i^1\eta_i^2 r_i^2\Big)^2 \le 2$$

provided $\sum_{i\le N} r_i^4 \le N/2$. In that case we have proved that $\mathsf{E}V_N^2 \le 2(\mathsf{E}V_N)^2$. It then follows from the Paley-Zygmund inequality (A.61) that

$$\mathsf{P}\Big(V_N \ge \frac{1}{2}\mathsf{E}V_N\Big) \ge \frac{1}{8} . \tag{11.118}$$

Now, when $\sum_{i\le N} r_i^4 \le N/2$ the Cauchy-Schwarz inequality implies

$$\Big(\frac{1}{\sqrt{N}}\sum_{i\le N} r_i^2 r_j^2\Big)^{1/2} \le \frac{1}{\sqrt{2N}}\sum_{i\le N} r_i^2 \le \sqrt{\frac{1}{2}}\Big(\sum_{i\le N} r_i^4\Big)^{1/2} \le \frac{\sqrt{N}}{2} .$$

Therefore by Proposition 1.3.5 we have

$$\mathsf{P}\Big(\frac{1}{N}\log V_N \ge \frac{1}{N}\mathsf{E}\log V_N + t\Big) \le 2\exp(-Nt^2) , \tag{11.119}$$

On the other hand (11.118) yields

$$\frac{1}{8} \le \mathsf{P}\Big(\frac{1}{N}\log V_N \ge \frac{1}{N}\log\Big(\frac{1}{2}\mathsf{E}V_N\Big)\Big) .$$

Taking $t = 2/\sqrt{N}$, so that for this value of t the right-hand side of (11.119) is $< 1/8$, we obtain

$$t + \frac{1}{N}\mathsf{E}\log V_N \ge \frac{1}{N}\log\Big(\frac{1}{2}\mathsf{E}V_N\Big) \ge \frac{1}{2N^2}\sum_{i<j} r_i^2 r_j^2 - \frac{L}{N} ,$$

and thus, using again that $\sum_{i\le N} r_i^4 \le N/2$,

$$\frac{1}{N}\mathsf{E}\log V_N \ge \frac{1}{4N^2}\Big(\sum_{i\le N} r_i^2\Big)^2 - \frac{L}{\sqrt{N}} . \qquad \square$$

Proof of Proposition 11.6.12. Let

$$A = \left\{ (\sigma_1, \ldots, \sigma_N) \in \mathbb{S}_d^N \;;\; \frac{\beta^2}{N} \sum_{i \le N} \sigma_{i,1}^4 \le \frac{1}{2} \right\} .$$

Using (11.117) for $r_i = \sqrt{\beta}\sigma_{i,1}$ and combining with (11.116) we see that, denoting by K a number that does not depend on N,

$$\frac{1}{N} \mathsf{E} \log Z_N \ge \frac{d\beta^2}{4} \int_A \left(\frac{1}{N} \sum_{i \le N} \sigma_{i,1}^2 \right)^2 \mathrm{d}\mu(\sigma_1) \cdots \mathrm{d}\mu(\sigma_N) - \frac{K}{\sqrt{N}} .$$

Since $\int \sigma_{i,1}^2 \mathrm{d}\mu(\sigma_i) = 1$, and using symmetry,

$$\begin{aligned} \int_A \left(\frac{1}{N} \sum_{i \le N} \sigma_{i,1}^2 \right)^2 \mathrm{d}\mu(\sigma_1) \cdots \mathrm{d}\mu(\sigma_N) &\ge \left(\int_A \frac{1}{N} \sum_{i \le N} \sigma_{i,1}^2 \mathrm{d}\mu(\sigma_1) \cdots \mathrm{d}\mu(\sigma_N) \right)^2 \\ &= \left(1 - \int_{A^c} \sigma_{i,1}^2 \mathrm{d}\mu(\sigma_1) \cdots \mathrm{d}\mu(\sigma_N) \right)^2 . \end{aligned}$$

Now, by the Cauchy-Schwarz inequality,

$$\int_{A^c} \sigma_{1,1}^2 \mathrm{d}\mu(\sigma_1) \cdots \mathrm{d}\mu(\sigma_N) \le \mu^{\otimes N}(A^c)^{1/2} \left(\int \sigma_{1,1}^4 \mathrm{d}\mu(\sigma_1) \right)^2 .$$

It follows from (11.102) that $\int \sigma_{1,1}^4 \mathrm{d}\mu(\sigma_1) \le L$. Since (11.103) also implies that $\int \sigma_{i,1}^8 \mathrm{d}\mu(\sigma_i) \le L$, the function $f = N^{-1} \sum_{i \le N} \sigma_{i,1}^4$ satisfies

$$\int \left(f - \int f \mathrm{d}\mu^{\otimes N} \right)^2 \mathrm{d}\mu^{\otimes N} \le \frac{L}{N}$$

and thus $\mu^{\otimes N}(A^c) \le L/N$ if β is small enough that $\beta^2 \int f \mathrm{d}\mu^{\otimes N} \le 1/4$. The previous estimates then imply

$$\frac{1}{N} \mathsf{E} \log Z_N \ge \frac{\beta^2 d}{4} - \frac{K}{\sqrt{N}} . \qquad \square$$

It is possible to go quite beyond Theorem 11.6.1.

Theorem 11.6.15. *There exists a constant L with the following property. Consider a probability measure μ' on $\mathbb{S}_d$, and assume that μ' has a density f with respect to the uniform measure μ. Then, if $|f-1| \le 1/L$, the replica-symmetric solution corresponding to μ' holds for $\beta \le 1/L$.*

What we mean here is that if $\rho_{u,v}$ and $q_{u,v}$ are given by (1.372) and (1.371) respectively, then

$$\lim_{N \to \infty} \sum_{u,v \le d} \nu\big((R^{u,v} - \rho_{u,v})^2 \big) = 0$$

$$\lim_{N\to\infty} \sum_{u,v\le d} \nu\big((R_{1,2}^{u,v} - q_{u,v})^2\big) = 0$$

and that (1.373) holds.

We could not prove this result using the methods of the present section so we refer the interested reader to the original paper [100].

11.7 A Research Problem: The Transition at $\beta = 1$

We understand well the SK model without external field for $\beta < 1$. On the other hand (as will become apparent later) the structure of this model for $\beta > 1$ is very complicated. It is thus natural to try to study in detail the case $\beta = 1$. Another natural idea is to consider a temperature $\beta = \beta_N < 1$ depending on N, and to find the slowest rate at which β_N may approach 1 as matters get complicated.

Theorem 11.7.1. *Assume that*

$$\lim_{N\to\infty} N^{1/3}(1-\beta_N^2) = \infty\ . \tag{11.120}$$

Then, given integers $k(\ell,\ell')$ for $1 \le \ell < \ell' \le n$, we have

$$\lim_{N\to\infty} \nu\left(\prod_{\ell<\ell'} \big((N(1-\beta_N^2))^{1/2} R_{\ell,\ell'}\big)^{k(\ell,\ell')}\right) = \prod_{\ell<\ell'} a(k(\ell,\ell'))\ , \tag{11.121}$$

where $a(k) = \mathsf{E} g^k$, for g a standard Gaussian r.v.

Of course, in (11.121), we have $\nu(f) = \mathsf{E}\langle f\rangle$, where the Gibbs measure is computed for $\beta = \beta_N$.

In words, the content of Theorem 11.7.1 is that if we rescale the overlaps $R_{\ell,\ell'}$ by the factor $\sqrt{N(1-\beta_N^2)}$, as long as $1-\beta_N^2 \gg N^{-1/3}$, they behave like independent Gaussian r.v.s. This nice picture breaks down in the case $1-\beta_N^2 \approx N^{-1/3}$. which we study now.

Theorem 11.7.2. *Assume that*

$$\lim_{N\to\infty} N^{1/3}(1-\beta_N^2) = c > 0\ . \tag{11.122}$$

Then, for each $k \ge 1$ we have

$$\sup_N \nu\big((N(1-\beta_N^2)R_{1,2}^2)^k\big) < \infty \tag{11.123}$$

but (11.121) fails. It even fails if we replace the normalizing factor $N(1-\beta_N^2)$ by $\nu(R_{1,2}^2)^{-1}$.

The situation (11.122) is exactly the situation where, in performing our usual expansions with the cavity method, the "error terms" become exactly of the same size as the "main terms". This will be explained precisely in (11.158) below. In some sense the situation (11.122) is canonical. Yet it gives rise to a really mysterious object. In particular (as we shall demonstrate later), in this setting the cavity method yields strange-looking relations.

Research Problem 11.7.3. (Level ≥ 2) In the situation (11.122), compute the limits in the left-hand side of (11.121).

Research Problem 11.7.4. (Level ≥ 2) Understand the SK model for $\beta = 1$. In particular, what is the order of $\nu(R_{1,2}^2)$?

Conjecture 11.7.5. If $\beta = 1$, the limit

$$a = \lim_{N \to \infty} N^{2/3} \nu(R_{1,2}^2) \tag{11.124}$$

exists for some $0 < a < \infty$.

This conjecture is largely motivated by a conversation with G. Parisi, who mentioned to the author that "he had no doubts" that $\nu(R_{1,2}^2)$ was of order $N^{-2/3}$. The following shows that $\nu(R_{1,2}^2)$ is not likely to be of smaller order.

Proposition 11.7.6. *(S. Chatterjee) For all $\beta \leq 1$ we have*

$$\left| (1 - \beta^2) \nu(R_{1,2}^2) - \frac{1}{N} \right| \leq \frac{3}{N^2} + L\nu(|R_{1,2}|^3) \, . \tag{11.125}$$

In particular for $\beta = 1$ we have

$$\nu(|R_{1,2}|^3) \geq \frac{1}{LN} \, . \tag{11.126}$$

Intuitively it seems that $\nu(R_{1,2}^2)$ increases with β (at least as $\beta \leq 1$). If Conjecture 11.7.5 is correct, one should be able to prove the following.

Research Problem 11.7.7. (Level ≥ 2) If $\beta_N^2 = 1 - cN^{-1/3}$, prove that

$$\limsup_{N \to \infty} N^{2/3} \nu(R_{1,2}^2)$$

is bounded independently of $c > 0$.

The best bound we could obtain for $\nu(R_{1,2}^2)$ is the following.

Proposition 11.7.8. *If* $\beta \le 1$ *we have*

$$\sqrt{N}\nu(R_{1,2}^2) \le L\ . \tag{11.127}$$

It even seems to be unknown whether

$$N\nu(R_{1,2}^4) \le L\ . \tag{11.128}$$

The proofs of all these results are based on the cavity method. Throughout the present section the notation ν_t refers to the Hamiltonian (11.41).

Proof of Proposition 11.7.6. As usual, we write

$$\nu(R_{1,2}^2) = \nu(\varepsilon_1\varepsilon_2 R_{1,2})$$

and, for $\varphi(t) = \nu_t(\varepsilon_1\varepsilon_2 R_{1,2})$ we use that, by integration by parts,

$$\varphi(1) = \varphi(0) + \int_0^1 \varphi'(t)\mathrm{d}t = \varphi(0) + \varphi'(1) - \int_0^1 t\varphi''(t)\mathrm{d}t\ . \tag{11.129}$$

Now (1.151) implies

$$\begin{aligned}\varphi'(t) = \nu_t'(\varepsilon_1\varepsilon_2 R_{1,2}) = \beta^2\big(&\nu_t(R_{1,2}^2) - 4\nu_t(\varepsilon_1\varepsilon_3 R_{1,2}R_{1,3}) \\ &+ 3\nu_t(\varepsilon_1\varepsilon_2\varepsilon_3\varepsilon_4 R_{1,2}R_{3,4})\big)\ .\end{aligned} \tag{11.130}$$

A further differentiation brings out an extra factor $R_{\ell,\ell'}$ in each term. Uses of (1.151), (1.153) and of Hölder's inequality show that

$$|\varphi''(t)| \le L\nu(|R_{1,3}|^3)\ .$$

Since $\varphi(0) = 1/N$, we get from (11.130) that

$$\begin{aligned}\nu(R_{1,2}^2) = \frac{1}{N} &+ \beta^2\nu(R_{1,2}^2) - 4\beta^2\nu(\varepsilon_2\varepsilon_3 R_{1,2}R_{1,3}) \\ &+ 3\beta^2\nu(\varepsilon_1\varepsilon_2\varepsilon_3\varepsilon_4 R_{1,2}R_{3,4}) + \mathcal{R}\ ,\end{aligned}$$

where $|\mathcal{R}| \le L\nu(|R_{1,3}|^2)$. Now, using symmetry among sites,

$$\nu(\varepsilon_2\varepsilon_3 R_{1,2}R_{1,3}) = \nu(R_{2,3}R_{1,2}R_{1,3})$$

and $|\nu(R_{2,3}R_{1,2}R_{1,3})| \le \nu(|R_{1,2}|^3)$. Moreover,

$$\nu_1(\varepsilon_1\varepsilon_2\varepsilon_3\varepsilon_4 R_{1,2}R_{3,4}) = \nu_0(\varepsilon_1\varepsilon_2\varepsilon_3\varepsilon_4 R_{1,2}R_{3,4}) + \int_0^1 \nu_t'(\varepsilon_1\varepsilon_2\varepsilon_3\varepsilon_4 R_{1,2}R_{3,4})\mathrm{d}t\ .$$

The first term is $1/N^2$, and the second one is as before $\le L\nu(|R_{1,2}|^3)$. Finally we get

$$(1-\beta^2)\nu(R_{1,2}^2) - \frac{1}{N} = \frac{3\beta^2}{N^2} + \mathcal{R}$$

where $|\mathcal{R}| \le L\nu(|R_{1,2}|^3)$. □

We provided the previous self-contained argument for the enjoyment of the reader, but the proof of Proposition 1.8.5 already shows that, in the case of the SK model with external field, we have, with the notation of Section 1.9,

$$(1-\beta^2(1-2q+\widehat{q}))\nu\left(\left(\frac{(\boldsymbol{\sigma}^1-\boldsymbol{\sigma}^2)\cdot(\boldsymbol{\sigma}^3-\boldsymbol{\sigma}^4)}{N}\right)^2\right) = \frac{4}{N}(1-2q+\widehat{q}) + \mathcal{R} ,$$

where $|\mathcal{R}| \le K(\beta,h)(N^{-3/2} + \nu(|R_{1,2}-q|^3))$. Therefore, on the AT line $1-\beta^2(1-2q+\widehat{q}) = 0$, we always have $\nu(|R_{1,2}-q|^3) \ge 1/NK(\beta,h)$.

Each time we take a derivative of $\nu_t(f)$, we bring in each term a factor $R_{\ell,\ell'}$. In order to use successfully the cavity method in the present setting, it is very useful to know a priori that these factors are small. The purpose of the next result is precisely to show this.

Proposition 11.7.9. *If $\beta < 1$ and $x > 0$ we have*

$$\nu(\mathbf{1}_{\{|R_{1,2}|\ge x\}}) \le 3N\exp\left(-\frac{Nx^4}{48}\right) + \frac{L}{1-\beta^2}\exp\left(-\frac{(1-\beta^2)N^2x^8}{L\log\frac{2}{1-\beta^2}}\right) . \qquad (11.131)$$

The rather awkward form of the previous bound is most likely an indication that this result is not optimal. But, it is the best we could achieve (in the range of β that concerns us here; see (11.140) below for smaller β, and [103], Proposition 2.14.5 for $\beta = 1$).

We shall use (11.131) via the following consequence. If $|f| \le 1$ we have

$$\nu(|fR_{1,2}|) \le x\nu(|f|) + \mathcal{R} , \qquad (11.132)$$

where $\mathcal{R} = \nu(\mathbf{1}_{\{|R_{1,2}|\ge x\}})$. If x is not too small this term will be small. It starts to be small for x such that

$$\frac{(1-\beta^2)N^2x^8}{L\log\frac{2}{1-\beta^2}} \simeq \log\frac{L}{1-\beta^2} ,$$

that is for x about

$$\frac{1}{N^{1/4}}\frac{\left(\log\frac{2}{1-\beta^2}\right)^{1/4}}{(1-\beta^2)^{1/8}} . \qquad (11.133)$$

Since we are interested in the case where $1-\beta^2$ is much smaller than $1/\log N$, the first term of (11.131) is already very small for these values of x. Forgetting for the moment about the logarithmic terms, we can say that Proposition

11.7.9 implies that "each factor $R_{\ell,\ell'}$ counts as $N^{-1/4}(1-\beta^2)^{-1/8}$". The main difficulty we face in this section is that this quantity is not very small. To understand this difficulty, let us try to use (11.125), written as

$$(1-\beta^2)\nu(R_{1,2}^2) = \frac{1}{N} + \mathcal{R}\ , \tag{11.134}$$

where $\mathcal{R} \le 3/N^2 + \nu(|R_{1,2}|^3)$. The best we can do (ignoring logarithmic terms) is to use one of the factors $|R_{1,2}|$ to write

$$|\mathcal{R}| \le \frac{3}{N^2} + \frac{L}{N^{1/4}(1-\beta^2)^{1/8}}\nu(R_{1,2}^2)\ .$$

Unfortunately, with this bound (11.134) yields information only when

$$\frac{1}{N^{1/4}(1-\beta^2)^{1/8}} \ll 1-\beta^2\ , \tag{11.135}$$

a relation that fails when $1-\beta^2 \simeq N^{-1/3}$. It is precisely for this reason that we cannot use order 2 expansions to study $\nu_t(f)$ for $f = \varepsilon_1\varepsilon_2 R_{1,2}$. Rather, we will have to use order 3 expansions. Doing this, we can use (neglecting again logarithmic factors)

$$|\nu_t^{(3)}(f)| \le L\nu(R_{1,2}^4) \le L\left(\frac{1}{N^{1/4}(1-\beta^2)^{1/8}}\right)^2 \nu(R_{1,2}^2)\ ,$$

and we may think of $|\nu_t^{(3)}(f)|$ as a lower order term as soon as

$$\left(\frac{1}{N^{1/4}(1-\beta^2)^{1/8}}\right)^2 \ll 1-\beta^2\ , \tag{11.136}$$

a relation that is satisfied when $1-\beta^2 \simeq N^{-1/3}$. Of course we will not be able to control $\nu_0''(f)$ through its absolute value, and the whole approach succeeds because we will be able to control this quantity through its sign.

The following prepares for the proof of Proposition 11.7.9.

Lemma 11.7.10. *When $\beta \le 1$ we have*

$$\mathsf{E}\left(Z_N^2\left\langle \exp\frac{NR_{1,2}^4}{12}\right\rangle\right) \le 2N(\mathsf{E}Z_N)^2\ . \tag{11.137}$$

Proof. Proceeding as in (11.6) we get

$$\mathsf{E}\left(Z_N^2\left\langle \exp\frac{NR_{1,2}^4}{12}\right\rangle\right) \le (\mathsf{E}Z_N)^2 2^{-2N}\sum_{\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2}\exp\left(\frac{\beta^2 N}{2}R_{1,2}^2 + \frac{NR_{1,2}^4}{12}\right)\ .$$

Now, setting $X = N^{-1}\sum_{i\le N}\sigma_i$, we have

$$2^{-2N} \sum_{\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2} \exp\left(\frac{\beta^2 N}{2} R_{1,2}^2 + \frac{N R_{1,2}^4}{12}\right) = 2^{-N} \sum_{\boldsymbol{\sigma}} \exp\left(\frac{\beta^2 N X^2}{2} + \frac{N X^4}{12}\right) ,$$

because at given $\boldsymbol{\sigma}_1$, the variable $\boldsymbol{\sigma}_2 \mapsto R_{1,2}$ is distributed as X. Recalling the function $\mathcal{I}(t)$ of (A.22) we obtain

$$\mathcal{I}(t) \geq \frac{t^2}{2} + \frac{t^4}{12}$$

by (A.23). Therefore by (A.24) we have

$$2^{-N} \mathrm{card}\{\boldsymbol{\sigma} \in \Sigma_N \; ; \; |X| \geq t\} \leq \exp\left(-N\left(\frac{t^2}{2} + \frac{t^4}{12}\right)\right) .$$

We then use (A.31) for the uniform probability P_0 on Σ_N, and we denote by E_0 the corresponding expectation. Then since $|X| \leq 1$,

$$\begin{aligned} 2^{-N} \sum_{\boldsymbol{\sigma}} \exp\left(\frac{\beta^2 N X^2}{2} + \frac{N X^4}{12}\right) &= \mathsf{E}_0 \exp\left(\frac{\beta^2 N X^2}{2} + \frac{N X^4}{12}\right) \\ &\leq 1 + \int_0^1 \left(\beta^2 N t + \frac{N t^3}{3}\right) \mathrm{d}t \leq 2N . \end{aligned}$$

This finishes the proof. □

Proof of Proposition 11.7.9. Our starting point is the relation

$$\langle \mathbf{1}_{\{|R_{1,2}| \geq x\}} \rangle \leq \exp\left(-\frac{N x^4}{12}\right) \left\langle \exp \frac{N R_{1,2}^4}{12} \right\rangle ,$$

so that if Ω denotes the event

$$\left\langle \exp \frac{N R_{1,2}^4}{12} \right\rangle \geq \exp \frac{3 N x^4}{48} ,$$

we have

$$\langle \mathbf{1}_{\{|R_{1,2}| \geq x\}} \rangle \leq \exp\left(-\frac{N x^4}{48}\right) + \mathbf{1}_\Omega .$$

Taking expectation yields

$$\nu(\mathbf{1}_{\{|R_{1,2}| \geq x\}}) \leq \exp\left(-\frac{N x^4}{48}\right) + \mathsf{P}\left(\left\langle \exp \frac{N R_{1,2}^4}{12} \right\rangle \geq \exp \frac{3 N x^4}{48}\right) . \tag{11.138}$$

Now

$$\left\langle \exp \frac{N R_{1,2}^4}{12} \right\rangle = \frac{1}{Z_N^2} Z_N^2 \left\langle \exp \frac{N R_{1,2}^4}{12} \right\rangle$$

and therefore

$$\mathsf{P}\left(\left\langle \exp \frac{NR_{1,2}^4}{12} \right\rangle \geq \exp \frac{3Nx^4}{48}\right) \leq \mathsf{P}\left(Z_N^2 \left\langle \exp \frac{NR_{1,2}^4}{12} \right\rangle \geq (\mathsf{E}Z_N)^2 \exp \frac{Nx^4}{48}\right) + \mathsf{P}\left(Z_N \leq \mathsf{E}Z_N \exp\left(-\frac{Nx^4}{48}\right)\right) .$$

Combining (11.137) with Markov inequality we obtain

$$\mathsf{P}\left(Z_N^2 \left\langle \exp \frac{NR_{1,2}^4}{12} \right\rangle \geq (\mathsf{E}Z_N)^2 \exp \frac{Nx^4}{48}\right) \leq 2N \exp\left(-\frac{Nx^4}{48}\right) .$$

We appeal to (11.11) with $t = Nx^4/48$ to see that

$$\mathsf{P}\left(Z_N \leq \mathsf{E}Z_N \exp\left(-\frac{Nx^4}{48}\right)\right) \leq \frac{L}{1-\beta^2} \exp\left(-\frac{(1-\beta^2)N^2x^8}{L \log \frac{2}{1-\beta^2}}\right) .$$

Combining with (11.138) this proves (11.131). □

It is difficult to find clean and natural bounds for $\nu(\mathbf{1}_{\{|R_{1,2}| \geq x\}})$ because there are many possible variations to the previous proof. Let us indicate one such variation (that will not use). Instead of (11.137) one can use that (11.8) implies

$$\mathsf{E}\left(Z_N^2 \left\langle \exp \frac{1-\beta^2}{4} NR_{1,2}^2 \right\rangle\right) \leq \frac{L}{\sqrt{1-\beta^2}} (\mathsf{E}Z_N)^2 . \tag{11.139}$$

One can then mimic the previous argument, writing now

$$\langle \mathbf{1}_{\{|R_{1,2}| \geq x\}} \rangle \leq \exp\left(-\frac{1-\beta^2}{4} Nx^2\right) \left\langle \exp \frac{1-\beta^2}{4} NR_{1,2}^2 \right\rangle$$

and one obtains the bound

$$\nu(\mathbf{1}_{\{|R_{1,2}| \geq x\}}) \leq \frac{L}{\sqrt{1-\beta^2}} \left(\exp\left(-\frac{1-\beta^2}{16} Nx^2\right) + \exp\left(-\frac{(1-\beta^2)^6}{L \log \frac{2}{1-\beta^2}} N^2 x^4\right)\right) . \tag{11.140}$$

The right-hand side starts to be small for x about

$$\frac{1}{N^{1/2}} \frac{\sqrt{\log \frac{2}{1-\beta^2}}}{(1-\beta^2)^{3/2}} . \tag{11.141}$$

Certainly (11.141) is better than (11.133) for $\beta \ll 1$, but not in the range of interest here.

As a first step toward Theorem 11.7.1 we prove the following:

Proposition 11.7.11. *There exists a number L_0 with the following property. If $N \geq 2$, and $k \geq 1$, then*

$$\beta^2 \leq 1 - L_0 k \left(\frac{\log N}{N}\right)^{2/5} \quad \Rightarrow \quad \nu(R_{1,2}^{2k}) \leq \left(\frac{16k}{N(1-\beta^2)}\right)^k . \tag{11.142}$$

Since $2/5 > 1/3$, this will be usable under (11.120) and (11.122).

The power $2/5$ in the left-hand side of (11.142) simply occurs from (11.136). We do not know what happens for larger values of β, i. e. $0 \leq 1 - \beta^2 \ll N^{-2/5}$. If Conjecture 11.7.5 is true, one expects anyway that (11.142) is not sharp for $0 \leq 1 - \beta^2 \ll N^{-1/3}$, but rather that one has $\nu(R_{1,2}^{2k}) \leq K(k) N^{-2k/3}$.

As already explained, the essential new feature of the proof is that certain terms will be controlled through their signs because we do not know how to properly control their absolute values. We recall the notation $R^-_{\ell,\ell'} = N^{-1} \sum_{i<N} \sigma_i^\ell \sigma_i^{\ell'}$.

Lemma 11.7.12. *For each k we have*

$$\nu((R^-_{1,2})^k R^-_{2,3} R^-_{1,3}) \geq 0\ ;\ \nu((R_{1,2})^k R_{2,3} R_{1,3}) \geq 0\ . \tag{11.143}$$

Proof. First we observe that for any function f we have

$$\begin{aligned}\langle f(\boldsymbol{\sigma}^1) R^-_{1,3} f(\boldsymbol{\sigma}^2) R^-_{2,3}\rangle &= \int \left(\int f(\boldsymbol{\sigma}^1) R^-(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^3) \mathrm{d}G_N(\boldsymbol{\sigma}^1)\right)^2 \mathrm{d}G_N(\boldsymbol{\sigma}^3) \\ &\geq 0\ .\end{aligned}$$

Then, expanding $(R^-_{1,2})^k$ we see that it is a sum of terms $\sigma^1_{i_1} \cdots \sigma^1_{i_k} \sigma^2_{i_1} \cdots \sigma^2_{i_k}$, each of which is of the type $f(\boldsymbol{\sigma}^1) f(\boldsymbol{\sigma}^2)$. The proof of the second inequality is identical. □

Here is another simple fact.

Lemma 11.7.13. *We have*

$$\nu((R^-_{1,2})^k) \leq \nu(R^k_{1,2})\ . \tag{11.144}$$

Proof. Expanding the power,

$$\nu((R^-_{1,2})^k) = N^{-k} \sum \nu(\sigma^1_{i_1}\sigma^2_{i_1} \cdots \sigma^1_{i_k}\sigma^2_{i_k}) = N^{-1} \sum \mathsf{E}\langle \sigma_{i_1} \cdots \sigma_{i_k}\rangle^2\ ,$$

where the sum is over all choices of $i_1, \ldots, i_k \leq N-1$. And $\nu(R^k_{1,2})$ is given by a similar expression, with the sum over all choices $i_1, \ldots, i_k \leq N$. □

Proof of Proposition 11.7.11. We will prove by induction that

$$\nu(R_{1,2}^{2k}) \le \left(\frac{16k}{N(1-\beta^2)}\right)^k . \tag{11.145}$$

The induction starts with $k = 0$, for which (11.145) holds. For the induction from k to $k+1$, symmetry between sites entails

$$\nu(R_{1,2}^{2k+2}) = \nu(\varepsilon_1 \varepsilon_2 R_{1,2}^{2k+1}) . \tag{11.146}$$

Using the inequality

$$|x^{2k+1} - y^{2k+1}| \le (2k+1)|x-y|(x^{2k} + y^{2k})$$

for $x = R_{1,2}$ and $y = R_{1,2}^-$, we obtain that if we define $f = \varepsilon_1\varepsilon_2(R_{1,2}^-)^{2k+1}$, we have

$$\begin{aligned} |\nu(\varepsilon_1\varepsilon_2 R_{1,2}^{2k+1}) - \nu(f)| &\le \frac{2k+1}{N}\nu((R_{1,2}^{2k}) + (R_{1,2}^-)^{2k}) \\ &\le \frac{4(k+1)}{N}\left(\frac{16k}{N(1-\beta^2)}\right)^k , \end{aligned} \tag{11.147}$$

using (11.144) and (11.145). We write

$$\nu(f) \le \nu_0(f) + \nu_0'(f) + \frac{1}{2}\nu_0''(f) + \sup_{0\le t\le 1} |\nu_t^{(3)}(f)| . \tag{11.148}$$

We have $\nu_0(f) = 0$. Using (1.150), we get

$$\begin{aligned} \nu_t'(f) = \beta^2\big(&\nu_t((R_{1,2}^-)^{2k+2}) - 4\nu_t(\varepsilon_1\varepsilon_3 R_{2,3}^-(R_{1,2}^-)^{2k+1}) \\ &+ 3\nu_t(\varepsilon_1\varepsilon_2\varepsilon_3\varepsilon_4 R_{3,4}^-(R_{1,2}^-)^{2k+1})\big) , \end{aligned} \tag{11.149}$$

so that

$$\nu_0'(f) = \beta^2\nu_0((R_{1,2}^-)^{2k+2}) .$$

The crucial fact is that, using (1.150) to compute the derivative in (11.149) we have

$$\nu_0''(f) = -4\beta^4\nu_0(R_{1,3}^- R_{2,3}^- (R_{1,2}^-)^{2k+1}) \le 0 ,$$

by Lemma 11.7.12. Therefore we deduce from (11.148) that

$$\nu(f) \le \beta^2\nu_0((R_{1,2}^-)^{2k+2}) + \sup_{0\le t\le 1} |\nu_t^{(3)}(f)| . \tag{11.150}$$

Using (1.151) three times, we see that $|\nu_t^{(3)}(f)|$ is bounded by a sum of terms of the type

$$\nu(|R_{1,2}^-|^{2k+1}|R_{\ell_1,\ell_2}||R_{\ell_3,\ell_4}||R_{\ell_5,\ell_6}|)$$

and using (11.132) for the last two factors, we get

$$|\nu_t^{(3)}(f)| \le L\big(x^2\nu(|R_{1,2}^-|^{2k+1}|R_{\ell_1,\ell_2}|) + \nu(\mathbf{1}_{\{|R_{1,2}|\ge x\}})\big)$$
$$\le L\big(x^2\nu(R_{1,2}^{2k+2}) + \nu(\mathbf{1}_{\{|R_{1,2}|\ge x\}})\big)\ , \tag{11.151}$$

using Hölder's inequality and (11.144).

In (11.150) we would like to have $\nu((R_{1,2}^-)^{2k+2})$ rather than $\nu_0((R_{1,2}^-)^{2k+2})$; to relate these two quantities, we write

$$\left|\nu((R_{1,2}^-)^{2k+2}) - \nu_0((R_{1,2}^-)^{2k+2}) - \nu_0'((R_{1,2}^-)^{2k+2})\right| \le \sup_{0\le t\le 1}\left|\nu_t''((R_{1,2}^-)^{2k+2})\right|\ .$$

We compute $\nu_t'((R_{1,2}^-)^{2k+2})$ through (1.150) and we see that $\nu_0'((R_{1,2}^-)^{2k+2}) = 0$. Moreover $\nu_t''((R_{1,2}^-)^{2k+2})$ is bounded by a sum of terms of the type $\nu_t((R_{1,2}^-)^{2k+2}|R_{\ell_1,\ell_2}||R_{\ell_3,\ell_4}|)$; using (11.132), this is bounded as in (11.151). Finally we get the relation

$$\nu(f) \le \beta^2\nu(R_{1,2}^{2k+2}) + L_0x^2\nu(R_{1,2}^{2k+2}) + L\nu(\mathbf{1}_{\{|R_{1,2}|\ge x\}})\ ,$$

and combining with (11.146) and (11.147) yields

$$\nu(R_{1,2}^{2k+2}) \le (\beta^2+L_0x^2)\nu(R_{1,2}^{2k+2}) + \frac{4(k+1)}{N}\left(\frac{16k}{N(1-\beta^2)}\right)^k + L\nu(\mathbf{1}_{\{|R_{1,2}|\ge x\}})\ . \tag{11.152}$$

Obviously it is a good idea to choose x so that

$$L_0x^2 = \frac{1-\beta^2}{2} \tag{11.153}$$

and $1-\beta^2 - L_0x^2 = (1-\beta^2)/2$, and (11.152) yields

$$\frac{1}{2}(1-\beta^2)\nu(R_{1,2}^{2k+2}) \le \frac{4(k+1)}{N}\left(\frac{16k}{N(1-\beta^2)}\right)^k + L\nu(\mathbf{1}_{\{|R_{1,2}|\ge x\}})$$

and

$$\nu(R_{1,2}^{2k+2}) \le \frac{1}{2}\left(\frac{16(k+1)}{N(1-\beta^2)}\right)^{k+1} + \frac{L}{1-\beta^2}\nu(\mathbf{1}_{\{|R_{1,2}|\ge x\}})\ .$$

This finishes the proof of the induction provided the second term on the right is not larger than the first term. We will show that this is the case as soon as

$$\beta^2 \le 1 - L(k+1)\left(\frac{\log N}{N}\right)^{2/5}\ .$$

For clarity let us assume that

$$\beta^2 \le 1 - A(k+1)\left(\frac{\log N}{N}\right)^{2/5}\ , \tag{11.154}$$

where A is a parameter. It suffices to show that when A is a large enough constant (that does not depend on k), (11.154) implies

$$\nu(\mathbf{1}_{\{|R_{1,2}|\geq x\}}) \leq \frac{1}{2N^{k+1}} . \tag{11.155}$$

Now, keeping in mind the value (11.153) of x and using (11.131) we obtain

$$\nu(\mathbf{1}_{\{|R_{1,2}|\geq x\}}) \leq 3N \exp\Big(-\frac{N}{L}(1-\beta^2)^2\Big) + \frac{L}{1-\beta^2} \exp\Big(-\frac{N^2(1-\beta^2)^5}{L \log \frac{2}{1-\beta^2}}\Big) . \tag{11.156}$$

Moreover (11.154) implies that if A is large enough we have $1-\beta^2 \geq 2/N$, so that

$$(1-\beta^2)^5 \geq A^5(k+1)^5 \frac{\log^2 N}{N^2} \geq \log \frac{2}{1-\beta^2} A^5(k+1)^5 \frac{\log N}{N^2} ,$$

and (11.156) now implies that

$$\begin{aligned}\nu(\mathbf{1}_{\{|R_{1,2}|\geq x\}}) \leq LN \Bigg(& \exp\Big(-\frac{N^{1/5}A^2(k+1)^2 \log^{4/5} N}{L}\Big) \\ & + \exp\Big(-\frac{A^5(k+1)^5}{L} \log N\Big)\Bigg) ,\end{aligned}$$

from which (11.155) indeed follows if A is a large enough constant. (The dependence on k could be improved, but it is unimportant.) □

Now that we have proved (11.142), we have a new way to conduct calculations, using this bound. In fact, if we denote by $O(k)$ a quantity such that

$$\sup_N |O(k)|(N(1-\beta_N^2))^{k/2} < \infty ,$$

the proof of Theorem 11.7.1 is really business as usual. To explain why this is the case, let us compute $\nu(R_{1,2}^2)$. We write

$$\nu(R_{1,2}^2) = \nu(\varepsilon_1\varepsilon_2 R_{1,2}) = \frac{1}{N} + \nu(\varepsilon_1\varepsilon_2 R_{1,2}^-)$$

and, if $f = \varepsilon_1\varepsilon_2 R_{1,2}^-$,

$$\nu(f) = \nu_0(f) + \nu_0'(f) + O(3),$$

since by (11.145) we have $\nu(|R_{1,2}^-|^3) = O(3)$ and consequently $\nu_t''(f) = O(3)$. Now $\nu_0(f) = 0$,

$$\nu_0'(f) = \beta^2 \nu_0((R_{1,2}^-)^2)$$

and since the property $\nu(|R_{1,2}^-|^3) = O(3)$ implies $\nu'((R_{1,2}^-)^2) = O(3)$, we get

$$\nu_0((R_{1,2}^-)^2) = \nu((R_{1,2}^-)^2) + O(3) = \nu(R_{1,2}^2) + O(3)$$

using $1/N = O(3)$ in the last inequality. Finally,

$$N(1-\beta^2)\nu(R_{1,2}^2) = 1 + NO(3)\ . \tag{11.157}$$

The key point is that

$$|NO(3)| \le N\frac{K}{(N(1-\beta^2))^{3/2}} = \frac{K}{N^{1/2}(1-\beta^2)^{3/2}} \to 0 \tag{11.158}$$

when $N^{1/3}(1-\beta^2) \to 0$ as $N \to \infty$. We then leave to the reader the easy task of completing the proof of Theorem 11.7.1.

On the other hand, (11.158) fails to be true under (11.122). In that case, the cavity method yields mysterious relations, that we now sample.

Proposition 11.7.14. *Under (11.122) we have*

$$cN^{2/3}\nu(R_{1,2}^2) = 1 - 2N\nu(R_{1,2}R_{2,3}R_{3,1}) + O(N^{-1/3}) \tag{11.159}$$

$$\begin{aligned} cN\nu(R_{1,2}R_{2,3}R_{3,1}) &= N^{4/3}\nu(R_{1,3}^2R_{2,3}^2) - 3N^{4/3}\nu(R_{1,3}R_{1,4}R_{2,3}R_{2,4}) \\ &\quad + O(N^{-1/3})\ . \end{aligned} \tag{11.160}$$

There, as well as in the rest of the section, the notation $O(N^{-1/3})$ has its traditional meaning: a quantity A with $N^{1/3}|A|$ bounded independently of N.

To understand these relations, let us first observe that by (11.145) all five quantities involving ν remain bounded as $N \to \infty$. To make immediately the point that these relations are challenging we note that by Lemma 11.7.12 we have $\nu(R_{1,2}R_{1,3}R_{2,3}) \ge 0$. Since $\nu(R_{1,2}^2) \ge 0$, we see from (11.159) that

$$0 \le N\nu(R_{1,2}R_{1,3}R_{3,1}) \le \frac{1}{2} + O(N^{-1/3})\ .$$

Also, if Research problem 11.7.7 has a positive answer, for c small we must have

$$N\nu(R_{1,2}R_{2,3}R_{1,3}) \simeq \frac{1}{2}\ ,$$

a very unexpected fact.

Proof of Theorem 11.7.2. We have already proved (11.123). Assume for contradiction that there exists a normalizing sequence b_N such that, given numbers $k(\ell,\ell')$ for $1 \le \ell < \ell' \le 4$, and $k = \sum k(\ell,\ell')$, we have

$$\lim_{N\to\infty} \nu\Bigg(\prod_{1\le \ell<\ell'\le 4} (b_N R_{\ell,\ell'})^{k(\ell,\ell')}\Bigg) = \prod_{\ell<\ell'} a(k(\ell,\ell'))\ . \tag{11.161}$$

Since by Lemma 11.7.12 we have $\nu(R_{1,2}R_{2,3}R_{3,1}) \ge 0$, and by (11.159) we get

$$\nu(R_{1,2}^2) \le \frac{L}{cN^{2/3}}\ ,$$

and since by (11.161) we have,

$$\lim_{N\to\infty} b_N^2 \nu(R_{1,2}^2) = 1 \tag{11.162}$$

it follows that

$$b_N^2 \geq \frac{cN^{2/3}}{L} \ . \tag{11.163}$$

By (11.161), we get

$$\lim_{N\to\infty} b_N^3 \nu(R_{1,2}R_{2,3}R_{1,3}) = 0$$

and thus by (11.163),

$$\lim_{N\to\infty} N\nu(R_{1,2}R_{2,3}R_{1,3}) = 0 \ . \tag{11.164}$$

Therefore (11.159) implies

$$\lim_{N\to\infty} cN^{2/3}\nu(R_{1,2}^2) = 1 \ , \tag{11.165}$$

and comparing with (11.162) we can improve (11.163) into

$$\frac{b_N^2}{cN^{2/3}} \to 1. \tag{11.166}$$

On the other hand, (11.161) implies

$$\lim_{N\to\infty} N^{4/3}\nu(R_{1,3}R_{1,4}R_{2,3}R_{2,4}) = 0 \tag{11.167}$$

and combining with (11.160) and (11.164) yields

$$\lim_{N\to\infty} N^{4/3}\nu(R_{1,3}^2 R_{2,3}^2) = 0 \ , \tag{11.168}$$

so that (11.166) yields

$$0 = \lim_{N\to\infty} b_N^4 \nu(R_{1,2}^2 R_{2,3}^2) \neq a(2)^2 = 1$$

and this contradicts (11.161). Therefore there exists no normalizing sequence that satisfies (11.161). □

The relations (11.159) and (11.160) are part of a family of such relations. One can obtain similar relations with any quantity $cN^{k/3}\nu\left(\prod R_{\ell,\ell}^{k(\ell,\ell')}\right)$ on the left-hand side, with $k = (\sum k(\ell,\ell')$, the right-hand side being a combination of similar quantities "one order above", with $k+1$ instead of k, but we have been unable to get information out of these.

Proof of Proposition 11.7.14. We prove only (11.160); (11.159) is similar but simpler. First, we observe that (11.142) implies

$$\nu(|R_{1,2}|^k) = O(N^{-k/3}) \ . \tag{11.169}$$

We have

$$\nu(R_{1,2}R_{2,3}R_{1,3}) = \nu(f)$$

for $f = \varepsilon_1\varepsilon_2 R_{2,3}R_{1,3}$ and, using (11.142),

$$\nu(f) = \nu_0(f) + \nu_0'(f) + \frac{1}{2}\nu_0''(f) + O(N^{-5/3}) \ . \tag{11.170}$$

Now, using $R_{\ell,\ell'} = R^-_{\ell,\ell'} + \varepsilon_\ell\varepsilon_{\ell'}/N$, we get $\nu_0(f) = 1/N^2 = O(N^{-5/3})$. Also, (1.151) implies

$$\begin{aligned}\nu_t'(f) = \beta^2\big(&\nu_t(R_{1,2}R_{1,3}R_{2,3}) + 2\nu_t(\varepsilon_1\varepsilon_3 R_{1,3}R_{2,3}^2)\\ &- 6\nu_t(\varepsilon_1\varepsilon_4 R_{1,3}R_{2,3}R_{2,4}) - 3\nu_t(\varepsilon_1\varepsilon_2\varepsilon_3\varepsilon_4 R_{1,3}R_{2,3}R_{3,4})\\ &+ 6\nu_t(\varepsilon_1\varepsilon_2\varepsilon_4\varepsilon_5 R_{1,3}R_{2,3}R_{4,5})\big) \ .\end{aligned}$$

We claim that

$$\nu_0'(f) = \beta^2\nu_0(R_{1,2}R_{1,3}R_{2,3}) + O(N^{-5/3}) \ . \tag{11.171}$$

To see this we simply replace $R_{\ell,\ell'}$ by $R^-_{\ell,\ell'} + \varepsilon_\ell\varepsilon_{\ell'}/N$ to get relations such as

$$\nu_0(\varepsilon_1\varepsilon_4 R_{1,3}R_{2,3}R_{2,4}) = \frac{1}{N^3} \ ,$$

and

$$\nu_0(\varepsilon_1\varepsilon_3 R_{1,3}R_{2,3}^2) = \frac{1}{N}\nu_0((R^-_{2,3})^2) = O(N^{-5/3}) \ .$$

Using (1.151) again, and proceeding in the same manner we obtain

$$\nu_0''(f) = 2\beta^4\big(\nu_0(R_{1,2}^2R_{2,3}^2) - 3\nu_0(R_{1,3}R_{1,4}R_{2,3}R_{2,4})\big) + O(N^{-5/3}) \ . \tag{11.172}$$

Defining $f' = R_{1,2}R_{1,3}R_{2,3}$ we see by the same method that $\nu_0'(f') = O(N^{-5/3})$ so that

$$\nu_0(R_{1,2}R_{1,3}R_{2,3}) = \nu(R_{1,2}R_{1,3}R_{2,3}) + O(N^{-5/3}) \ ,$$

and similarly for the terms in (11.172). Since $\beta^4 = 1 + O(N^{-1/3})$ we have

$$\begin{aligned}&2\beta^4\big(\nu_0(R_{1,2}^2R_{2,3}^2) - 3\nu_0(R_{1,3}R_{1,4}R_{2,3}R_{2,4})\big)\\ &= 2\big(\nu(R_{1,2}^2R_{2,3}^2) - 3\nu(R_{1,3}R_{1,4}R_{2,3}R_{2,4})\big) + O(N^{-5/3}) \ .\end{aligned}$$

Combining these with (11.170) we have proved the relation

$$(1-\beta^2)\nu(R_{1,2}R_{1,3}R_{2,3}) = \nu(R_{1,2}^2R_{2,3}^2) - 3\nu(R_{1,2}R_{1,4}R_{2,3}R_{2,4}) + O(N^{-5/3})$$

which implies (11.160) because $1-\beta^2 = cN^{-1/3}$. □

We turn to the proof of Proposition 11.7.8. As will soon be apparent, we have to control the situation where $1-\beta^2 = N^{-1/2}$. In that case, even (11.136) fails, and this implies in some sense that we cannot work with order 3 expansions, but rather that we will have to use an order 4 expansion. Unfortunately we do not really know how to control third derivatives, and the proof will go through using some seemingly magical coincidences. It would of course be very desirable to find a more robust approach. This seems required for further progress, e.g. to prove (11.128).

Lemma 11.7.15. *For all values of β we have*

$$\frac{\mathrm{d}}{\mathrm{d}\beta}\nu(R_{1,2}^2) = \beta N\big(\nu(R_{1,2}^4) - 4\nu(R_{1,2}^2R_{1,3}^2) + 3\nu(R_{1,2}^2R_{3,4}^2)\big) \le \frac{1}{\beta}\ . \quad (11.173)$$

Proof. The equality is obtained by differentiation and integration by parts. It is an avatar of (1.89). The function $p_N(\beta) = N^{-1}\mathsf{E}\log Z_N(\beta)$ is a convex function of β, and (1.83) entails

$$p_N'(\beta) = \frac{\beta}{2}(1-\nu(R_{1,2}^2))\ ,$$

so that

$$0 \le p_N''(\beta) = \frac{1}{2}(1-\nu(R_{1,2}^2)) - \frac{\beta}{2}\frac{\mathrm{d}}{\mathrm{d}\beta}\nu(R_{1,2}^2)$$

and thus

$$\frac{\mathrm{d}}{\mathrm{d}\beta}\nu(R_{1,2}^2) \le \frac{1}{\beta}(1-\nu(R_{1,2}^2)) \le \frac{1}{\beta}\ .$$ □

The key to Proposition 11.7.8 is the following.

Lemma 11.7.16. *Consider the function $\psi(\beta) = \nu(R_{1,2}^2)$. Then for N large enough and all $\beta \le 1 - 1/\sqrt{N}$ we have*

$$\frac{1}{2}(1-\beta^2)\psi(\beta) \le \frac{2}{N} - \frac{1}{N}\Big(\frac{\beta^3}{2} - \frac{\beta^5}{6}\Big)\psi'(\beta)\ . \quad (11.174)$$

Proof of Proposition 11.7.8. Let us fix N and consider

$$\gamma = \sup\left\{\beta \le 1 - \frac{1}{\sqrt{N}}\ ;\ (1-\beta^2)\psi(\beta) \le \frac{5}{N}\right\}\ .$$

We first prove that $\gamma = 1 - 1/\sqrt{N}$. Suppose, otherwise, that $\gamma < 1 - 1/\sqrt{N}$. Then $(1-\gamma^2)\psi(\gamma) = 5/N$, and, obviously, $\psi'(\gamma) \ge 0$, for otherwise we could find $\gamma < \beta < 1 - 1/\sqrt{N}$ with $(1-\beta^2)\psi(\beta) < (1-\beta^2)\psi(\gamma) < (1-\gamma^2)\psi(\gamma)$. But then (11.174) implies

$$\frac{1}{2}(1-\gamma^2)\psi(\gamma) \le \frac{2}{N} ,$$

a contradiction.

Thus we have shown that $(1-\beta^2)\psi(\beta) \le 5/N$ for $\beta = 1 - 1/\sqrt{N}$, and therefore $\psi(1-1/\sqrt{N}) \le L/\sqrt{N}$. Now (11.173) shows that $\psi'(\beta) \le 2$ for $1 - 1/\sqrt{N} \le \beta \le 1$, so $\psi(1) \le \psi(1-1/\sqrt{N}) + 2/\sqrt{N} \le L/\sqrt{N}$. □

Proof of Lemma 11.7.16. We have $\nu(R_{1,2}^2) = \nu(f)$ for $f = \varepsilon_1\varepsilon_2 R_{1,2}$, so that we may apply

$$\nu(f) = \nu_0(f) + \int_0^1 \nu_t'(f)\mathrm{d}t$$

where

$$\nu_t'(f) = \beta^2(\nu_t(R_{1,2}^2) - 4\nu_t(\varepsilon_2\varepsilon_3 R_{1,2}R_{1,3}) + 3\nu_t(\varepsilon_1\varepsilon_2\varepsilon_3\varepsilon_4 R_{1,2}R_{3,4})) .$$

Now

$$\begin{aligned}
\nu_t(\varepsilon_2\varepsilon_3 R_{1,2}R_{1,3}) &= \nu_t(\varepsilon_1\varepsilon_2\varepsilon_1\varepsilon_3 R_{1,2}R_{1,3}) \\
&= \mathsf{E}\int\left(\int \varepsilon_1\varepsilon_2 R_{1,2}\mathrm{d}G_t(\boldsymbol{\sigma}^2)\right)^2 \mathrm{d}G_t(\boldsymbol{\sigma}^1) \\
&\ge \mathsf{E}\left(\int \varepsilon_1\varepsilon_2 R_{1,2}\mathrm{d}G_t(\boldsymbol{\sigma}^2)\mathrm{d}G_t(\boldsymbol{\sigma}^1)\right)^2 \\
&= \mathsf{E}\langle\varepsilon_1\varepsilon_2 R_{1,2}\rangle_t^2 = \nu_t(\varepsilon_1\varepsilon_2\varepsilon_3\varepsilon_4 R_{1,2}R_{3,4}) \ge 0
\end{aligned}$$

and therefore

$$\nu_t'(f) \le \beta^2\nu_t(R_{1,2}^2) .$$

Thus, setting $\varphi(t) = \nu_t(R_{1,2}^2)$, and since $\nu_0(f) = 1/N$, we obtain

$$\nu(R_{1,2}^2) = \varphi(1) = \varphi(0) + \int_0^1 \varphi'(t)\mathrm{d}t \le \frac{1}{N} + \beta^2\int_0^1 \varphi(t)\mathrm{d}t . \tag{11.175}$$

Now, by successive integration by parts,

$$\int_0^1 \varphi(t)\mathrm{d}t = \varphi(1) - \frac{1}{2}\varphi'(1) + \frac{1}{6}\varphi''(1) - \int_0^1 \frac{t^2}{6}\varphi^{(3)}(t)\mathrm{d}t . \tag{11.176}$$

Since $1-\beta^2 \ge 2/\sqrt{N}$, it follows from (11.131) and (11.132) that "each factor $R_{\ell,\ell'}$ counts at most as $(\log N)^{1/4}N^{-3/16}$" and since there are five such factors in each term of $\varphi^{(3)}(t)$, we may use three of these factors to obtain

$$\begin{aligned}
\left|\int_0^1 \frac{t^3}{16}\varphi^{(3)}(t)\mathrm{d}t\right| &\le L\frac{\log N}{N^{9/16}}\nu(R_{1,2}^2) + L\nu(\mathbf{1}_{\{|R_{1,2}|\ge L(\log N)^{1/4}N^{-3/16}\}}) \\
&\le L\frac{\log N}{N^{9/16}}\nu(R_{1,2}^2) + \frac{L}{N^2} ,
\end{aligned} \tag{11.177}$$

the crucial fact being that $9/16 > 1/2$. Now,

$$\varphi'(t) = \beta^2\big(\nu_t(\varepsilon_1\varepsilon_2 R_{1,2}^3) - 4\nu_t(\varepsilon_1\varepsilon_3 R_{1,2}^2 R_{1,3}) + 3\nu_t(\varepsilon_3\varepsilon_4 R_{1,2}^2 R_{3,4})\big) \ , \quad (11.178)$$

and, using symmetry between sites,

$$\varphi'(1) = \beta^2\big(\nu(R_{1,2}^4) - 4\nu(R_{1,2}^2 R_{1,3}^2) + 3\nu(R_{1,2}^2 R_{3,4}^2)\big) \ .$$

What makes the proof work is that this expression resembles (11.173). This seems now to be magical, but of course it probably will seem natural when the situation is better understood.

We compute $\varphi''(t)$ from (11.178) and (1.151); we see that

$$\varphi''(1) = \beta^4\big(\nu(R_{1,2}^4) - 4\nu(R_{1,2}^2 R_{1,3}^2) + 3\nu(R_{1,2}^2 R_{3,4}^2)\big) + \mathrm{I} + \mathrm{II} \ .$$

Here I is a sum of terms $\beta^4\nu(\varepsilon_{\ell_1}\varepsilon_{\ell_2}\overline{f})$, where $\overline{f}$ is the product of four terms $R_{\ell,\ell'}$ and $\ell_1 \neq \ell_2$ and II is a sum of terms $\beta^4\nu(\varepsilon_{\ell_1}\varepsilon_{\ell_2}\varepsilon_{\ell_3}\varepsilon_{\ell_4}\overline{f})$ for $\overline{f}$ of the same type and $\ell_1, \ldots, \ell_4$ are all different. Using symmetry among sites, the terms I satisfy the bound (11.177). To handle the other terms, letting $f' = \varepsilon_{\ell_1}\varepsilon_{\ell_2}\varepsilon_{\ell_3}\varepsilon_{\ell_4} f$, we observe simply that $\nu_t'(f')$ obeys the bound (11.177) and that $\nu_0(f') \le L/N^2$. Combining with (11.175) and (11.176), we have proved that

$$\begin{aligned}\nu(R_{1,2}^2) \le \frac{1}{N} &+ \left(\beta^2 + L\frac{\log N}{N^{9/16}}\right)\nu(R_{1,2}^2) \\ &- \left(\frac{\beta^4}{2} - \frac{\beta^6}{6}\right)\big(\nu(R_{1,2}^4) - 4\nu(R_{1,2}^2 R_{1,3}^2) + 3\nu(R_{1,2}^2 R_{3,4}^2)\big) + \frac{L}{N^2} \ ,\end{aligned}$$

and using (11.173) this implies

$$\Big(1 - \beta^2 - L\frac{\log N}{N^{9/16}}\Big)\varphi(\beta) \le \frac{1}{N} - \frac{1}{N}\Big(\frac{\beta^3}{2} - \frac{\beta^5}{6}\Big)\varphi'(\beta) + \frac{L}{N^2} \ .$$

Now for large N we have $L/N^2 \le 1/N$ and

$$1 - \beta^2 - L\frac{\log N}{N^{9/16}} \ge \frac{1-\beta^2}{2}$$

since $1 - \beta^2 \ge 1/\sqrt{N}$ and $9/16 > 1/2$. □

Part II

Low Temperature

12. The Ghirlanda-Guerra Identities

12.1 The Identities

It is a general principle of statistical mechanics that nearly all configurations have nearly the same energy. In the case of mean field models it will turn out that this energy is also non random, a fact that will have powerful consequences.

To clarify matters we shall give a general statement. Let us consider a Hamiltonian of the type

$$H^x(\boldsymbol{\sigma}) = H(\boldsymbol{\sigma}) + xH'(\boldsymbol{\sigma}) \;, \tag{12.1}$$

where $\boldsymbol{\sigma} \in \Sigma_N = \{-1,1\}^N$. Here of course the dependence on N is kept implicit, and H and H' do not depend on x. In the left-hand side of (12.1) the index x is a superscript in order to avoid conflict with the notation H_0 below.

We assume that $H = H_0 + H_1$, where H_0 is non random and where H_1 is a centered Gaussian Hamiltonian. We assume that H' is a Gaussian Hamiltonian independent of H_1. That is, the families $(H_1(\boldsymbol{\sigma}))_{\boldsymbol{\sigma}}$ and $(H'(\boldsymbol{\sigma}))_{\boldsymbol{\sigma}}$ are independent jointly Gaussian families of centered r.v.s. To control the size of these Hamiltonians we assume that for a certain number A independent of N we have

$$\forall \boldsymbol{\sigma} \in \Sigma_N \;, \quad \mathsf{E}H'(\boldsymbol{\sigma})^2 \le A^2 N \;;\; \mathsf{E}H_1(\boldsymbol{\sigma})^2 \le A^2 N. \tag{12.2}$$

Theorem 12.1.1. *Given $a > 0$ we have*

$$\int_{-a}^{a} \nu\left(\left|\frac{H'(\boldsymbol{\sigma})}{N} - \nu\left(\frac{H'(\boldsymbol{\sigma})}{N}\right)\right|\right) \mathrm{d}x \le \frac{K(a)A}{N^{1/4}} \tag{12.3}$$

where $K(a)$ depends on a only.

Here of course $\nu(f) = \mathsf{E}\langle f\rangle$ where the bracket is computed for the Hamiltonian (12.1), for which the dependence on x is kept implicit. The reason for writing H'/N in (12.3) is that (often) this quantity is of order 1 (when the configuration is drawn for the Gibbs' measure) and (12.3) implies that for the typical value of x, the fluctuations of H'/N, both with respect to the

M. Talagrand, *Mean Field Models for Spin Glasses*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys in Mathematics 55, DOI 10.1007/978-3-642-22253-5_5,

Gibbs' measure and the disorder, are of lower order. Note however that (12.3) remains of interest even when H'/N is of smaller order than 1 (but not too small).

Research Problem 12.1.2. (Level 2) Is it true that if A is independent of N, the left-hand side of (12.3) decays as $K/\sqrt{N}$ rather than $KN^{-1/4}$?

Consider

$$Z_N(x) = \sum_{\boldsymbol{\sigma}} \exp(-H^x(\boldsymbol{\sigma})) ,$$

and, keeping the dependence in N implicit,

$$\theta(x) = \frac{1}{N} \log Z_N(x) \; ; \qquad p(x) = \mathsf{E}\theta(x) \; .$$

Theorem 12.1.3. *(D. Panchenko [73]) Assume that $\mathcal{P}(y) = \lim_{N\to\infty} p(y)$ exists everywhere and is differentiable at $y = x$. Then*

$$\lim_{N\to\infty} \nu\left(\left|\frac{H'(\boldsymbol{\sigma})}{N} - \nu\left(\frac{H'(\boldsymbol{\sigma})}{N}\right)\right|\right) = 0 \; .$$

It is not the least remarkable feature of this theorem that by a seemingly simple argument it solves several problems that were previously rated at level 3.

We first sketch the proof of Theorem 12.1.1. We shall control separately the integrals

$$\int_{-a}^{a} \nu\left(\left|\frac{H'(\boldsymbol{\sigma})}{N} - \left\langle\frac{H'(\boldsymbol{\sigma})}{N}\right\rangle\right|\right) \mathrm{d}x \tag{12.4}$$

and

$$\int_{-a}^{a} \mathsf{E}\left|\left\langle\frac{H'(\boldsymbol{\sigma})}{N}\right\rangle - \nu\left(\frac{H'(\boldsymbol{\sigma})}{N}\right)\right| \mathrm{d}x \; . \tag{12.5}$$

As explained in Section 1.3, if f is any positive (measurable) function on a space provided with a positive measure μ, the map $\beta \mapsto \log \int f^\beta \, \mathrm{d}\mu$ is convex, as follows from Hölder's inequality. As a consequence, the function $x \mapsto \theta(x)$ is convex, a fact that will be proved again in equation (12.8) below. Moreover, by Theorem 1.3.4 the fluctuations of the r.v. $\theta(x)$ are small for every x. The idea underlying (12.5) is that (as a kind of quantitative version of Griffiths' lemma; Griffiths' lemma is explained right after the proof of Theorem 1.3.9) the fluctuations of $\theta'(x)$ are also small for the typical value of x, so that $\mathsf{E}|\theta'(x) - p'(x)|$ is small for such a value. Now

$$\theta'(x) = \left\langle -\frac{H'(\boldsymbol{\sigma})}{N}\right\rangle \tag{12.6}$$

so that

$$\int_{-a}^{a} \mathsf{E}\left|\left\langle \frac{H'(\boldsymbol{\sigma})}{N} \right\rangle - \nu\left(\frac{H'(\boldsymbol{\sigma})}{N}\right)\right| \mathrm{d}x = \int_{-a}^{a} \mathsf{E}|\theta'(x) - \mathsf{E}\theta'(x)|\mathrm{d}x$$
$$= \int_{-a}^{a} \mathsf{E}|\theta'(x) - p'(x)|\mathrm{d}x \;, \quad (12.7)$$

which will be shown to be small by the previous argument.

To control the integral (12.4), we differentiate (12.6) to get

$$\theta''(x) = \frac{1}{N}(\langle H'^2(\boldsymbol{\sigma})\rangle - \langle H'(\boldsymbol{\sigma})\rangle^2)$$
$$= \frac{1}{N}\langle (H'(\boldsymbol{\sigma}) - \langle H'(\boldsymbol{\sigma})\rangle)^2\rangle \quad (12.8)$$

and taking expectation

$$p''(x) = N\nu\left(\left(\frac{H'(\boldsymbol{\sigma})}{N} - \frac{\langle H'(\boldsymbol{\sigma})\rangle}{N}\right)^2\right) \;,$$

so by integration

$$\int_{-a}^{a} \nu\left(\left(\frac{H'(\boldsymbol{\sigma})}{N} - \frac{\langle H'(\boldsymbol{\sigma})\rangle}{N}\right)^2\right) \mathrm{d}x = \frac{1}{N}(p'(a) - p'(-a)) \;. \quad (12.9)$$

We then use the Cauchy-Schwarz inequality to obtain, using (12.9) in the last inequality

$$\int_{-a}^{a} \nu\left(\left|\frac{H'(\boldsymbol{\sigma})}{N} - \frac{\langle H'(\boldsymbol{\sigma})\rangle}{N}\right|\right) \mathrm{d}x$$
$$\leq \int_{-a}^{a} \left(\nu\left(\left(\frac{H'(\boldsymbol{\sigma})}{N} - \frac{\langle H'(\boldsymbol{\sigma})\rangle}{N}\right)^2\right)\right)^{1/2} \mathrm{d}x$$
$$\leq \sqrt{2a}\left(\int_{-a}^{a} \nu\left(\left(\frac{H'(\boldsymbol{\sigma})}{N} - \frac{\langle H'(\boldsymbol{\sigma})\rangle}{N}\right)^2\right) \mathrm{d}x\right)^{1/2}$$
$$\leq \sqrt{\frac{2a}{N}(p'(a) - p'(-a))} \;. \quad (12.10)$$

We shall show that $p'(a)$ is bounded independently of N, so that the left-hand side is small, and this concludes the scheme of proof of Theorem 12.1.1.

We now turn to the details of the proof of Theorem 12.1.1.

Lemma 12.1.4. *We have*

$$|p'(x)| \leq 2|x|A^2 \quad (12.11)$$
$$|p'(x)| \leq LA \;. \quad (12.12)$$

Here as usual L denotes a universal constant.

Proof. For two replicas ℓ, ℓ' we set

$$U_{\ell,\ell'} = \frac{1}{N}\mathsf{E}H'(\boldsymbol{\sigma}^\ell)H'(\boldsymbol{\sigma}^{\ell'}) \ .$$

From (12.6) we obtain, using integration by parts as in Lemma 1.3.11 in the second equality,

$$p'(x) = \mathsf{E}\theta'(x) = \mathsf{E}\left\langle -\frac{H'(\boldsymbol{\sigma})}{N}\right\rangle = x\left(\mathsf{E}\langle U_{1,1}\rangle - \mathsf{E}\langle U_{1,2}\rangle\right) \ ,$$

and (12.11) follows since 12.2 implies that $|U_{\ell,\ell'}| \le A^2$. Also we have

$$p'(x) = \mathsf{E}\left\langle -\frac{H'(\boldsymbol{\sigma})}{N}\right\rangle \le \frac{1}{N}\mathsf{E}\max_{\boldsymbol{\sigma}}(-H'(\boldsymbol{\sigma})) \le LA \ ,$$

where we have used (A.7) with $M = 2^N$ in the last inequality. □

Therefore (12.10) and (12.11) imply

$$\int_{-a}^{a} \nu\left(\left|\frac{H'(\boldsymbol{\sigma})}{N} - \frac{\langle H'(\boldsymbol{\sigma})\rangle}{N}\right|\right)\mathrm{d}x \le K(a)\frac{A}{\sqrt{N}} \ . \tag{12.13}$$

Here and below the number $K(a)$ depends on a only and need not be the same at each occurrence. We now turn to the details of controlling the integral (12.7).

Lemma 12.1.5. *Consider* $b > 0$ *and*

$$W = W(x,b) := \frac{1}{b}(|\theta(x+b) - p(x+b)| + |\theta(x-b) - p(x-b)| + |\theta(x) - p(x)|) \ .$$

Then

$$|\theta'(x) - p'(x)| \le p'(x+b) - p'(x-b) + W \ . \tag{12.14}$$

Proof. Since θ and p are convex functions, we have

$$\theta'(x) \le \frac{\theta(x+b) - \theta(x)}{b} \le W + \frac{p(x+b) - p(x)}{b} \le W + p'(x+b)$$

so

$$\theta'(x) - p'(x) \le W + p'(x+b) - p'(x) \le W + p'(x+b) - p'(x-b) \ .$$

Similarly,

$$\theta'(x) \ge \frac{\theta(x) - \theta(x-b)}{b} \ge \frac{p(x) - p(x-b)}{b} - W \ge -W + p'(x-b)$$

and thus

$$\theta'(x) - p'(x) \ge -W + p'(x-b) - p'(x) \ge -(W + p'(x+b) - p'(x-b)) \ . \ □$$

The next lemma offers a control of the size of W.

Lemma 12.1.6. *We have*

$$\mathsf{E}|W(x,b)| \leq \frac{LA}{b\sqrt{N}}(1+|x|+b)\ . \tag{12.15}$$

Proof. Since $(H'(\boldsymbol{\sigma}))_{\boldsymbol{\sigma}}$ and $(H_1(\boldsymbol{\sigma}))_{\boldsymbol{\sigma}}$ are independent jointly Gaussian families of r.v.s there exists a representation

$$H_1(\boldsymbol{\sigma}) + xH'(\boldsymbol{\sigma}) = \mathbf{g}\cdot\mathbf{a}(\boldsymbol{\sigma})\ ,$$

where, for a certain integer M (e.g. $M = 2^{N+1}$), $\mathbf{g} \in \mathbb{R}^M$ is a standard Gaussian vector, and $\mathbf{a}(\boldsymbol{\sigma}) \in \mathbb{R}^M$, $\|\mathbf{a}(\boldsymbol{\sigma}))\| = (\mathsf{E}(H_1(\boldsymbol{\sigma}) + xH'(\boldsymbol{\sigma}))^2)^{1/2} \leq A(1+|x|)$. The argument is sketched in Section A.2, but is not crucial since this representation is obvious in all the models we consider. Proposition 1.3.5 then implies

$$\mathsf{E}|\theta(x) - p(x)| \leq \frac{LA(1+|x|)}{\sqrt{N}}\ ,$$

from which the result follows. □

Proof of Theorem 12.1.1. Combining (12.14), (12.15) and taking expectation we get that, given $b > 0$,

$$\mathsf{E}|\theta'(x) - p'(x)| \leq p'(x+b) - p'(x-b) + \frac{LA}{b\sqrt{N}}(1+|x|+b)\ . \tag{12.16}$$

Integrating we get

$$\int_{-a}^{a} \mathsf{E}|\theta'(x) - p'(x)|\mathrm{d}x \leq \int_{-a}^{a}(p'(x+b) - p'(x-b))\mathrm{d}x + \frac{LAa}{b\sqrt{N}}(1+a+b)\ .$$

Now, (12.12) implies

$$\begin{aligned}
&\int_{-a}^{a}(p'(x+b) - p'(x-b))\mathrm{d}x \\
&= p(a+b) - p(-a+b) - p(a-b) + p(-a+b) \\
&= \int_{a-b}^{a+b} p'(x)\mathrm{d}x - \int_{-a-b}^{-a+b} p'(x)\mathrm{d}x \leq LbA\ .
\end{aligned}$$

Taking $b = N^{-1/4}$ and combining with (12.7) completes the proof. □

We now turn to the proof of Theorem 12.1.3

Lemma 12.1.7. *Consider the function*

$$\psi(x) := \frac{1}{N}\langle|H'(\boldsymbol{\sigma}^1) - H'(\boldsymbol{\sigma}^2)|\rangle\ .$$

Then the following hold:

$$\psi(x)^2 \leq \frac{4}{N}\theta''(x) \tag{12.17}$$

$$|\psi'(x)| \leq 8\theta''(x)\ . \tag{12.18}$$

Proof. To prove (12.17) we simply observe that

$$\psi(x) \le \frac{2}{N}\langle |H'(\boldsymbol{\sigma}) - \langle H'(\boldsymbol{\sigma})\rangle|\rangle \ ,$$

and we use (12.8) and the Cauchy-Schwarz inequality. To prove (12.18) we observe that, setting $V_\ell = H'(\boldsymbol{\sigma}^\ell) - \langle H'(\boldsymbol{\sigma})\rangle$, and using the Cauchy-Schwarz inequality in the second line and (12.8) in the last line,

$$\begin{aligned}\psi'(x) &= \frac{1}{N}\langle |H'(\boldsymbol{\sigma}^1) - H'(\boldsymbol{\sigma}^2)|(H'(\boldsymbol{\sigma}^2) + H'(\boldsymbol{\sigma}^2) - 2H'(\boldsymbol{\sigma}^3))\rangle \\ &\le \frac{1}{N}\langle |V_1 - V_2||V_1 + V_2 - 2V_3|\rangle \\ &\le \frac{8}{N}\langle V_1^2\rangle = 8\theta''(x) \ .\end{aligned}$$

This concludes the proof. □

Lemma 12.1.8. *Given $b > 0$ let us define $D(x,b) = p'(x+b) - p'(x-b)$. Then*

$$\mathsf{E}|\psi(x)| \le \sqrt{\frac{2}{bN}}D(x,b)^{1/2} + 8D(x,b) \ . \tag{12.19}$$

Proof. We start by writing the identity

$$2b\psi(x) = \int_{x-b}^{x+b}\psi(y)\mathrm{d}y + \int_{x-b}^{x+b}(\psi(x) - \psi(y))\mathrm{d}y \ . \tag{12.20}$$

We observe that for $x - b \le y \le x + b$ we have

$$|\psi(y) - \psi(x)| \le \int_{x-b}^{x+b}|\psi'(t)|\mathrm{d}t \ ,$$

so that

$$\left|\int_{x-b}^{x+b}(\psi(x) - \psi(y))\mathrm{d}y\right| \le 2b\int_{x-b}^{x+b}|\psi'(t)|\mathrm{d}t \ .$$

Combining with (12.20) and Lemma 12.1.7 yields

$$|\psi(x)| \le \frac{1}{b\sqrt{N}}\int_{x-b}^{x+b}\sqrt{\theta''(y)}\mathrm{d}y + 8\int_{x-b}^{x+b}\theta''(y)\mathrm{d}y \ .$$

To conclude, we use the Cauchy-Schwarz inequality and the relation

$$\int_{x-b}^{x+b}\theta''(y)\mathrm{d}y = \theta'(x+b) - \theta'(x-b) \ ,$$

we take expectation and we use the Cauchy-Schwarz inequality again, together with the relation $\mathsf{E}(\theta'(x+b) - \theta'(x-b)) = D(x,b)$. □

Lemma 12.1.9. *Under the conditions of Theorem 12.1.3, that is when $\mathcal{P}(y) = \lim_{N\to\infty} p(y)$ exists everywhere and is differentiable at $y = x$, we have*

$$\lim_{b\to 0} \limsup_{N\to\infty} D(x,b) = 0 \,. \tag{12.21}$$

Proof. We use the convexity of the function p to get

$$p'(x+b) \le \frac{1}{b}(p(x+2b) - p(x+b)) \;;\; p'(x-b) \ge \frac{1}{b}(p(x-b) - p(x-2b)) \,,$$

so that

$$D(x,b) \le \frac{1}{b}(p(x+2b) - p(x+b)) - p(x-b) + p(x-2b)) \,,$$

and consequently

$$\limsup_{N\to\infty} D(x,b) \le \frac{1}{b}(\mathcal{P}(x+2b) - \mathcal{P}(x+b) - \mathcal{P}(x-b) + \mathcal{P}(x-2b)) \,.$$

Finally, the differentiability of $\mathcal{P}$ at x implies that the limit of the right-hand side as $b \to 0$ is 0. □

Proof of Theorem 12.1.3. Recalling that by Jensen's inequality we have

$$\frac{1}{N}\mathsf{E}\langle|H'(\boldsymbol{\sigma}) - \langle H'(\boldsymbol{\sigma})\rangle|\rangle = \frac{1}{N}\mathsf{E}\langle|H'(\boldsymbol{\sigma}^1) - \langle H'(\boldsymbol{\sigma}^2)\rangle|\rangle \le \psi(x)$$

we deduce from (12.19) and Lemma 12.1.9 (letting first $N \to \infty$ and then $b \to 0$) that

$$\lim_{N\to\infty} \frac{1}{N}\mathsf{E}\langle|H'(\boldsymbol{\sigma}) - \langle H'(\boldsymbol{\sigma})\rangle|\rangle = 0 \,.$$

Now, (12.16) means

$$\frac{1}{N}\mathsf{E}|\langle H'(\boldsymbol{\sigma})\rangle - \mathsf{E}\langle H'(\boldsymbol{\sigma})\rangle| = \mathsf{E}|\theta'(x) - p'(x)| \le D(x,b) + \frac{LA}{b\sqrt{N}}(1 + |x| + b) \,,$$

and same limiting procedure yields

$$\lim_{N\to\infty} \frac{1}{N}\mathsf{E}|\langle H'(\boldsymbol{\sigma})\rangle - \mathsf{E}\langle H'(\boldsymbol{\sigma})\rangle| = 0 \,.$$

The result follows. □

We now turn to the consequences of Theorem 12.1.1.

Theorem 12.1.10. (The Ghirlanda-Guerra identities) *Consider a function f on Σ_N^n with $|f| \le 1$. Assuming (12.2), and if $\mathsf{E}H'(\boldsymbol{\sigma})^2$ is independent of $\boldsymbol{\sigma}$, we have*

$$\nu(U_{1,n+1}f) = \frac{1}{n}\nu(U_{1,2})\nu(f) + \frac{1}{n}\sum_{2\le \ell\le n}\nu(U_{1,\ell}f) + \delta\ , \tag{12.22}$$

where

$$U_{\ell,\ell'} = \frac{1}{N}\mathsf{E}H'(\boldsymbol{\sigma}^{\ell})H'(\boldsymbol{\sigma}^{\ell'})$$

and where, for any $a > 0$ *we have*

$$\int_{-a}^{a}|\delta|\mathrm{d}x \le K(a)N^{-1/8}A^{3/2}\ . \tag{12.23}$$

Moreover, under the conditions of Theorem 12.1.3, and if $x \ne 0$, *(12.22) holds with* $\lim_{N\to\infty}\delta = 0$.

In (12.22), the dependence on x is kept implicit. It would be more formal to say that $f = f_N$ also depends on N, but we always choose clarity over formality.

Proof. Let

$$\varDelta = \nu\left(\left|\frac{H'(\boldsymbol{\sigma})}{N} - \nu\left(\frac{H'(\boldsymbol{\sigma})}{N}\right)\right|\right)$$

so that

$$\left|\nu\left(\frac{H'(\boldsymbol{\sigma}^1)}{N}f\right) - \nu\left(\frac{H'(\boldsymbol{\sigma}^1)}{N}\right)\nu(f)\right| = \left|\nu\left(\left(\frac{H'(\boldsymbol{\sigma}^1)}{N} - \nu\left(\frac{H'(\boldsymbol{\sigma}^1)}{N}\right)\right)f\right)\right| \le \varDelta$$

and hence

$$\nu\left(\frac{H'(\boldsymbol{\sigma}^1)}{N}f\right) = \nu\left(\frac{H'(\boldsymbol{\sigma}^1)}{N}\right)\nu(f) + \delta' \tag{12.24}$$

where $|\delta'| \le \varDelta$. Setting $U_{1,1} = N^{-1}\mathsf{E}H'(\boldsymbol{\sigma}^1)^2$, by integration by parts we get

$$\nu\left(\frac{H'(\boldsymbol{\sigma}^1)}{N}\right) = -x(\nu(U_{1,1}) - \nu(U_{1,2}))$$

$$\nu\left(\frac{H'(\boldsymbol{\sigma}^1)}{N}f\right) = -x\left(\sum_{\ell\le n}\nu(U_{1,\ell}f) - n\nu(U_{1,n+1}f)\right)$$

and (12.24) becomes

$$\nu(U_{1,n+1}f) = \frac{1}{n}(\nu(U_{1,1}f) - \nu(U_{1,1})\nu(f)) + \frac{1}{n}\nu(U_{1,2})\nu(f) + \frac{1}{n}\sum_{2\le\ell\le n}\nu(U_{1,\ell}f) + \delta\ , \tag{12.25}$$

where $\delta = \delta'/n|x|$. Since $|U_{\ell,\ell'}| \le A^2$, we see from (12.25) that $|\delta| \le 2A^2$ so that for any $\varepsilon > 0$ we have

$$\int_{-a}^{a} |\delta| \mathrm{d}x \le \int_{|x|\le\varepsilon} |\delta| \mathrm{d}x + \int_{\varepsilon\le|x|\le a} \frac{|\delta'|}{|x|} \mathrm{d}x \le 4\varepsilon A^2 + \frac{K(a)}{\varepsilon} \frac{A}{N^{1/4}}$$

and optimization over ε yields (12.23). When we assume that $U_{1,1}$ does not depend on $\boldsymbol{\sigma}$, (12.25) yields (12.22), and moreover $\delta \to 0$ under the conditions of Theorem 12.1.3. □

It is not difficult to improve upon the rate (12.23) but there is little point in doing that. Throughout this chapter, we write explicit rates, but we make little effort to reach the best possible rates that could be obtained through our methods, since these do not seem optimal in any case, see Research Problem 12.1.2.

To provide a first illustration of the power of the Ghirlanda-Guerra identities we consider the situation where

$$-H'(\boldsymbol{\sigma}) = \sum_{i \le N} g_i \sigma_i \ , \tag{12.26}$$

in which case $U_{\ell,\ell'} = R_{\ell,\ell'}$, and we assume for simplicity that we are in the situation of Theorem 12.1.3. Let us denote by δ a quantity depending on N with $\lim_{N\to\infty} \delta = 0$. Then, when $n = 2$ and $f = R_{1,2}$, (12.22) implies that

$$\nu(R_{1,3}R_{1,2}) = \frac{1}{2}\nu(R_{1,2})^2 + \frac{1}{2}\nu(R_{1,2}^2) + \delta \ ,$$

and when $n = 3$, $f = R_{2,3}$, since $\nu(R_{1,3}R_{2,3}) = \nu(R_{1,2}R_{2,3})$ it means that

$$\nu(R_{1,4}R_{2,3}) = \frac{1}{3}\nu(R_{1,2})^2 + \frac{2}{3}\nu(R_{1,2}R_{2,3}) + \delta \ .$$

Combining these equations yields, since $\nu(R_{1,2}R_{2,3}) = \nu(R_{1,3}R_{1,2})$,

$$\nu(R_{1,4}R_{2,3}) = \frac{2}{3}\nu(R_{1,2})^2 + \frac{1}{3}\nu(R_{1,2}^2) + \delta \ ,$$

and since $\nu(R_{1,4}R_{2,3}) = \mathsf{E}\langle R_{1,2}\rangle^2$, this is equivalent to

$$\nu\big((R_{1,2} - \nu(R_{1,2}))^2\big) = 3\mathsf{E}(\langle R_{1,2}\rangle - \mathsf{E}\langle R_{1,2}\rangle)^2 + \delta \ . \tag{12.27}$$

This means that (when $x \ne 0$) the r.v. $\langle R_{1,2}\rangle$ fluctuates unless $R_{1,2}$ is nearly constant (i.e. $R_{1,2} \simeq \nu(R_{1,2})$). One must be cautious. This result does **not** hold for $x = 0$. For example if in (12.1) we take

$$-H(\boldsymbol{\sigma}) = \frac{\beta}{\sqrt{N}} \sum_{i<j} g_{ij}\sigma_i\sigma_j$$

then for $x = 0$ by symmetry we have $\langle R_{1,2}\rangle = 0$, so $\nu(R_{1,2}) = 0$ and the relation

$$\nu\big((R_{1,2} - \nu(R_{1,2}))^2\big) \simeq 3\mathsf{E}(\langle R_{1,2}\rangle - \mathsf{E}\langle R_{1,2}\rangle)^2$$

becomes $\nu(R_{1,2}^2) \simeq 0$ and we will see later that this is not true for $\beta > 1$.

12.2 The Extended Identities

In this section we prove one of the most amazing facts of mean-field spin glasses theory. Given any (non-random) Hamiltonian, we can find a "small random perturbation" of this Hamiltonian such that given a function f on Σ_N^n with $|f| \le 1$ and a continuous function ψ on $\mathbb{R}$ we have

$$\nu(\psi(R_{1,n+1})f) = \frac{1}{n}\nu(\psi(R_{1,2}))\nu(f) + \frac{1}{n}\sum_{2\le\ell\le n}\nu(\psi(R_{1\,\ell})f) + \delta\ , \tag{12.28}$$

where δ is typically small. These relations are called the extended Ghirlanda-Guerra identities. That these identities seem to appear out of nowhere probably means that they represent the "generic case". Consider an Hamiltonian $H_N(\boldsymbol{\sigma})$. We assume that $H_N = H_{0,N} + H_{1,N}$ where $H_{0,N}$ is non-random and $H_{1,N}$ is centered Gaussian. We assume that for a certain number B independent of N we have

$$\mathsf{E}H_{1,N}^2(\boldsymbol{\sigma}) \le NB\ . \tag{12.29}$$

For $s \ge 1$, $i_1, \ldots, i_s \ge 1$ consider independent standard Gaussian r.v.s $g_{i_1\cdots i_s}$, that are independent of the randomness of $H_{1,N}$. Let

$$-H_{N,s}(\boldsymbol{\sigma}) = \frac{1}{N^{(s-1)/2}}\sum_{i_1,\ldots,i_s} g_{i_1\cdots i_s}\sigma_{i_1}\cdots\sigma_{i_s}\ , \tag{12.30}$$

where the sum is over $1 \le i_1, \ldots, i_s \le N$. Therefore,

$$\mathsf{E}H_{N,s}(\boldsymbol{\sigma}^1)H_{N,s}(\boldsymbol{\sigma}^2) = NR_{1,2}^s\ . \tag{12.31}$$

Given numbers β_s, $|\beta_s| \le 1$, consider the "perturbing Hamiltonian"

$$-H_{N,\boldsymbol{\beta}}^{per}(\boldsymbol{\sigma}) = c_N\sum_{s\ge1}\beta_s 2^{-s}H_{N,s}(\boldsymbol{\sigma})\ , \tag{12.32}$$

where $c_N = N^{-1/32}$, and $\boldsymbol{\beta} = (\beta_s)_{s\ge1}$. Since $\mathsf{E}H_{N,s}^2(\boldsymbol{\sigma}) \le N$ these series converge, and since $c_N \to 0$ we should think of this Hamiltonian as a lower order perturbation term. Let

$$H_{N,\boldsymbol{\beta}}(\boldsymbol{\sigma}) = H_N(\boldsymbol{\sigma}) + H_{N,\boldsymbol{\beta}}^{per}(\boldsymbol{\sigma}) \tag{12.33}$$

be the Hamiltonian H_N perturbed by the lower order term. Let

$$p_N(\boldsymbol{\beta}) = \frac{1}{N}\mathsf{E}\log\sum_{\boldsymbol{\sigma}}\exp(-H_{N,\boldsymbol{\beta}}(\boldsymbol{\sigma}))\ . \tag{12.34}$$

The following expresses that, at least as far as the computation of p_N is concerned, the perturbation term has indeed small influence.

Lemma 12.2.1. *We have*

$$p_N(\mathbf{0}) \le p_N(\boldsymbol{\beta}) \le p_N(\mathbf{0}) + c_N^2 \; . \tag{12.35}$$

Proof. Let us denote by E' expectation in the r.v.s $g_{i_1 \cdots i_s}$, so that

$$\mathsf{E}'(H^{per}_{N,\boldsymbol{\beta}}(\boldsymbol{\sigma}))^2 = c_N^2 N \sum_{s \ge 1} \beta_s^2 2^{-2s} \le 2c_N^2 N \; ,$$

and, using (A.1) we have

$$\mathsf{E}' \exp(-H_{N,\boldsymbol{\beta}}(\boldsymbol{\sigma})) \le (\exp N c_N^2) \exp(-H_N(\boldsymbol{\sigma})) \; .$$

Using Jensen's inequality by taking the expectation E' inside the logarithm rather than outside then yields the right-hand side of (12.35).

To prove the left-hand side inequality we observe that

$$p_N(\boldsymbol{\beta}) - p_N(\mathbf{0}) = \frac{1}{N} \mathsf{E} \log \langle \exp(-H^{per}_{N,\boldsymbol{\beta}}) \rangle \; , \tag{12.36}$$

where the bracket denotes an average for the Gibbs measure with Hamiltonian H_N. Now, Jensen's inequality implies

$$\langle \exp(-H^{per}_{N,\boldsymbol{\beta}}) \rangle \ge \exp \langle -H^{per}_{N,\boldsymbol{\beta}} \rangle$$

so that the quantity (12.36) is $\ge N^{-1} \mathsf{E} \langle -H^{per}_{N,\boldsymbol{\beta}} \rangle = 0$. □

One should be cautious. The previous lemma shows that adding the perturbation term to the Hamiltonian has a limited influence on p_N. This does **not** say that this has also a limited influence on the structure of the Gibbs measure.

Our next result asserts that for the perturbed Hamiltonian (12.33), then (12.28) holds for a quantity δ that is small in average over $\boldsymbol{\beta}$. We find it convenient to restrict the range of $\boldsymbol{\beta}$ to $[-1,1]^{\mathbb{N}}$.

Theorem 12.2.2. *Denote by $\langle \cdot \rangle$ an average for the Hamiltonian (12.33), the dependence on $\boldsymbol{\beta}$ being implicit. Then (12.28) holds with $\delta = \delta_N(\boldsymbol{\beta})$ satisfying*

$$\lim_{N \to \infty} \int |\delta| \mathrm{d}\boldsymbol{\beta} = 0 \; , \tag{12.37}$$

where the integral is for the uniform probability over $[-1,1]^{\mathbb{N}}$.

Unfortunately this says nothing about the most interesting value of $\boldsymbol{\beta}$, that is $\boldsymbol{\beta} = \mathbf{0}$! Let us also mention that in Chapter 14 we will show that a particularly important class of models satisfies the extended Ghirlanda-Guerra identities even without the perturbation term.

Proof. Since the function ψ can be uniformly approximated over $[-1,1]$ by a polynomial, it suffices to prove that for each $s \geq 1$ we have

$$\nu(R_{1,n+1}^s f) = \frac{1}{n}\nu(R_{1,2}^s)\nu(f) + \frac{1}{n}\sum_{2\leq \ell \leq n} \nu(R_{1,\ell}^s f) + \delta \,, \tag{12.38}$$

where δ is as in (12.37). To see this we will use (12.22) in the case $H' = 2^{-s}c_N H_{N,s}$, β_s instead of x, $H_0 = H_{0,N}$ and

$$-H_1 = -H_{N,\boldsymbol{\beta}} - \beta_s 2^{-s} c_N H_{N,s} + H_{0,N} = -H_{1,N} + \sum_{p\neq s} \beta_p 2^{-p} H_{N,p}$$

(which does not depend on β_s), so that, recalling (12.29) we see that (12.2) holds for $A = 1 + B$. Then $U_{\ell,\ell'} = N^{-1}\mathsf{E}H'(\boldsymbol{\sigma}^\ell)H'(\boldsymbol{\sigma}^{\ell'}) = 2^{-2s}c_N^2 R_{\ell,\ell'}^s$. Therefore (12.2) implies (12.38) where $\delta = 2^{2s}c_N^{-2}\delta'$, for a quantity δ' such that

$$\int_{-1}^{1} |\delta'| \mathrm{d}\beta_s \leq L(1+B)N^{-1/8}$$

and thus

$$\int_{-1}^{1} |\delta| \mathrm{d}\beta_s \leq L(1+B)2^{2s}c_N^{-2}N^{-1/8} \leq L(1+B)2^{s/2}N^{-1/32} \,.$$

That δ satisfies (12.37) then follows from Fubini's theorem. □

In the next two sections we will prove some amazing consequences of (12.28). Other applications will be given in Chapter 16. It is quite interesting that, as of today, all these applications use (12.28) only in the special case where f is of the form $f(\boldsymbol{\sigma}^1,\ldots,\boldsymbol{\sigma}^n) = f^\sim((R_{\ell,\ell'})_{1\leq \ell<\ell'\leq n})$ for a certain continuous function $f^\sim : \mathbb{R}^{n(n-1)/2} \to \mathbb{R}$.

12.3 A Positivity Principle

In this section we prove a remarkably general principle: if a system satisfies the extended Ghirlanda-Guerra identities, then the overlap $R_{1,2}$ is essentially non-negative.

Theorem 12.3.1. *Consider a Hamiltonian depending both on N and a parameter $\boldsymbol{\beta} \in [-1,1]^{\mathbb{N}}$; assume that given any n, any function f on Σ_N^n with $|f| \leq 1$ and any continuous function ψ on $[-1,1]$, we have*

$$\nu(\psi(R_{1,n+1})f) = \frac{1}{n}\nu(\psi(R_{1,2}))\nu(f) + \frac{1}{n}\sum_{2\leq \ell\leq n}\nu(\psi(R_{1,\ell})f) + \delta \tag{12.39}$$

where $\delta = \delta_N(\boldsymbol{\beta})$ satisfies

$$\lim_{N\to\infty}\int|\delta|\mathrm{d}\boldsymbol{\beta}=0\,. \tag{12.40}$$

Then, for each $\varepsilon>0$ we have

$$\lim_{N\to\infty}\int\nu(\mathbf{1}_{\{R_{1,2}\le-\varepsilon\}})\mathrm{d}\boldsymbol{\beta}=0\,. \tag{12.41}$$

Both integrals in (12.40) are performed for the uniform measure on $[-1,1]^{\mathbb{N}}$. It is convenient to write δ for any function of N and $\boldsymbol{\beta}$ that satisfies (12.40). This function might not be the same at each occurrence. Therefore (12.41) can be written as $\nu(\mathbf{1}_{\{R_{1,2}\le-\varepsilon\}})=\delta$.

Theorem 12.3.1 illustrates how a (comparatively small) perturbation of the Hamiltonian can change the Gibbs measure in a dramatic way. It is certainly not true in general that the overlap is essentially positive. Indeed, for the ordinary SK model without external field, the distribution of the overlap is symmetric around 0. The perturbation term breaks this symmetry.

The proof of Theorem 12.3.1 has two parts. Until the end of this proof, we fix once and for all $0<\varepsilon\le 1$ and we set

$$D_n=\left\{(\boldsymbol{\sigma}^1,\dots,\boldsymbol{\sigma}^n)\in\Sigma_N^n\;;\;\forall\ell\;,\;2\le\ell\le n\;,\;R_{1,\ell}\le-\varepsilon\right\}\,.$$

The first part is purely deterministic.

Proposition 12.3.2. *Consider $n\ge 3$ and a probability measure G on Σ_N. Then we have*

$$G^{\otimes n}(D_n)\le\frac{5\log n}{\varepsilon n}\,. \tag{12.42}$$

Using (12.42) for Gibbs' measure, and taking expectation, this implies that

$$\nu(D_n)\le\frac{5\log n}{\varepsilon n}\,. \tag{12.43}$$

The second part of the proof is to observe that the relations

$$\ell\le n\quad\Rightarrow\quad D_n=D_n\cap\{R_{1,\ell}\le-\varepsilon\} \tag{12.44}$$

make it possible to recursively estimate $\nu(D_n)$ through (12.39) as a function of $\nu(D_2)$ and to show that (12.43) implies $\nu(D_2)=\delta$.

We prepare the proof of Proposition 12.3.2 with the following, where we remind the reader that $0<\varepsilon<1$ has been fixed once and for all.

Lemma 12.3.3. *Consider a probability measure G on Σ_N, a number $0<c<1$, and let*

$$U=\{\boldsymbol{\sigma}^1\;;\;G(\{\boldsymbol{\sigma}^2\;;\;R_{1,2}\le-\varepsilon\})>1-c\}\,. \tag{12.45}$$

Then

$$G(U)\le\frac{2c}{\varepsilon}\,. \tag{12.46}$$

Proof. For any probability measure μ on Σ_N we have

$$\int R_{1,2}\mathrm{d}\mu(\boldsymbol{\sigma}^1)\mathrm{d}\mu(\boldsymbol{\sigma}^2) = \frac{1}{N}\int \boldsymbol{\sigma}^1\cdot\boldsymbol{\sigma}^2\mathrm{d}\mu(\boldsymbol{\sigma}^1)\mathrm{d}\mu(\boldsymbol{\sigma}^2) = \frac{1}{N}\left\|\int \boldsymbol{\sigma}\mathrm{d}\mu(\boldsymbol{\sigma})\right\|^2 \geq 0\,.$$

Let $A = \{(\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2)\ ;\ R_{1,2} > -\varepsilon\}$. Then, since $R_{1,2} \leq 1$ we get

$$0 \leq \int R_{1,2}\mathrm{d}\mu(\boldsymbol{\sigma}^1)\mathrm{d}\mu(\boldsymbol{\sigma}^2) \leq -\varepsilon\mu^{\otimes 2}(A^c) + \mu^{\otimes 2}(A) = -\varepsilon + (1+\varepsilon)\mu^{\otimes 2}(A)$$

and therefore

$$\mu^{\otimes 2}(A) \geq \frac{\varepsilon}{1+\varepsilon} \geq \frac{\varepsilon}{2}\,. \tag{12.47}$$

Consider the probability measure μ on Σ_N given by

$$\mu(C) = G(C\cap U)/G(U)\,.$$

For $\boldsymbol{\sigma}^1 \in U$, by (12.45) we have $G(\{\boldsymbol{\sigma}^2\ ;\ R_{1,2} > -\varepsilon\}) \leq c$, and therefore

$$\mu(\{\boldsymbol{\sigma}^2\ ;\ R_{1,2} > -\varepsilon\}) \leq \frac{1}{G(U)}G(\{\boldsymbol{\sigma}^2\ ;\ R_{1,2} > -\varepsilon\}) \leq \frac{c}{G(U)}\,.$$

Thus, by Fubini theorem and since μ is supported on U we have

$$\begin{aligned}\mu^{\otimes 2}(A) &= \mu^{\otimes 2}(\{(\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2)\ ;\ R_{1,2} > -\varepsilon\}) \\ &= \int_U \mu(\{\boldsymbol{\sigma}^2\ ;\ R_{1,2} > -\varepsilon\})\mathrm{d}\mu(\boldsymbol{\sigma}^1) \leq \frac{c}{G(U)}\,,\end{aligned}$$

and comparing with (12.47) proves the result (12.46). □

Proof of Proposition 12.3.2. We define

$$f(\boldsymbol{\sigma}^1) = G(\{\boldsymbol{\sigma}^2\ ;\ R_{1,2} \leq -\varepsilon\})\,,$$

so that by Fubini theorem we have

$$\begin{aligned}G^{\otimes n}(D_n) &= G^{\otimes n}(\{(\boldsymbol{\sigma}^1,\dots,\boldsymbol{\sigma}^n)\ ;\forall\ell\ ;\ 2\leq\ell\leq n\ ,\ R_{1,\ell}\leq -\varepsilon\}) \\ &= \int G^{\otimes(n-1)}(\{(\boldsymbol{\sigma}^2,\dots,\boldsymbol{\sigma}^n)\ ;\forall\ell\ ;\ 2\leq\ell\leq n\ ,\ R_{1,\ell}\leq -\varepsilon\})\mathrm{d}G(\boldsymbol{\sigma}^1) \\ &= \int f^{n-1}(\boldsymbol{\sigma})\mathrm{d}G(\boldsymbol{\sigma})\,.\end{aligned}$$

Given $0 < c < 1$, the set U defined in (12.45) is the set $\{f > 1-c\}$, so that using (12.46) and since $f \leq 1$ we get

$$\begin{aligned}\int f^{n-1}(\boldsymbol{\sigma})\mathrm{d}G(\boldsymbol{\sigma}) &= \int_{U^c} f^{n-1}(\boldsymbol{\sigma})\mathrm{d}G(\boldsymbol{\sigma}) + \int_U f^{n-1}(\boldsymbol{\sigma})\mathrm{d}G(\boldsymbol{\sigma}) \\ &\leq (1-c)^{n-1} + G(U) \\ &\leq \exp(-c(n-1)) + \frac{2c}{\varepsilon}\,.\end{aligned}$$

We then take $c = (\log n)/(n-1) \le 2\log n/n$ to obtain (since $\varepsilon < 1 < \log 3 \le \log n$)

$$G^{\otimes n}(D_n) \le \frac{1}{n} + \frac{4\log n}{\varepsilon n} \le \frac{5\log n}{\varepsilon n} \,. \qquad \square$$

The next step it to estimate $\nu(D_n)$ from below. The idea is very simple. Even though we are permitted to use (12.39) only when ψ is continuous, let us suppose for a moment that (12.39) holds for the function $\psi(x) = \mathbf{1}_{\{x\le-\varepsilon\}}$. When $f = \mathbf{1}_{D_n}$, this relation read as

$$\nu(D_{n+1}) = \frac{n-1+\nu(D_2)}{n}\nu(D_n) + \delta$$

because $f = \mathbf{1}_{D_n}\mathbf{1}_{\{R_{1,\ell}\le-\varepsilon\}}$ for $\ell \le n$ by (12.44). This relation is then easy to iterate. We will show later that the previous scheme of proof can be adapted to prove the following.

Proposition 12.3.4. *Let $a = \nu(\mathbf{1}_{\{R_{1,2}\le-2\varepsilon\}})$. Then*

$$\nu(D_n) \ge \frac{a}{Ln^{1-a}} + \delta \,. \tag{12.48}$$

Proof of Theorem 12.3.1. Let us repeat that $0<\varepsilon<1$ is fixed once and for all. Comparing (12.43) and (12.48) we have, given n,

$$\frac{a}{Ln^{1-a}} \le \frac{5\log n}{\varepsilon n} + \delta$$

so that

$$an^a \le \frac{L_1}{\varepsilon}\log n + \delta \,. \tag{12.49}$$

For each $n \ge 3$ consider the smallest number $d(n)$ such that

$$a \ge d(n) \quad\Rightarrow\quad an^a \ge 1 + \frac{L_1}{\varepsilon}\log n$$

so that by (12.49),

$$a \ge d(n) \quad\Rightarrow\quad \delta \ge 1$$

and thus

$$\lim_{N\to\infty}\int \mathbf{1}_{\{a\ge d(n)\}}\mathrm{d}\boldsymbol{\beta} \le \lim_{N\to\infty}\int |\delta|\mathrm{d}\boldsymbol{\beta} = 0 \,.$$

Since $a \le 1$, we have

$$\int a\mathrm{d}\boldsymbol{\beta} \le d(n) + \int \mathbf{1}_{\{a\ge d(n)\}}\mathrm{d}\boldsymbol{\beta}$$

so that

$$\limsup_{N\to\infty}\int a\mathrm{d}\boldsymbol{\beta} \le d(n) \,.$$

This is true for each n; since, as is obvious from the definition of $d(n)$ we have $\lim_{n\to\infty} d(n) = 0$, we have proved that $\lim_{N\to\infty} \int a \mathrm{d}\boldsymbol{\beta} = 0$, i.e. $a = \delta$. □

Proof of Proposition 12.3.4. Let us fix numbers

$$-2\varepsilon = b_1 < b_2 < \cdots < b_n = -\varepsilon$$

and for $2 \le r \le n$ let us consider a continuous function $\psi_r : \mathbb{R} \to [0,1]$ which satisfies

$$x \le b_{r-1} \quad \Rightarrow \quad \psi_r(x) = 1\ ; \quad x \ge b_r \quad \Rightarrow \quad \psi_r(x) = 0\ .$$

Thus

$$x < -2\varepsilon = b_1 \Rightarrow \psi_2(x) \ge 1$$

and therefore

$$\nu(\psi_2(R_{1,2})) \ge \nu(\mathbf{1}_{\{R_{1,2} \le -2\varepsilon\}}) = a\ . \tag{12.50}$$

For $r \ge 2$, define on Σ_N^r the function

$$f_r = \prod_{2 \le \ell \le r} \psi_\ell(R_{1,\ell})\ .$$

For $2 \le \ell \le r$ we observe the identity

$$\psi_{r+1}(R_{1,\ell}) f_r = f_r\ , \tag{12.51}$$

because if $f_r \ne 0$, then for each $\ell \le r$ we have $\psi_\ell(R_{1,\ell}) \ne 0$ and hence $R_{1,\ell} \le b_\ell \le b_r$ so that $\psi_{r+1}(R_{1,\ell}) = 1$. Using (12.50), (12.51) and (12.39) for r rather than n we get

$$\nu(f_{r+1}) \ge \frac{1}{r} a \nu(f_r) + \frac{r-1}{r} \nu(f_r) + \delta$$

and thus

$$\nu(f_{r+1}) \ge \frac{r-1+a}{r} \nu(f_r) + \delta\ . \tag{12.52}$$

Since $\log(1-x) \ge -x - x^2$ for $x \le 1/2$, we have for $r \ge 2$:

$$\frac{r-1-a}{r} = 1 - \frac{1-a}{r} \ge \exp\left(-\frac{1-a}{r} - \frac{1}{r^2}\right)$$

and since $1/2 + \cdots + 1/(n-1) \le \log n$, iteration of (12.52) from $r = 2$ to $r = n-1$ yields

$$\nu(f_n) \ge \frac{1}{L n^{1-a}} \nu(f_2) + \delta\ .$$

This implies (12.48) because $\nu(f_2) \ge a$ and $\nu(f_n) \le \nu(D_n)$. □

12.4 The Distribution of the Overlaps at Given Disorder

In this section we prove remarkable general principle. Roughly speaking, if a system satisfies the generalized Ghirlanda-Guerra identities, then, at a given disorder, the distribution of the overlap $R_{1,2}$ under Gibbs' measure charges only a few points. Typically, given $0 < \eta < 1$, all but a proportion $1 - \eta$ of this distribution is carried by about η^{-2} points (possibly depending on the disorder) in the interval $[0,1]$. Moreover, the support of the random distribution of $R_{1,2}$ under Gibbs' measure is non-random.

Consider the distribution μ of $R_{1,2}$ under Gibbs' measure. Consider the average $\overline{\mu}$ of μ under the disorder, and the support S of $\overline{\mu}$. Then S is also the support of μ and thus one can picture at given disorder μ as being carried by a (random) sequence of points that is dense in S.

The author constructed an example (for the "spherical model" [110] for which computations are easier) where the average $\overline{\mu}$ over the disorder of the distribution μ of $R_{1,2}$ under Gibbs' measure is itself continuous. However, according to the numerical simulations of physicists, it seems that often $\overline{\mu}$ has a point mass at the rightmost point of its support S.

The above very general description of the results of this section implicitly takes place "in the limit $N \to \infty$". To formulate a result at a given N, we introduce the following definition.

Definition 12.4.1. *For a probability measure μ on $[-1,1]$, an integer n and $\varepsilon > 0$, we define $A(\mu, n, \varepsilon)$ as the maximum amount of mass of μ that can be carried by the union of n intervals, each of length at most 2ε, when these intervals are optimally chosen i.e.*

$$A(\mu, n, \varepsilon) = \sup\big\{\mu(B)\ ;\ B \text{ is the union of } n \text{ sub-intervals of } [0,1] \\ \text{each of length} \le 2\varepsilon\big\}\ . \tag{12.53}$$

Theorem 12.4.2. *Consider a Hamiltonian depending on N and on a parameter $\boldsymbol{\beta} \in [-1,1]^{\mathbb{N}}$. Assume that given any n, any function f on Σ_N^n and any continuous function ψ on $[-1,1]$, conditions (12.39) and (12.40) hold, that is*

$$\nu(\psi(R_{1,n+1})f) = \frac{1}{n}\nu(\psi(R_{1,2}))\nu(f) + \frac{1}{n}\sum_{2\le\ell\le n}\nu(\psi(R_{1,\ell})f) + \delta\ ,$$

where $\lim_{N\to\infty}\int|\delta|\mathrm{d}\boldsymbol{\beta} = 0$. Then, if μ denotes the (random) law of $R_{1,2}$ under Gibbs' measure, for any n and any $\varepsilon > 0$ we have

$$\mathsf{E}\left(A\left(\mu, \frac{n(n-1)}{2}, \varepsilon\right)\right) \ge 1 - \frac{2}{n+1} + \delta\ . \tag{12.54}$$

Moreover, if φ is a continuous function with $0 \le \varphi \le 1$, for each n we have

$$\mathsf{P}\left(\int \varphi \mathrm{d}\mu \geq \frac{1}{n}\right) \geq 1 - \frac{8}{(\log n)\nu(\varphi(R_{1,2}))} + \delta \ . \tag{12.55}$$

Here and everywhere in this section δ satisfies $\lim_{N\to\infty} \int |\delta| \mathrm{d}\boldsymbol{\beta} = 0$.

To interpret (12.54), we fix n and $\varepsilon > 0$. For N large and the typical value of $\boldsymbol{\beta}$, δ is small, say $\delta \leq 1/(n+1)$. Then

$$\mathsf{E}\left(A\left(\mu, \frac{n(n-1)}{2}, \varepsilon\right)\right) \geq 1 - \frac{3}{n+1} \ ,$$

that is, typically all the mass of μ but the small proportion $3/(n+1)$ is carried by $n(n-1)/2$ intervals of length $\leq 2\varepsilon$. This is true however small ε is (but the smaller ε, the larger one has to take N), so "in the limit $N \to \infty$" all but at most a proportion of $3/(n+1)$ (or even $2/(n+1)$) of the mass of μ is carried by $n(n-1)/2$ points.

We turn now to the interpretation of (12.55). Let $\overline{\mu}$ denote the average of μ over the disorder, and S the support of $\overline{\mu}$ (which is conceivably a rather small set). We first show that for very general reasons the support of μ is a.s. contained in S. For any continuous function φ, it holds

$$\nu(\varphi(R_{1,2})) = \mathsf{E}\langle \varphi(R_{1,2})\rangle = \mathsf{E}\int \varphi(x) \mathrm{d}\mu(x) = \int \varphi(x) \mathrm{d}\overline{\mu}(x) \ .$$

(In other words, $\overline{\mu}$ is the law of $R_{1,2}$ under ν). If $\varphi \geq 0$, $\varphi > 0$ outside S and $\varphi = 0$ on S, then $\int \varphi \mathrm{d}\overline{\mu} = 0$ and therefore $\mathsf{E}\int \varphi \mathrm{d}\mu = 0$, so that $\int \varphi \mathrm{d}\mu = 0$ a.e. Now, when $\int \varphi \mathrm{d}\mu = 0$, μ is supported by the set $\{\varphi = 0\} = S$. Thus, the support of μ is a.s. contained in the support S of $\overline{\mu}$. Assuming now that, as expected, $\overline{\mu}$ has a limit as $N \to \infty$, we argue that (12.55) can be interpreted as saying that "in the limit $N = \infty$ the support of μ is a.s. the support S of $\overline{\mu}$". Consider an interval I with $S \cap I \neq \emptyset$, and $0 \leq \varphi \leq 1$ such that $\varphi = 0$ outside I and $\int \varphi \mathrm{d}\overline{\mu} > 0$. Given $\varepsilon > 0$, we may take n large enough so that (12.55) implies

$$\mathsf{P}\left(\int \varphi \mathrm{d}\mu \geq \frac{1}{n}\right) \geq 1 - \varepsilon + \delta \ .$$

Thus for large N and the typical value of δ we have $\mathsf{P}\left(\int \varphi \mathrm{d}\mu > 0\right) \geq 1 - 2\varepsilon$. As ε is arbitrarily small, this means in a sense that "a.s. $\int \varphi \mathrm{d}\mu > 0$" so that the support of μ meets I; thus the support of μ cannot be smaller than the support of $\overline{\mu}$.

Lemma 12.4.3. *Assuming (12.39), given any function f on Σ_N^n with $|f| \leq 1$ and any continuous function ψ we have*

$$\begin{aligned}\nu(\psi(R_{n+1,n+2})f) = {} & \frac{2}{n+1}\nu(\psi(R_{1,2}))\nu(f) \\ & + \frac{1}{n(n+1)} \sum_{k \neq \ell,\ k,\ell \leq n} \nu(\psi(R_{k,\ell})f) + \delta \ . \end{aligned} \tag{12.56}$$

Proof. We observe that (12.39) implies (by exchanging 1 and n) that

$$\nu(\psi(R_{n,n+1})f) = \frac{1}{n}\nu(\psi(R_{1,2}))\nu(f) + \frac{1}{n}\sum_{1\le\ell\le n-1}\nu(\psi(R_{n,\ell})f) + \delta$$

and we use this with $n+1$ rather than n to get

$$\begin{aligned}\nu(\psi(R_{n+1,n+2})f) &= \frac{1}{n+1}\nu(\psi(R_{1,2}))\nu(f)\\ &\quad + \frac{1}{n+1}\sum_{\ell\le n}\nu(\psi(R_{n+1,\ell})f) + \delta\,. \end{aligned} \tag{12.57}$$

We use (12.40) again to obtain for each $\ell \le n$ that

$$\nu(\psi(R_{n+1,\ell})f) = \frac{1}{n}\nu(\psi(R_{1,2}))\nu(f) + \frac{1}{n}\sum_{k\le n,\ k\ne\ell}\nu(\psi(R_{k,\ell})f) + \delta\,,$$

and we substitute in (12.57). □

Proof of (12.55). We use (12.56) for $n = 2m$, taking $\psi = \varphi$ and

$$f_m = \prod_{1\le k\le m}(1 - \varphi(R_{2k-1,2k}))$$

so that, since $f_m \ge 0$,

$$\nu(\varphi(R_{2m+1,2m+2})f_m) \ge \frac{2}{2m+1}\nu(\varphi(R_{1,2}))\nu(f_m) + \delta\,. \tag{12.58}$$

Now, by definition of μ and independence,

$$\langle f_m\rangle = \left(1 - \int\varphi \mathrm{d}\mu\right)^m \quad;\qquad \langle\varphi(R_{2m+1,2m+2})f_m\rangle = \int\varphi \mathrm{d}\mu\left(1 - \int\varphi \mathrm{d}\mu\right)^m$$

and since for $0 < a < 1$ it holds true that $\sum_{m\ge1} a(1-a)^m = 1 - a \le 1$, we have

$$\sum_{m\ge1}\langle\varphi(R_{2m+1,2m+2})f_m\rangle = \sum_{m\ge1}\int\varphi \mathrm{d}\mu\left(1 - \int\varphi \mathrm{d}\mu\right)^m \le 1\,.$$

Summation of (12.58) for $m \le n$ yields

$$\begin{aligned}1 &\ge \nu(\varphi(R_{1,2}))\sum_{m\le n}\frac{2}{2m+1}\mathsf{E}\left(1 - \int\varphi \mathrm{d}\mu\right)^m + \delta\\ &\ge \frac{1}{2}\log n\ \nu(\varphi(R_{1,2}))\mathsf{E}\left(1 - \int\varphi \mathrm{d}\mu\right)^n + \delta\,, \end{aligned} \tag{12.59}$$

using that $\sum_{m\leq n} 2/(2m+1) \geq (1/2)\log n$ and that $\nu(f_m) = \mathsf{E}\langle f_m\rangle = \mathsf{E}\left(1-\int\varphi\mathrm{d}\mu\right)^m$. Since $(1-x)^n \geq 1/4$ for $x \leq 1/n$, we have

$$\mathsf{E}\left(1-\int\varphi\mathrm{d}\mu\right)^n \geq \frac{1}{4}\mathsf{P}\left(\int\varphi\mathrm{d}\mu \leq \frac{1}{n}\right) ,$$

and combining this with (12.59) we get (12.55). □

We now turn to the proof of (12.54). We write $C_n = \mathbb{R}^{n(n-1)/2}$ and we denote by $\mathbf{x} = (x_{k,\ell})_{1\leq k<\ell\leq n}$ the generic point of C_n. Consider the (random) law γ of $(R_{k,\ell})_{1\leq k<\ell\leq n}$ in C_n, and $\overline{\gamma}$ the average of γ over the disorder.

Lemma 12.4.4. *For a continuous function φ on $\mathbb{R}\times C_n$, one has*

$$\begin{aligned}\mathsf{E}\int\varphi(x,\mathbf{x})\mathrm{d}\mu(x)\mathrm{d}\gamma(\mathbf{x}) = &\frac{2}{n+1}\int\varphi(x,\mathbf{x})\mathrm{d}\overline{\mu}(x)\mathrm{d}\overline{\gamma}(\mathbf{x}) \\ &+\frac{1}{n(n+1)}\sum_{k\neq\ell,\ k,\ell\leq n}\int\varphi(x_{k,\ell},\mathbf{x})\mathrm{d}\overline{\gamma}(\mathbf{x})+\delta\ .\end{aligned} \tag{12.60}$$

Proof. Since φ can be approximated arbitrarily well by a sum of functions of the type $\psi(x)g(\mathbf{x})$ it suffices to consider a function φ of this type. Let us then define $f = g((R_{k,\ell})_{1\leq k<\ell\leq n})$. Then

$$\begin{aligned}\int\varphi(x,\mathbf{x})\mathrm{d}\overline{\mu}(x)\mathrm{d}\overline{\gamma}(\mathbf{x}) &= \int\psi(x)g(\mathbf{x})\mathrm{d}\overline{\mu}(x)\mathrm{d}\overline{\gamma}(\mathbf{x}) \\ &= \int\psi(x)\mathrm{d}\overline{\mu}(x)\int g(\mathbf{x})\mathrm{d}\overline{\gamma}(\mathbf{x}) \\ &= \mathsf{E}\int\psi(x)\mathrm{d}\mu(x)\mathsf{E}\int g(\mathbf{x})\mathrm{d}\gamma(\mathbf{x}) \\ &= \mathsf{E}\langle\psi(R_{1,2})\rangle\mathsf{E}\langle f\rangle \\ &= \nu(\psi(R_{1,2}))\nu(f)\ .\end{aligned}$$

Moreover,

$$\begin{aligned}\int\varphi(x_{k,\ell},\mathbf{x})\mathrm{d}\overline{\gamma}(\mathbf{x}) &= \mathsf{E}\int\psi(x_{k,\ell})g(\mathbf{x})\mathrm{d}\gamma(\mathbf{x}) \\ &= \mathsf{E}\langle\psi(R_{k,\ell})g\rangle = \nu(\psi(R_{k,\ell})g)\ .\end{aligned}$$

Therefore (12.60) reduces to (12.56) when $f = g((R_{k,\ell})_{1\leq k<\ell\leq n})$. □

Proof of (12.54). We use (12.60) for the function

$$\varphi(x,\mathbf{x}) = \min\left(1, \frac{1}{\varepsilon}\min_{1\leq k<\ell\leq n}|x-x_{k,\ell}|\right)$$

so that $\varphi(x_{k,\ell},\mathbf{x}) = 0$ and by (12.60)

$$\mathsf{E}\int \varphi(x,\mathbf{x})\mathrm{d}\mu(x)\mathrm{d}\gamma(\mathbf{x}) \le \frac{2}{n+1} + \delta \; . \tag{12.61}$$

Now, given $\mathbf{x}$

$$\forall k < \ell \; , \quad |x - x_{k,\ell}| \ge \varepsilon \quad \Rightarrow \quad \varphi(x,\mathbf{x}) = 1 \; ,$$

so that

$$\begin{aligned}\int \varphi(x,\mathbf{x})\mathrm{d}\mu(x) &\ge \mu(\{x \; ; \; \forall k < \ell \; , \; |x - x_{k,\ell}| \ge \varepsilon\}) \\ &= 1 - \mu(\{x \; ; \; \exists k < \ell \; , \; |x - x_{k,\ell}| \le \varepsilon\}) \\ &\ge 1 - A\left(\mu, \frac{n(n-1)}{2}, \varepsilon\right) \; ,\end{aligned}$$

by definition of this latter quantity. Therefore

$$\int \varphi(x,\mathbf{x})\mathrm{d}\mu(x)\mathrm{d}\gamma(\mathbf{x}) \ge 1 - A\left(\mu, \frac{n(n-1)}{2}, \varepsilon\right) \; .$$

Comparing with (12.61) this completes the proof. □

12.5 Large Deviations

Given a random Hamiltonian H_N (say with Σ_N as configuration space) one can argue that the fundamental question is the computation of

$$p_N = \frac{1}{N}\mathsf{E}\log Z_N \; , \tag{12.62}$$

where $Z_N = \sum_{\boldsymbol{\sigma}} \exp(-H_N(\boldsymbol{\sigma}))$ is the partition function. Given a number a, it is also of interest to evaluate

$$p_{N,a} = \frac{1}{Na}\log \mathsf{E}Z_N^a \; , \tag{12.63}$$

of which (12.62) is "the case $a = 0$" because

$$\lim_{a\to 0}\frac{1}{a}\log \mathsf{E}Z_N^a = \lim_{a\to 0}\frac{1}{a}\log \mathsf{E}\exp a\log Z_N = \mathsf{E}\log Z_N \; .$$

In some sense, this gives information on the system conditioned on the fact that Z_N is large (if $a > 0$) or small (if $a < 0$).

The tools we have developed (say in Chapter 1) adapt very well to the study of (12.63) "at high temperature". The main difference is that when taking expectation in the disorder one has to make a change of density $Z_N^a/\mathsf{E}Z_N^a$. In other words, for a function f on Σ_N^n, we now define (keeping the dependence in a implicit)

$$\nu(f) = \mathsf{E}\left(\frac{Z_N^a}{\mathsf{E}Z_N^a}\langle f\rangle\right) , \tag{12.64}$$

where $\langle\cdot\rangle$ denotes as usual an average for the Gibbs measure.

In this section we will investigate what happens to the Ghirlanda-Guerra identities in the present setting. An unexpected fact is that these inequalities become *stronger* as a increases. For $a > 1$ they become so strong that in some sense they imply that

$$\nu\big((R_{1,2} - \nu(R_{1,2}))^2\big) \simeq 0 . \tag{12.65}$$

One can interpret this relation by saying that for $a > 1$, there is no such thing as a low-temperature phase; all systems are at high temperature. One could hope that we will then understand everything in detail; this is not yet the case (and shows how little we really understand). For example, in the situation of the SK model with Hamiltonian (1.10), it is not difficult to show that $\nu(R_{1,2})$ must be a near-solution of the self-consistency equation

$$q = \mathsf{E}\frac{\mathrm{ch}^a Y}{\mathsf{E}\mathrm{ch}^a Y}\mathrm{th}^2 Y$$

for $Y = \beta z\sqrt{q} + h$. But how does one show that this solution is unique? And, even if this calculus problem can be solved, how can one improve the information that $\nu((R_{1,2} - q)^2) \simeq 0$ into an exponential inequality? We refer to [111] and to Panchenko's beautiful paper [68] for more on this topic.

Keeping the dependence on N implicit we consider a (possibly random) Hamiltonian H and we introduce a "perturbed" version of this Hamiltonian as follows:

$$-H^x(\boldsymbol{\sigma}) = -H(\boldsymbol{\sigma}) + xc_N \sum_{i\le N} g_i\sigma_i , \tag{12.66}$$

where $(g_i)_{i\le N}$ are independent standard Gaussian r.v.s, independent of the randomness of H. As in (12.33), the idea is that the second term on the right-hand side of (12.66) is a lower order perturbation of the original Hamiltonian. We denote by $\nu = \nu_x$ the average (12.64) relative to the Hamiltonian (12.66), keeping the dependence on x implicit.

Theorem 12.5.1. *If $a > 1$ we have*

$$\int_0^1 x\nu\big((R_{1,2} - \nu(R_{1,2}))^2\big)\mathrm{d}x \le \frac{K(a)}{c_N\sqrt{N}} ,$$

where $K(a)$ depends on a only.

Thus, as soon as $c_N^2 N \to \infty$, given N large, $\nu((R_{1,2} - \nu(R_{1,2}))^2)$ is small for the typical value of x.

We consider the function

$$\varphi(x) = \frac{1}{Na} \log \mathsf{E} Z_{N,x}^a \ , \tag{12.67}$$

where of course $Z_{N,x} = \sum_{\boldsymbol{\sigma}} \exp(-H^x(\boldsymbol{\sigma}))$. To lighten notation we define

$$H'(\boldsymbol{\sigma}) := \sum_{i \le N} g_i \sigma_i \ .$$

(Note that there is no minus sign here.)

Lemma 12.5.2. *We have*

$$\int_0^1 \nu\left(\left(\frac{H'}{N} - \nu\left(\frac{H'}{N}\right)\right)^2\right) \mathrm{d}x \le \frac{a}{N} \ . \tag{12.68}$$

Proof. We note that

$$\frac{\mathrm{d}}{\mathrm{d}x} Z_{N,x} = c_N Z_{N,x} \langle H' \rangle \tag{12.69}$$

$$\frac{\mathrm{d}}{\mathrm{d}x} Z_{N,x}^a = c_N a Z_{N,x}^a \langle H' \rangle \tag{12.70}$$

$$\varphi'(x) = \frac{c_N}{N} \frac{1}{\mathsf{E} Z_{N,x}^a} \mathsf{E}(Z_{N,x}^a \langle H' \rangle) = c_N \nu\left(\frac{H'}{N}\right) \ . \tag{12.71}$$

We differentiate (12.71), using that

$$\frac{\mathrm{d}}{\mathrm{d}x} \langle H' \rangle = c_N(\langle H'^2 \rangle - \langle H' \rangle^2)$$

and, using the notation

$$\mathsf{E}'(A) = \frac{1}{\mathsf{E} Z_{N,x}^a} \mathsf{E}(Z_{N,x}^a A) \ ,$$

we get by a straightforward computation that

$$\varphi''(x) = \frac{c_N^2}{N}\big(\mathsf{E}'(\langle H'^2 \rangle) + (a-1)\mathsf{E}'(\langle H' \rangle^2) - a(\mathsf{E}'\langle H' \rangle)^2\big) \ .$$

Since $\mathsf{E}'(\langle H' \rangle^2) \ge (\mathsf{E}'\langle H' \rangle)^2$ and $a > 1$, we have

$$\begin{aligned}
\varphi''(x) &\ge \frac{c_N^2}{N}(\mathsf{E}'\langle H'^2 \rangle - (\mathsf{E}'\langle H' \rangle)^2) \\
&= \frac{c_N^2}{N}(\nu(H'^2) - \nu(H')^2) \\
&= \frac{c_N^2}{N}\nu((H' - \nu(H'))^2) \\
&= N c_N^2 \nu\left(\left(\frac{H'}{N} - \nu\left(\frac{H'}{N}\right)\right)^2\right) \ ,
\end{aligned} \tag{12.72}$$

and consequently

$$Nc_N^2 \int_0^1 \nu\left(\left(\frac{H'}{N} - \nu\left(\frac{H'}{N}\right)\right)^2\right) \mathrm{d}x \le \int_0^1 \varphi''(x)\mathrm{d}x$$
$$= \varphi'(1) - \varphi'(0) \ . \tag{12.73}$$

Now,

$$\nu\left(\frac{H'}{N}\right) = \frac{1}{N\mathsf{E}Z_N^a}\mathsf{E}\left(\left(\sum_{\boldsymbol{\tau}} \exp(-H^x(\boldsymbol{\tau}))\right)^{a-1} \sum_{\boldsymbol{\sigma}} H'(\boldsymbol{\sigma})\exp(-H^x(\boldsymbol{\sigma}))\right) .$$

Integration by parts in the r.v.s $H'(\boldsymbol{\sigma})$ yields the identity

$$\nu\left(\frac{H'}{N}\right) = xc_N(1 + (a-1)\nu(R_{1,2})) \ , \tag{12.74}$$

so that by (12.71) we have $\varphi'(0) = 0$ and $\varphi'(1) = c_N\nu(H'/N) \le ac_N^2$ since $|R_{1,2}| \le 1$. Combining with (12.73) proves the result. □

Proof of Theorem 12.5.1. Consider

$$f(x) = \nu\left(\frac{H'(\boldsymbol{\sigma}^1)}{N}R_{1,2}\right) - \nu\left(\frac{H'}{N}\right)\nu(R_{1,2}) \ .$$

Since $|R_{1,2}| \le 1$, the Cauchy-Schwarz inequality implies

$$|f(x)| \le \left|\nu\left(\left(\frac{H'(\boldsymbol{\sigma}^1)}{N} - \nu\left(\frac{H'}{N}\right)\right)R_{1,2}\right)\right| \le \nu\left(\left(\frac{H'}{N} - \nu\left(\frac{H'}{N}\right)\right)^2\right)^{1/2}$$

and therefore

$$\int_0^1 |f(x)|\mathrm{d}x \le \sqrt{\frac{a}{N}} \ , \tag{12.75}$$

using (12.68) and the Cauchy-Schwarz inequality again. To conclude, we prove that

$$f(x) \ge xc_N \min(1, a-1)(\nu(R_{1,2}^2) - \nu(R_{1,2})^2) \ . \tag{12.76}$$

We have

$$N\mathsf{E}Z_N^a\nu\left(\frac{H'(\boldsymbol{\sigma}^1)}{N}R_{1,2}\right)$$
$$= \mathsf{E}\left(\left(\sum_{\boldsymbol{\tau}} \exp(-H^x(\boldsymbol{\tau}))\right)^{a-1} \sum_{\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2} H'(\boldsymbol{\sigma}^1)R_{1,2}\exp(-H^x(\boldsymbol{\sigma}^1) - H^x(\boldsymbol{\sigma}^2))\right) .$$

Integration by parts in the r.v.s $H'(\boldsymbol{\sigma})$ yields, after a few lines of computation,

$$\nu\left(\frac{H'(\boldsymbol{\sigma}^1)}{N}R_{1,2}\right) = xc_N\left(\nu(R_{1,2}) + \nu(R_{1,2}^2) + (a-2)\nu(R_{1,2}R_{1,3})\right) ,$$

and combining with (12.74) we get

$$f(x) = xc_N\big(\nu(R_{1,2}^2) + (1-a)\nu(R_{1,2})^2 + (a-2)\nu(R_{1,2}R_{1,3})\big) \ . \qquad (12.77)$$

Next, we prove that

$$\langle R_{1,2}^2\rangle \geq \langle R_{1,2}R_{1,3}\rangle \geq \langle R_{1,2}\rangle^2 \ .$$

The first inequality follows from the Cauchy-Schwarz inequality, the second from

$$\begin{aligned}\int R_{1,2}R_{1,3}\mathrm{d}G(\boldsymbol{\sigma}^1)\mathrm{d}G(\boldsymbol{\sigma}^2)\mathrm{d}G(\boldsymbol{\sigma}^3) &= \int\left(\int R(\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2)\mathrm{d}G(\boldsymbol{\sigma}^2)\right)^2\mathrm{d}G(\boldsymbol{\sigma}^1)\\ &\geq \left(\int R_{1,2}\mathrm{d}G(\boldsymbol{\sigma}^1)\mathrm{d}G(\boldsymbol{\sigma}^2)\right)^2 \ .\end{aligned}$$

Taking expectation and using again the Cauchy-Schwarz inequality yields

$$\nu(R_{1,2}^2) \geq \nu(R_{1,2}R_{1,3}) \geq \nu(R_{1,2})^2 \ .$$

When $a \leq 2$, we use that $1 = a - 1 + 2 - a$ and

$$\nu(R_{1,2}^2) \geq (a-1)\nu(R_{1,2}^2) + (2-a)\nu(R_{1,2}R_{1,3}) \ ,$$

and thus by (12.77) that $f(x) \geq xc_N(a-1)(\nu(R_{1,2}^2) - \nu(R_{1,2})^2)$. When $a \geq 2$, we use that $\nu(R_{1,2}R_{1,3}) \geq \nu(R_{1,2})^2$ to get

$$f(x) \geq xc_N(\nu(R_{1,2}^2) - \nu(R_{1,2})^2) \ . \qquad \square$$

13. The High-Temperature Region of the SK Model

13.1 The Poisson-Dirichlet Distribution and the REM

Consider a parameter $0 < m < 1$. Throughout this chapter we denote by

$$\mu_m \quad \text{the positive measure on } \mathbb{R}^+ \text{ of density } x^{-m-1} \text{ with respect to Lebesgue's measure.} \tag{13.1}$$

Since $m > 0$, for every $\varepsilon > 0$ we have

$$\mu_m([\varepsilon, \infty]) < \infty \,, \tag{13.2}$$

and since $m < 1$,

$$\int_0^1 x \mathrm{d}\mu_m(x) = \int_0^1 x^{-m} \mathrm{d}x < \infty \,. \tag{13.3}$$

Consider a Poisson point process Π of intensity measure μ_m (see Section A.9). This is simply a random countable subset of $\mathbb{R}^+$ such that for every (measurable) subset A of $\mathbb{R}^+$ with $\mu_m(A) < \infty$ the r.v. $\mathrm{card} \Pi \cap A$ is Poisson of expectation $\mu_m(A)$; moreover $\mathrm{card} \Pi \cap A$ and $\mathrm{card} \Pi \cap B$ are independent r.v.s when $A \cap B = \emptyset$. Some probabilists like to think of Π as the set of jumps of a m-stable subordinator (i.e. with a Lévy measure μ_m), but we shall not use this point of view here.

By (13.2) for every $\varepsilon > 0$ the set $\Pi \cap [\varepsilon, \infty)$ is a.s. finite, so we can enumerate the points of Π as a non-increasing sequence $(u_\alpha)_{\alpha \geq 1}$. The notation α for an integer is somewhat unusual; we adopt it in reference to physics (where the letter α indexes the "pure states" of a spin system). We have

$$\mathsf{E}\Big(\sum_{\alpha \geq 1} u_\alpha \mathbf{1}_{\{u_\alpha \leq 1\}}\Big) = \int_0^1 x \mathrm{d}\mu_m(x) < \infty \,, \tag{13.4}$$

and since $\mathsf{E}\mathrm{card}\{\alpha \,;\, u_\alpha \geq 1\} = \mu([1, \infty])$ is finite, the sum $S = \sum_{\gamma \geq 1} u_\gamma$ is finite a.s. We can then consider the sequence

$$v_\alpha = \frac{u_\alpha}{S} = \frac{u_\alpha}{\sum_{\gamma \geq 1} u_\gamma} \,.$$

M. Talagrand, *Mean Field Models for Spin Glasses*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys in Mathematics 55, DOI 10.1007/978-3-642-22253-5_6,

This is a non-negative, non-increasing random sequence that sums to 1. The notations u_α and v_α will be used throughout this chapter, and should be carefully distinguished.

Let us denote by $\mathcal{S}$ the set of non-negative, non-increasing sequences with sum at most 1. This set is provided with a natural topology (the weakest that makes the maps $(x_\alpha)_{\alpha\geq 1} \mapsto x_\gamma$ continuous for each γ) for which it is compact. The law of $(v_\alpha)_{\alpha\geq 1}$ is a probability measure Λ_m on $\mathcal{S}$ that we call the Poisson-Dirichlet distribution of parameter m. The paper [83], where references to earlier work can be found, explores the multiple facets of this object. In that paper the Poisson-Dirichlet distribution depends on 2 parameters, and with its notation we have $\Lambda_m = PD(m,0)$. Let us also mention that this distribution was in fact introduced by J.F.C. Kingman [55]. The "second parameter" of the Poisson-Dirichlet distribution occurs naturally in spin glasses in the setting of Section 12.5, but here we will only consider the case where this second parameter is 0, and this motivates the use of another notation for $PD(m,0)$.

The Poisson-Dirichlet distribution seems to be fundamental in the study of the low temperature phase of spin glass models. We will motivate its introduction by a heuristic study of Derrida's Random Energy Model (REM). The REM is a toy model, where the space of configurations is Σ_N as usual, and where the Hamiltonian $H_N(\boldsymbol{\sigma})$ is such that the family $(H_N(\boldsymbol{\sigma}))_{\boldsymbol{\sigma}}$ is an i.i.d. sequence of Gaussian r.v.s with

$$\mathsf{E}H_N^2(\boldsymbol{\sigma}) = N \,. \tag{13.5}$$

This is a toy model, because we have removed all correlations between the energies of different configurations, an over simplification. Probabilists *do* understand i.i.d. r.v.s so it is really nothing more than an exercise to rigorously prove all the statements we will make below. As these statements serve only as motivation, there seems to be no point to give their proofs (which can be found e.g. in [103]).

Given a measurable subset A of $\mathbb{R}$, we have

$$\mathsf{P}(H_N(\boldsymbol{\sigma}) \in A) = \frac{1}{\sqrt{2\pi N}} \int_A \exp\left(-\frac{t^2}{2N}\right) \mathrm{d}t \,. \tag{13.6}$$

Let us define the number a_N by

$$\exp\left(-\frac{a_N^2}{2N}\right) = 2^{-N}\sqrt{2\pi N} \,.$$

Thus, by change of variables,

$$\begin{aligned}\mathsf{P}(H_N(\boldsymbol{\sigma}) - a_N \in A) &= \frac{1}{\sqrt{2\pi N}} \int_{a_N + A} \exp\left(-\frac{t^2}{2N}\right) \mathrm{d}t \\ &= \frac{1}{\sqrt{2\pi N}} \int_A \exp\left(-\frac{(t+a_N)^2}{2N}\right) \mathrm{d}t\end{aligned}$$

$$= 2^{-N} \int_A \exp\left(-\frac{t^2}{2N} - \frac{a_N t}{N}\right) \mathrm{d}t \; .$$

Now, $a_N/N \simeq \sqrt{2\log 2}$, so that when A is bounded below we get

$$\mathsf{P}(H_N(\boldsymbol{\sigma}) - a_N \in A) \simeq 2^{-N} \int_A \exp(-t\sqrt{2\log 2})\mathrm{d}t \; .$$

The r.v.s $H_N(\boldsymbol{\sigma}) - a_N$ are independent as $\boldsymbol{\sigma}$ varies. Consequently the r.v. $\mathrm{card}\{\boldsymbol{\sigma}; H_N(\boldsymbol{\sigma}) - a_N \in A\}$ should approximately be Poisson with expectation $\int_A \exp(-t\sqrt{2\log 2})\mathrm{d}t$. Consider the sequence $(h_\alpha)_{1\le\alpha\le 2^N}$ which is the non-increasing reordering of the 2^N numbers $(H_N(\boldsymbol{\sigma}) - a_N)_{\boldsymbol{\sigma}}$. Then, given any number b, for large N, the trace of the sequence (h_α) on the interval $[b,\infty)$ should look like a Poisson point process of intensity measure $\exp(-x\sqrt{2\log 2})\mathrm{d}x$ on this interval i.e. $h_\alpha \simeq d_\alpha$, where $(d_\alpha)_{\alpha\ge 1}$ denotes a non-increasing enumeration of the realization of a Poisson point process of intensity measure $\exp(-x\sqrt{2\log 2})\mathrm{d}x$. In particular for α not too large we should have

$$h_\alpha \simeq d_\alpha \simeq c_\alpha \; ,$$

where c_α is defined by $\int_{c_\alpha}^\infty \exp(-x\sqrt{2\log 2})\mathrm{d}x = \alpha$, so that

$$c_\alpha = -\frac{1}{\sqrt{2\log 2}} \log(\alpha\sqrt{2\log 2}) \; .$$

Thus we expect that whenever $\beta > \sqrt{2\log 2}$, the sequence $(\exp\beta h_\alpha)_{\alpha\ge 1}$ is "summable", in the sense that the sum of all the terms is essentially the sum of the first few terms, independently of the value of N, and thus the Gibbs measure should be determined by the first few terms of the sequence (h_α). That is, for $\beta > \sqrt{2\log 2}$, the Gibbs weights for the Hamiltonian H_N at given temperature β, i.e. the numbers $G(\{\boldsymbol{\sigma}\}) = Z_N^{-1}\exp(-\beta H_N(\boldsymbol{\sigma}))$, where $Z_N = \sum_{\boldsymbol{\sigma}} \exp(-\beta H_N(\boldsymbol{\sigma}))$ should resemble, once rearranged in non-increasing order, the sequence

$$\left(\frac{\exp\beta d_\alpha}{\sum_\gamma \exp\beta d_\gamma}\right)_{\alpha\ge 1} \; . \tag{13.7}$$

We turn to the computation of the distribution of this sequence. A very useful property of Poisson point process is as follows. Consider a positive measure μ, a continuous map f, and the positive measure $f(\mu)$ given by $f(\mu)(A) = \mu(f^{-1}(A))$. Assume that $f(\mu)$ has no atoms. Then if Π a Poisson point process of intensity measure μ, $f(\Pi)$ is a Poisson point process of intensity measure $f(\mu)$. When μ has density $\exp(-x\sqrt{2\log 2})$ and $f(x) = c\exp\beta x$ (for some $c > 0$), then $f(\mu)$ has density $\beta^{-1}c^m x^{-m-1}\mathrm{d}x$ for $m = \sqrt{2\log 2}/\beta$. This is because one has

$$f(\mu)(A) = \mu(f^{-1}(A)) = \int_{f^{-1}(A)} \exp(-x\sqrt{2\log 2})\mathrm{d}x = c^m\beta^{-1}\int_A y^{-m-1}\mathrm{d}y \; ,$$

as is seen by setting $x = f^{-1}(y) = \beta^{-1}\log(y/c)$. Therefore, when $c = \beta^{-1/m}$, the sequence $(c\exp\beta d_\alpha)_{\alpha\geq 1}$ is, in distribution, the non-increasing enumeration of a Poisson point process of intensity measure μ_m, where $\mathrm{d}\mu_m = x^{-m-1}\mathrm{d}x$. Consequently, the sequence (13.7) has distribution Λ_m; one should therefore expect that this is the limiting distribution of the Gibbs' weights for the REM. One should also expect that the distribution Λ_m will inherit some of the remarkable properties of Gibbs' measure, such as the Ghirlanda-Guerra identities, an idea that will be explored in Chapter 16. While probabilists had discovered long ago many remarkable properties of Λ_m, it is striking that the above intuition does shed some new light on this object.

A useful property of Poisson point process is as follows. Consider a positive measure μ (on $\mathbb{R}$) and a probability measure η on $\mathbb{R}^n$ and consider the positive measure $\mu\otimes\eta$ on $\mathbb{R}\times\mathbb{R}^n$. Consider a random sequence (u_α), and assume that the set $\{u_\alpha\ ;\ \alpha\geq 1\}$ is distributed like a Poisson point process of intensity measure μ. Consider an independent identically distributed sequence (U_α) distributed like η. Then the set consisting of the points $(u_\alpha, U_\alpha)_{\alpha\geq 1}$ is distributed like a Poisson point process on $\mathbb{R}\times\mathbb{R}^n$ of intensity measure $\mu\otimes\eta$. (one may, if one wishes, assume that the sequence (u_α) is non-increasing, but this is not necessary.)

The measure μ_m of (13.1) has the following remarkable stability property.

Lemma 13.1.1. *Consider a probability measure ν on $\mathbb{R}^+$ and assume that $c_\nu := \int y^m \mathrm{d}\nu(y) < \infty$. Then the image of $\mu_m\otimes\nu$ under the map $(x,y)\mapsto xy$ is $c_\nu\mu_m$.*

Proof. By definition the image μ' of $\mu_m\otimes\nu$ under the map $(x,y)\mapsto xy$ is such that

$$\begin{aligned}\mu'(A) = \int_{\{(x,y)\ ;\ xy\in A\}} \mathrm{d}\mu(x)\mathrm{d}\nu(y) &= \int_{\{xy\in A\}} x^{-m-1}\mathrm{d}x\mathrm{d}\nu(y)\\ &= \int \mathrm{d}\nu(y)\int_{\{xy\in A\}} x^{-m-1}\mathrm{d}x\ .\end{aligned}$$

Making the change of variable $x = y^{-1}z$ yields

$$\int_{\{xy\in A\}} x^{-m-1}\mathrm{d}x = y^m\int_A x^{-m-1}\mathrm{d}x\ ,$$

so that

$$\mu'(A) = \int y^m\mathrm{d}\nu(y)\int_A x^{-m-1}\mathrm{d}x = c_\nu\int_A x^{-m-1}\mathrm{d}x = c_\nu\mu_m(A)\ . \qquad \square$$

Corollary 13.1.2. *Consider a non-increasing enumeration $(u_\alpha)_{\alpha\geq 1}$ of a Poisson point process of intensity measure μ_m, and i.i.d. copies $(V_\alpha)_{\alpha\geq 1}$ of r.v. V with $V > 0$ and $\mathsf{E}V^m = 1$, that are independent of the sequence u_α. Then there exists a random permutation σ such that the sequence $(u_{\sigma(\alpha)}V_{\sigma(\alpha)})_{\alpha\geq 1}$ has the same distribution as the sequence $(u_\alpha)_{\alpha\geq 1}$.*

Proof. The set of points $(u_\alpha, V_\alpha)_{\alpha\geq 1}$ is distributed like a Poisson point process of intensity measure $\mu_m \otimes \nu$, where ν is the law of V, so that $\int y^m \mathrm{d}\nu(y) = 1$ since $\mathsf{E}v^m = 1$. By Lemma 13.1.1, the set $(u_\alpha V_\alpha)_{\alpha\geq 1}$ is distributed like a Poisson point process of intensity measure μ_m, so that it suffices to chose the permutation σ such that the sequence $(u_{\sigma(\alpha)}V_{\sigma(\alpha)})_{\alpha\geq 1}$ is non-increasing. □

Proposition 13.1.3. *Consider a non-increasing enumeration $(u_\alpha)_{\alpha\geq 1}$ of a Poisson point process of intensity measure μ_m. Then if $0 < m' < m$ we have*

$$\mathsf{E}\Big(\sum_{\alpha\geq 1} u_\alpha\Big)^{m'} < \infty\,. \tag{13.8}$$

Consider i.i.d copies $(V_\alpha)_{\alpha\geq 1}$ of a r.v. $V > 0$ with $\mathsf{E}V^m < \infty$, that are independent of the sequence $(u_\alpha)_{\alpha\geq 1}$. Then we have

$$\mathsf{E}\Big(\sum_{\alpha\geq 1} u_\alpha V_\alpha\Big)^{m'} = \mathsf{E}\Big(\sum_{\alpha\geq 1} u_\alpha\Big)^{m'} (\mathsf{E}V^m)^{m'/m} \tag{13.9}$$

and

$$\mathsf{E}\log\sum_{\alpha\geq 1} u_\alpha V_\alpha = \mathsf{E}\log\sum_{\alpha\geq 1} u_\alpha + \frac{1}{m}\log\mathsf{E}V^m\,. \tag{13.10}$$

Proof. Since $m' < 1$ we have

$$\Big(\sum_{\alpha\geq 1} u_\alpha \mathbf{1}_{\{u_\alpha\geq 1\}}\Big)^{m'} \leq \sum_{\alpha\geq 1} u_\alpha^{m'} \mathbf{1}_{\{u_\alpha\geq 1\}}\,.$$

Now we use that if $f \geq 0$, we have $\mathsf{E}\sum_\alpha f(u_\alpha) = \int f \mathrm{d}\mu_m$, as is seen by approximating f by step functions. Therefore, taking expectation in (13.1), we get

$$\mathsf{E}\Big(\sum_{\alpha\geq 1} u_\alpha \mathbf{1}_{\{u_\alpha\geq 1\}}\Big)^{m'} \leq \int_1^\infty x^{m'-1-m}\mathrm{d}x < \infty\,, \tag{13.11}$$

and (13.8) follows from (13.4).

We turn to the proof of (13.9) and (13.10). By homogeneity, we may assume that $\mathsf{E}V^m = 1$. Thus Corollary 13.1.2 implies

$$\mathsf{E}\Big(\sum_{\alpha\geq 1} u_\alpha V_\alpha\Big)^{m'} = \mathsf{E}\Big(\sum_{\alpha\geq 1} u_\alpha\Big)^{m'}$$

and

$$\mathsf{E}\log\sum_{\alpha\geq 1} u_\alpha V_\alpha = \mathsf{E}\log\sum_{\alpha\geq 1} u_\alpha\ . \qquad \square$$

Comment. A really formal proof should show that $\mathsf{E}\big|\log\sum_{\alpha\geq 1} u_\alpha\big| < \infty$. We will elaborate on this later.

Exercise 13.1.4. Prove that

$$\mathsf{E}\Big(\sum_{\alpha\geq 1} u_\alpha\Big)^{m} = \infty\ .$$

(Hint: consider the largest term u_1 of this sum.)

Theorem 13.1.5. *Consider $0 < m < 1$. Consider i.i.d copies $(V_\alpha)_{\alpha\geq 1}$ of a r.v. $V > 0$ with $\mathsf{E}V^m < \infty$, that are independent of a sequence $(v_\alpha)_{\alpha\geq 1}$ of distribution Λ_m. Then*

$$\mathsf{E}\log\sum_{\alpha\geq 1} v_\alpha V_\alpha = \frac{1}{m}\log \mathsf{E}V^m\ . \tag{13.12}$$

One can gain some intuition concerning this fundamental formula as follows. When m is close to 1, the individual weights v_α are very small, and since the r.v.s V_α are independent, the law of large numbers shows that with high probability one has $\sum_{\alpha\geq 1} v_\alpha V_\alpha \simeq \mathsf{E}V$, which means that the left-hand side of (13.12) is nearly $\mathsf{E}V$, and so of course is the right-hand side. On the contrary, when m is close to 0, then $v_1 \simeq 1$ and the remainder of the v_α are nearly 0, which means that the left-hand side of (13.12) is nearly $\mathsf{E}\log V$. The right-hand side is approximately the same quantity since for $m \to 0$ we have

$$\frac{1}{m}\log \mathsf{E}V^m = \frac{1}{m}\log \mathsf{E}e^{m\log V} \simeq \frac{1}{m}\log(1+m\mathsf{E}\log V) \simeq \mathsf{E}\log V\ .$$

Proof of Theorem 13.1.5. Since we may assume that $v_\alpha = u_\alpha/\sum_{\gamma\geq 1} u_\gamma$ where $(u_\alpha)_{\alpha\geq 1}$ is an enumeration of a Poisson point process of intensity measure μ_m, this is simply a reformulation of (13.10). $\square$

Theorem 13.1.6. *Consider $0 < m < 1$. Consider a triple (U, V, W) of r.v.s with $V \geq 1$ and assume that $\mathsf{E}V^m < \infty$, $\mathsf{E}U^2 + \mathsf{E}W^2 < \infty$. Consider independent copies $(U_\alpha, V_\alpha, W_\alpha)$ of this triple, which are independent of a sequence $(v_\alpha)_{\alpha\geq 1}$ with distribution Λ_m. Then we have*

$$\mathsf{E}\,\frac{\sum_{\alpha\geq 1} v_\alpha U_\alpha}{\sum_{\alpha\geq 1} v_\alpha V_\alpha} = \frac{\mathsf{E}\,UV^{m-1}}{\mathsf{E}\,V^m} \tag{13.13}$$

$$\mathsf{E}\,\frac{\sum_{\alpha\geq 1} v_\alpha^2 U_\alpha W_\alpha}{\left(\sum_{\alpha\geq 1} v_\alpha V_\alpha\right)^2} = (1-m)\frac{\mathsf{E}\,UWV^{m-2}}{\mathsf{E}\,V^m} \tag{13.14}$$

$$\mathsf{E}\,\frac{\sum_{\alpha\neq\gamma} v_\alpha v_\gamma U_\alpha W_\gamma}{\left(\sum_{\alpha\geq 1} v_\alpha V_\alpha\right)^2} = m\frac{\mathsf{E}UV^{m-1}\mathsf{E}\,WV^{m-1}}{(\mathsf{E}\,V^m)^2}\ . \tag{13.15}$$

Proof. Assume first that $\varepsilon|U|\leq V$ for some $\varepsilon>0$. We write (13.12) for $V+tU$ when $|t|<\varepsilon$, and we take the derivative at $t=0$ to obtain (13.13). Replacing V by $V+\varepsilon|U|$ and letting $\varepsilon\to 0$ yields (13.13) in full generality.

Using the same method to reduce to the case where $\varepsilon|W|\leq V$, replacing V by $V+tW$ in (13.13) and taking the derivative at $t=0$ yields

$$\mathsf{E}\,\frac{\left(\sum_{\alpha\geq 1} v_\alpha U_\alpha\right)\left(\sum_{\alpha\geq 1} v_\alpha W_\alpha\right)}{\left(\sum_{\alpha\geq 1} v_\alpha V_\alpha\right)^2} = (1-m)\frac{\mathsf{E}UWV^{m-2}}{\mathsf{E}V^m} + m\frac{\mathsf{E}UV^{m-1}\mathsf{E}WV^{m-1}}{(\mathsf{E}V^m)^2}\ . \tag{13.16}$$

Replacing U by rU and W by rW where r is a random sign independent of U,V,W with $\mathsf{P}(r=1)=\mathsf{P}(r=-1)=1/2$ yields (13.14), and combining with (13.16) proves (13.15). □

Of course the previous proof is incomplete: one has to justify the fact that one can differentiate in these infinite series. It is tedious to give the details, but the idea to prove this is completely straightforward. The basic fact is that the number X_k of points u_α that are contained in a given interval $[2^{-k-1},2^{-k}[$ has expected value

$$a_k = \int_{2^{-k-1}}^{2^{-k}} x^{-m-1}\mathrm{d}x\ ,$$

of order 2^{km} and is sharply concentrated around this value, e.g. $\mathsf{P}(|X_k-a_k|\geq a_k/2)\leq \exp(-2^{km}/K)$, where K is independent of k, as is shown in (A.60). It seems safe to assume that any reader having reached this point is either not interested in these details, or else can complete them herself. Therefore there is no point to write them down here.

Taking $U=V=W=1$ in (13.14) yields the fundamental relation

$$\mathsf{E}\sum_{\alpha\geq 1} v_\alpha^2 = 1-m\ . \tag{13.17}$$

13.2 The 1-RSB Bound for the SK Model

We recall the Hamiltonian of the SK model,

$$-H_N(\boldsymbol{\sigma}) = \frac{\beta}{\sqrt{N}} \sum_{i<j} g_{ij}\sigma_i\sigma_j + \sum_{i\le N} h_i\sigma_i \;, \tag{13.18}$$

where (g_{ij}) are i.i.d standard Gaussian r.v.s and (h_i) are i.i.d. r.v.s. For simplicity we will assume throughout this section that the r.v.s. h_i are independent copies of a Gaussian (not necessarily centered) r.v. h and we recall the notation

$$p_N(\beta, h) = \frac{1}{N}\mathsf{E}\log\sum_{\boldsymbol{\sigma}}\exp(-H_N(\boldsymbol{\sigma})) \;.$$

Theorem 13.2.1. *(F. Guerra) Consider* $0 \le q \le q' \le 1$ *and* $0 < m \le 1$. *We set*

$$Y' = \beta z\sqrt{q} + \beta z'\sqrt{q'-q} + h \;, \tag{13.19}$$

where z *and* z' *are independent standard Gaussian r.v.s independent of* h, *and we denote by* E' *expectation in* z' *only. Then*

$$p_N(\beta,h) \le \log 2 + \frac{\beta^2}{4}(1-q')^2 - \frac{\beta^2}{4}m(q'^2-q^2) + \frac{1}{m}\mathsf{E}\log\mathsf{E}'\mathrm{ch}^m Y' \;. \tag{13.20}$$

Taking $q = q'$ one sees that this improves upon (1.72), which we recall below:

$$p_N(\beta,h) \le \mathrm{SK}(\beta,h) = \log 2 + \mathsf{E}\log\mathrm{ch}(\beta z\sqrt{q}+h) + \frac{\beta^2}{4}(1-q)^2 \;. \tag{1.72}$$

The bound (13.20) is the "first step" of a fundamental general bound that we will study in Chapter 14. The acronym 1-RSB stands for "one step of replica-symmetry breaking".

Proof. Consider independent standard Gaussian r.v.s $z_i, z_i', z_{i,\alpha}'$ for $i \le N$ and $\alpha \ge 1$. Consider a sequence $(v_\alpha)_{\alpha\ge 1}$ with Poisson-Dirichlet distribution Λ_m. Assume that these are all independent of each other and of all the other randomness. For $0 \le t \le 1$ we define

$$\varphi(t) = \frac{1}{N}\mathsf{E}\log\sum_{\alpha\ge 1} v_\alpha \sum_{\boldsymbol{\sigma}}\exp(-H_t(\boldsymbol{\sigma},\alpha)) \;, \tag{13.21}$$

where

$$\begin{aligned}-H_t(\boldsymbol{\sigma},\alpha) = {}& \beta\sqrt{\frac{t}{N}}\sum_{i<j} g_{ij}\sigma_i\sigma_j + \beta\sqrt{1-t}\sum_{i\le N}\sigma_i(z_i\sqrt{q} + z_{i,\alpha}'\sqrt{q'-q}) \\ & + \sum_{i\le N} h_i\sigma_i \;.\end{aligned} \tag{13.22}$$

Thus

$$\varphi(1) = p_N(\beta, h) \tag{13.23}$$

and

$$\varphi(0) = \frac{1}{N} \mathsf{E} \log \sum_{\alpha \geq 1} v_\alpha \sum_{\boldsymbol{\sigma}} \exp \sum_{i \leq N} \sigma_i (\beta z_i \sqrt{q} + \beta z'_{i,\alpha} \sqrt{q' - q} + h_i) \,. \tag{13.24}$$

To compute this, we use (13.12) given the r.v.s z_i and h_i so that

$$\varphi(0) = \frac{1}{Nm} \mathsf{E} \log \mathsf{E}' \left(\sum_{\boldsymbol{\sigma}} \exp \sum_{i \leq N} \sigma_i (\beta z_i \sqrt{q} + \beta z'_i \sqrt{q' - q} + h_i) \right)^m ,$$

where E' denotes expectation in the r.v.s z'_i only. Now

$$\sum_{\boldsymbol{\sigma}} \exp \sum_{i \leq N} \sigma_i (\beta z_i \sqrt{q} + \beta z'_i \sqrt{q' - q} + h_i) = 2^N \prod_{i \leq N} \mathrm{ch}(\beta z_i \sqrt{q} + \beta z'_i \sqrt{q' - q} + h_i)$$

so that, using independence,

$$\begin{aligned} &\mathsf{E}' \left(\sum_{\boldsymbol{\sigma}} \exp \sum_{i \leq N} \sigma_i (\beta z_i \sqrt{q} + \beta z'_i \sqrt{q' - q} + h_i) \right)^m \\ &= 2^{Nm} \prod_{i \leq N} \mathsf{E}' \mathrm{ch}^m (\beta z_i \sqrt{q} + \beta z'_i \sqrt{q' - q} + h_i) \,, \end{aligned}$$

and therefore

$$\begin{aligned} \varphi(0) &= \log 2 + \frac{1}{Nm} \mathsf{E} \log \prod_{i \leq N} \mathsf{E}' \mathrm{ch}^m (\beta z_i \sqrt{q} + \beta z'_i \sqrt{q' - q} + h_i) \\ &= \log 2 + \frac{1}{m} \mathsf{E} \log \mathsf{E}' \mathrm{ch}^m Y' \,. \end{aligned} \tag{13.25}$$

We turn to the computation of φ'. Given a function f on $\Sigma_N \times \mathbb{N}$, we define

$$\langle f \rangle_t = \frac{1}{Z(t)} \sum_{\boldsymbol{\sigma}, \alpha} v_\alpha f(\boldsymbol{\sigma}, \alpha) \exp(-H_t(\boldsymbol{\sigma}, \alpha)) \,,$$

where $Z(t)$ is the normalizing factor. Differentiation of (13.21) proves that

$$2\varphi'(t) = \mathrm{I} + \mathrm{II} + \mathrm{III} \tag{13.26}$$

where

$$\begin{aligned} \mathrm{I} &= \frac{\beta}{N\sqrt{t}} \mathsf{E} \left\langle \frac{1}{\sqrt{N}} \sum_{i<j} g_{ij} \sigma_i \sigma_j \right\rangle_t \\ \mathrm{II} &= -\frac{\beta}{N\sqrt{1-t}} \mathsf{E} \left\langle \sum_{i \leq N} \sigma_i z_i \sqrt{q} \right\rangle_t \\ \mathrm{III} &= -\frac{\beta}{N\sqrt{1-t}} \mathsf{E} \left\langle \sum_{i \leq N} \sigma_i z'_{i,\alpha} \sqrt{q' - q} \right\rangle_t \,. \end{aligned}$$

We want to integrate by parts in these formulas; for this we need to use two replicas of the configuration $(\boldsymbol{\sigma}, \alpha)$. We denote them by $(\boldsymbol{\sigma}^1, \alpha)$ and $(\boldsymbol{\sigma}^2, \gamma)$. The quantity $R_{1,2}$ is as usual $N^{-1} \sum_{i \le N} \sigma_i^1 \sigma_i^2$. Integrating by parts we obtain, denoting again $\langle \cdot \rangle_t$ an average on two replicas,

$$\begin{aligned} \mathrm{I} &= \frac{\beta^2}{2}(1 - \mathsf{E}\langle R_{1,2}^2 \rangle_t) \\ \mathrm{II} &= -\beta^2 (q - q\mathsf{E}\langle R_{1,2} \rangle_t) \\ \mathrm{III} &= -\beta^2 \big((q' - q) - (q' - q)\mathsf{E}\langle R_{1,2} \mathbf{1}_{\{\alpha = \gamma\}} \rangle_t\big) \ . \end{aligned}$$

We observe that

$$\mathsf{E}\langle R_{1,2} \rangle_t = \mathsf{E}\langle R_{1,2} \mathbf{1}_{\{\alpha \neq \gamma\}} \rangle_t + \mathsf{E}\langle R_{1,2} \mathbf{1}_{\{\alpha = \gamma\}} \rangle_t \ ,$$

and therefore

$$\mathrm{II} + \mathrm{III} = -\beta^2 (q' - q\mathsf{E}\langle R_{1,2} \mathbf{1}_{\{\alpha \neq \gamma\}} \rangle_t - q'\mathsf{E}\langle R_{1,2} \mathbf{1}_{\{\alpha = \gamma\}} \rangle_t) \ .$$

Using that

$$\mathsf{E}\langle R_{1,2}^2 \rangle_t = \mathsf{E}\langle R_{1,2}^2 \mathbf{1}_{\{\alpha \neq \gamma\}} \rangle_t + \mathsf{E}\langle R_{1,2}^2 \mathbf{1}_{\{\alpha = \gamma\}} \rangle_t \ ,$$

we get the identity

$$\begin{aligned} 2\varphi'(t) = \frac{\beta^2}{2} \big((1 - 2q') &+ q^2 \mathsf{E}\langle \mathbf{1}_{\{\alpha \neq \gamma\}} \rangle_t + q'^2 \mathsf{E}\langle \mathbf{1}_{\{\alpha = \gamma\}} \rangle_t \\ &- \mathsf{E}\langle (R_{1,2} - q)^2 \mathbf{1}_{\{\alpha \neq \gamma\}} \rangle_t - \mathsf{E}\langle (R_{1,2} - q')^2 \mathbf{1}_{\{\alpha = \gamma\}} \rangle_t\big) \ , \end{aligned} \tag{13.27}$$

and thus

$$2\varphi'(t) \le \frac{\beta^2}{2} \big((1 - 2q') + q^2 \mathsf{E}\langle \mathbf{1}_{\{\alpha \neq \gamma\}} \rangle_t + q'^2 \mathsf{E}\langle \mathbf{1}_{\{\alpha = \gamma\}} \rangle_t\big) \ . \tag{13.28}$$

Suppose that we can prove

$$\mathsf{E}\langle \mathbf{1}_{\{\alpha = \gamma\}} \rangle_t = 1 - m \ . \tag{13.29}$$

Then (13.28) yields

$$2\varphi'(t) \le \frac{\beta^2}{2} \big((1 - q')^2 - m(q'^2 - q^2)\big) \ ,$$

which combined with (13.23) and (13.24) completes the proof. Finally, to prove (13.29) we simply use (13.14) for $U_\alpha = V_\alpha = W_\alpha = \sum_{\boldsymbol{\sigma}} \exp(-H_t(\boldsymbol{\sigma}, \alpha))$ given the randomness of g_{ij}, h_i and z_i. □

13.3 Toninelli's Theorem

Throughout this section q is the solution of (1.73) i.e.

$$q = \mathsf{E}\mathrm{th}^2(\beta z\sqrt{q}+h)\ .$$

We recall the notation

$$\mathrm{SK}(\beta,h) = \log 2 + \frac{\beta^2}{4}(1-q)^2 + \mathsf{E}\log\mathrm{ch}(\beta z\sqrt{q}+h)\ . \tag{13.30}$$

We denote by $\mathcal{P}_1(\beta,h)$ the infimum of the right-hand side of (13.20) over all choices of q' and m, that is

$$\mathcal{P}_1(\beta,h) = \inf_{m,q'} \Phi(m,q')\ , \tag{13.31}$$

where the infimum is over $q \le q' \le 1$ and $0 < m \le 1$ and where

$$\Phi(m,q') = \log 2 + \frac{\beta^2}{4}(1-q')^2 - \frac{\beta^2}{4}m(q'^2-q^2) + \frac{1}{m}\mathsf{E}\log\mathsf{E}'\mathrm{ch}^m Y'\ . \tag{13.32}$$

The dependence of Φ on β and h is kept implicit. The choice $q = q'$ proves that

$$\mathcal{P}_1(\beta,h) \le \mathrm{SK}(\beta,h)\ , \tag{13.33}$$

and (13.20) implies

$$p_N(\beta,h) \le \mathcal{P}_1(\beta,h)\ . \tag{13.34}$$

Theorem 13.3.1. (F. Toninelli [114]) *Beyond the A-T line, i.e. if*

$$\beta^2\mathsf{E}\frac{1}{\mathrm{ch}^4(\beta z\sqrt{q}+h)} > 1\ , \tag{13.35}$$

we have

$$\mathcal{P}_1(\beta,h) < \mathrm{SK}(\beta,h)\ . \tag{13.36}$$

Consequently from (13.20) we have

$$\lim_{N\to\infty} p_N(\beta,h) < \mathrm{SK}(\beta,h)\ .$$

Definition 13.3.2. *The* high-temperature region *of the SK model as the region of* (β,h) *where*

$$\lim_{N\to\infty} p_N(\beta,h) = \mathrm{SK}(\beta,h)\ .$$

Then we may express Toninelli's theorem as follows: "the high-temperature region is entirely on the high-temperature side of the A-T line".

Throughout the proof of Theorem 13.3.1 we write

$$\begin{aligned} Y &= \beta z\sqrt{q} + h \\ Y' &= \beta z\sqrt{q} + \beta z'\sqrt{q'-q} + h = Y + \beta z'\sqrt{q'-q}\,. \end{aligned}$$

Lemma 13.3.3. *For any $q \le q' \le 1$ we have*

$$\Phi(1,q') = \mathrm{SK}(\beta,h)\,. \tag{13.37}$$

Proof. The identity

$$\mathsf{E}'\mathrm{ch}Y' = \exp\frac{\beta^2}{2}(q'-q)\mathrm{ch}Y \tag{13.38}$$

implies

$$\begin{aligned} \Phi(1,q') &= \log 2 + \frac{\beta^2}{4}(1-q')^2 - \frac{\beta^2}{4}(q'^2-q^2) + \frac{\beta^2}{2}(q'-q) + \mathsf{E}\log\mathrm{ch}Y \\ &= \log 2 + \frac{\beta^2}{4}(1-q)^2 + \mathsf{E}\log\mathrm{ch}Y = \mathrm{SK}(\beta,h)\,, \end{aligned}$$

and this concludes the proof. □

Lemma 13.3.4. *Let us define*

$$V(q') = V(\beta,h,q') = \frac{\partial\Phi}{\partial m}(m,q')\Big|_{m=1}\,. \tag{13.39}$$

Then

$$V(q') = -\frac{\beta^2}{4}(q'^2-q^2) - \mathsf{E}\log\mathsf{E}'\mathrm{ch}Y' + \mathsf{E}\frac{\mathsf{E}'\log(\mathrm{ch}Y')\mathrm{ch}Y'}{\mathsf{E}'\mathrm{ch}Y'} \tag{13.40}$$

$$V'(q') = \frac{\beta^2}{2}\left(\mathsf{E}\frac{\mathsf{E}'\mathrm{sh}^2Y'\mathrm{ch}^{-1}Y'}{\mathsf{E}'\mathrm{ch}Y'} - q'\right) \tag{13.41}$$

$$V(q) = V'(q) = 0 \tag{13.42}$$

$$V''(q) = \frac{\beta^2}{2}\left(\beta^2\mathsf{E}\frac{1}{\mathrm{ch}^4 Y} - 1\right)\,. \tag{13.43}$$

One observes that V'' brings in the A-T criteria, that was obtained in Section 1.8 in an apparently unrelated manner. Many more facts which look like striking coincidences will occur in the next chapter. Of course, the author does not believe that they are mere coincidences, but rather that there is some underlying structure yet to be understood.

Proof of Theorem 13.3.1. Since $V(q) = V'(q) = 0$ by (13.42) and $V''(q) > 0$ by (13.43), there exists $q' > q$ such that $V(q') > 0$. Thus there exists $0 < m < 1$ such that

$$\Phi(m,q') < \Phi(1,q') = \mathrm{SK}(\beta,h)$$

using (13.37). □

Research Problem 13.3.5. If

$$\beta^2 \mathsf{E}\frac{1}{\mathrm{ch}^4 Y} < 1 ,$$

is it true that $\mathcal{P}_1(\beta,h) = \mathrm{SK}(\beta,h)$?

To solve the above problem we must in particular show that $V(q') \le 0$ for all $q < q' \le 1$. The difficulty is to control $V(q')$ for q' far from q. For q' close to q, it follows from (13.42) and (13.43) that $V(q') < 0$.

Lemma 13.3.6. *Consider*

$$W(q') = \mathsf{E}\frac{\mathsf{E}'\mathrm{sh}^2 Y'\mathrm{ch}^{-1}Y'}{\mathsf{E}'\mathrm{ch}Y'} - q' . \tag{13.44}$$

Then

$$W(q) = 0 \tag{13.45}$$

$$W'(q) = \beta^2 \mathsf{E}\frac{1}{\mathrm{ch}^4 Y} - 1 \tag{13.46}$$

$$W''(q) \le 4\beta^4 . \tag{13.47}$$

Proof. We use (13.38) to write

$$W(q') = \exp\left(-\frac{\beta^2}{2}(q'-q)\right)\mathsf{E}\frac{\mathsf{E}'\mathrm{sh}^2 Y'\mathrm{ch}^{-1}Y'}{\mathrm{ch}Y} - q' .$$

When $q' = q$ we have $Y' = Y$, that does not depend on z', so that

$$\frac{\mathsf{E}'\mathrm{sh}^2 Y'\mathrm{ch}^{-1}Y'}{\mathrm{ch}Y} = \mathrm{th}^2 Y ,$$

and therefore we have $W(q) = 0$ since $q = \mathsf{E}\mathrm{th}^2 Y$. For a function f, integration by parts yields

$$\frac{\mathrm{d}}{\mathrm{d}q'}\mathsf{E}' f(Y') = \mathsf{E}'\frac{\beta z'}{2\sqrt{q'-q}} f'(Y') = \frac{\beta^2}{2}\mathsf{E}' f''(Y') , \tag{13.48}$$

and using this for $f(x) = \mathrm{sh}^2 x\,\mathrm{ch}^{-1} x$, we get

$$W'(q') = \frac{\beta^2}{2}\exp\left(-\frac{\beta^2}{2}(q'-q)\right)\mathsf{E}\frac{\mathsf{E}'(f''(Y') - f(Y'))}{\mathrm{ch}Y} - 1 .$$

Now $f(x) = \mathrm{ch}x - \mathrm{ch}^{-1}x$, so that $f'(x) = \mathrm{sh}x + \mathrm{sh}x\mathrm{ch}^{-2}x$, $f''(x) = \mathrm{ch}x + \mathrm{ch}^{-1}x - 2\mathrm{sh}^2x\mathrm{ch}^{-3}x$, and hence $f''(x) - f(x) = 2\mathrm{ch}^{-3}x$. It follows that

$$W'(q') = \beta^2 \exp\left(-\frac{\beta^2}{2}(q'-q)\right) \mathsf{E}\frac{\mathsf{E}'\mathrm{ch}^{-3}Y'}{\mathrm{ch}Y} - 1 \ ,$$

which proves (13.46). Finally, repeating this procedure with now $f(x) = \mathrm{ch}^{-3}x$, so that $f'(x) = -3\mathrm{sh}x\mathrm{ch}^{-4}x$, $f''(x) = -3\mathrm{ch}^{-3}x + 12\mathrm{sh}^2x\mathrm{ch}^{-5}x$, and hence $f''(x) - f(x) = 8\mathrm{ch}^{-3}x - 12\mathrm{ch}^{-5}x$, we get

$$W''(q) = 2\beta^4 \exp\left(-\frac{\beta^2}{2}(q'-q)\right) \mathsf{E}\frac{\mathsf{E}'(2\mathrm{ch}^{-3}Y' - 3\mathrm{ch}^{-5}Y')}{\mathrm{ch}Y} \ ,$$

from which (13.47) follows. □

Proof of Lemma 13.3.4. Straightforward differentiation of (13.32) yields (13.40). When $q' = q$, we have $Y' = Y$, that does not depend on z', so that the expectation E' disappears and $V(q) = 0$. Using (13.38) twice, we obtain the formula

$$\begin{aligned} V(q') = &-\frac{\beta^2}{4}(q'^2 - q^2) - \frac{\beta^2}{2}(q'-q) - \mathsf{E}\log\mathrm{ch}Y \\ &+ \exp\left(-\frac{\beta^2}{2}(q'-q)\right) \mathsf{E}\frac{\mathsf{E}'(\mathrm{ch}Y')\log\mathrm{ch}Y'}{\mathrm{ch}Y} \ . \end{aligned} \tag{13.49}$$

Now, using (13.48) again, this time for $f(x) = \mathrm{ch}x\log\mathrm{ch}x$, we get

$$V'(q') = -\frac{\beta^2}{2}q' - \frac{\beta^2}{2} + \frac{\beta^2}{2}\exp\left(-\frac{\beta^2}{2}(q'-q)\right) \mathsf{E}\frac{\mathsf{E}'(f''(Y') - f(Y'))}{\mathrm{ch}Y} \ . \tag{13.50}$$

Now $f'(x) = \mathrm{sh}x + \mathrm{sh}x\log\mathrm{ch}x$, $f''(x) = \mathrm{ch}x + \mathrm{ch}x\log\mathrm{ch}x + \mathrm{sh}^2x\mathrm{ch}^{-1}x$ so that $f''(x) - f(x) = \mathrm{ch}x + \mathrm{sh}^2x\mathrm{ch}^{-1}x$, and, using (13.38),

$$V'(q') = \frac{\beta^2}{2}\left(\mathsf{E}\frac{\mathsf{E}'(\mathrm{ch}Y' + \mathrm{sh}^2Y'\mathrm{ch}^{-1}Y')}{\mathsf{E}'\mathrm{ch}Y'} - q' - 1\right) \ .$$

This proves (13.41). Thus $V'(q) = \beta^2(\mathsf{E}\mathrm{th}^2Y - q)/2 = 0$ and (13.43) follows from Lemma 13.3.6. □

13.4 Overview of Proof

Let us recall that we defined the high-temperature region of the SK model as the set of parameters (β, h) such that

$$\lim_{N\to\infty} p_N(\beta, h) = \mathrm{SK}(\beta, h) \ , \tag{13.51}$$

and that we defined the quantity $\mathcal{P}_1(\beta, h)$ as the infimum of the right-hand side of (13.20) over all choices of q, q' and m. The goal of the rest of this chapter is to prove the following.

Theorem 13.4.1. *The high-temperature region of the SK model is the region where*

$$\mathrm{SK}(\beta, h) = \mathcal{P}_1(\beta, h) . \tag{13.52}$$

It is obvious from (13.34) that if (13.51) holds one must have $\mathrm{SK}(\beta, h) \le \mathcal{P}_1(\beta, h)$ and therefore by (13.33) that $\mathrm{SK}(\beta, h) = \mathcal{P}_1(\beta, h)$. The converse is the content of the theorem.

Toninelli's theorem proves that the high-temperature region of the SK model is entirely located inside "high-temperature side of the A-T line". Therefore it is natural to ask whether the high-temperature region of the SK model is exactly the "high-temperature side of the A-T line". Theorem 13.4.1 reduces this question to Research Problem 13.3.5, which is strictly a question of *calculus*. Numerical evidence seems to show that the answer is yes. Whether or not this is the case must be viewed as an *incidental question* rather than a central one. Indeed D. Panchenko [65] has discovered that for a similar model (where the spins can take continuous values) the answer to the corresponding question is no. Therefore the fruitful characterization of the high-temperature region is most probably the one given in Theorem 13.4.1.

Our approach to Theorem 13.4.1 is based on (1.66), that we recall now as (13.54) below. Let

$$-H_t(\boldsymbol{\sigma}) = \frac{\beta\sqrt{t}}{\sqrt{N}} \sum_{i<j} g_{ij}\sigma_i\sigma_j + \beta\sqrt{1-t}\sum_{i\le N} z_i\sigma_i\sqrt{q} + \sum_{i\le N} h_i\sigma_i \tag{13.53}$$

and

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

Of course φ depends on N but the dependence is kept implicit. Then

$$\varphi'(t) = \frac{\beta^2}{4}(1-q)^2 - \frac{\beta^2}{4}\mathsf{E}\langle (R_{1,2}-q)^2\rangle_t , \tag{13.54}$$

where $\langle\cdot\rangle_t$ denotes an average for the Gibbs measure with Hamiltonian (13.53). Moreover $\varphi(1) = p_N(\beta, h)$. Let us consider the function

$$\psi(t) = \log 2 + \mathsf{E}\log\operatorname{ch}(\beta z\sqrt{q} + h) + \frac{\beta^2}{4}t(1-q)^2 , \tag{13.55}$$

which is simply the value of the functional SK for the interpolating system with Hamiltonian 13.53. Thus

$$\psi(0) = \varphi(0) \tag{13.56}$$

$$\psi'(t) - \varphi'(t) = \frac{\beta^2}{4}\mathsf{E}\langle (R_{1,2}-q)^2\rangle_t . \tag{13.57}$$

Assuming (13.52), we would like to show that $\psi'(t) - \varphi'(t)$ is small, so that $\varphi(1) \simeq \psi(1)$, i.e. $p_N(\beta, h) \simeq \mathrm{SK}(\beta, h)$. The basic idea is as follows. For $|u| \le 1$ let us define

$$\psi(t,u) = \frac{1}{N}\mathsf{E}\log \sum_{R_{1,2}=u} \exp(-H_t(\boldsymbol{\sigma}^1) - H_t(\boldsymbol{\sigma}^2)) \,. \tag{13.58}$$

When considering such a quantity involving a summation over values of $\boldsymbol{\sigma}^1$ and $\boldsymbol{\sigma}^2$ with $R_{1,2} = u$, we will always assume that such values of $\boldsymbol{\sigma}^1$ and $\boldsymbol{\sigma}^2$ exist, or, equivalently, that $N(1-u) \in 2\mathbb{N}$.

The basic idea of the approach is as follows. If $\psi(t,u) \le 2\varphi(t) - \varepsilon$, where ε does not depend on N, we will show that for large N it is very rare for the Gibbs measure that $R_{1,2} = u$. Thus if it holds true that $\psi(t,u) \le 2\varphi(t) - \varepsilon$ whenever u is not very close to q, we have $\psi'(t) - \varphi'(t) = \mathsf{E}\langle (R_{1,2} - q)^2 \rangle_t \simeq 0$. When $\varphi(t) \simeq \psi(t)$, to prove that $\psi(t,u) \le 2\varphi(t) - \varepsilon$ it suffices to prove that $\psi(t,u) \le 2\psi(t) - \varepsilon/2$. Thus **if** one can show that $\psi(t,u) < 2\psi(t)$ for $u \ne q$, this should imply that

$$\varphi(t) \simeq \psi(t) \quad \Rightarrow \quad \psi'(t) \simeq \varphi'(t) \,.$$

More formally, we will obtain a differential inequality, the integration of which yields the result.

The central part of our approach to Theorem 13.4.1 is therefore to prove in a suitable form that $\psi(t,u) < 2\psi(t)$ for $u \ne q$. This is the content of the next theorem.

Theorem 13.4.2. *Assume (13.52) and consider $t_0 < 1$. Then there exists a number K_0 independent of N such that for $t \le t_0$ we have*

$$\psi(t,u) \le 2\psi(t) - \frac{(u-q)^2}{K_0} \,. \tag{13.59}$$

The number K_0 might of course depend on β, h and t_0.

Theorem 13.4.2 is at the core of the proof of Theorem 13.4.1, but its proof is far from being immediate. Consequently we delay it until Section 13.6. Assuming Theorem 13.4.2, we complete the proof of Theorem 13.4.1.

Our next result explicits the idea that if $\psi(t,u) < 2\varphi(t) - \varepsilon$, it is very rare that $R_{1,2} \simeq u$.

Proposition 13.4.3. *Let us assume that for some $\varepsilon > 0$, we have*

$$\psi(t,u) \le 2\varphi(t) - \varepsilon \,. \tag{13.60}$$

Then we have

$$\mathsf{E}\langle \mathbf{1}_{\{R_{1,2}=u\}} \rangle_t \le K(\varepsilon) \exp\left(-\frac{N}{K(\varepsilon)} \right) \,, \tag{13.61}$$

where $K(\varepsilon)$ does not depend on N, u or t.

Proof. This is a consequence of concentration of measure. In the proof $K(\varepsilon)$ denotes a number independent of N, u or t, that need not be the same at each occurrence. Let

$$Z_t = \sum_{\boldsymbol{\sigma}} \exp(-H_t(\boldsymbol{\sigma}))$$

so that by Theorem 1.3.4 we have

$$\begin{aligned}\mathsf{P}\left(Z_t \geq \exp N\left(\varphi(t) - \frac{\varepsilon}{6}\right)\right) &= \mathsf{P}\left(\frac{1}{N}\log Z_t \geq \varphi(t) - \frac{\varepsilon}{6}\right) \\ &= \mathsf{P}\left(\frac{1}{N}\log Z_t \geq \frac{1}{N}\mathsf{E}\log Z_t - \frac{\varepsilon}{6}\right) \\ &\geq 1 - K(\varepsilon)\exp\left(-\frac{N}{K(\varepsilon)}\right) .\end{aligned}$$

In a similar manner we have

$$\mathsf{P}\left(\sum_{R_{1,2}=u} \exp(-H_t(\boldsymbol{\sigma}^1) - H_t(\boldsymbol{\sigma}^2)) \leq \exp N\left(\psi(t,u) + \frac{\varepsilon}{6}\right)\right)$$
$$\geq 1 - K(\varepsilon)\exp\left(-\frac{N}{K(\varepsilon)}\right) .$$

Now, under (13.60), assuming

$$Z_t \geq \exp N\left(\varphi(t) - \frac{\varepsilon}{6}\right) ;$$

$$\sum_{R_{1,2}=u} \exp(-H_t(\boldsymbol{\sigma}^1) - H_t(\boldsymbol{\sigma}^2)) \leq \exp N\left(\psi(t,u) + \frac{\varepsilon}{6}\right) ,$$

we have

$$\begin{aligned}\langle \mathbf{1}_{\{R_{1,2}=u\}}\rangle_t &= \frac{1}{Z_t^2}\sum_{R_{1,2}=u} \exp(-H_t(\boldsymbol{\sigma}^1) - H_t(\boldsymbol{\sigma}^2)) \\ &\leq \exp N\Big(-2\varphi(t) + \frac{\varepsilon}{3} + \psi(t,u) + \frac{\varepsilon}{6}\Big) \\ &\leq \exp\left(-\frac{N\varepsilon}{2}\right) ,\end{aligned}$$

from which (13.61) follows. □

It is important to *fully understand* the previous argument. The same principle will be used again and again, but the argument will not be repeated.

Proposition 13.4.4. *Assume (13.59) and consider $t_0 < 1$ and $\varepsilon > 0$. Then there exists a number K that does not depend on N, t or ε and such that for N large enough one has for each $t \leq t_0$ that*

$$\mathsf{E}\langle (R_{1,2} - q)^2\rangle_t \leq K(\psi(t) - \varphi(t) + \varepsilon) . \tag{13.62}$$

Proof. Consider the number K_0 of (13.59). Then if

$$(u-q)^2 \geq 2K_0(\psi(t) - \varphi(t) + \varepsilon) ,$$

by (13.59) we have $\psi(t,u) \leq 2\varphi(t) - \varepsilon$, so that (13.61) implies

$$\mathsf{E}\langle \mathbf{1}_{\{R_{1,2}=u\}}\rangle_t \leq K(\varepsilon) \exp\left(-\frac{N}{K(\varepsilon)}\right) . \tag{13.63}$$

Since $NR_{1,2} \in \mathbb{Z}$, we have

$$\langle \mathbf{1}_{\{(R_{1,2}-q)^2 \geq c\}}\rangle_t \leq \sum_u \langle \mathbf{1}_{\{R_{1,2}=u\}}\rangle_t , \tag{13.64}$$

where the summation is over the values of u such that $(u-q)^2 \geq c$, $|u| \leq 1$, $Nu \in \mathbb{Z}$. There are at most $2N+1$ such values. Thus if we choose $c = 2K_0(\psi(t) - \varphi(t) + \varepsilon)$, (13.63) implies

$$\mathsf{E}\langle \mathbf{1}_{\{(R_{1,2}-q)^2 \geq c\}}\rangle_t \leq (2N+1)K(\varepsilon) \exp\left(-\frac{N}{K(\varepsilon)}\right) \leq \frac{\varepsilon}{4} \tag{13.65}$$

for N large enough. Finally, we note that since $|R_{1,2} - q| \leq 2$ we have

$$\langle (R_{1,2}-q)^2\rangle_t \leq c + 4\langle \mathbf{1}_{\{(R_{1,2}-q)^2 \geq c\}}\rangle_t$$

and we take expectation and use (13.65). This proves (13.62) with $K = 2K_0 + 1$. □

Proof of Theorem 13.4.1. From (13.57) we know that

$$(\psi(t) - \varphi(t))' = \frac{\beta^2}{4} \mathsf{E}\langle (R_{1,2}-q)^2\rangle_t . \tag{13.66}$$

Proposition 13.4.4 shows that, given $t_0 < 1$ and $\varepsilon > 0$, for N large enough and $t \leq t_0$ one has

$$(\psi(t) - \varphi(t))' \leq K(\psi(t) - \varphi(t) + \varepsilon) , \tag{13.67}$$

where K does not depend on N or ε. Integrating this relation (e.g. as in Lemma A.11.1) we get, since $\psi(0) = \varphi(0)$, that for N large enough and all $t \leq t_0$ we have $\psi(t) - \varphi(t) \leq K\varepsilon \exp Kt$. Since ε is arbitrary, and K does not depend on ε, it follows that

$$t \leq t_0 \quad \Rightarrow \quad \lim_{N\to\infty} (\psi(t) - \varphi(t)) = 0 . \tag{13.68}$$

Moreover, (13.66) entails $(\psi(t) - \varphi(t))' \leq \beta^2$, so that since (13.68) holds for any t_0 we have $\lim_{N\to\infty}(\psi(1) - \varphi(1)) = 0$. □

It is probably worth repeating the fundamental new idea of this chapter. A *lower* bound for $\lim_N p_N(\beta, h)$ is deduced from the *upper* bound (13.59).

13.5 A Bound for Coupled Copies

Theorem13.4.2 asserts a bound for

$$\psi(t,u)=\frac{1}{N}\mathsf{E}\log\sum_{R_{1,2}=u}\exp(-H_t(\boldsymbol{\sigma}^1)-H_t(\boldsymbol{\sigma}^2))\ .$$

We first must learn how to bound such quantities. The interpolating Hamiltonian H_t is the Hamiltonian of a certain SK model. This motivates the following definition. Recalling the Hamiltonian (13.18) we set

$$r_N(\beta,h,u)=\frac{1}{N}\mathsf{E}\log\sum_{R_{1,2}=u}\exp(-H_N(\boldsymbol{\sigma}^1)-H_N(\boldsymbol{\sigma}^2))\ . \tag{13.69}$$

(Let us recall that we assume that u is such that there are pairs $(\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2)$ with $R_{1,2}=u$.) In this section we prove a bound for $r_N(\beta,h,u)$, from which we will deduce Theorem 13.4.2 in the next section. The words "coupled copies" refer to the fact that in (13.69) the summation is only over the pairs for which $R_{1,2}=u$ rather than over all pairs.

It is a fundamental unsolved problem to know what are the best possible bounds for $r_N(\beta,h,u)$. This will be discussed in Chapter 16. Fortunately the bounds will be able to prove will be sufficient to obtain Theorem 13.4.2. A rather notable feature is that, even though we are trying to control the high-temperature region of the SK model (a "replica-symmetric" region), it is very useful to use a bound of the 1-RSB type for $r_N(\beta,h,u)$ (rather than a "replica-symmetric" bound).

Let us consider two numbers $0\le q_1\le q_2\le 1$, and standard normal r.v.s z_1,z_2,z_1',z_2'. We assume that the pairs (z_1,z_2) and (z_1',z_2') are independent of each other. We define

$$c=\mathsf{E}z_1z_2\quad ;\quad c'=\mathsf{E}z_1'z_2'\ . \tag{13.70}$$

Theorem 13.5.1. *Assume that*

$$u=cq_1+c'(q_2-q_1)\ . \tag{13.71}$$

For $j=1,2$, let us set

$$Y_j=\beta z_j\sqrt{q_1}+\beta z_j'\sqrt{q_2-q_1}+h \tag{13.72}$$

and let us denote by E' expectation in z_1' and z_2' only. Then if $0\le m\le 1$, for any λ,

$$\begin{aligned}r_N(\beta,h,u)\le{}&2\log 2+\frac{\beta^2}{2}\big((1-q_2)^2-m(q_2^2-q_1^2)-m(u^2-c^2q_1^2)\big)\\&+\frac{1}{m}\mathsf{E}\log\mathsf{E}'(\mathrm{ch}Y_1\mathrm{ch}Y_2\mathrm{ch}\lambda+\mathrm{sh}Y_1\mathrm{sh}Y_2\mathrm{sh}\lambda)^m-\lambda u\ .\end{aligned} \tag{13.73}$$

When $m = 0$, the term before the last has to be interpreted as

$$\mathsf{E}\log(\mathrm{ch}Y_1\mathrm{ch}Y_2\mathrm{ch}\lambda + \mathrm{sh}Y_1\mathrm{sh}Y_2\mathrm{sh}\lambda) .$$

This bound looks complicated at first sight, but we will learn how to choose parameters efficiently. We will always use the case $c \in \{-1, 1\}$, $c' \in \{-1, 0, 1\}$, but these assumptions do not simplify the proof of Theorem 13.5.1.

Proof. Overall, the proof is a 2 dimensional version of Guerra's bound (1.72).

The essential new feature of the argument is rather subtle. It is fundamental that we deal only with pairs of configurations $(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2)$ such that $R_{1,2}$ has a given value ($= u$). Otherwise, the interaction between these 2 replicas would create a term $\beta^2\mathsf{E}\langle R_{1,2}^2\rangle$ with the *wrong sign*, that we could not control.

It is enough to prove (13.73) for $0 < m < 1$, since the result for $m = 0$ or $m = 1$ then follows by continuity. We consider independent copies $(z_{i,1}, z_{i,2})$ of the pair (z_1, z_2) and independent copies $(z'_{i,1}, z'_{i,2})$ and $(z'_{i,\alpha,1}, z'_{i,\alpha,2})$ of the pair (z'_1, z'_2). These are independent of each other and of the randomness of H_N. We define

$$\begin{aligned} -H_t(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2, \alpha) = \beta\sqrt{\frac{t}{N}}&\sum_{i<j} g_{ij}(\sigma_i^1\sigma_j^1 + \sigma_i^2\sigma_j^2) \\ &+ \beta\sqrt{1-t}\sum_{i\le N}\sum_{j=1,2}\sigma_i^j(z_{i,j}\sqrt{q_1} + z'_{i,\alpha,j}\sqrt{q_2-q_1}) \\ &+ \sum_{i\le N} h_i(\sigma_i^1 + \sigma_i^2) . \end{aligned} \tag{13.74}$$

We consider a sequence (v_α) of law Λ_m (the Poisson Dirichlet distribution of parameter m) and we set

$$\varphi(t) = \frac{1}{N}\mathsf{E}\log\sum_{\alpha\ge1} v_\alpha \sum_{R_{1,2}=u}\exp(-H_t(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2, \alpha)) , \tag{13.75}$$

so that

$$\varphi(1) = r_N(\beta, h, u) . \tag{13.76}$$

As usual we will bound $\varphi(0)$ and $\varphi'(t)$. We will prove that

$$\varphi(0) \le 2\log 2 + \frac{1}{m}\mathsf{E}\log\mathsf{E}(\mathrm{ch}Y_1\mathrm{ch}Y_2\mathrm{ch}\lambda + \mathrm{sh}Y_1\mathrm{sh}Y_2\mathrm{sh}\lambda)^m - \lambda u \tag{13.77}$$

and

$$2\varphi'(t) \le \beta^2\big((1-q_2)^2 - m(q_2^2 - q_1^2) - m(u^2 - c^2q_1^2)\big) .$$

Let

$$\begin{aligned} Y_{i,j} &= \beta z_{i,j}\sqrt{q_1} + \beta z'_{i,j}\sqrt{q_2-q_1} + h_i \\ Y_{i,\alpha,j} &= \beta z_{i,j}\sqrt{q} + z'_{i,\alpha,j}\sqrt{q_2-q_1} + h_i , \end{aligned} \tag{13.78}$$

so that, if E' denotes expectation in the r.v. $z'_{i,1}, z'_{i,2}$ only

$$\begin{aligned}\varphi(0) &= \frac{1}{N}\mathsf{E}\log\sum_{\alpha} v_\alpha\left(\sum_{R_{1,2}=u}\exp\sum_{i\le N}\sum_{j=1,2}\sigma_i^j Y_{i,\alpha,j}\right)\\ &= \frac{1}{Nm}\mathsf{E}\log\mathsf{E}'\left(\sum_{R_{1,2}=u}\exp\sum_{i\le N}\sum_{j=1,2}\sigma_i^j Y_{i,j}\right)^m, \qquad (13.79)\end{aligned}$$

using (13.12) given the randomness of $z_{i,j}$ and h_i. Since $NR_{1,2} = \sum_{i\le N}\sigma_i^1\sigma_i^2$, for any λ we have

$$\begin{aligned}&\sum_{R_{1,2}=u}\exp\sum_{i\le N}\sum_{j=1,2}\sigma_i^j Y_{i,j}\\ &= \exp(-Nu\lambda)\sum_{R_{1,2}=u}\exp\left(\sum_{i\le N}\sum_{j=1,2}\sigma_i^j Y_{i,j} + \lambda N R_{1,2}\right)\\ &\le \exp(-Nu\lambda)\sum_{\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2}\exp\sum_{i\le N}(\sigma_i^1 Y_{i,1} + \sigma_i^2 Y_{i,2} + \lambda\sigma_i^1\sigma_i^2)\,. \qquad (13.80)\end{aligned}$$

Now

$$\begin{aligned}&\sum_{\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2}\exp\sum_{i\le N}(\sigma_i^1 Y_{i,1} + \sigma_i^2 Y_{i,2} + \lambda\sigma_i^1\sigma_i^2)\\ &= \prod_{i\le N}\left(\sum_{\sigma_i^1,\sigma_i^2=\pm1}\exp(\sigma_i^1 Y_{i,1} + \sigma_i^2 Y_{i,2} + \lambda\sigma_i^1\sigma_i^2)\right)\\ &= 4^N\prod_{i\le N}(\mathrm{ch}Y_{i,1}\mathrm{ch}Y_{i,2}\mathrm{ch}\lambda + \mathrm{sh}Y_{i,1}\mathrm{sh}Y_{i,2}\mathrm{sh}\lambda)\,,\end{aligned}$$

and combining with (13.79) and (13.80) yields (13.77).

To compute $\varphi'(t)$ we need to define an average $\langle\cdot\rangle_t$ on the configuration space

$$\{(\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2)\ ;\ R_{1,2}=u\}\times\mathbb{N}\,.$$

This average is given by

$$\langle f(\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2,\alpha)\rangle_t = \frac{1}{Z_t}\sum_{\alpha} v_\alpha\sum_{R_{1,2}=u} f(\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2,\alpha)\exp(-H_t(\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2,\alpha))$$

where Z_t is the normalization factor. By straightforward differentiation, we get

$$2\varphi'(t) = \mathrm{I} + \mathrm{II} + \mathrm{III}$$

where

$$\mathrm{I} = \frac{\beta}{\sqrt{tN}} \mathsf{E} \left\langle \sum_{i<j} g_{ij}(\sigma_i^1 \sigma_j^1 + \sigma_i^2 \sigma_j^2) \right\rangle_t$$

$$\mathrm{II} = -\frac{\beta}{\sqrt{1-t}} \mathsf{E} \left\langle \sum_{i \le N} \sum_{j=1,2} \sigma_i^j z_{i,j} \sqrt{q_1} \right\rangle_t$$

$$\mathrm{III} = -\frac{\beta}{\sqrt{1-t}} \mathsf{E} \left\langle \sum_{i \le N} \sum_{j=1,2} \sigma_i^j z'_{i,\alpha,j} \sqrt{q_2 - q_1} \right\rangle_t .$$

We need to integrate by parts. For this we denote by $(\boldsymbol{\tau}^1, \boldsymbol{\tau}^2, \gamma)$ a replica of the configuration $(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2, \alpha)$; and we use obvious notation such as

$$R(\boldsymbol{\sigma}^j, \boldsymbol{\tau}^{j'}) = \frac{1}{N} \sum_{i \le N} \sigma_i^j \tau_i^{j'} .$$

To perform integration by parts we remember that $R_{1,2} = u$ on the configuration space, and we find

$$\mathrm{I} = \frac{\beta^2}{2}(2 + 2u^2) - \frac{\beta^2}{2} \sum_{j,j'=1,2} \mathsf{E}\langle R(\boldsymbol{\sigma}^j, \boldsymbol{\tau}^{j'})^2 \rangle_t . \tag{13.81}$$

It is at this stage that the relation $R_{1,2} = u$ allows to control the term with the wrong sign by making it constant.

The rest of the computation is certainly more complicated that one wishes, so it might be useful to explain the strategy. When computing the quantities II and III there will occur terms of the type $a\langle R(\boldsymbol{\sigma}^j, \boldsymbol{\tau}^{j'}) \mathbf{1}_{\{\alpha \ne \gamma\}} \rangle_t$ or $a\langle R(\boldsymbol{\sigma}^j, \boldsymbol{\tau}^j) \mathbf{1}_{\{\alpha = \gamma\}} \rangle_t$, where a is a quantity that depends on q_1, q_2 and c. These will be gathered with the last terms of (13.81), completing the squares. One will then observe that terms such as $\langle (R(\boldsymbol{\sigma}^j, \boldsymbol{\tau}^{j'}) - b)^2 \mathbf{1}_{\{\alpha \ne \gamma\}} \rangle_t$ are ≥ 0 to obtain the required bound for φ'. The complication is purely algebraic.

To compute II and III we integrate by parts using the relations (13.70) and (13.72). Defining

$$A = R(\boldsymbol{\sigma}^1, \boldsymbol{\tau}^1) + R(\boldsymbol{\sigma}^2, \boldsymbol{\tau}^2) + cR(\boldsymbol{\sigma}^1, \boldsymbol{\tau}^2) + cR(\boldsymbol{\sigma}^2, \boldsymbol{\tau}^1)$$

and A' similarly with c' instead of c, we find

$$\mathrm{II} = -\beta^2(2q_1 + 2cq_1 u - q_1 \mathsf{E}\langle A \rangle_t)$$
$$\mathrm{III} = -\beta^2(2(q_2 - q_1) + 2c'(q_2 - q_1)u - (q_2 - q_1)\mathsf{E}\langle A' \mathbf{1}_{\{\alpha = \gamma\}} \rangle_t) .$$

To gather these terms we write

$$\mathsf{E}\langle A \rangle_t = \mathsf{E}\langle A \mathbf{1}_{\{\alpha \ne \gamma\}} \rangle_t + \mathsf{E}\langle A \mathbf{1}_{\{\alpha = \gamma\}} \rangle_t ,$$

and similarly for A', so that

$$\mathrm{II} + \mathrm{III} = -\beta^2(2q_2 + 2(cq_1 + c'(q_2 - q_1))u - \mathsf{E}\langle B1_{\{\alpha\neq\gamma\}}\rangle_t - \mathsf{E}\langle B'\mathbf{1}_{\{\alpha=\gamma\}}\rangle_t)$$

where

$$B = q_1 A = q_1 R(\boldsymbol{\sigma}^1, \boldsymbol{\tau}^2) + q_1 R(\boldsymbol{\sigma}^2, \boldsymbol{\tau}^2) + cq_1(R(\boldsymbol{\sigma}^1, \boldsymbol{\tau}^2) + R(\boldsymbol{\sigma}^2, \boldsymbol{\tau}^1))$$

and B' is defined by a similar formula replacing q_1 by q_2 and cq_1 by $cq_1 + c'(q_2 - q_1)$. Now we gather the terms $\mathrm{II} + \mathrm{III}$ with the term I; to do this we write

$$\langle R(\boldsymbol{\sigma}^j, \boldsymbol{\tau}^{j'})^2\rangle_t = \langle R(\boldsymbol{\sigma}^j, \boldsymbol{\tau}^{j'})^2\mathbf{1}_{\{\alpha=\gamma\}}\rangle_t + \langle R(\boldsymbol{\sigma}^j, \boldsymbol{\tau}^{j'})^2\mathbf{1}_{\{\alpha\neq\gamma\}}\rangle_t$$

for each values of j and j'. Writing $d_1 = cq_1$ and $d_2 = cq_1 + c'(q_2 - q_1)$ and completing the squares yields

$$\begin{aligned} 2\varphi'(t) = \beta^2(1 + u^2 - 2q_2 - 2d_2 u) + \frac{\beta^2}{2}(2q_1^2 + 2d_1^2)\mathsf{E}\langle\mathbf{1}_{\{\alpha\neq\gamma\}}\rangle_t \\ + \frac{\beta^2}{2}(2q_2^2 + 2d_2^2)\mathsf{E}\langle\mathbf{1}_{\{\alpha=\gamma\}}\rangle_t - \frac{\beta^2}{2}C\,, \end{aligned}$$

where

$$\begin{aligned} C = \mathsf{E}\big\langle\mathbf{1}_{\{\alpha\neq\gamma\}}\big((R(\boldsymbol{\sigma}^1, \boldsymbol{\tau}^1) - q_1)^2 + (R(\boldsymbol{\sigma}^2, \boldsymbol{\tau}^2) - q_1)^2 \\ + (R(\boldsymbol{\sigma}^1, \boldsymbol{\tau}^2) - d_1)^2 + (R(\boldsymbol{\sigma}^2, \boldsymbol{\tau}^1) - d_1)^2\big)\big\rangle_t \\ + \mathsf{E}\big\langle\mathbf{1}_{\{\alpha=\gamma\}}\big((R(\boldsymbol{\sigma}^1, \boldsymbol{\tau}^1) - q_2)^2 + (R(\boldsymbol{\sigma}^2, \boldsymbol{\tau}^2) - q_2)^2 \\ + (R(\boldsymbol{\sigma}^1, \boldsymbol{\tau}^2) - d_2)^2 + (R(\boldsymbol{\sigma}^2, \boldsymbol{\tau}^1) - d_2)^2\big)\big\rangle_t\,, \end{aligned}$$

so that $C \geq 0$. Now, using (13.14) as in the proof of Theorem 13.2.1 we have $\mathsf{E}\langle\mathbf{1}_{\{\alpha=\gamma\}}\rangle_t = 1 - m$ so that we have proved that

$$\begin{aligned} 2\varphi'(t) &\leq \beta^2(1 + u^2 - 2q_2 - 2d_2 u) - \beta^2 m(q_2^2 - q_1^2) - \beta^2 m(d_2^2 - d_1^2) \\ &\quad + \beta^2 q_2^2 + \beta^2 d_2^2 \\ &= \beta^2\big((1 - q_2)^2 - m(q_2^2 - q_1^2) - m(u^2 - c^2 q_1^2)\big) \end{aligned}$$

because $d_2 = cq_1 + c'(q_2 - q_1) = u$ by (13.71) and $d_1 = cq_1$ by definition of d_1. □

13.6 The Main Estimate

In this section we complete the proof of Theorem 13.4.1 by proving Theorem 13.4.2.

One difficulty in proving Theorem 13.4.1 is that under (13.52) (i.e. when $\mathrm{SK}(\beta, h) = \mathcal{P}_1(\beta, h)$) the parameters (β, h) can be "at the boundary of the high-temperature region", while in order to prove nice estimates, one would rather be "in the interior of the high-temperature region". Our first task is to prove that (as expected) the interpolation (13.53) brings us immediately "in the interior of the high-temperature region". This interpolation amounts to replace β by

$$\beta_t = \sqrt{t}\beta$$

and h by

$$h_t = h + \sqrt{1-t}\beta z''\sqrt{q}$$

(where z'' is standard Gaussian independent of h). We recall that

$$\beta_t z\sqrt{q} + h_t \overset{\mathcal{D}}{=} \beta z\sqrt{q} + h \;, \tag{13.82}$$

where $\overset{\mathcal{D}}{=}$ means equality in distribution. Let us also recall the function $V(\beta, h, q')$ of (13.39),

$$V(\beta, h, q') = \frac{\partial \Phi}{\partial m}(m, q')\Big|_{m=1} \;,$$

where Φ is given by (13.32).

Lemma 13.6.1. *Assume (13.52), that is* $\mathrm{SK}(\beta, h) = \mathcal{P}_1(\beta, h)$. *Then, for* $q \le q' \le 1$ *and* $0 \le t \le 1$ *we have*

$$V(\beta_t, h_t, q') \le -\frac{\beta^2}{4} t(1-t)(q'-q)^2 \tag{13.83}$$

$$\beta_t^2 \mathsf{E} \frac{1}{\mathrm{ch}^4(\beta_t z\sqrt{q} + h_t)} \le t \;. \tag{13.84}$$

Proof. We first prove (13.84). By Theorem 13.3.1, under (13.52) we have

$$\beta^2 \mathsf{E} \frac{1}{\mathrm{ch}^4(\beta z\sqrt{q} + h)} \le 1$$

and since $\beta_t^2 = t\beta$ and using (13.82),

$$\beta_t^2 \mathsf{E} \frac{1}{\mathrm{ch}^4(\beta_t z\sqrt{q} + h_t)} = t\beta^2 \mathsf{E} \frac{1}{\mathrm{ch}^4(\beta z\sqrt{q} + h)} \le t \;.$$

This proves (13.84). We now turn to the proof of (13.83). The relations (13.52), (13.37) and (13.39) imply

$$\forall q' \ge q \;, \quad V(\beta, h, q') \le 0 \;. \tag{13.85}$$

Indeed, otherwise we can choose m close to 1 with

$$\mathcal{P}_1(\beta, q) \le \Phi(m, q') < \Phi(1, q') = \mathrm{SK}(\beta, h) \;.$$

Let us rewrite (13.40) as

$$V(\beta, h, q') = -\frac{\beta^2}{4}(q'^2 - q^2) + C(\beta, h, q') \tag{13.86}$$

where

$$C(\beta, h, q') = -\mathsf{E}\log \mathsf{E}'\mathrm{ch}Y' + \mathsf{E}\frac{\mathsf{E}'\log(\mathrm{ch}Y')\mathrm{ch}Y'}{\mathsf{E}'\mathrm{ch}Y'} ,$$

for $Y' = \beta z\sqrt{q} + \beta z'\sqrt{q'-q} + h$. Then it follows from (13.82) that the identity

$$C(\beta, h, q'_t) = C(\beta_t, h_t, q') \tag{13.87}$$

holds whenever $\beta_t\sqrt{q'-q} = \beta\sqrt{q'_t - q}$, i.e. $q'_t = q + t(q'-q)$.

It then follows from (13.86) and (13.87) that

$$V(\beta, h, q'_t) + \frac{\beta^2}{4}({q'_t}^2 - q^2) = V(\beta_t, h_t, q') + \frac{\beta_t^2}{4}(q'^2 - q^2) ,$$

and using (13.85) we get

$$\begin{aligned} V(\beta_t, h_t, q') &\le \frac{\beta^2}{4}({q'_t}^2 - q^2) - \frac{\beta_t^2}{4}(q'^2 - q^2) \\ &= \frac{\beta^2}{4}\big((q + t(q'-q))^2 - q^2 - t(q'^2 - q^2)\big) \\ &= -\frac{\beta^2}{4}t(1-t)(q'-q)^2 . \end{aligned}$$

This conclude the proof of (13.83). □

Theorem 13.6.2. *Assume either $\beta < 1/10$ or else that the following two conditions hold:*

$$q' > q \quad \Rightarrow \quad V(\beta, h, q') < 0 \tag{13.88}$$

$$\beta^2\mathsf{E}\frac{1}{\mathrm{ch}^4(\beta z\sqrt{q} + h)} < 1 . \tag{13.89}$$

Then there exists K such that for $u \ne q$ we have

$$\forall N , \quad r_N(\beta, h, u) \le 2\mathrm{SK}(\beta, h) - \frac{(u-q)^2}{K} . \tag{13.90}$$

The strict inequalities in (13.88) and (13.89) are fundamentally important in the proof. These inequalities mean that we are "inside the interior of the high-temperature region", and are made possible by Lemma 13.6.1. Since

$$\psi(t, u) = r_N(\beta_t, h_t, u) \ ; \quad \psi(t) = \mathrm{SK}(\beta_t, h_t) ,$$

combining Theorem 13.6.2 and Lemma 13.6.1 we have proved under (13.52) that if $t < 1$ we have

$$\psi(t,u) \le 2\psi(t) - \frac{(u-q)^2}{K(t)} \ .$$

The fact that this holds uniformly over $t \le t_0$, i.e. that $K(t) \le K$ where K depends only on t_0, β and h should be obvious from the arguments that we will give. The reason for considering separately the case $\beta < 1/10$ in Theorem 13.6.2 is that (13.83) does not prove that (13.88) holds "uniformly for $t > 0$"; but the case of small t is covered by the case $\beta_t = \beta\sqrt{t} < 1/10$.

The principle of the proof of Theorem 13.6.2 is simply to make an appropriate choice of parameters in Theorem 13.5.1.

It is rather useful to think of (13.90) as made up of two different statements. First, there is the statement that (13.90) holds for u close to q, and this is a consequence of (13.89) alone. The second statement is that $\sup_N r_N(\beta,h,u) < 2\mathrm{SK}(\beta,h)$ for $u \ne q$. It is the case $u > q$ that is the trickiest and requires (13.88).

It seems a rather difficult problem to optimize the bound (13.73) over the choice of parameters. The fundamental observation that will allow us to bypass this problem is that there is a natural choice of these parameters for which the bound (13.73) is exactly $2\mathrm{SK}(\beta,h)$; and then we will show that we can improve this bound by a small variation of either q or m. The important situation is really the case $u \ge 0$ (as should be expected from Theorem 12.3.1). The arguments are easier but less "canonical" if $u < 0$.

We find it clearer to assume that $q \ne 0$, i.e. $h \ne 0$. Theorem 13.4.1 has already been proved in Chapter 11 when $h = 0$. The present approach can certainly be adapted to recover this result. This is left to the enterprising reader.

We will split the proof of Theorem 13.6.2 in three parts.

Proposition 13.6.3. *The conditions of Theorem 13.6.2 imply*

$$\forall N\ , \quad |u| \le q \quad \Rightarrow \quad r_N(\beta,h,u) \le 2\mathrm{SK}(\beta,h) - \frac{(u-q)^2}{K} \ . \tag{13.91}$$

Proof. We use (13.73) for $q_1 = |u|$, $q_2 = q$, $c' = 0$, $c = \mathrm{sign}(u)$, $m = 0$ to get

$$\begin{aligned} r_N(\beta,h,u) &\le 2\log 2 + \frac{\beta^2}{2}(1-q)^2 + \mathsf{E}\log(\mathrm{ch}Y_1\mathrm{ch}Y_2\mathrm{ch}\lambda + \mathrm{sh}Y_1\mathrm{sh}Y_2\mathrm{sh}\lambda) - \lambda u \\ &=: G(u,\lambda)\ . \end{aligned}$$

Since both Y_1 and Y_2 are distributed like $\beta z\sqrt{q} + h$, we have

$$G(u,0) = 2\mathrm{SK}(\beta,h)\ . \tag{13.92}$$

Let

$$G_1(u) := \frac{\partial}{\partial\lambda}G(u,\lambda)\Big|_{\lambda=0} = \mathsf{E}\mathrm{th}Y_1\mathrm{th}Y_2 - u\ . \tag{13.93}$$

We will prove that

$$G_1(q) = 0 \tag{13.94}$$

$$G_1'(u) \le G_1'(q) < 0 \,. \tag{13.95}$$

We first show that these relations imply (13.91). It is obvious (by writing the explicit expression) that $|\partial^2 G(u,\lambda)/\partial\lambda^2| \le K$, and Taylor's formula implies

$$G(u,\lambda) \le G(u,0) + \lambda G_1(u) + \frac{\lambda^2}{2} K \,.$$

The choice $\lambda = -G_1(u)/K$ yields

$$\inf_\lambda G(u,\lambda) \le 2\mathrm{SK}(\beta,h) - \frac{G_1(u)^2}{K} \,. \tag{13.96}$$

Now, (13.94) and (13.95) show that $G_1(u)^2 \ge (q-u)^2/K$, and this proves (13.91).

To prove (13.94) we observe that for $u = q$ we have $c = 1$, so $z_1 = z_2$ and $Y_1 = Y_2 \stackrel{\mathcal{D}}{=} \beta z\sqrt{q} + h$. Therefore

$$G_1(q) = \mathsf{E}\mathrm{th}^2(\beta z\sqrt{q} + h) - q = 0 \,, \tag{13.97}$$

and this proves (13.94).

To prove (13.95) we compute $G_1'(u)$. We first assume that $u > 0$. Then, in distribution, we have

$$Y_1 \stackrel{\mathcal{D}}{=} h + \beta z\sqrt{u} + \beta z_1\sqrt{q-u} \,; \quad Y_2 \stackrel{\mathcal{D}}{=} h + \beta z\sqrt{u} + \beta z_2\sqrt{q-u} \,,$$

where z, z_1, z_2 are independent standard Gaussian r.v.s. Let

$$Y_j' = \frac{\mathrm{d}Y_j}{\mathrm{d}u} = \frac{\beta z}{2\sqrt{u}} - \frac{\beta z_j}{2\sqrt{q-u}}$$

so that $\mathsf{E}Y_j Y_j' = 0$, $\mathsf{E}Y_1' Y_2 = \mathsf{E}Y_1 Y_2' = \beta^2/2$. Therefore using integration by parts,

$$G_1'(u) = \mathsf{E}\left(Y_1' \frac{\mathrm{th}Y_2}{\mathrm{ch}^2 Y_1} + Y_2' \frac{\mathrm{th}Y_1}{\mathrm{ch}^2 Y_2}\right) - 1 = \beta^2 \mathsf{E}\frac{1}{\mathrm{ch}^2 Y_1}\frac{1}{\mathrm{ch}^2 Y_2} - 1 \,,$$

and, using the Cauchy-Schwarz inequality and the fact that Y_1 and Y_2 are distributed like $\beta z\sqrt{q} + h$,

$$G_1'(u) \le \beta^2 \mathsf{E}\frac{1}{\mathrm{ch}^4(\beta z\sqrt{q} + h)} - 1 = G_1'(q) \,.$$

Thus (13.89) implies

$$G_1'(u) \le G_1'(q) < 0 \,,$$

and using (13.97) we have $G_1(u) > 0$ for $0 \le u < q$. Let us now assume $u < 0$. Then, since $c = -1$, we have $z_2 = -z_1$, so that, in distribution, we have

$$Y_1 \stackrel{\mathcal{D}}{=} h + \beta z\sqrt{-u} + \beta z_1'\sqrt{q+u}\ ; \quad Y_2 \stackrel{\mathcal{D}}{=} h - \beta z\sqrt{-u} + \beta z_2'\sqrt{q+u}$$

and again

$$\mathsf{E} Y_j Y_j' = 0\ , \quad \mathsf{E} Y_1' Y_2 = \mathsf{E} Y_1 Y_2' = \frac{\beta^2}{2}\ ,$$

so as before

$$G_1'(u) = \beta^2 \mathsf{E} \frac{1}{\mathrm{ch}^2 Y_1}\frac{1}{\mathrm{ch}^2 Y_2} - 1 \le \beta^2 \mathsf{E} \frac{1}{\mathrm{ch}^4(\beta z\sqrt{q} + h)} - 1 = G_1'(q)\ . \qquad \square$$

Lemma 13.6.4. *Consider u with $|u| \ge q$, and let $c = \mathrm{sign}(u)$. Then*

$$\begin{aligned} r_N(\beta, h, u) \le G(u, m, \lambda) := 2\log 2 &+ \frac{\beta^2}{2}(1-|u|)^2 - \beta^2 m(u^2 - q^2) \\ &+ \frac{1}{m}\mathsf{E}\log \mathsf{E}'(\mathrm{ch}Y_1\mathrm{ch}Y_2\mathrm{ch}\lambda + \mathrm{sh}Y_1\mathrm{sh}Y_2\mathrm{sh}\lambda)^m \\ &- \lambda u \end{aligned} \tag{13.98}$$

where

$$Y_1 = \beta z\sqrt{q} + \beta z'\sqrt{|u| - q} + h\ ; \quad Y_2 = c\beta z\sqrt{q} + c\beta z'\sqrt{|u| - q} + h\ ,$$

and where z and z' are independent standard Gaussian r.v.s.

Proof. Use (13.73) with $c' = c$, $q_1 = q$ and $q_2 = |u|$. $\square$

Proposition 13.6.5. *Under the conditions of Theorem 13.6.2, if $u < -q$ we have*

$$r_N(\beta, h, u) \le G\left(u, \frac{1}{2}, 0\right) < 2\mathrm{SK}(\beta, h)\ .$$

Proof. The Cauchy-Schwarz inequality implies

$$\mathsf{E}'(\mathrm{ch}Y_1\mathrm{ch}Y_2)^{1/2} \le (\mathsf{E}'\mathrm{ch}Y_1\mathsf{E}'\mathrm{ch}Y_2)^{1/2}\ . \tag{13.99}$$

We can have equality only if for some number A we have $\mathrm{ch}Y_1 = A\mathrm{ch}Y_2$ a.e. Since Y_1 and Y_2 have the same distribution this can happen only for $A = 1$, but this is not the case since $c = -1$ and h is not 0 a.e. Therefore we cannot have equality in (13.99), i.e. we have

$$\mathsf{E}'(\mathrm{ch}Y_1\mathrm{ch}Y_2)^{1/2} < (\mathsf{E}'\mathrm{ch}Y_1\mathsf{E}'\mathrm{ch}Y_2)^{1/2}\ .$$

Now

$$\mathsf{E}'\mathrm{ch}Y_1 = \exp\frac{\beta^2}{2}(|u|-q)\mathrm{ch}(\beta z\sqrt{q}+h)$$
$$\mathsf{E}'\mathrm{ch}Y_2 = \exp\frac{\beta^2}{2}(|u|-q)\mathrm{ch}(-\beta z\sqrt{q}+h)\ ,$$

and thus

$$\mathsf{E}\log\mathsf{E}'(\mathrm{ch}Y_1\mathrm{ch}Y_2)^{1/2} < \frac{\beta^2}{2}(|u|-q)+\mathsf{E}\log\mathrm{ch}(\beta z\sqrt{q}+h)$$

because $\beta z\sqrt{q}+h$ and $-\beta z\sqrt{q}+h$ have the same distribution. Therefore (13.98) implies

$$\begin{aligned} G\left(u,\frac{1}{2},0\right) &< 2\log 2+\frac{\beta^2}{2}(1-|u|)^2-\frac{\beta^2}{2}(u^2-q^2)+\frac{\beta^2}{2}(|u|-q) \\ &\quad + 2\mathsf{E}\log\mathrm{ch}(\beta z\sqrt{q}+h) \\ &= 2\mathrm{SK}(\beta,h)\ , \end{aligned}$$

and this concludes the proof. □

If one wishes to extend the present argument to the case where $h=0$, one can find a different argument (along the lines of Proposition 13.6.5 below). The following completes the proof of Theorem 13.6.2.

Proposition 13.6.6. *Under the conditions of Theorem 13.6.2, for $q\le u\le 1$, we have*

$$r_N(\beta,h,u)\le 2\mathrm{SK}(\beta,h)-\frac{(u-q)^2}{K}\ . \tag{13.100}$$

Proof. By (13.98) it suffices to prove that

$$\inf_{m,\lambda} G(u,m,\lambda) < 2\mathrm{SK}(\beta,h)-\frac{(u-q)^2}{K}\ . \tag{13.101}$$

Since $c=\mathrm{sign}(u)=1$, we have

$$Y_1=Y_2=\beta z\sqrt{q}+\beta z'\sqrt{u-q}+h$$

so that

$$G(u,m,\lambda) = 2\log 2+\frac{\beta^2}{2}(1-u)^2-\beta^2 m(u^2-q^2) \tag{13.102}$$
$$+\frac{1}{m}\mathsf{E}\log\mathsf{E}'(\mathrm{ch}^2Y_1\mathrm{ch}\lambda+\mathrm{sh}^2Y_1\mathrm{sh}\lambda)^m-\lambda u\ . \tag{13.103}$$

Now

$$\mathsf{E}'\mathrm{ch}Y_1 = \exp\frac{\beta^2}{2}(u-q)\mathrm{ch}(\beta z\sqrt{q}+h)$$

from which it is straightforward to obtain

$$G\left(u,\frac{1}{2},0\right) = 2\mathrm{SK}(\beta,h)\ . \tag{13.104}$$

Next, let us define

$$W(u) := \frac{\partial G}{\partial\lambda}\left(u,\frac{1}{2},\lambda\right)\Big|_{\lambda=0}\ .$$

Since (as is obvious from the explicit formula) $\partial^2 G/\partial\lambda^2$ is bounded, as in (13.96) we have

$$\inf_\lambda G\left(u,\frac{1}{2},\lambda\right) \le G\left(u,\frac{1}{2},0\right) - \frac{W(u)^2}{K} = 2\mathrm{SK}(\beta,h) - \frac{W(u)^2}{K}\ .$$

A straightforward computation yields

$$W(u) = \mathsf{E}\frac{\mathsf{E}'\mathrm{sh}^2Y_1\mathrm{ch}^{-1}Y_1}{\mathsf{E}'\mathrm{ch}Y_1} - u\ , \tag{13.105}$$

so W is the function of Lemma 13.3.6. When $\beta < 1/10$, (13.46) and (13.47) show that $W'(u) < -1/2$ for all u and since $W(q) = 0$ by (13.45) this shows that $|W(u)| \ge |u-q|/2$ for all $u < 1$. Thus (13.101) is proved for all u in that case.

Therefore it remains only to consider the case where (13.89) and (13.90) hold. Then by (13.45), (13.46) and (13.89) we have $|W(u)| \ge |u-q|/K$ for $|u-q| \le 1/K$ and we have proved (13.101) in that case. To finish the proof we observe that since

$$G(u,m,0) = 2\log 2 + \frac{\beta^2}{2}(1-u)^2 - \beta^2 m(u^2-q^2) + \frac{1}{m}\mathsf{E}\log\mathsf{E}'\mathrm{ch}^{2m}Y_1,$$

it is straightforward that (recalling the notation (13.39)) we have

$$\frac{\partial G}{\partial m}(u,m,0)\Big|_{m=1/2} = 4V(\beta,h,u)\ ,$$

so that, if $u > q$, by (13.88) we have

$$\frac{\partial G}{\partial m}(u,m,0) < 0$$

and

$$\inf_m G(u,m,0) < 2\mathrm{SK}(\beta,h)\ .$$

It is for this part of the argument that (13.88) is really needed and that (13.89) does not suffice. □

13.7 Exponential Inequalities

We have been able to compute $\lim_{N\to\infty} p_N(\beta,h)$ $(= \mathrm{SK}(\beta,h))$ in the entire high-temperature region. In this section we prove that in the "interior" of the high-temperature region, we can achieve a control as good as that of Proposition 1.4.8. Quite naturally we will define the "interior" of the high-temperature region by the conditions (13.88) and (13.89).

Theorem 13.7.1. *Assume that (13.88) and (13.89) hold; then for some number K independent of N we have*

$$\nu\left(\exp\frac{N(R_{1,2}-q)^2}{K}\right)\le 2\,. \tag{13.106}$$

Here and through the rest of the section K denotes a number independent of N, that need not be the same at each occurrence.

The basic idea is to use the cavity method to compute (or at least bound) recursively the quantities $\nu((R_{1,2}-q)^{2k}))$ as in Sections 1.8 and 1.10, and to try to prove

$$\nu\big((R_{1,2}-q)^{2k}\big)\le\left(\frac{Kk}{N}\right)^k\,. \tag{13.107}$$

In order to control the error terms arising in the cavity method we shall use the a priori knowledge that $R_{1,2}-q$ is suitably small, as provided by the next result, which is a therefore a central ingredient in the proof of Theorem 13.7.1.

Proposition 13.7.2. *Under the conditions of Theorem 13.7.1, given $\varepsilon>0$ (independent of N) there is a number K such that*

$$\nu\big(\mathbf{1}_{\{|R_{1,2}-q|\ge\varepsilon\}}\big)\le K\exp\left(-\frac{N}{K}\right)\,. \tag{13.108}$$

Proof. Let us denote by K_0 the constant of (13.90), so that

$$r_N(\beta,h,u)\le 2\mathrm{SK}(\beta,h)-\frac{(u-q)^2}{K_0}\,.$$

Theorem 13.4.1 implies that, for N large enough , $p_N(\beta,h)\ge \mathrm{SK}(\beta,h)-\varepsilon^2/4K_0$, so that, for these values of N, we have

$$|u-q|\ge\varepsilon\quad\Rightarrow\quad r_N(\beta,h,u)\le 2p_N(\beta,h)-\frac{(u-q)^2}{2K_0}\,.$$

We copy the proof of (13.61) to obtain

$$\nu\big(\mathbf{1}_{\{R_{1,2}=u\}}\big)\le K\exp\left(-\frac{N}{K}\right)\,,$$

and we use that there are at most $2N+1$ values of u to consider. □

The use of (13.108) is that, unless k is very large, we have

$$\nu\big(|R_{1,2}-q|^{2k+1}\big) \ll \nu\big((R_{1,2}-q)^{2k}\big) .$$

To see this, we note that since $|R_{1,2}-q| \le 2$ we have

$$\nu\big((R_{1,2}-q)^{2k+1}\big) \le 2^{2k+1}\nu\big(\mathbf{1}_{\{|R_{1,2}-q|\ge\varepsilon\}}\big) + \varepsilon\nu\big((R_{1,2}-q)^{2k}\big)$$

and that by (13.108) the first term is very small (unless k is of order N). In words "higher order terms" in $R_{1,2}-q$ are really smaller.

The next result is technical. It allows, using (13.107) to control some auxiliary error terms. We recall the notation $R^-_{1,2} = \sum_{i\le N-1}\sigma_i^1\sigma_i^2$.

Lemma 13.7.3. *If it is true that*

$$\forall \ell \le k\ , \quad \nu\big((R_{1,2}-q)^{2\ell}\big) \le \left(\frac{K_0\ell}{N}\right)^{\ell} \tag{13.109}$$

and if $K_0 \ge 4$, *then we have*

$$\forall\, j \le 2k\ , \quad \nu(|R_{1,2}-q|^j) \le \left(\frac{K_0(j+1)}{2N}\right)^{j/2} \tag{13.110}$$

$$\nu\big((R^-_{1,2}-q)^{2k}\big) \le 3\left(\frac{K_0(k+1)}{N}\right)^{k} . \tag{13.111}$$

Proof. To prove (13.110), if j is even we use (13.109) for $\ell = j/2$. If j is odd, say $j = 2\ell-1$, we use Hölder's inequality and (13.109) to obtain

$$\begin{aligned}\nu(|R_{1,2}-q|^j) &\le (\nu((R_{1,2}-q)^{2\ell}))^{j/2\ell} \le \left(\left(\frac{K_0\,\ell}{N}\right)^{\ell}\right)^{j/2\ell} \\ &= \left(\frac{K_0\,\ell}{N}\right)^{j/2} = \left(\frac{K_0(j+1)}{2N}\right)^{j/2} .\end{aligned}$$

To prove (13.111), we expand the expression

$$(R^-_{1,2}-q)^{2k} = \Big(R_{1,2}-q-\frac{\varepsilon_1\varepsilon_2}{N}\Big)^{2k}$$

to get, using (13.110),

$$\nu\big((R^-_{1,2}-q)^{2k}\big) \le \sum_{0\le j\le 2k}\binom{2k}{j}\frac{1}{N^{2k-j}}\nu(|R_{1,2}-q|^j)$$

$$\leq \sum_{0\leq j\leq 2k} \binom{2k}{j} \frac{1}{N^{2k-j}} \left(\frac{K_0(j+1)}{2N}\right)^{j/2}$$

$$\leq \sum_{0\leq j\leq 2k} \binom{2k}{j} \frac{1}{N^{2k-j}} \left(\frac{K_0(2k+1)}{2N}\right)^{j/2}$$

$$= \left(\frac{1}{N} + \sqrt{\frac{K_0(2k+1)}{2N}}\right)^{2k}$$

$$= \left(\frac{K_0(2k+1)}{2N}\right)^{k} \left(1 + \sqrt{\frac{2}{NK_0(2k+1)}}\right)^{2k} . \quad (13.112)$$

Finally since $1+x \leq \exp x$, we have

$$\left(1 + \sqrt{\frac{2}{NK_0(2k+1)}}\right)^{2k} \leq \exp\left(2\sqrt{\frac{k}{NK_0}}\right) \leq e \leq 3$$

provided that we assume $K_0 \geq 4$ and since $k \leq N$. □

Proof of Theorem 13.7.1. We try to prove by induction on k that (13.109) holds for a suitable constant K_0. By symmetry among sites,

$$\nu\big((R_{1,2}-q)^{2k+2}\big) = \nu\big((\varepsilon_1\varepsilon_2 - q)(R_{1,2}-q)^{2k+1}\big) . \quad (13.113)$$

Using the inequality

$$|x^{p+1} - y^{p+1}| \leq (p+1)|x-y|(|x|^p + |y|^p) \quad (13.114)$$

for $p = 2k$, $x = R_{1,2} - q$ and $y = R^-_{1,2} - q$, we obtain (since $|\varepsilon_1\varepsilon_2 - q| \leq 2$ and $|x - y| = |\varepsilon_1\varepsilon_2| = 1$)

$$\nu\big((\varepsilon_1\varepsilon_2 - q)(R_{1,2}-q)^{2k+1}\big) = \nu\big((\varepsilon_1\varepsilon_2 - q)(R^-_{1,2}-q)^{2k+1}\big) + \mathcal{R} ,$$

where

$$|\mathcal{R}| \leq \frac{2(2k+1)}{N}\big(\nu\big((R_{1,2}-q)^{2k}\big) + \nu\big((R^-_{1,2}-q)^{2k}\big)\big)$$

$$\leq \frac{L(k+1)}{N}\left(\frac{K_0(k+1)}{N}\right)^k , \quad (13.115)$$

using (13.109) and (13.111).

Let $f = (\varepsilon_1\varepsilon_2 - q)(R^-_{1,2}-q)^{2k+1}$. Recalling ν_t as in Lemma 1.6.3 we have

$$|\nu(f) - \nu_0(f) - \nu'_0(f)| \leq \sup |\nu''_t(f)| ,$$

and iterating (1.150) and using Hölder's inequality we obtain

$$|\nu_t''(f)| \le K\nu\big(|R_{1,2}^- - q|^{2k+3}\big) .$$

We recall the numbers

$$b(2) = b^2(1-q^2)\ ;\ b(1) = \beta^2(q-q^2)\ ;\ b(0) = \beta^2(\widehat{q}-q^2) .$$

We recall that $\nu_0(f) = 0$, and we compute $\nu_0'(f)$ using Lemma 1.8.2 to get

$$\nu_0'(f) = b(2)\nu_0\big((R_{1,2}^- - q)^{2k+2}\big) - 4b(1)\nu_0\big((R_{1,3}^- - q)(R_{1,2}^- - q)^{2k+1}\big)$$
$$+ 3b(0)\nu_0\big((R_{3,4}^- - q)(R_{1,2}^- - q)^{2k+1}\big) .$$

Using now that

$$|\nu(f') - \nu_0(f')| \le \sup|\nu_t'(f')|$$

for $f' = (R_{\ell,\ell'}^- - q)(R_{1,2}^- - q)^{2k+1}$, computing $\nu_t'(f)$ using Lemma 1.8.2 and using again Hölder's inequality we get that $\nu(f') = \nu_0(f') + \mathcal{R}$ where $|\mathcal{R}| \le K\nu(|R_{1,2}^- - q|^{2k+3})$. Consequently,

$$\nu_0'(f) = b(2)\nu\big((R_{1,2}^- - q)^{2k+2}\big) - 4b(1)\nu\big((R_{1,3}^- - q)(R_{1,2}^- - q)^{2k+1}\big)$$
$$+ 3b(0)\nu\big((R_{3,4}^- - q)(R_{1,2}^- - q)^{2k+1}\big) + \mathcal{R} ,$$

where $|\mathcal{R}| \le K\nu(|R_{1,2}^- - q|^{2k+3})$. Combining these estimates yields

$$\begin{aligned} &\nu\big((R_{1,2} - q)^{2k+2}\big) \\ &= b(2)\nu\big((R_{1,2}^- - q)^{2k+2}\big) - 4b(1)\nu\big((R_{1,3}^- - q)(R_{1,2}^- - q)^{2k+1}\big) \\ &+ 3b(0)\nu\big((R_{3,4}^- - q)(R_{1,2}^- - q)^{2k+1}\big) + \mathcal{R} , \end{aligned} \tag{13.116}$$

where

$$|\mathcal{R}| \le K\nu(|R_{1,2}^- - q|^{2k+3}) + \frac{L(k+1)}{N}\left(\frac{K_0(k+1)}{N}\right)^k .$$

Using (13.114) for $p = 2k+1$, and since $|R_{1,2} - q| \le 2$ and $|R_{1,2}^- - q| \le 2$, we get

$$\begin{aligned} &\big|\nu\big((R_{1,2}^- - q)^{2k+2}\big) - \nu\big((R_{1,2} - q)^{2k+2}\big)\big| \\ &\le \frac{2k+2}{N}\big(\nu\big(|R_{1,2} - q|^{2k+1}\big) + \nu\big(|R_{1,2}^- - q|^{2k+1}\big)\big) \\ &\le 2\frac{2k+2}{N}\big(\nu\big((R_{1,2} - q)^{2k}\big) + \nu\big((R_{1,2}^- - q)^{2k}\big)\big) \\ &\le \frac{L(k+1)}{N}\left(\frac{K_0(k+1)}{N}\right)^k , \end{aligned} \tag{13.117}$$

using (13.111) in the last line. Performing similar near-trivial bounds for the other two terms of (13.116) we have reached the relation

$$U = b(2)U - 4b(1)V + 3b(0)W + \mathcal{R}_1 , \tag{13.118}$$

where

$$U = \nu\big((R_{1,2} - q)^{2k+2}\big)\ ,\quad V = \nu\big((R_{1,3} - q)(R_{1,2} - q)^{2k+1}\big)$$

$$W = \nu\big((R_{3,4} - q)(R_{1,2} - q)^{2k+1}\big)$$

and

$$|\mathcal{R}_1| \le K\nu\big(|R_{1,2} - q|^{2k+3}\big) + \frac{L(k+1)}{N}\left(\frac{K_0(k+1)}{N}\right)^k\ . \tag{13.119}$$

In a similar manner (as in (1.231) and (1.232)) we prove the relations

$$V = b(1)U + (b(2) - 2b(1) - 3b(0))V + (6b(0) - 3b(1))W + \mathcal{R}_2 \tag{13.120}$$

$$W = b(0)U + (4b(1) - 8b(0))V + (b(2) - 8b(1) + 10b(0))W + \mathcal{R}_3\ , \tag{13.121}$$

where $\mathcal{R}_2$ and $\mathcal{R}_3$ satisfy bounds as (13.119). In Section 1.8 we have seen that the transpose of the matrix (1.233) given by (13.118), (13.120) and (13.121) has two eigenvalues (one of which with multiplicity 2), that are given by the formulas (1.234) and (1.235). They are both $\neq 1$. For the second eigenvalue this follows from Lemma 1.9.3. For the first eigenvalue this follows from (13.89) and the fact that this eigenvalue is

$$\beta^2 \mathsf{E}\frac{1}{\mathrm{ch}^4(\beta z\sqrt{q} + h)}\ .$$

Therefore the equations (13.118), (13.120) and (13.121) imply that

$$U = \nu\big((R_{1,2} - q)^{2k+2}\big) \le K_1\nu\big(|R^-_{1,2} - q|^{2k+3}\big) + \frac{K(k+1)}{N}\left(\frac{K_0(k+1)}{N}\right)^k\ . \tag{13.122}$$

Finally let us choose $\varepsilon = 1/4K_1$, so that (13.108) implies

$$\nu\big(\mathbf{1}_{\{|R_{1,2}-q|\ge\varepsilon\}}\big) \le K\exp\left(-\frac{N}{K}\right)$$

and for N large, say $N \ge N_0$,

$$\nu\big(\mathbf{1}_{\{|R^-_{1,2}-q|\ge 2\varepsilon\}}\big) \le K\exp\left(-\frac{N}{K}\right)\ .$$

Therefore, for such N,

$$\nu\big(|R^-_{1,2} - q|^{2k+3}\big) \le 2^{2k+3}K\exp\left(-\frac{N}{K}\right) + 2\varepsilon\nu\big((R^-_{1,2} - q)^{2k+2}\big)$$

and using (13.117) we have shown that

$$\nu\big(|R^-_{1,2}-q|^{2k+3}\big)\le 2\varepsilon U+2^{2k+2}K\exp\left(-\frac{N}{K}\right)+\frac{K(k+1)}{N}\left(\frac{K_0(k+1)}{N}\right)^k .$$

Recalling that $\varepsilon=1/4K_1$ and combining with (13.122) we see that

$$U\le\frac{1}{2}U+2^{2k+3}K\exp\left(-\frac{N}{K}\right)+\frac{K(k+1)}{N}\left(\frac{K_0(k+1)}{N}\right)^k ,$$

i.e.

$$U\le 2^{2k+3}K\exp\left(-\frac{N}{K}\right)+\frac{K_2(k+1)}{N}\left(\frac{K_0(k+1)}{N}\right)^k .$$

Let us now assume, as we may, that $K_0\ge 2K_2$. This yields

$$U\le\frac{1}{2}\left(\frac{K_0(k+1)}{N}\right)^{k+1}+2^{2k}K_3\exp\left(-\frac{N}{K_3}\right) .$$

This proves that $U\le(K_0(k+1)/N)^{k+1}$ (and completes the induction) provided k is small enough that

$$K_3\exp\left(-\frac{N}{K_3}\right)\le\left(\frac{K_0(k+1)}{4N}\right)^{k+1} . \tag{13.123}$$

Let us denote by k_0 the smallest integer k for which (13.123) fails, so we have proved by induction that (13.109) holds for $k\le k_0$. As usual, the case $k\ge k_0$ is obtained by a near trivial argument. By definition of k_0, we have

$$\frac{K_0(k_0+1)}{N}\le 4K_3\exp\left(-\frac{N}{K_3(k_0+1)}\right) ,$$

so that $(k_0+1)/N\ge 1/K_4$, and then, for $k\ge k_0$,

$$\nu\big((R_{1,2}-q)^{2k+2}\big)\le 2^{2k+2}\le\left(\frac{4K_4(k_0+1)}{N}\right)^{k+1} .$$

It then follows that for all k we have

$$\nu\big((R_{1,2}-q)^{2k}\big)\le\left(\frac{K_5k}{N}\right)^k ,$$

where $K_5=\max(4K_4,K_0)$, and this proves Theorem 13.7.1. □

Theorem 13.7.4. *Under (13.88) and (13.89), we have*

$$|p_N(\beta,h)-\mathrm{SK}(\beta,h)|\le\frac{K}{N} .$$

Proof. The proof of Theorem 13.7.1 shows that we have $\nu_t((R_{1,2}-q)^2)\le K/N$ uniformly on $t\le 1$. □

14. The Parisi Formula

14.1 Introduction

In this chapter we obtain a considerable extension of the results of Chapter 13: we compute in the limit the quantity $p_N(\beta, h)$ for any values of β and h. The result, called the Parisi formula, is very beautiful but is not immediate to explain, and its proof is rather involved. A large part of the motivation of Chapter 13 is to present a simpler special case of the main arguments of the present chapter, and we advise the reader to first master the first six sections of Chapter 13 before attempting to read anything at all here. Starting with Section 14.7 matters get a bit technical, but let us insist that the reader certainly does not need to master the details of the computations in order to enjoy the next chapter, which attempts to describe the structure underlying the Parisi formula.

14.2 Poisson-Dirichlet Cascades

In this section we briefly discuss certain objects that we call Poisson-Dirichlet cascades, which seem intrinsically connected to the low-temperature phase of spin-glass systems. These objects are very pretty, but we will refrain from the temptation of studying them for their own sake, and will present only the results that are really helpful in the sequel.

Consider an integer k and the set $\mathbb{N}^{*k}$, where $\mathbb{N}^* = \{1, 2, \ldots, \}$. We will denote by α a sequence $(j_1, \ldots, j_k)$ in $\mathbb{N}^{*k}$. It could be useful to think of $\mathbb{N}^{*k}$ as a tree. As we scan the integers $j_1, j_2, \ldots$ we discover which branch we follow at each node. For $p \le k$ we write $\alpha|p = (j_1, \ldots, j_p) \in \mathbb{N}^{*p}$. Let us fix a sequence $0 < m_1 < m_2 < \ldots < m_k < 1$. Given this sequence we are going to construct random quantities (u^*_α), $\alpha \in \mathbb{N}^{*k}$. Let us first consider a non-increasing rearrangement $(u_j)_{j\ge 1}$ of a Poisson point process of intensity measure $x^{-m_1-1}\mathrm{d}x$. For each integer j_1, we consider a non-increasing rearrangement $(u_{j_1,j})_{j\ge 1}$ of a Poisson point process of intensity measure $x^{-m_2-1}\mathrm{d}x$. These are all independent of each other and of the sequence (u_j). More generally, for each $1 \le p \le k$ and each integers $j_1, \ldots, j_{p-1}$ we consider a non-increasing rearrangement $(u_{j_1,\ldots,j_{p-1},j})_{j\ge 1}$ of a Poisson point process of intensity measure $x^{-m_p-1}\mathrm{d}x$. All these are independent of each other. For $\alpha \in \mathbb{N}^{*k}$, we define

M. Talagrand, *Mean Field Models for Spin Glasses*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys in Mathematics 55, DOI 10.1007/978-3-642-22253-5_7,

$$\begin{aligned} u^*_\alpha &= u_{\alpha|1} u_{\alpha|2} \cdots u_{\alpha|k-1} u_\alpha \\ &= u_{j_1} u_{j_1 j_2} \cdots u_{j_1 \ldots j_k} \ . \end{aligned} \tag{14.1}$$

It will be shown in (14.9) below (taking $F = 0$ there) that $\sum_\gamma u^*_\gamma < \infty$ a.s. The family of weights $(v_\alpha)_{\alpha \in \mathbb{N}^{*k}}$ where

$$v_\alpha = \frac{u^*_\alpha}{\sum u^*_\gamma} \tag{14.2}$$

will be called the *Poisson-Dirichlet cascade associated with the sequence* $m_1, \ldots, m_k$. When $k = 1$, the sequence v_α has the Poisson-Dirichlet distribution Λ_{m_1} first defined in Section 13.1 page 314.

In this chapter and the next, we will often consider random weights. Weights associated to a Poisson-Dirichlet cascade will be denoted by (v_α) while we will denote by (w_α) weights that need not be of this type.

Before we can describe some remarkable properties of this object, we must explain another procedure that will be fundamental, maybe even more so than Poisson-Dirichlet cascades. Consider a metric space T. (The case $T = \mathbb{R}^N$ or $T = \mathbb{R}$ will be the most useful.) Consider a function $F : T^k \to \mathbb{R}$. Consider also independent random maps $\mathbf{z}_1, \ldots, \mathbf{z}_k$ valued in T, and define the r.v.

$$F_{k+1} = F(\mathbf{z}_1, \ldots, \mathbf{z}_k) \ . \tag{14.3}$$

We will assume that

$$\mathsf{E} \exp F_{k+1} < \infty \ ; \ \mathsf{E}|F_{k+1}| < \infty \ . \tag{14.4}$$

For $1 \le p \le k$, let us denote by E_p expectation in the r.v.s $\mathbf{z}_p, \ldots, \mathbf{z}_k$, and let us define recursively the r.v.s

$$F_p = \frac{1}{m_p} \log \mathsf{E}_p \exp m_p F_{p+1} \ , \tag{14.5}$$

so that F_p depends only on the r.v.s $\mathbf{z}_1, \ldots, \mathbf{z}_{p-1}$, and in particular F_1 is a number.

For each $p \le k$ and integers $j_1, \ldots, j_p$, let us consider independent copies $\mathbf{z}_{p, j_1, \ldots, j_p}$ of $\mathbf{z}_p$. These are all independent of each other. The reader should observe the similarity with the procedure by which we define the r.v.s u^*_α. To lighten notation for $\alpha = (j_1, \ldots, j_k) \in \mathbb{N}^{*k}$ we write

$$\mathbf{z}_{p,\alpha} = \mathbf{z}_{p, j_1, \ldots, j_p} \ . \tag{14.6}$$

This variable depends only on p and $\alpha|p$. The procedure of defining the r.v.s $\mathbf{z}_{p,\alpha}$ from the r.v.s $\mathbf{z}_p$ will occur a great many times, and the notation (14.6) remains in force through the chapter.

Let us then define

$$F(\alpha) = F(\mathbf{z}_{1,\alpha}, \ldots, \mathbf{z}_{k,\alpha}) \ . \tag{14.7}$$

Let us attract the attention of the reader on this unusual (but convenient) notation: the quantity $F(\alpha)$ is a r.v. indexed by α. The crucial fact is as follows.

Theorem 14.2.1. *Assuming (14.4) we have*

$$\mathsf{E}\log\sum_{\alpha} v_\alpha \exp F(\alpha) = F_1 \ , \tag{14.8}$$

where v_α is defined in (14.2).

First we prove that (14.8) holds when $k = 1$. In that case $\alpha \in \mathbb{N}^*$. When $(v_\alpha)_{\alpha\geq 1}$ is a sequence with distribution Λ_m and the r.v.s $(F(\alpha))_{\alpha\geq 1}$ are i.i.d. copies of a r.v. X, then the r.v.s $\exp F(\alpha)$ are i.i.d. copies of $V = \exp X$ and (13.12) entails that

$$\mathsf{E}\log\sum_{\alpha} v_\alpha \exp F(\alpha) = \frac{1}{m}\log \mathsf{E}V^m = \frac{1}{m}\log\mathsf{E}\exp mX = F_1 \ ,$$

so that (14.8) holds when $k = 1$.

Theorem 14.2.1 links Poisson-Dirichlet cascades and the recursive construction of the r.v.s F_p. For the time being, it might be correct to think that the truly fundamental procedure is the recursive construction of the r.v.s F_p, and that (14.8) is a "one-step" method to compute the complicated quantity F_1.

The secret about Poisson Dirichlet cascades is to be unimpressed by the definition, and to work by induction over k, the case $k = 1$ being always the crucial case.

Proposition 14.2.2. *Consider $0 < m_0 < m_1 < \ldots < m_k < 1$, and (u^*_α) as in (14.1). Then we have*

$$\mathsf{E}\Big(\sum_{\alpha} u^*_\alpha \exp F(\alpha)\Big)^{m_0} = (\exp m_0 F_1)\prod_{1\leq p\leq k}\bigg(\mathsf{E}\Big(\sum_j u_j^{(p)}\Big)^{m_{p-1}}\bigg)^{m_0/m_{p-1}} \ , \tag{14.9}$$

where $(u_j^{(p)})_{j\geq 1}$ is an enumeration of a Poisson point process with intensity measure $x^{-m_p-1}\mathrm{d}x$.

Lemma 14.2.3. *Consider a r.v. Y with $\mathsf{E}\exp Y < \infty$ and $\mathsf{E}|Y| < \infty$. Then*

$$\lim_{m\to 0^+}\frac{1}{m}\log\mathsf{E}\exp mY = \mathsf{E}Y \ .$$

Proof. First, we note that Jensen's inequality implies

$$\frac{1}{m}\log \mathsf{E}\exp mY \geq \mathsf{E}Y \; .$$

Next, defining $\varphi(x) = \exp x - 1 - x$ it holds that

$$\mathsf{E}\exp mY = 1 + m\mathsf{E}Y + \mathsf{E}\varphi(mY) \; ,$$

and since $\log(1+x) \leq x$ we have

$$\frac{1}{m}\log \mathsf{E}\exp mY \leq \mathsf{E}Y + \frac{1}{m}\mathsf{E}\varphi(mY) \; . \tag{14.10}$$

Since the function φ is convex with $\varphi(0) = \varphi'(0) = 0$, for each y the function $m \mapsto \varphi(ym)/m$ is increasing and has limit 0 as $m \to 0$. Moreover $\varphi \geq 0$ and $\mathsf{E}\varphi(Y) < \infty$ since $\varphi(x) \leq L(\exp x + |x|)$. Therefore as $m \to 0$ the last term of (14.10) goes to zero by dominated convergence. □

Proof of Theorem 14.2.1. We write (14.9) as

$$\left(\mathsf{E}\left(\sum_\alpha u^*_\alpha \exp F(\alpha)\right)^{m_0}\right)^{1/m_0} = (\exp F_1)\prod_{1\leq p\leq k}\left(\mathsf{E}\left(\sum_j u^{(p)}_j\right)^{m_{p-1}}\right)^{1/m_{p-1}} .$$

Taking logarithm, we get that for a certain number $C(m_0,\ldots,m_k)$ independent of F it holds that

$$\frac{1}{m_0}\log \mathsf{E}\left(\sum_\alpha u^*_\alpha \exp F(\alpha)\right)^{m_0} = F_1 + C(m_0,\ldots,m_k) \; . \tag{14.11}$$

We now use Lemma 14.2.3 with $Y = \log\sum_\alpha u^*_\alpha \exp F(\alpha)$. It follows from (14.4) that $\mathsf{E}\exp Y < \infty$. To see that $\mathsf{E}|Y| < \infty$, we have only to control $\mathsf{E}Y^-$ where $Y^- = \max(-Y, 0)$. For this we simply bound $\sum_\alpha u^*_\alpha \exp F(\alpha)$ from below by the term where $\alpha = (1,\ldots,1)$, and we observe that if u_1 is the largest point of a Poisson point process of intensity measure x^{-m-1} then $P(u_1 < t) = \exp(-t^{-m}/m)$ so that in particular $\mathsf{E}(\log u_1)^- < \infty$ and hence $\mathsf{E}(\log u^*_\alpha)^- < \infty$ for $\alpha = \{1,\ldots,1\}$.

Letting $m_0 \to 0$ in (14.11) yields

$$\mathsf{E}\log\sum_a u^*_\alpha \exp F(\alpha) = F_1 + C \; ,$$

where C depends only on $m_1,\ldots,m_k$. Using this for $F(\alpha) \equiv 0$, we obtain

$$\mathsf{E}\log\sum_\alpha u^*_\alpha = C \; ,$$

and subtraction of these relations yields Theorem 14.2.1. □

Proof of Proposition 14.2.2. We proceed by induction over k. For $k = 1$ this follows from (13.9). For the induction step from $k-1$ to k, we consider the quantity

$$W_{j_1} = \sum u_{j_1,j_2} \cdots u_{j_1,j_2,\ldots,j_k} \exp F(\mathbf{z}_{1,j_1}, \mathbf{z}_{1,j_1,j_2}, \ldots, \mathbf{z}_{k,j_1,j_2,\ldots,j_k}) ,$$

where the sum is over all values of $j_2, \ldots, j_k$. The sequence $(W_j)_{j\geq 1}$ is i.i.d, and moreover

$$\sum_\alpha u^*_\alpha \exp F(\alpha) = \sum_j u_j W_j .$$

We use (13.9) for $m' = m_0$ and $m = m_1$ to get

$$\mathsf{E}\Big(\sum u_j W_j\Big)^{m_0} = (\mathsf{E}W_1^{m_1})^{m_0/m_1} \mathsf{E}\Big(\sum u_j\Big)^{m_0} . \tag{14.12}$$

We observe that $\mathsf{E}W_1^{m_1} = \mathsf{E}\mathsf{E}_2 W_1^{m_1}$ and we compute $\mathsf{E}_2 W_1^{m_1}$ through the induction hypothesis:

$$\mathsf{E}_2 W_1^{m_1} = (\exp m_1 F_2) \prod_{2\leq p\leq k} \left(\mathsf{E}\Big(\sum_j u_j^{(p)}\Big)^{m_{p-1}}\right)^{m_1/m_{p-1}} .$$

We also observe from (14.5) that $\mathsf{E}\exp m_1 F_2 = \exp m_1 F_1$, so that, taking expectation in the previous inequality yields

$$\mathsf{E}W_1^{m_1} = (\exp m_1 F_1) \prod_{2\leq p\leq k} \left(\mathsf{E}\Big(\sum_j u_j^{(p)}\Big)^{m_{p-1}}\right)^{m_1/m_{p-1}} .$$

We then substitute this relation in (14.12) to conclude the induction. □

14.3 Fundamental Identities

In this section we prove an extension of Theorem 13.1.6, Theorem 14.3.5 below.

We consider a bounded function $U : T^k \to \mathbb{R}$, and, keeping the notation of Section 14.2 and in particular (14.6) we define the r.v.s

$$U(\alpha) = U(\mathbf{z}_{1,\alpha}, \ldots, \mathbf{z}_{k,\alpha})$$

so that $U(\alpha)$ is to the function U what $F(\alpha)$ is to the function F. We define

$$F_t(\alpha) = F(\alpha) + tU(\alpha) .$$

The basic idea is that we will obtain remarkable identities by differentiating once or twice in t the identity (14.8) written for F_t rather than F. We define

$$U_{k+1} = U(\mathbf{z}_1, \ldots, \mathbf{z}_k) \ ,$$

so that U_{k+1} is to the function U what F_{k+1} is to the function F. We further define $F_{k+1,t} = F_{k+1} + tU_{k+1}$, and the functions $F_{p,t}$ through the relation (14.5), that is

$$F_{p,t} = \frac{1}{m_p} \log \mathsf{E}_p \exp m_p F_{p+1,t} \ , \tag{14.13}$$

so that by (14.8)

$$F_{1,t} = \mathsf{E} \log \sum_\alpha v_\alpha \exp F_t(\alpha) \tag{14.14}$$

and, since $U(\alpha) = \partial F_t(\alpha)/\partial t$,

$$\frac{\partial}{\partial t} F_{1,t} = \mathsf{E} \frac{\sum_\alpha v_\alpha U(\alpha) \exp F_t(\alpha)}{\sum_\alpha v_\alpha \exp F_t(\alpha)} \ . \tag{14.15}$$

Here, and at many places below some work is needed to show that one can exchange the expectation and the derivative. It does not seem appropriate to spend the reader's energy on this since if she is able to follow the present chapter till the end she will no doubt have the skills to fill in such details.

For a (possibly random) function $U(\alpha)$ of α it is convenient to define

$$\langle U \rangle_t = \frac{\sum_\alpha v_\alpha U(\alpha) \exp F_t(\alpha)}{\sum_\alpha v_\alpha \exp F_t(\alpha)} \ . \tag{14.16}$$

When $t = 0$ we omit the subscript 0 and we write

$$\langle U \rangle = \frac{\sum_\alpha v_\alpha U(\alpha) \exp F(\alpha)}{\sum_\alpha v_\alpha \exp F(\alpha)} \ . \tag{14.17}$$

It will also be convenient to consider "replicas". In (14.17) we make an average of a function of α. Throughout the chapter we denote by γ a replica of α. Thus, if $U^\sim(\alpha, \gamma)$ is a function of α and γ we define

$$\langle U^\sim \rangle = \frac{\sum_{\alpha,\gamma} v_\alpha v_\gamma U^\sim(\alpha,\gamma) \exp F(\alpha) \exp F(\gamma)}{\left(\sum_\alpha v_\alpha \exp F(\alpha)\right)^2} \ . \tag{14.18}$$

To compute $\partial F_{1,t}/\partial t$, we start with the relation

$$\frac{\partial}{\partial t} F_{k+1,t} = U_{k+1} \ , \tag{14.19}$$

and we aim to compute $\partial F_{p,t}/\partial t$ by differentiating the relation (14.13); this gives

$$\frac{\partial}{\partial t} F_{p,t} = \frac{\mathsf{E}_p \frac{\partial F_{p+1,t}}{\partial t} \exp m_p F_{p+1,t}}{\mathsf{E}_p \exp m_p F_{p+1,t}} \ . \tag{14.20}$$

Now, (14.13) implies

$$\mathsf{E}_p \exp m_p F_{p+1,t} = \exp m_p F_{p,t}$$

and since $F_{p,t}$ does not depend on $\mathbf{z}_p, \ldots, \mathbf{z}_k$ we get from (14.20) that

$$\begin{aligned} \frac{\partial}{\partial t} F_{p,t} &= \frac{\mathsf{E}_p \frac{\partial F_{p+1,t}}{\partial t} \exp m_p F_{p+1,t}}{\exp m_p F_{p,t}} \\ &= \mathsf{E}_p \frac{\frac{\partial F_{p+1,t}}{\partial t} \exp m_p F_{p+1,t}}{\exp m_p F_{p,t}} \\ &= \mathsf{E}_p \frac{\partial}{\partial t} F_{p+1,t} W_{p,t} \ , \end{aligned} \tag{14.21}$$

where

$$W_{p,t} = \exp m_p (F_{p+1,t} - F_{p,t}) \ . \tag{14.22}$$

Let us observe two important properties of the quantity $W_{p,t}$. It does not depend on the r.v.s $\mathbf{z}_{p+1}, \ldots, \mathbf{z}_k$, but only in the r.v.s $\mathbf{z}_1, \ldots, \mathbf{z}_p$; and, as follows from (14.13) it satisfies

$$\mathsf{E}_p W_{p,t} = 1 \ . \tag{14.23}$$

Now we prove by decreasing induction over p that

$$\frac{\partial}{\partial t} F_{p,t} = \mathsf{E}_p (W_{p,t} \cdots W_{k,t} U_{k+1}) \ . \tag{14.24}$$

For $p = k+1$ this is simply (14.19). For $p = k$ this follows from (14.21). Using (14.24) for $p+1$ instead of p, and substituting in (14.21) we get

$$\frac{\partial}{\partial t} F_{p,t} = \mathsf{E}_p \big(W_{p,t} \mathsf{E}_{p+1} (W_{p+1,t} \cdots W_{k,t} U_{k+1}) \big)$$

and since $W_{p,t}$ does not depend on $\mathbf{z}_{p+1}, \ldots, \mathbf{z}_k$ we can move this term inside the expectation E_{p+1} and use that $\mathsf{E}_p = \mathsf{E}_p \mathsf{E}_{p+1}$ to complete the induction. For $p = 1$ we obtain

$$\frac{\partial}{\partial t} F_{1,t} = \mathsf{E}(W_{1,t} \cdots W_{k,t} U_{k+1}) \ , \tag{14.25}$$

and comparing with (14.15) yields

$$\mathsf{E}(W_{1,t} \cdots W_{k,t} U_{k+1}) = \mathsf{E}\langle U \rangle_t \ , \tag{14.26}$$

and in particular, for $t = 0$, with the notation (14.17),

$$\mathsf{E}(W_1 \cdots W_k U_{k+1}) = \mathsf{E}\langle U \rangle \ . \tag{14.27}$$

Exercise 14.3.1. Despite the fact that the notations are absolutely incompatible, prove that the case $k = 1$ of this formula recovers (13.13).

Let us now differentiate (14.26) at $t = 0$. From (14.22) we observe that

$$\frac{\partial}{\partial t} W_{p,t} = m_p \left(\frac{\partial}{\partial t} F_{p+1,t} - \frac{\partial}{\partial t} F_{p,t} \right) W_{p,t} \; . \tag{14.28}$$

To lighten notation we write

$$\partial F_p = \frac{\partial}{\partial t} F_{p,t} \Big|_{t=0} \; ,$$

and differentiation of (14.26) at $t = 0$ yields

$$\mathsf{E}(\langle U^2 \rangle - \langle U \rangle^2) = \mathsf{E}\Bigl(W_1 \cdots W_k U_{k+1} \sum_{1 \le p \le k} m_p (\partial F_{p+1} - \partial F_p) \Bigr) \; . \tag{14.29}$$

It is convenient to set

$$m_0 = 0 \; , \quad m_{k+1} = 1$$

so that, since $\partial F_{k+1} = U_{k+1}$,

$$\sum_{1 \le p \le k} m_p (\partial F_{p+1} - \partial F_p) = - \sum_{1 \le p \le k+1} (m_p - m_{p-1}) \partial F_p + U_{k+1} \; . \tag{14.30}$$

Using the identity (14.27) for U^2 rather than U, we obtain

$$\mathsf{E}\langle U^2 \rangle = \mathsf{E}(W_1 \cdots W_k U_{k+1}^2) \; ,$$

and combining with (14.29) and (14.30) yields

$$\mathsf{E}\langle U \rangle^2 = \sum_{1 \le p \le k+1} (m_p - m_{p-1}) \mathsf{E}(\partial F_p W_1 \cdots W_k U_{k+1}) \; . \tag{14.31}$$

Now, using (14.24) at $t = 0$ we get

$$\partial F_p = \mathsf{E}_p (W_p \cdots W_k U_{k+1}) \; .$$

Since $W_1, \ldots, W_{p-1}$ and $\mathsf{E}_p(W_1 \cdots W_k U_{k+1})$ do not depend on the r.v.s $\mathbf{z}_p, \ldots, \mathbf{z}_k$, we have

$$\begin{aligned}
\mathsf{E}(\partial F_p W_1 \cdots W_k U_{k+1}) &= \mathsf{E}\bigl(W_1 \cdots W_k U_{k+1} \mathsf{E}_p (W_p \cdots W_k U_{k+1}) \bigr) \\
&= \mathsf{E}\mathsf{E}_p \bigl(W_1 \cdots W_k U_{k+1} \mathsf{E}_p (W_p \cdots W_k U_{k+1}) \bigr) \\
&= \mathsf{E}\bigl(\mathsf{E}_p (W_1 \cdots W_k U_{k+1}) \mathsf{E}_p (W_p \cdots W_k U_{k+1}) \bigr) \\
&= \mathsf{E}\bigl(W_1 \cdots W_{p-1} (\mathsf{E}_p W_p \cdots W_k U_{k+1})^2 \bigr) \; ,
\end{aligned}$$

and therefore (14.31) implies

$$\mathsf{E}\langle U \rangle^2 = \sum_{1 \le p \le k+1} (m_p - m_{p-1}) \mathsf{E}\bigl(W_1 \cdots W_{p-1} (\mathsf{E}_p W_p \cdots W_k U_{k+1})^2 \bigr) \; . \tag{14.32}$$

The left-hand side is the average

$$\mathsf{E}\frac{\sum_{\alpha,\gamma} v_\alpha v_\gamma U(\alpha)U(\gamma)\exp(F(\alpha)+F(\gamma))}{\left(\sum_\alpha v_\alpha \exp F(\alpha)\right)^2}\ ,$$

that is therefore expressed as a weighted sum of certain averages. It would be nice if each of these pieces should correspond to an average over a natural set. There is a simple trick that achieves this.

Let us fix an integer $1 \le r \le k$ and let us replace $U_{k+1} = U(\mathbf{z}_1, \ldots, \mathbf{z}_k)$ by

$$U'_{k+1} = \eta_r U(\mathbf{z}_1, \ldots, \mathbf{z}_k) = \eta_r U_{k+1}\ ,$$

where η_r is a random sign (independent of all the $\mathbf{z}_p$) that "goes with $\mathbf{z}_r$". That is, if $\eta_{j_1,\ldots,j_r}$ are independent random signs, then

$$U'(\alpha) = \eta_{\alpha|r} U(\mathbf{z}_{1,\alpha}, \ldots, \mathbf{z}_{k,\alpha})\ .$$

To be more formal, we replace the map $U : T^k \to \mathbb{R}$ by the map $U' : (T^k \times \{-1,1\})^k \to \mathbb{R}$ given by

$$U'((\mathbf{x}_1, \varepsilon_1), \ldots, (\mathbf{x}_k, \varepsilon_k)) = \varepsilon_r U(\mathbf{x}_1, \ldots, \mathbf{x}_k)\ ,$$

and we replace the r.v.s $\mathbf{z}_p$ by the r.v.s $\mathbf{z}'_p = (\mathbf{z}_p, \eta_p)$ where η_p are independent random signs independent of the r.v.s $(\mathbf{z}_p)$. We observe that

$$\mathsf{E}\eta_{\alpha|r}\eta_{\gamma|r} = 1 \text{ if } \alpha|r = \gamma|r$$
$$\mathsf{E}\eta_{\alpha|r}\eta_{\gamma|r} = 0 \text{ if } \alpha|r \ne \gamma|r\ ,$$

so that, taking first expectation in the r.v.s η_α,

$$\mathsf{E}\langle U'\rangle^2 = \mathsf{E}\frac{\sum_{\alpha|r=\gamma|r} v_\alpha v_\gamma U(\alpha)U(\gamma)\exp(F(\alpha)+F(\gamma))}{\left(\sum_\alpha v_\alpha \exp F(\alpha)\right)^2}\ ,$$

and, in accordance to (14.18) we lighten notation by writing this quantity as

$$\mathsf{E}\langle \mathbf{1}_{\{\alpha|r=\gamma|r\}} U(\alpha)U(\gamma)\rangle\ .$$

Moreover, since $U'_{k+1} = \eta_r U_{k+1}$ and since $\mathsf{E}_p\eta_r = 0$ if $r \ge p$ and $\mathsf{E}_p\eta_r = \eta_r$ if $r < p$, we have

$$\mathsf{E}_p(W_p \cdots W_k U'_{k+1}) = 0 \text{ if } p \le r$$
$$\mathsf{E}_p(W_p \cdots W_{k+1} U'_{k+1}) = \eta_r \mathsf{E}_p(W_p \cdots W_k U_{k+1}) \text{ if } p > r$$

so that using (14.32) for U' rather than U yields

$$\mathsf{E}\langle \mathbf{1}_{\{\alpha|r=\gamma|r\}} U(\alpha)U(\gamma)\rangle = \sum_{r<p\le k+1} (m_p - m_{p-1})\mathsf{E}\big(W_1 \cdots W_{p-1}(\mathsf{E}_p W_p \cdots W_k U_{k+1})^2\big)\ . \quad (14.33)$$

Let us make throughout this chapter the convention that for all α, γ in $\mathbb{N}^{*k}$ we have

$$\alpha|0 = \gamma|0\ ; \quad \alpha|(k+1) \neq \gamma|(k+1)\ . \tag{14.34}$$

We then observe that (14.33) remains true for $r = 0$, since it then coincides with (14.31), and remains true for $r = k+1$, where it simply means that $0 = 0$. We observe moreover that (14.34) implies that given α and γ in $\mathbb{N}^{*k}$, there exists a unique integer $1 \leq r \leq k+1$ such that

$$\alpha|r \neq \gamma|r\ ; \quad \alpha|(r-1) = \gamma|(r-1)\ , \tag{14.35}$$

and we define

$$(\alpha, \gamma) = r\ . \tag{14.36}$$

This is "the first coordinate on which the sequences α and γ differ". This notation will be used throughout this entire chapter.

Given $1 \leq r \leq k+1$, let us subtract from (14.33) the corresponding equality for $r-1$ rather than r. Since

$$\mathbf{1}_{\{\alpha|(r-1)=\gamma|(r-1)\}} - \mathbf{1}_{\{\alpha|r=\gamma|r\}} = \mathbf{1}_{\{(\alpha,\gamma)=r\}}\ ,$$

we have obtained the following identity.

Proposition 14.3.2. *For any $1 \leq r \leq k+1$ we have*

$$\begin{aligned} &\mathsf{E}\langle \mathbf{1}_{\{(\alpha,\gamma)=r\}} U(\alpha) U(\gamma)\rangle \\ &= (m_r - m_{r-1})\mathsf{E}\big(W_1 \cdots W_{r-1}(\mathsf{E}_r W_r \cdots W_k U_{k+1})^2\big)\ . \end{aligned} \tag{14.37}$$

For $r = 1$, there is no term in the product $W_1 \cdots W_{r-1}$, while for $r = k+1$, the right-hand side is $(1 - m_k)\mathsf{E}(W_1 \cdots W_k U_{k+1}^2)$.

Since $\mathsf{E}_p W_p = 1$, by (14.23) we obtain recursively that

$$\mathsf{E}_p(W_p \cdots W_k) = \mathsf{E}_p(W_p \mathsf{E}_{p+1}(W_{p+1} \cdots W_k)) = 1\ ,$$

and similarly $\mathsf{E}(W_1 \cdots W_{r-1}) = 1$. Using (14.37) for $U = 1$, we have proved the following.

Proposition 14.3.3. *For $1 \leq r \leq k+1$ we have*

$$\mathsf{E}\langle \mathbf{1}_{\{(\alpha,\gamma)=r\}}\rangle = m_r - m_{r-1}\ . \tag{14.38}$$

It is worth spelling out the special case of (14.38) corresponding to the case $F = 0$:

$$\mathsf{E} \sum_{(\alpha,\gamma)=r} v_\alpha v_\gamma = m_r - m_{r-1}\ . \tag{14.39}$$

Let us now explain the real miracle that occurs in (14.37), and which will be absolutely essential in the entire chapter. We think of $1 \le r \le k+1$ as fixed once and for all. We consider two copies $(\mathbf{z}_p^\ell)$ for $\ell = 1$ and $\ell = 2$ of the sequence $(\mathbf{z}_p)_{1 \le p \le k}$. These copies are coupled in the following manner

$$\begin{cases} \mathbf{z}_p^1 = \mathbf{z}_p^2 \ \text{ if } \ p < r \\ \\ \mathbf{z}_p^1 \text{ and } \mathbf{z}_p^2 \text{ are independent if } p \ge r \ . \end{cases} \tag{14.40}$$

Recalling (14.3) we define $F_{k+1}^\ell = F(\mathbf{z}_1^\ell, \ldots, \mathbf{z}_k^\ell)$ and the variables F_p^ℓ through the recursion relation (14.5), where of course E_p denotes now expectation in the r.v.s $\mathbf{z}_n^\ell$ for $n \ge p$. We define

$$W_p^\ell = \exp m_p(F_{p+1}^\ell - F_p^\ell) \ . \tag{14.41}$$

This quantity depends only on $\mathbf{z}_1^\ell, \ldots, \mathbf{z}_p^\ell$. It allows for simpler notation if we assume (without loss of generality) that

$$p < r \quad \Rightarrow \quad \mathbf{z}_p^1 = \mathbf{z}_p^2 = \mathbf{z}_p \ .$$

With this convention, thinking of W_p as a function $W_p(\mathbf{z}_1, \ldots, \mathbf{z}_p)$, we have

$$W_p^\ell = W_p(\mathbf{z}_1^\ell, \ldots, \mathbf{z}_p^\ell)$$

so that $W_p^1 = W_p^2 = W_p$ for $p \le r-1$. Now, since the r.v.s $(\mathbf{z}_p^1)_{p \ge r}$ are independent from the r.v.s $(\mathbf{z}_p^2)_{p \ge r}$,

$$(\mathsf{E}_r W_r \cdots W_k U_{k+1})^2 = \mathsf{E}_r(W_r^1 W_r^2 \cdots W_k^1 W_k^2 U_{k+1}^1 U_{k+1}^2) \ ,$$

where $U_{k+1}^\ell = U_{k+1}(\mathbf{z}_1^\ell, \ldots, \mathbf{z}_{k+1}^\ell)$. Since $W_1 \cdots W_{r-1}$ do not depend on the r.v.s $\mathbf{z}_p^\ell$ for $p \ge r$ we obtain

$$\begin{aligned} &\mathsf{E}\big(W_1 \cdots W_{r-1}(\mathsf{E}_r W_r \cdots W_k U_{k+1})^2\big) \\ &= \mathsf{E}\big(W_1 \cdots W_{r-1} \mathsf{E}_r(W_r^1 W_r^2 \cdots W_k^1 W_k^2 U_{k+1}^1 U_{k+1}^2)\big) \\ &= \mathsf{E}\mathsf{E}_r\big(W_1 \cdots W_{r-1} W_r^1 W_r^2 \cdots W_k^1 W_k^2 U_{k+1}^1 U_{k+1}^2\big) \\ &= \mathsf{E}(W_1 \cdots W_{r-1} W_r^1 W_r^2 \cdots W_k^1 W_k^2 U_{k+1}^1 U_{k+1}^2) \ . \end{aligned} \tag{14.42}$$

Therefore, comparing with (14.37) yields the relation

$$\begin{aligned} &\mathsf{E}\langle \mathbf{1}_{\{(\alpha,\gamma)=r\}} U(\alpha) U(\gamma) \rangle \\ &= (m_r - m_{r-1}) \mathsf{E}(W_1 \cdots W_{r-1} W_r^1 W_r^2 \cdots W_k^1 W_k^2 U_{k+1}^1 U_{k+1}^2) \ . \end{aligned}$$

we may use the standard "polarization argument" (replacing U by $U + U'$) to obtain that if U' is another function $T^k \to \mathbb{R}$ and if we define $U_{k+1}'^2 = U'(\mathbf{z}_1^2, \ldots, \mathbf{z}_k^2)$ then

$$\begin{aligned} &\mathsf{E}\langle \mathbf{1}_{\{(\alpha,\gamma)=r\}} U(\alpha) U'(\gamma) \rangle \\ &= (m_r - m_{r-1}) \mathsf{E}(W_1 \cdots W_{r-1} W_r^1 W_r^2 \cdots W_k^1 W_k^2 U_{k+1}^1 U_{k+1}'^2) \ . \end{aligned} \tag{14.43}$$

Exercise 14.3.4. Prove that when $k = 1$ the case $r = 1$ of the previous formula recovers (13.15) while the case $r = 2$ recovers (13.14).

Consider now a continuous bounded function $U^\sim : T^k \times T^k \to \mathbb{R}$, say

$$U^\sim((\mathbf{x}_1^1, \mathbf{x}_1^2), \dots, (\mathbf{x}_k^1, \mathbf{x}_k^2)) \tag{14.44}$$

and let

$$U_{k+1}^\sim = U^\sim((\mathbf{z}_1^1, \mathbf{z}_1^2), \dots, (\mathbf{z}_k^1, \mathbf{z}_k^2)) \tag{14.45}$$

$$U^\sim(\alpha, \gamma) = U^\sim((\mathbf{z}_{1,\alpha}, \mathbf{z}_{1,\gamma}), \dots, (\mathbf{z}_{k,\alpha}, \mathbf{z}_{k,\gamma})) . \tag{14.46}$$

The reader should observe carefully the formula (14.46). There are no upper indices there. The idea of replicas is implemented only through the lower indices (α, γ).

Theorem 14.3.5. *We have*

$$\mathsf{E}\langle \mathbf{1}_{\{(\alpha,\gamma)=r\}} U^\sim(\alpha,\gamma)\rangle = (m_r - m_{r-1})\mathsf{E}(W_1 \cdots W_{r-1} W_r^1 W_r^2 \cdots W_k^1 W_k^2 U_{k+1}^\sim) . \tag{14.47}$$

Proof. When the function $U^\sim$ is of the type

$$U^\sim((\mathbf{x}_1^1, \mathbf{x}_1^2), \dots, (\mathbf{x}_k^1, \mathbf{x}_k^2)) = U(\mathbf{x}_1^1, \dots, \mathbf{x}_k^1) U'(\mathbf{x}_1^2, \dots, \mathbf{x}_k^2) ,$$

then (14.47) coincides with (14.43). The general case follows by approximating an arbitrary function by a sum of functions of the previous type. $\square$

The amazing fact is that the expression on the right-hand side above is of the same nature as the left-hand side of (14.27), as we demonstrate now. Let us consider the sequence

$$m_1^* = \frac{m_1}{2} \ , \ \dots \ , \ m_{r-1}^* = \frac{m_{r-1}}{2} \ , \ m_r^* = m_r \ , \ \dots \ , m_k^* = m_k \tag{14.48}$$

(that will be used many times). Let us define

$$J_{k+1} = F_{k+1}^1 + F_{k+1}^2$$

and then J_p recursively by

$$J_p = \frac{1}{m_p^*} \log \mathsf{E}_p \exp m_p^* J_{p+1} . \tag{14.49}$$

Lemma 14.3.6. *a) For each p we have*

$$J_p = F_p^1 + F_p^2 , \tag{14.50}$$

b) Let

$$V_p = \exp m_p^* (J_{p+1} - J_p) . \tag{14.51}$$

Then

$$p < r \quad \Rightarrow \quad V_p = W_p \ ; \qquad p \geq r \quad \Rightarrow \quad V_p = W_p^1 W_p^2 .$$

Consequently, the expression on the right of (14.47) is $\mathsf{E}(V_1 \cdots V_k U^{\sim}_{k+1})$. Let us state this for further reference.

Corollary 14.3.7. *With the notation (14.51) we have*

$$\mathsf{E}\langle \mathbf{1}_{\{(\alpha,\gamma)=r\}} U^{\sim}(\alpha,\gamma)\rangle = (m_r - m_{r-1})\mathsf{E}(V_1 \cdots V_k U^{\sim}_{k+1}) \ . \tag{14.52}$$

The right-hand side of (14.52) is of the same nature as the left-hand side of (14.27), since the quantities V_p are defined by a similar procedure as the quantities W_p. To illustrate this point, let us write down the formula corresponding to (14.27). Instead of the function $F(\mathbf{x}_1, \ldots, \mathbf{x}_k)$ we use the function

$$\widehat{F}((\mathbf{x}_1^1, \mathbf{x}_1^2), \ldots, (\mathbf{x}_k^1, \mathbf{x}_k^2)) = F(\mathbf{x}_1^1, \ldots, \mathbf{x}_k^1) + F(\mathbf{x}_1^2, \ldots, \mathbf{x}_k^2) \ ,$$

instead of the sequence of r.v.s $\mathbf{z}_1, \ldots, \mathbf{z}_k$ we use the sequence of r.v.s $(\mathbf{z}_1^1, \mathbf{z}_1^2), \ldots, (\mathbf{z}_k^1, \mathbf{z}_k^2)$, and, recalling (14.44) we consider (defining $(\mathbf{z}_{p,\alpha}^1, \mathbf{z}_{p,\alpha}^2)$ in the obvious manner)

$$\begin{aligned} \widehat{F}(\alpha) &= \widehat{F}((\mathbf{z}_{1,\alpha}^1, \mathbf{z}_{1,\alpha}^2), \ldots, (\mathbf{z}_{k,\alpha}^1, \mathbf{z}_{k,\alpha}^2)) \\ \widehat{U}(\alpha) &= U^{\sim}((\mathbf{z}_{1,\alpha}^1, \mathbf{z}_{1,\alpha}^2), \ldots, (\mathbf{z}_{k,\alpha}^1, \mathbf{z}_{k,\alpha}^2)) \ . \end{aligned} \tag{14.53}$$

The formula (14.53) should be compared with the formula (14.14).

Then the formula corresponding to (14.27) is

$$\mathsf{E}(V_1 \cdots V_k U^{\sim}_{k+1}) = \mathsf{E}\frac{\sum_\alpha v_\alpha \widehat{U}(\alpha) \exp \widehat{F}(\alpha)}{\sum_a v_\alpha \exp \widehat{F}(\alpha)} \ , \tag{14.54}$$

where the weights v_α form a Poisson-Dirichlet cascade associated with the sequence (14.48).

Proof of Lemma 14.3.6. The argument of this lemma is simple but essential. If will be used again and again. We prove (14.50) by decreasing induction on p. It holds for $p = k+1$ by definition of J_{k+1}. If $p \geq r$, by independence and since F_{p+1}^ℓ satisfies (14.5),

$$\begin{aligned} J_p &= \frac{1}{m_p} \log \mathsf{E}_p \exp m_p J_{p+1} \\ &= \frac{1}{m_p} \log \mathsf{E}_p \exp m_p (F_{p+1}^1 + F_{p+1}^2) \\ &= \frac{1}{m_p} \log\big((\mathsf{E}_p \exp m_p F_{p+1}^1)(\mathsf{E}_p \exp m_p F_{p+1}^2)\big) \\ &= \frac{1}{m_p} \log \exp m_p (F_p^1 + F_p^2) = F_p^1 + F_p^2 \ . \end{aligned}$$

For $p = r$, since F_r^ℓ depends only on $(\mathbf{z}_1^\ell, \ldots, \mathbf{z}_{r-1}^\ell) = (\mathbf{z}_1, \ldots, \mathbf{z}_{r-1})$ we have $F_r^1 = F_r^2 = F_r$, and therefore $F_p^1 = F_p^2 = F_p$ for $p \leq r$. If $p < r$, we write

$$J_p = \frac{2}{m_p} \log \mathsf{E}_p \exp \frac{m_p}{2}(F^1_{p+1} + F^2_{p+1})$$
$$= \frac{2}{m_p} \log \mathsf{E}_p \exp m_p F_{p+1} = 2F_p = F^1_p + F^2_p \ .$$

This proves (14.50). Now

$$V_p = \exp m^*_p (J_{p+1} - J_p) = \exp m^*_p (F^1_{p+1} - F^1_p) \exp m^*_p (F^2_{p+1} - F^2_p)$$

so if $p \geq r$ we have $V_p = W^1_p W^2_p$, while if $p < r$ we have

$$V_p = (W^1_p W^2_p)^{1/2} = W_p \ . \qquad \square$$

14.4 Guerra's Broken Replica-Symmetry Bound

In this chapter we study rather general Gaussian Hamiltonians. We consider Hamiltonians H_N such that $(H_N(\boldsymbol{\sigma}))$ is a jointly centered Gaussian family of r.v.s such that, for a suitable function ξ,

$$\forall \boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2 \ , \qquad \frac{1}{N} \mathsf{E} H_N(\boldsymbol{\sigma}^1) H_N(\boldsymbol{\sigma}^2) = \xi(R_{1,2}) \ . \tag{14.55}$$

We assume

$$\xi'(0) = 0 \ . \tag{14.56}$$

An important example is the case of the p-spin interaction model, that is

$$- H_N(\boldsymbol{\sigma}) = \frac{\beta_p}{N^{(p-1)/2}} \sum_{i_1,\ldots,i_p \leq N} g_{i_1,\ldots,i_p} \sigma_{i_1} \cdots \sigma_{i_p} \ , \tag{14.57}$$

where $g_{i_1,\ldots,i_p}$ are independent standard Gaussian r.v.s and the summation is over all choices of $i_1, \ldots, i_p \leq N$. Thus

$$\frac{1}{N} \mathsf{E} H_N(\boldsymbol{\sigma}^1) H_N(\boldsymbol{\sigma}^2) = \frac{\beta_p^2}{N^p} \sum_{i_1,\ldots,i_p} \sigma^1_{i_1} \sigma^2_{i_1} \sigma^1_{i_2} \sigma^2_{i_2} \cdots \sigma^1_{i_p} \sigma^2_{i_p}$$
$$= \beta_p^2 R^p_{1,2} \ .$$

In that case (14.55) holds for $\xi(x) = \beta_p^2 x^p$. Sums of independent terms as in (14.57) for $p \geq 2$ yield functions of the type

$$\xi(x) = \sum_{p \geq 1} \beta_p^2 x^p \ , \tag{14.58}$$

which satisfy (14.56). Such a function is always convex on $\mathbb{R}^+$. When only terms with p even are involved, it is convex on $\mathbb{R}$. We will only consider sequences β_p such that the series in (14.55) converges for all x.

The expert reader might object that (14.58) is not the traditional definition of the p-spin interaction model, which is given by

$$-H_N(\boldsymbol{\sigma}) = \beta_p \left(\frac{p!}{N^{p-1}}\right)^{1/2} \sum_{1\le i_1<i_2<\ldots<i_p\le N} g_{i_1,\ldots,i_p}\sigma_{i_1}\cdots\sigma_{i_p} \ . \tag{14.59}$$

This does not quite satisfy (14.55). It is however rather straightforward in our analysis to replace (14.55) by the condition

$$\left|\frac{1}{N}\mathsf{E}H_N(\boldsymbol{\sigma}^1)H_N(\boldsymbol{\sigma}^2) - \xi(R_{1,2})\right| \le c_N$$

where $c_N \to 0$. We will not do this to lighten the exposition.

Associated to the function ξ is the function

$$\theta(x) = x\xi'(x) - \xi(x) \ . \tag{14.60}$$

This notation will be used throughout this chapter and the next. When ξ is convex on $\mathbb{R}$,

$$\xi(x) - x\xi'(q) + \theta(q) = \xi(x) - \xi(q) - (x-q)\xi'(q) \ge 0 \ ,$$

and thus

$$\xi(x) - x\xi'(q) \ge -\theta(q) \tag{14.61}$$

and when ξ is convex on $\mathbb{R}^+$

$$x, q \ge 0 \quad \Rightarrow \quad \xi(x) - x\xi'(q) \ge -\theta(q) \ . \tag{14.62}$$

Fixing the Hamiltonian H_N as in (14.57), we consider now a countable set A and another Gaussian family $(H(\boldsymbol{\sigma},\alpha))$ for $\boldsymbol{\sigma} \in \Sigma_N$, $\alpha \in A$. We assume that for $\alpha, \gamma \in A$ we have

$$\frac{1}{N}\mathsf{E}H(\boldsymbol{\sigma}^1,\alpha)H(\boldsymbol{\sigma}^2,\gamma) = R_{1,2}\xi'(q_{\alpha,\gamma}) \ , \tag{14.63}$$

where $q_{\alpha,\gamma}$ is a number depending on α and γ that satisfies for a certain number $\overline{q}$

$$\forall \alpha \ , \qquad q_{\alpha,\alpha} = \overline{q} \ . \tag{14.64}$$

We consider the Hamiltonian

$$-H_t(\boldsymbol{\sigma},\alpha) = -\sqrt{t}H_N(\boldsymbol{\sigma}) - \sqrt{1-t}H(\boldsymbol{\sigma},\alpha) + \sum_{i\le N}\sigma_i h_i, \tag{14.65}$$

where as usual $(h_i)_{i\le N}$ are i.i.d. copies of a r.v. h. Consider now random weights $(w_\alpha)_{\alpha\in A}$. As the notation indicates, for the time being, we do not assume that they form a Poisson-Dirichlet cascade, although this will be the most interesting situation.

For a function $f(\boldsymbol{\sigma}, \alpha)$ we define

$$\langle f(\boldsymbol{\sigma}, \alpha)\rangle_t = \frac{\sum_{\boldsymbol{\sigma},\alpha} w_\alpha f(\boldsymbol{\sigma}, \alpha) \exp(-H_t(\boldsymbol{\sigma}, \alpha))}{\sum_{\boldsymbol{\sigma},\alpha} w_\alpha \exp(-H_t(\boldsymbol{\sigma}, \alpha))} \ . \tag{14.66}$$

Lemma 14.4.1. *Consider the function*

$$\varphi(t) = \frac{1}{N} \mathsf{E} \log \sum_{\boldsymbol{\sigma},\alpha} w_\alpha \exp(-H_t(\boldsymbol{\sigma}, \alpha)) \ . \tag{14.67}$$

Then

$$\varphi'(t) = \frac{1}{2}(\xi(1) - \xi'(\overline{q})) + \frac{1}{2}\mathsf{E}\langle \theta(q_{\alpha,\gamma})\rangle_t - \frac{1}{2}\mathsf{E}\langle \xi(R_{1,2}) - R_{1,2}\xi'(q_{\alpha,\gamma}) + \theta(q_{\alpha,\gamma})\rangle_t \ . \tag{14.68}$$

The function φ will be used throughout this chapter.

Proof. We recall that γ is a replica of α, so that $(\boldsymbol{\sigma}^2, \gamma)$ is a replica of $(\boldsymbol{\sigma}^1, \alpha)$. We define

$$U(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2, \alpha, \gamma) = \frac{1}{N}(\mathsf{E} H_N(\boldsymbol{\sigma}^1) H_N(\boldsymbol{\sigma}^2) - \mathsf{E} H(\boldsymbol{\sigma}^1, \alpha) H(\boldsymbol{\sigma}^2, \gamma)) \ .$$

Differentiation and integration by parts yield the formula

$$\varphi'(t) = \frac{1}{2}\big(\mathsf{E}\langle U(\boldsymbol{\sigma}, \boldsymbol{\sigma}, \alpha, \alpha)\rangle_t - \mathsf{E}\langle U(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2, \alpha, \gamma)\rangle_t\big) \ .$$

Then (14.55) and (14.63) imply

$$U(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2, \alpha, \gamma) = \xi(R_{1,2}) - R_{1,2}\xi'(q_{\alpha,\gamma}) \ ,$$

and in particular from (14.64) we obtain

$$U(\boldsymbol{\sigma}, \boldsymbol{\sigma}, \alpha, \alpha) = \xi(1) - \xi'(\overline{q}) \ ,$$

so that

$$\begin{aligned} \varphi'(t) &= \frac{1}{2}(\xi(1) - \xi'(\overline{q})) - \frac{1}{2}\mathsf{E}\langle \xi(R_{1,2}) - R_{1,2}\xi'(q_{\alpha,\gamma})\rangle_t \\ &= \frac{1}{2}(\xi(1) - \xi'(\overline{q})) + \frac{1}{2}\mathsf{E}\langle \theta(q_{\alpha,\gamma})\rangle_t \\ &\quad - \frac{1}{2}\mathsf{E}\langle \xi(R_{1,2}) - R_{1,2}\xi'(q_{\alpha,\gamma}) + \theta(q_{\alpha,\gamma})\rangle_t \ , \end{aligned}$$

and the proof is complete. □

The reason for writing $\varphi'(t)$ as in (14.68) is that when ξ is convex by (14.61) we have $\xi(R_{1,2}) - R_{1,2}\xi'(q_{\alpha,\gamma}) + \theta(q_{\alpha,\gamma}) \geq 0$ so that (14.68) yields the clean bound

$$\varphi'(t) \le \frac{1}{2}(\xi(1) - \xi'(\overline{q})) + \frac{1}{2}\mathsf{E}\langle\theta(q_{\alpha,\gamma})\rangle_t \ .$$

We now turn to what will eventually appear as the optimal choice for the Hamiltonian $H(\boldsymbol{\sigma}, \alpha)$ and the weights w_α. We consider an integer $k \ge 0$, a sequence

$$0 = m_0 < m_1 < \ldots < m_{k-1} < m_k < 1 = m_{k+1} \ , \tag{14.69}$$

and a sequence

$$0 = q_0 \le q_1 \le \ldots \le q_{k+1} \le q_{k+2} = 1 \ . \tag{14.70}$$

It is very useful to think of these sequences as defining a non-decreasing step function $x(q)$ with $x(q) = m_p$ for $q_p \le q < q_{p+1}$. This function is equal to 1 in the interval $[q_{k+1}, 1[$. Somehow this interval requires a special treatment, the reason being that the distribution Λ_m does not exist for $m = 1$.

To the sequences (14.69) and (14.70) we will attach a Hamiltonian $H(\boldsymbol{\sigma}, \alpha)$ and random weights, that we will denote by (v_α) since they form a Poisson-Dirichlet cascade. In the remainder of this chapter,

The function φ will be the function of (14.67) for these choices. (14.71)

Consider independent Gaussian r.v.s $(z_p)_{0 \le p \le k}$ and assume that

$$\mathsf{E}z_p^2 = \xi'(q_{p+1}) - \xi'(q_p) \ . \tag{14.72}$$

For $i \le N$ we consider independent copies $(z_{i,0})$ of z_0. For $p \ge 1$, we consider independent copies $z_{i,p,j_1,\ldots,j_p}$ of z_p, for $i \le N$, $j_1, \ldots, j_p \in \mathbb{N}^*$. We set $A = \mathbb{N}^{*k}$, and for $\alpha = (j_1, \ldots, j_k) \in A$ we set

$$z_{i,p,\alpha} = z_{i,p,j_1,\ldots,j_p} \ ; \qquad z_{i,0,\alpha} = z_{i,0} \ .$$

Finally we define

$$-H(\boldsymbol{\sigma}, \alpha) = \sum_{i \le N} \sigma_i \sum_{0 \le p \le k} z_{i,p,\alpha} \ . \tag{14.73}$$

We denote by (v_α) weights that form a Poisson-Dirichlet cascade associated with the sequence (14.69), so from now on as pointed in (14.71) the letter φ refers to the function (14.67) for the Hamiltonian (14.73) and these weights, and the notation $\langle\cdot\rangle_t$ refers to (14.66) with the same choices. Recalling the notation (α, γ) of Section 14.3, we have

$$\mathsf{E}\left(\sum_{0 \le p \le k} z_{i,p,\alpha} \sum_{0 \le p \le k} z_{i,p,\gamma}\right) = \sum_{0 \le p < (\alpha,\gamma)} \mathsf{E}z_p^2 \tag{14.74}$$

because $z_{i,p,\alpha}$ and $z_{i,p,\gamma}$ are equal if $p < (\alpha, \gamma)$ and are independent otherwise. Now, since $q_0 = 0$ and $\xi'(0) = 0$,

$$\sum_{0 \le p < (\alpha,\gamma)} \mathsf{E}z_p^2 = \sum_{p < (\alpha,\gamma)} (\xi'(q_{p+1}) - \xi'(q_p)) = \xi'(q_{(\alpha,\gamma)}) \ .$$

Therefore since $z_{i,p,\alpha}$ and $z_{i',p',\gamma}$ are independent when $i \neq i'$, (14.73) implies

$$\frac{1}{N}\mathsf{E}H(\boldsymbol{\sigma}^1,\alpha)H(\boldsymbol{\sigma}^2,\gamma) = R_{1,2}\xi'(q_{(\alpha,\gamma)}) \ .$$

This is (14.63) with $q_{\alpha,\gamma} = q_{(\alpha,\gamma)}$. In particular (14.64) holds for $\overline{q} = q_{k+1}$.

Now we would like to compute

$$\mathsf{E}\langle\theta(q_{\alpha,\gamma})\rangle_t = \mathsf{E}\langle\theta(q_{(\alpha,\gamma)})\rangle_t = \sum_{1\le p\le k+1}\theta(q_p)\mathsf{E}\langle\mathbf{1}_{\{(\alpha,\gamma)=p\}}\rangle_t \ . \tag{14.75}$$

We expect from (14.38) that

$$\mathsf{E}\langle\mathbf{1}_{\{(\alpha,\gamma)=p\}}\rangle_t = m_p - m_{p-1} \ , \tag{14.76}$$

and let us first check that this is indeed the case.

For $p \le k$ let us consider the variable $\mathbf{x}_p = (x_{i,p})_{i\le N} \in T = \mathbb{R}^N$. Given t, the Hamiltonian H_N, the r.v.s $(h_i)_{i\le N}$ and the r.v.s $(z_{i,0})_{i\le N}$, let us define the quantity

$$-H(\boldsymbol{\sigma},\mathbf{x}_1,\dots,\mathbf{x}_k) = -\sqrt{t}H_N(\boldsymbol{\sigma}) + \sqrt{1-t}\sum_{i\le N}\sigma_i\Big(z_{i,0} + \sum_{1\le p\le k}x_{i,p}\Big) + \sum_{i\le N}\sigma_i h_i. \tag{14.77}$$

Let us then consider the function $F : T^k \to \mathbb{R}$ given by

$$F(\mathbf{x}_1,\dots,\mathbf{x}_k) = \log\sum_{\boldsymbol{\sigma}}\exp(-H(\boldsymbol{\sigma},\mathbf{x}_1,\dots,\mathbf{x}_k)) \ . \tag{14.78}$$

It satisfies (14.4) for reasons that have already been explained: it is obvious that $\mathsf{E}\exp F_{k+1} < \infty$ and to prove that $\mathsf{E}|F_{k+1}| < \infty$ we bound below the quantity (14.78) by replacing the summation in the right-hand side by a single term of this summation. For $p \le k$ let us consider independent copies $(z_{i,p})_{i\le N}$ of the r.v.s z_p and define $\mathbf{z}_p = (z_{i,p})_{i\le N}$. Consider the r.v.s $\mathbf{z}_{p,\alpha}$ as in (14.6). By (14.65) and the definition (14.73) of the Hamiltonian $H(\boldsymbol{\sigma},\alpha)$, recalling the notation (14.7), we obtain

$$\sum_{\boldsymbol{\sigma}}\exp(-H_t(\boldsymbol{\sigma},\alpha)) = \exp F(\mathbf{z}_{1,\alpha},\dots,\mathbf{z}_{k,\alpha}) = \exp F(\alpha) \ .$$

Moreover we have

$$\begin{aligned}\langle\mathbf{1}_{\{(\alpha,\gamma)=p\}}\rangle_t &= \frac{\sum_{(\alpha,\gamma)=p}v_\alpha v_\gamma\sum_{\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2}\exp(-H_t(\boldsymbol{\sigma}^1,\alpha)-H_t(\boldsymbol{\sigma}^2,\gamma))}{\big(\sum_{\alpha,\boldsymbol{\sigma}}v_\alpha\exp(-H_t(\boldsymbol{\sigma},\alpha))\big)^2} \\ &= \frac{\sum_{(\alpha,\gamma)=p}v_\alpha v_\gamma\exp(F(\alpha)+F(\gamma))}{\big(\sum_\alpha v_\alpha\exp F(\alpha)\big)^2} \ .\end{aligned}$$

The average on the right is the same as the average occurring in (14.38). Denoting by E_0 expectation given H_N, the r.v.s $(h_i)_{i\le N}$ and the r.v.s $(z_{i,0})_{i\le N}$, we then see from (14.38) that

$$\mathsf{E}_0\langle \mathbf{1}_{\{(\alpha,\gamma)=p\}}\rangle_t = m_p - m_{p-1} \ .$$

Taking expectation implies (14.76).

It is good to note that in the previous argument the r.v.s $z_{i,0}$ do not play the same role as the r.v.s $z_{i,p}$ for $p \geq 1$. This is simply because the r.v.s $z_{i,0}$ are associated with the value $m = 0$ for which the distribution Λ_m does not exist. This feature will occur repeatedly.

Since $q_{k+2} = 1$ and $m_0 = 0$, using (14.75) in the first line and (14.76) in the second line we obtain

$$\begin{aligned}\mathsf{E}\langle\theta(q_{\alpha,\gamma})\rangle_t &= \sum_{1\leq p\leq k+1} \theta(q_p)\mathsf{E}\langle \mathbf{1}_{\{(\alpha,\gamma)=p\}}\rangle_t \\ &= \sum_{1\leq p\leq k+1} \theta(q_p)(m_p - m_{p-1}) \\ &= -\sum_{1\leq p\leq k+1} m_p(\theta(q_{p+1}) - \theta(q_p)) + \theta(1)\end{aligned}$$

and since $\theta(1) = \xi'(1) - \xi(1)$ (14.68) yields

$$\begin{aligned}\varphi'(t) = \frac{1}{2}(\xi'(1) - \xi'(q_{k+1})) &- \frac{1}{2}\sum_{1\leq p\leq k+1} m_p(\theta(q_{p+1}) - \theta(q_p)) \\ &- \frac{1}{2}\mathsf{E}\langle \xi(R_{1,2}) - R_{1,2}\xi'(q_{(\alpha,\gamma)}) + \theta(q_{(\alpha,\gamma)})\rangle_t \ .\end{aligned} \tag{14.79}$$

Let us compute $\varphi(0)$. First, we observe that

$$\sum_{\boldsymbol{\sigma}} \exp\left(\sum_{i\leq N}\sigma_i\left(h_i + \sum_{0\leq p\leq k} z_{i,p,\alpha}\right)\right) = 2^N\prod_{i\leq N}\mathrm{ch}\left(h_i + \sum_{0\leq p\leq k} z_{i,p,\alpha}\right) .$$

Next, since $\sum_\alpha v_\alpha = 1$,

$$\begin{aligned}\varphi(0) &= \frac{1}{N}\mathsf{E}\log\sum_{\boldsymbol{\sigma},\alpha} v_\alpha \exp\left(-H(\boldsymbol{\sigma},\alpha) + \sum_{i\leq N} h_i\sigma_i\right) \\ &= \frac{1}{N}\mathsf{E}\log\sum_{\boldsymbol{\sigma},\alpha} v_\alpha \exp\left(\sum_{i\leq N}\sigma_i\left(h_i + \sum_{0\leq p\leq k} z_{i,p,\alpha}\right)\right) \\ &= \log 2 + \frac{1}{N}\mathsf{E}\log\sum_{\alpha} v_\alpha \prod_{i\leq N}\mathrm{ch}\left(h_i + \sum_{0\leq p\leq k} z_{i,p,\alpha}\right) .\end{aligned} \tag{14.80}$$

Let us define

$$F_{k+1} = \log\prod_{i\leq N}\mathrm{ch}\left(h_i + \sum_{0\leq p\leq k} z_{i,p}\right) \tag{14.81}$$

and, recursively, for $p \geq 1$,

$$F_p = \frac{1}{m_p} \log \mathsf{E}_p \exp m_p F_{p+1} ,$$

where E_p denotes expectation in the r.v.s $z_{i,n}$ for $n \geq p$. Theorem 14.1 implies

$$\mathsf{E} \log \sum_\alpha v_\alpha \prod_{i \leq N} \mathrm{ch}\Bigl(h_i + \sum_{0 \leq p \leq k} z_{i,p,\alpha}\Bigr) = \mathsf{E} F_1 . \tag{14.82}$$

Next, we get from (14.81) that

$$F_{k+1} = \sum_{i \leq N} F_{k+1,i} ,$$

where

$$F_{k+1,i} = \log \mathrm{ch}\Bigl(h_i + \sum_{0 \leq p \leq k} z_{i,p}\Bigr) .$$

The randomness appearing in the terms $F_{k+1,i}$ for $i \leq N$ are independent. Thus if we define recursively

$$F_{p,i} = \frac{1}{m_p} \log \mathsf{E}_p \exp m_p F_{p+1,i}$$

we obtain

$$F_p = \sum_{i \leq N} F_{p,i}$$

so that $\mathsf{E} F_1 = N \mathsf{E} F_{1,1}$ and therefore

$$\varphi(0) = \log 2 + \mathsf{E} F_{1,1} .$$

Let us describe again the quantity $\mathsf{E} F_{1,1}$. Letting

$$X'_{k+1} = \log \mathrm{ch}\Bigl(h + \sum_{0 \leq p \leq k} z_p\Bigr) ,$$

we define for $1 \leq p \leq k$

$$X'_p = \frac{1}{m_p} \log \mathsf{E}_p \exp m_p X'_{p+1}$$

and $X'_0 = \mathsf{E} X'_1$. Then $\mathsf{E} F_{1,1} = X'_0$.

It is more elegant "to incorporate the term $(\xi'(1) - \xi'(q_{k+1}))/2$ with X'_0". This term represents "the contribution of the interval $[q_{k+1}, 1[$" to which is associated the value $m_{k+1} = 1$, and on which the function $x(q)$ equals 1. To do that let us consider a new independent Gaussian r.v. z_{k+1} with $\mathsf{E} z_{k+1}^2 = \xi'(1) - \xi'(q_{k+1})$. Let us set

$$X_{k+2} = \log \operatorname{ch}\left(h + \sum_{0\le p\le k+1} z_p\right) , \tag{14.83}$$

and for $1 \le p \le k+1$ let us define recursively

$$X_p = \frac{1}{m_p} \log \mathsf{E}_p \exp m_p X_{p+1} \; .$$

Since $m_{k+1} = 1$ and E_{k+1} denotes expectation in z_{k+1} only, we have

$$X_{k+1} = \frac{1}{2}(\xi'(1) - \xi'(q_{k+1})) + X'_{k+1} \; .$$

By decreasing induction over p this implies that for $1 \le p \le k+1$

$$X_p = \frac{1}{2}(\xi'(1) - \xi'(q_{k+1})) + X'_p \; ,$$

and therefore

$$X_0 = \mathsf{E} X_1 = X'_0 + \frac{1}{2}(\xi'(1) - \xi'(q_{k+1})) \; . \tag{14.84}$$

Thus

$$\varphi(0) = \log 2 + X_0 - \frac{1}{2}(\xi'(1) - \xi'(q_{k+1})) \; . \tag{14.85}$$

Proposition 14.4.2. *Consider the function*

$$\psi(t) = \log 2 + X_0 - \frac{1-t}{2}(\xi'(1) - \xi'(q_{k+1})) - \frac{t}{2} \sum_{1\le p\le k+1} m_p(\theta(q_{p+1}) - \theta(q_p)) \; . \tag{14.86}$$

Then $\psi(0) = \varphi(0)$ *and*

$$\varphi'(t) = \psi'(t) - \frac{1}{2}\mathsf{E}\langle \xi(R_{1,2}) - R_{1,2}\xi'(q_{(\alpha,\gamma)}) + \theta(q_{(\alpha,\gamma)})\rangle_t \; . \tag{14.87}$$

Proof. Relation (14.87) is obvious from (14.79) and (14.85) entails that $\varphi(0) = \psi(0)$. □

Theorem 14.4.3. (Guerra's broken replica-symmetry bound). *Assume that ξ is convex. Then we have*

$$\begin{aligned} p_N &:= \frac{1}{N} \mathsf{E} \log \sum_{\boldsymbol{\sigma}} \exp\left(-H_N(\boldsymbol{\sigma}) + \sum_{i\le N} h_i\sigma_i\right) \\ &\le \log 2 + X_0 - \frac{1}{2} \sum_{1\le p\le k+1} m_p(\theta(q_{p+1}) - \theta(q_p)) \; , \end{aligned} \tag{14.88}$$

where θ is defined in (14.60) and X_0 in (14.84).

Proof. By (14.87) we have $\varphi'(t) \le \psi'(t)$, and since $\psi(0) = \varphi(0)$ we have $p_N = \varphi(1) \le \psi(1)$. □

Here and in the remainder of the chapter, the value of ξ and h is kept implicit. To simplify notation we write

$$\mathbf{q} = (q_1, \ldots, q_{k+1}) \quad ; \quad \mathbf{m} = (m_1, \ldots, m_k)$$

and (setting as usual $m_{k+1} = 1$) we denote by $\mathcal{P}_k(\mathbf{m}, \mathbf{q})$ the right-hand side of (14.88),

$$\mathcal{P}_k(\mathbf{m}, \mathbf{q}) = \log 2 + X_0 - \frac{1}{2} \sum_{1 \le p \le k+1} m_p(\theta(q_{p+1}) - \theta(q_p)) \;, \tag{14.89}$$

so that (14.88) rewrites as

$$p_N \le \mathcal{P}_k(\mathbf{m}, \mathbf{q}) \;. \tag{14.90}$$

In the proof of (14.88) we have assumed that $0 < m_1 < \ldots < m_k < m_{k+1} = 1$, but the quantity $\mathcal{P}_k(\mathbf{m}, \mathbf{q})$ is well defined under the weaker condition

$$0 \le m_1 \le \ldots \le m_k \le m_{k+1} = 1 \;, \tag{14.91}$$

provided we agree as usual that $m^{-1} \log \mathsf{E} \exp mX = \mathsf{E}X$ when $m = 0$. Then (14.90) remains true in that situation, since, using Lemma 14.2.3, $\mathcal{P}_k(\mathbf{m}, \mathbf{q})$ is a continuous function of $\mathbf{m}$.

For a physicist k is the "number of steps of replica-symmetric breaking", and the case $k = 0$ corresponds to the "replica-symmetric solution". Let us check this. When $k = 0$, we have only one parameter $q = q_1$. Then, with obvious notation,

$$\begin{aligned} X_2 &= \log \mathrm{ch}\bigl(h + z\sqrt{\xi'(q)} + z'\sqrt{\xi'(1) - \xi'(q)}\bigr) \\ X_1 &= \log \mathrm{ch}\bigl(h + z\sqrt{\xi'(q)}\bigr) + \frac{1}{2}(\xi'(1) - \xi'(q)) \\ X_0 &= \mathsf{E} \log \mathrm{ch}\bigl(h + z\sqrt{\xi'(q)}\bigr) + \frac{1}{2}(\xi'(1) - \xi'(q)) \end{aligned}$$

and, since now $m_1 = 1$,

$$\sum_{1 \le p \le k+1} m_p(\theta(q_{p+1}) - \theta(q_p)) = \theta(1) - \theta(q) \;,$$

so that, since $\xi'(1) - \theta(1) = \xi(1)$,

$$\mathcal{P}_0(\mathbf{m}, \mathbf{q}) = \log 2 + \frac{1}{2}(\xi(1) + \theta(q) - \xi'(q)) + \mathsf{E} \log \mathrm{ch}\bigl(h + z\sqrt{\xi'(q)}\bigr) \;,$$

which coincides with the right-hand side of (1.72) when $\xi(x) = \beta^2 x^2/2$.

Our next result asserts that for the validity of (14.88) it suffices in fact that ξ be convex on $\mathbb{R}^+$ rather than on $\mathbb{R}$. This allows in particular to cover the case of the p-spin interaction model when p is odd.

Theorem 14.4.4. *Assume that ξ is convex on $\mathbb{R}^+$. Then for each values of k, $\mathbf{m}$ and $\mathbf{q}$ we have*
$$\limsup_{N\to\infty} p_N \le \mathcal{P}_k(\mathbf{m},\mathbf{q}) \;. \tag{14.92}$$

Proof. This follows the proof of Theorem 14.4.3. The problem is that in (14.87) we no longer know that $\xi(R_{1,2}) - R_{1,2}\xi'(q_{(\alpha,\gamma)}) + \theta(q_{(\alpha,\gamma)}) \ge 0$. The idea is simply that if ξ is still convex on $\mathbb{R}^+$ and since $q_{(\alpha,\gamma)} \ge 0$ we still have

$$R_{1,2} \ge 0 \quad\Rightarrow\quad \xi(R_{1,2}) - R_{1,2}\xi'(q_{(\alpha,\gamma)}) + \theta(q_{(\alpha,\gamma)}) \ge 0 \;,$$

and that by Theorem 12.3.1 we may pretend that $R_{1,2} \ge 0$. More formally let us add the perturbation term (12.32) to the interpolating Hamiltonian (14.65). Then the identity (14.68) still holds true. The computation of $\varphi(0)$ in the limit is the same because "the perturbation term is of lower order" (as formalized by (12.35)), and all we need to have the proof of Theorem 14.4.3 carry over is to show that, uniformly over t, the average over the parameter $\boldsymbol{\beta}$ of (12.32) of the quantity

$$\mathsf{E}\langle \xi(R_{1,2}) - R_{1,2}\xi'(q_{(\alpha,\gamma)}) + \theta(q_{(\alpha,\gamma)}))\rangle_t$$

is "positive in the limit". To see this, we define

$$b(\varepsilon) = \inf\{\xi(x) - x\xi'(q) + \theta(q) \;;\; q \ge 0 \;,\; -\varepsilon \le x \le 1\} \;,$$

we observe that $b(\varepsilon) \le 0$ and we simply write that, given $\varepsilon > 0$,

$$\xi(R_{1,2}) - R_{1,2}\xi'(q_{(\alpha,\gamma)}) + \theta(q_{(\alpha,\gamma)}) \ge -\mathbf{1}_{\{R_{1,2}\le -\varepsilon\}}A + b(\varepsilon)$$

where
$$A = \sup\{|\xi(x) - x\xi'(q) + \theta(q)| \;;\; |x|, |q| \le 1\}$$
depends on ξ only. For each $\varepsilon > 0$, the term $A\mathsf{E}\langle \mathbf{1}_{\{R_{1,2}\le -\varepsilon\}}\rangle_t$ vanishes (in average over $\boldsymbol{\beta}$) in the limit by Theorem 12.3.1; and $\lim_{\varepsilon\to 0} b(\varepsilon) = 0$ since ξ is convex in $\mathbb{R}^+$ and hence $\xi(x) - x\xi'(q) + \theta(q) \ge 0$ for $x, q \ge 0$. □

14.5 Method of Proof

The remainder of this chapter is devoted to the proof of the fundamental fact that the bound (14.88) is optimal (under certain extra conditions on ξ), that is
$$\lim_{N\to\infty} p_N = \inf \mathcal{P}_k(\mathbf{m},\mathbf{q}) \;, \tag{14.93}$$
where the infimum is over all values of $k, \mathbf{m}, \mathbf{q}$. This is the "Parisi formula".

Before starting the detailed outline of our approach, let us wonder, assuming that this result is true, how it could ever be proved. Let us choose $k, \mathbf{m}$ and $\mathbf{q}$ so that $\mathcal{P}_k(\mathbf{m},\mathbf{q})$ is close to the infimum $\inf \mathcal{P}_k(\mathbf{m},\mathbf{q})$. Then, for large

N, recalling the functions φ of (14.71) and ψ of (14.86), we have $\varphi(0)=\psi(0)$ and therefore by (14.93) and the choice of $k, \mathbf{m}, \mathbf{q}$,

$$\varphi(1)=p_N \simeq \mathcal{P}_k(\mathbf{m},\mathbf{q})=\psi(1)\ .$$

(Here the quality of the approximation $\simeq$ does not increase with N.) Since $\xi(R_{1,2})-R_{1,2}\xi'(q_{(\alpha,\gamma)})+\theta(q_{(\alpha,\gamma)})\geq 0$, this means by (14.87) that the term

$$\mathsf{E}\langle \xi(R_{1,2})-R_{1,2}\xi'(q_{(\alpha,\gamma)})+\theta(q_{(\alpha,\gamma)})\rangle_t$$

must be typically small. Therefore for $r=1,\ldots,q_{k+1}$, the terms

$$\mathsf{E}\langle \mathbf{1}_{\{(\alpha,\gamma)=r\}}(\xi(R_{1,2})-R_{1,2}\xi'(q_r)+\theta(q_r))\rangle_t \tag{14.94}$$

should be typically small. Since the function $\xi(x)-x\xi'(q)+\theta(q)$ is likely to be >0 for $x\neq q$, a natural condition is that whenever $u\neq q_r$ the quantity

$$\mathsf{E}\langle \mathbf{1}_{\{(\alpha,\gamma)=r\}}\mathbf{1}_{\{R_{1,2}=u\}}\rangle_t \tag{14.95}$$

be small.

Let us fix $1\leq r\leq k+1$ once and for all. The formula (14.47) and the previous discussion motivate the following construction. For $0\leq p\leq k$ we consider independent pairs (z_p^1,z_p^2) of jointly Gaussian r.v.s with the following properties

$$0\leq p\leq k \quad\Rightarrow\quad \mathsf{E}(z_p^1)^2=\mathsf{E}(z_p^2)^2=\xi'(q_{p+1})-\xi'(q_p) \tag{14.96}$$

$$p<r \quad\Rightarrow\quad z_p^1=z_p^2\ ;\qquad p\geq r\quad\Rightarrow\quad z_p^1 \text{ and } z_p^2 \text{ are independent.} \tag{14.97}$$

Each of the sequences $(z_p^\ell)_{0\leq p\leq k}$ is a copy of the sequence $(z_p)_{0\leq p\leq k}$.

Let us define

$$J_{k+1}(u)=\log\sum_{R_{1,2}=u}\exp\Bigg(-\sqrt{t}H_N(\boldsymbol{\sigma}^1)-\sqrt{t}H_N(\boldsymbol{\sigma}^2) + \sum_{\ell=1,2}\sum_{i\leq N}\sigma_i^\ell\Big(h_i+\sqrt{1-t}\sum_{0\leq p\leq k}z_{i,p}^\ell\Big)\Bigg)\ , \tag{14.98}$$

where $(z_{i,p}^1,z_{i,p}^2)$ are i.i.d. copies of the pair (z_p^1,z_p^2). This quantity also depends on t, but the dependence is kept implicit. Consider the sequence (m_p^*) defined by (14.48) and for $1\leq p\leq k$ let us define recursively

$$J_p(u)=\frac{1}{m_p^*}\log \mathsf{E}_p\exp m_p^* J_{p+1}(u)\ , \tag{14.99}$$

where E_p denotes expectation in the r.v.s $(z_{i,n}^\ell)$, $\ell=1,2$, $n\geq p$. The intuition is that suitable upper bounds for

$$\Psi_r(t,u) = \frac{1}{N}\mathsf{E}J_1(u) \tag{14.100}$$

will allow us to show that the quantity (14.95) is very small.

Let us now state precise results. Consider again a jointly Gaussian Hamiltonian as in (14.55) and let us assume that the function ξ satisfies

$$\xi \text{ is convex}\ ; \quad \xi(x) = \xi(-x)\ ; \quad \xi''(x) > 0 \text{ if } x \neq 0\ ; \quad \xi^{(3)}(x) \geq 0 \text{ if } x > 0\ . \tag{14.101}$$

Let us recall the quantity $\mathcal{P}_k(\mathbf{m},\mathbf{q})$ of (14.89) and define

$$\mathcal{P}(\xi,h) = \inf \mathcal{P}_k(\mathbf{m},\mathbf{q})\ ,$$

where the infimum is over all choices of $k, \mathbf{m}$ and $\mathbf{q}$. Then (14.77) means that $p_N \leq \mathcal{P}(\xi,h)$.

Theorem 14.5.1. (The Parisi formula) *If the function ξ satisfies (14.101) then*

$$\lim_{N\to\infty} p_N = \mathcal{P}(\xi,h)\ . \tag{14.102}$$

For technical reasons we shall consider first the case where $h \neq 0$. (Since h is random this means that $\mathsf{E}h^2 > 0$.)

Theorem 14.5.2. *If $h \neq 0$ and if the function ξ satisfies (14.101) then (14.102) holds.*

Proof of Theorem 14.5.1. Theorem 14.5.1 follows from Theorem 14.5.2 simply by considering the case where h is a constant and letting $h \to 0$. To justify the interchange of the limits $h \to 0$ and $n \to \infty$ we simply observe that, with obvious notation, we have $|p_n(h) - p_n(0)| \leq h$. ☐

The basic idea to prove Theorem 14.5.2 is as follows. Given $t_0 < 1$, there is a "suitable choice" of k and of sequences $\mathbf{m}$ and $\mathbf{q}$ such that if φ and ψ denote the functions of Proposition 14.4.2, whenever $t \leq t_0$, we have $\lim_{N\to\infty} \varphi(t) = \psi(t)$. The first task is to give the technical definition of this "suitable choice". First, let us stress that when we consider a sequence $\mathbf{m}$ as in (14.69) we always define $m_0 = 0$ and $m_{k+1} = 1$, but we do not think of m_0 and m_{k+1} as being part of the sequence $\mathbf{m}$. The reason for this point of view is that we can remove any term we wish from the sequence $\mathbf{m}$, and, changing k into $k-1$, we get another sequence of interest; but we cannot remove the terms m_0 and $m_{k+1} = 1$. Similarly when we consider a sequence $\mathbf{q} = (q_1, \ldots, q_{k+1})$ we always define $q_0 = 0$ and $q_{k+2} = 1$, but these two values are not part of the sequence $\mathbf{q}$.

Definition 14.5.3. *Given $\varepsilon > 0$ we say that $k, \mathbf{m}, \mathbf{q}$ satisfy condition* $\mathrm{MIN}(\varepsilon)$ *if the following occurs. First the sequence $\mathbf{m} = (m_1, \ldots, m_k)$ satisfies (14.69) i.e.*

$$0 < m_1 < \ldots < m_k < 1 \,, \tag{14.103}$$

and the sequence $\mathbf{q} = (q_1, \ldots, q_{k+1})$ *satisfies*

$$0 \le q_1 < q_2 < \ldots < q_k < q_{k+1} \le 1 \,. \tag{14.104}$$

Next,

$$\mathcal{P}_k(\mathbf{m}, \mathbf{q}) \le \mathcal{P}(\xi, h) + \varepsilon \tag{14.105}$$

and

$$\mathcal{P}_k(\mathbf{m}, \mathbf{q}) \text{ realizes the minimum of } \mathcal{P}_k \text{ over all choices of } \mathbf{m} \text{ and } \mathbf{q}. \tag{14.106}$$

Theorem 14.5.4. *Assume that* $h \ne 0$. *Given* $t_0 < 1$, *there exists a number* $\varepsilon > 0$, *depending only on* t_0, ξ *and* h, *with the following property. Assume that* $k, \mathbf{m}, \mathbf{q}$ *satisfy condition* MIN(ε). *Then, if* φ *and* ψ *denote the functions of Proposition 14.4.2, we have*

$$\forall t \le t_0 \,, \lim_{N \to \infty} \varphi(t) = \psi(t) \,. \tag{14.107}$$

In order to be able to use Theorem 14.5.4 at all, we need to know that there exist $k, \mathbf{m}, \mathbf{q}$ that satisfy condition MIN(ε).

Lemma 14.5.5. *Given any* $\varepsilon > 0$ *we can find* $k, \mathbf{m}, \mathbf{q}$ *that satisfy condition* MIN(ε).

Proof. First, by definition of (14.105) we can find $k, \mathbf{m}$ and $\mathbf{q}$ such that $\mathcal{P}_k(\mathbf{m}, \mathbf{q}) \le \mathcal{P}(\xi, h) + \varepsilon$. Next we observe that the quantity $\mathcal{P}_k(\mathbf{m}, \mathbf{q})$ is well defined for $0 \le q_1 \le \ldots \le q_{k+1} \le 1$ and $0 \le m_1 \le \ldots \le m_k \le 1$, provided we define the relation

$$X_p = \frac{1}{m_p} \log \mathsf{E}_p \exp m_p X_{p+1}$$

as $X_p = \mathsf{E}_p X_{p+1}$ when $m_p = 0$. Given k, it follows from Lemma 14.2.3 that this is a continuous function of the parameters $m_1, \ldots, m_k, q_1, \ldots, q_{k+1}$. Therefore we may chose

$$0 \le m_1 \le m_2 \le \ldots \le m_k \le 1$$

and

$$0 \le q_1 \le q_2 \le \ldots \le q_k \le 1 \,,$$

such the quantity $\mathcal{P}_k(\mathbf{m}, \mathbf{q})$ is as small as possible among all possible choices of $\mathbf{m}$ and $\mathbf{q}$. The idea of the proof is that we will obtain the required triplet $(k, \mathbf{m}, \mathbf{q})$ by deleting if necessary certain terms of the sequences $\mathbf{m}$ and $\mathbf{q}$, and decreasing the value of k accordingly. The underlying idea is that the quantity $\mathcal{P}_k(\mathbf{m}, \mathbf{q})$ depends on the lists $\mathbf{m}$ and $\mathbf{q}$ only through the probability measure on $[0,1]$ that gives mass $m_p - m_{p-1}$ to each of the points q_p, $1 \le p \le k+1$.

We call the lists $\mathbf{m}$, $\mathbf{q}$ just constructed the *original lists*, and we show how to delete certain terms from these original lists to obtain condition (14.103). We set $m_{k+1} = 1 = q_{k+2}$ and we make the following observations. First, if for some $1 \le p \le k$ we have $m_p = m_{p+1}$, we may remove q_{p+1} from the list $q_1, \ldots, q_{k+1}$ and m_p from the list $m_1, \ldots, m_k$ and change k into $k-1$ without changing the value of $\mathcal{P}_k(\mathbf{m}, \mathbf{q})$. We "merge the intervals $[q_p, q_{p+1}[$ and $[q_{p+1}, q_{p+2}[$".(The easy formal argument is given in Lemma 14.7.1 below.) Similarly, if $m_1 = 0$ we may remove m_1 from the list $m_1, \ldots, m_k$ and q_1 from the list $q_1, \ldots, q_{k+1}$, and change k into $k-1$ without changing the value of $\mathcal{P}_k(\mathbf{m}, \mathbf{q})$. In this manner we may assume that (14.103) holds.

Next, if for some $1 \le p \le k$ we have $q_p = q_{p+1}$, we can remove q_{p+1} from the list $q_1, \ldots, q_{k+1}$ and m_p from the list $m_1, \ldots, m_k$ and replace k by $k-1$ without changing the value of $\mathcal{P}_k(\mathbf{m}, \mathbf{q})$. Thus we may assume that in the final list we have

$$0 = q_0 \le q_1 < q_2 < \ldots < q_k < q_{k+1} \le 1 = q_{k+2} \ ,$$

i.e. (14.104). □

Later on, we will show that in fact $q_1 > 0$ and $q_{k+1} < 1$.

Proof of Theorem 14.5.2. Consider $t_0 < 1$, $\varepsilon > 0$ as in Theorem 14.5.4 and $k, \mathbf{m}, \mathbf{q}$ that satisfy $\mathrm{MIN}(\varepsilon)$. By (14.107) we have

$$\lim_{N \to \infty} \varphi(t_0) = \psi(t_0). \tag{14.108}$$

By (14.87), there exists a number M depending on ξ only such that for $0 \le t \le 1$

$$\psi'(t) - M \le \varphi'(t) \le \psi'(t) \ , \tag{14.109}$$

and combining with (14.108) this implies

$$\liminf_{N \to \infty} \varphi(1) \ge \psi(1) - (1 - t_0)M \ .$$

Since $\varphi(1) = p_N$ and $\psi(1) = \mathcal{P}_k(\mathbf{m}, \mathbf{q})$, we get

$$\liminf_{N \to \infty} p_N \ge \mathcal{P}_k(\mathbf{m}, \mathbf{q}) - (1 - t_0)M \ge \mathcal{P}(\xi, h) - (1 - t_0)M \ ,$$

and the result since t_0 is arbitrary. □

We will deduce Theorem 14.5.4 from the following.

Proposition 14.5.6. *When $h \ne 0$, given $t_0 < 1$, there exists $\varepsilon > 0$, depending only on t_0, ξ and h, with the following properties. Assume that $k, \mathbf{m}, \mathbf{q}$ satisfy* $\mathrm{MIN}(\varepsilon)$*. Then we can find a number K with the following property. Given any $\varepsilon_1 > 0$ and any $1 \le r \le k+1$, for N large enough we have*

$$t \le t_0 \Rightarrow \mathsf{E}\langle \mathbf{1}_{\{(\alpha,\gamma)=r\}} \mathbf{1}_A \rangle_t \le \varepsilon_1 \ , \tag{14.110}$$

where

$$A = \{(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) \ ; \ (R_{1,2} - q_r)^2 \ge K(\psi(t) - \varphi(t) + \varepsilon_1)\} \ . \tag{14.111}$$

As the order of the quantifiers indicates, K might depend on $\xi, t_0, h, \varepsilon, \mathbf{q}$ and $\mathbf{m}$, but certainly does not depend on N, t or ε_1. In the remainder of the proof of this proposition we denote by K a constant depending on $\xi, t_0, h, \varepsilon, \mathbf{q}$ and $\mathbf{m}$ only, that may vary between occurrences. In particular, K never depends on N, t or ε_1.

Proof of Theorem 14.5.4. Since ξ is twice continuously differentiable, we have

$$|\xi(R_{1,2}) - R_{1,2}\xi'(q_r) + \theta(q_r)| \le K(R_{1,2} - q_r)^2 \ .$$

Now $|R_{1,2} - q_r| \le 2$, and therefore

$$(R_{1,2} - q_r)^2 \le 4\mathbf{1}_A + K(\psi(t) - \varphi(t) + \varepsilon_1) \ ,$$

so we deduce from (14.110) that for N large enough

$$\begin{aligned} &\mathsf{E}\big\langle \mathbf{1}_{\{(\alpha,\gamma)=r\}}(\xi(R_{1,2}) - R_{1,2}\xi'(q_r) + \theta(q_r))\big\rangle_t \\ &\le K\mathsf{E}\langle \mathbf{1}_{\{(\alpha,\gamma)=r\}}(R_{1,2} - q_r)^2\rangle_t \\ &\le K(4\mathsf{E}\langle \mathbf{1}_{\{(\alpha,\gamma)=r\}}\mathbf{1}_A\rangle_t + K(\psi(t) - \varphi(t) + \varepsilon_1)) \\ &\le K(\psi(t) - \varphi(t) + \varepsilon_1) \end{aligned}$$

and by summation over $1 \le r \le k+1$ (and using of course the fact that K might depend on k),

$$\mathsf{E}\langle \xi(R_{1,2}) - R_{1,2}\xi'(q_{(\alpha,\gamma)}) + \theta(q_{(\alpha,\gamma)})\rangle_t \le K(\varepsilon_1 + \psi(t) - \varphi(t))$$

so that (14.87) implies

$$(\psi(t) - \varphi(t))' \le K(\psi(t) - \varphi(t) + \varepsilon_1) \ .$$

Since $\psi(0) - \varphi(0) = 0$, and since K does not depend on ε_1 the result follows from integration using Lemma A.11.1. □

Consider sequences $\mathbf{m}$ and $\mathbf{q}$ as in (14.103) and (14.104). For $1 \le r \le k+1$ we recall the quantity $\Psi_r(t,u)$ of (14.100). We come to the central ingredient of the proof of the Parisi formula. It will be proved in Section 14.8.

Theorem 14.5.7. *Assume that $h \ne 0$. If $t_0 < 1$, there is a number $\varepsilon > 0$, depending only on t_0, ξ and h such that whenever $k, \mathbf{m}, \mathbf{q}$ satisfy* MIN(ε)*, then for all $t \le t_0$ and each $1 \le r \le k+1$ we have*

$$\Psi_r(t,u) \le 2\psi(t) - \frac{(u - q_r)^2}{K} \ . \tag{14.112}$$

Here as usual K does not depend on u, t or N (but might depend on ξ, h, t_0, k, $\mathbf{m}$ and $\mathbf{q}$.)

Proposition 14.5.8. *Assume that for some $\varepsilon_2 > 0$ (independent of N, u or t) we have*

$$\Psi_r(t,u) \le 2\varphi(t) - \varepsilon_2 \; . \tag{14.113}$$

Then we have

$$\mathsf{E}\langle \mathbf{1}_{\{(\alpha,\gamma)=r\}} \mathbf{1}_{\{R_{1,2}=u\}}\rangle_t \le K \exp\Big(-\frac{N}{K}\Big) \tag{14.114}$$

where K does not depend on u, N or t.

Proof of Proposition 14.5.6. Let us denote by K_0 the constant of (14.112). We prove (14.110) when the constant K of (14.111) is taken equal to $2K_0$. We observe that

$$\mathsf{E}\langle \mathbf{1}_{\{(\alpha,\gamma)=r\}} \mathbf{1}_A\rangle_t \le \sum \mathsf{E}\langle \mathbf{1}_{\{(\alpha,\gamma)=r\}} \mathbf{1}_{\{R_{1,2}=u\}}\rangle_t \; , \tag{14.115}$$

where the summation is over values of u of the type n/N ($n \in \mathbb{Z}$, $|n| \le N$) with $(u-q_r)^2 \ge 2K_0(\psi(t)-\varphi(t)+\varepsilon_1)$. But then for such a choice of u we have $\Psi_r(t,u) \le 2\varphi(t) - 2\varepsilon_1$ by (14.112). Therefore (14.113) holds for $\varepsilon_2 = 2\varepsilon_1$, so that by (14.114) and since there are at most $2N+1$ terms in the summation, the right-hand side of (14.115) is $\le KN\exp(-N/K)$, and this is $< \varepsilon_1$ for large N. □

Proof of Proposition 14.5.8. Let us recall the r.v.s z_p of (14.72) and consider i.i.d. copies $(z_{i,p})$ of these. Define $\mathbf{z}_p = (z_{i,p})_{i\le N}$. Recall the function $F(\mathbf{x}_1,\ldots,\mathbf{x}_k)$ of (14.78), and define

$$\begin{aligned} F_{k+1} &= F(\mathbf{z}_1,\ldots,\mathbf{z}_k) \\ &= \log \sum_{\boldsymbol{\sigma}} \exp\Big(-\sqrt{t} H_N(\boldsymbol{\sigma}) + \sqrt{1-t} \sum_{i\le N} \sigma_i \Big(z_{i,0} + \sum_{1\le p\le k} z_{i,p}\Big) + \sum_{i\le N} \sigma_i h_i\Big) \; . \end{aligned} \tag{14.116}$$

Then define recursively

$$F_p = \frac{1}{m_p} \log \mathsf{E}_p \exp m_p F_{p+1} \; .$$

Theorem 14.2.1 implies that

$$\varphi(t) = \frac{1}{N} \mathsf{E} F_1 \; .$$

Let us define, with the notation of (14.98)

$$\begin{aligned} J_{k+1} = \log \sum_{\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2} \exp\Big(&-\sqrt{t} H_N(\boldsymbol{\sigma}^1) - \sqrt{t} H_N(\boldsymbol{\sigma}^2) \\ &+ \sum_{\ell=1,2} \sum_{i\le N} \sigma_i^\ell \Big(h_i + \sqrt{1-t} \sum_{0\le p\le k} z_{i,p}^\ell\Big)\Big) \; , \end{aligned}$$

so the difference between J_{k+1} and the quantity $J_{k+1}(u)$ of (14.98) is that in the latter quantity the summation is restricted over the pairs $(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2)$ with $R_{1,2} = u$. Let us define recursively

$$J_p = \frac{1}{m_p^*} \log \mathsf{E}_p \exp m_p^* J_{p+1} \ ,$$

where the sequence m_p^* is as in (14.48). The formula (14.50) used for $p = 1$ and the function $F(\mathbf{x}_1, \ldots, \mathbf{x}_k)$ of (14.78) implies that

$$2\varphi(t) = \frac{1}{N} \mathsf{E} J_1 \ .$$

We strongly encourage the reader to repeat the argument of that proof, that is to show by decreasing induction over p that $J_p = F_p^1 + F_p^2$. Thus (14.113) means that $\mathsf{E} J_1(u) < \mathsf{E} J_1 - \varepsilon_2 N$.

Let us now consider the function

$$U^{\sim}((\mathbf{x}_1^1, \mathbf{x}_1^2), \ldots, (\mathbf{x}_k^1, \mathbf{x}_k^2)) = \frac{A((\mathbf{x}_1^1, \mathbf{x}_1^2), \ldots, (\mathbf{x}_k^1, \mathbf{x}_k^2))}{\exp F(\mathbf{x}_1^1, \ldots, \mathbf{x}_k^1) \exp F(\mathbf{x}_1^2, \ldots, \mathbf{x}_k^2)} \ ,$$

where

$$\begin{aligned} &A((\mathbf{x}_1^1, \mathbf{x}_1^2), \ldots, (\mathbf{x}_k^1, \mathbf{x}_k^2)) \\ &= \sum_{R_{1,2}=u} \exp\Big(-\sqrt{t} H_N(\boldsymbol{\sigma}^1) - \sqrt{t} H_N(\boldsymbol{\sigma}^2) \\ &+ \sum_{\ell=1,2} \sum_{i \le N} \sigma_i^\ell \Big(h_i + \sqrt{1-t} \Big(z_{i,0} + \sum_{1 \le p \le k} x_{i,p}^\ell \Big) \Big) \Big) \ . \end{aligned}$$

We recall that $\mathbf{z}_p^\ell = (z_{i,p}^\ell)_{i \le N}$, so that

$$A((\mathbf{z}_1^1, \mathbf{z}_1^2), \ldots, (\mathbf{z}_k^1, \mathbf{z}_k^2)) = \exp J_{k+1}(u)$$

and

$$\exp F(\mathbf{z}_1^1, \ldots, \mathbf{z}_k^1) \exp F(\mathbf{z}_1^2, \ldots, \mathbf{z}_k^2) = \exp J_{k+1} \ .$$

Thus, recalling the formula (14.45) i.e. $U_{k+1}^{\sim} = U^{\sim}((\mathbf{z}_1^1, \mathbf{z}_1^2), \ldots, (\mathbf{z}_k^1, \mathbf{z}_k^2))$ we have

$$U_{k+1}^{\sim} = \exp(J_{k+1}(u) - J_{k+1}) \ .$$

Recalling the formulas (14.7) and (14.46), we obtain that

$$\begin{aligned} U^{\sim}(\alpha, \gamma) \exp F(\alpha) \exp F(\gamma) &= A((\mathbf{z}_{1,\alpha}, \mathbf{z}_{1,\gamma}), \ldots, (\mathbf{z}_{k,\alpha}, \mathbf{z}_{k,\gamma})) \\ &= \sum_{R_{1,2}=u} \exp\Big(-\sqrt{t} H_N(\boldsymbol{\sigma}^1) - \sqrt{t} H_N(\boldsymbol{\sigma}^2) \\ &+ \sum_{i \le N} \sigma_i^1 \Big(h_i + \sqrt{1-t} \Big(z_{i,0} + \sum_{1 \le p \le k} z_{i,p,\alpha} \Big) \Big) \\ &+ \sum_{i \le N} \sigma_i^2 \Big(h_i + \sqrt{1-t} \Big(z_{i,0} + \sum_{1 \le p \le k} z_{i,p,\gamma} \Big) \Big) \Big) \ . \end{aligned}$$

If follows that

$$\langle \mathbf{1}_{\{(\alpha,\gamma)=r\}} U^{\sim}(\alpha,\gamma)\rangle = \frac{\sum_{(\alpha,\gamma)=r} v_\alpha v_\gamma U^{\sim}(\alpha,\gamma)\exp F(\alpha)\exp F(\gamma)}{(\sum v_\gamma \exp F(\gamma))^2}$$
$$= \langle \mathbf{1}_{\{(\alpha,\gamma)=r\}} \mathbf{1}_{\{R_{1,2}=u\}}\rangle_t \ .$$

Therefore if $V_p = \exp m_p^*(J_{p+1} - J_p)$, (14.52) implies that

$$\mathsf{E}\langle \mathbf{1}_{\{(\alpha,\gamma)=r\}} \mathbf{1}_{\{R_{1,2}=u\}}\rangle_t = (m_r - m_{r-1})\mathsf{E}(V_1 \cdots V_k B) \ ,$$

where

$$B := U_{k+1}^{\sim} = \exp(J_{k+1}(u) - J_{k+1}) \le 1 \ . \tag{14.117}$$

Consequently, to prove Proposition 14.5.8 it suffices to show that

$$\mathsf{E}J_1(u) \le \mathsf{E}J_1 - \varepsilon_2 N \quad \Rightarrow \quad \mathsf{E}(V_1 \cdots V_k B) \le K \exp\left(-\frac{N}{K}\right) \ . \tag{14.118}$$

So, we assume

$$\mathsf{E}J_1(u) \le \mathsf{E}J_1 - \varepsilon_2 N \ . \tag{14.119}$$

We prove first by decreasing induction on p that

$$J_{p+1}(u) \ge J_{p+1} + \frac{1}{m_{p+1}^*} \log \mathsf{E}_{p+1}(V_{p+1} \cdots V_k B) \ . \tag{14.120}$$

For $p = k$, we have $\mathsf{E}_{p+1}(V_{p+1} \cdots V_k B) = B$ and the last term in (14.120) is $\log B$ and since $m_{k+1}^* = m_{k+1} = 1$, (14.120) then follows from (14.117). Now, since

$$\mathsf{E}_p V_p = 1$$

and V_p does not depend on the r.v.s $z_{i,n}^\ell$ for $n > p$, we see recursively that

$$\mathsf{E}_p(V_p \cdots V_k) = \mathsf{E}_p(\mathsf{E}_{p+1}(V_p \cdots V_k)) = \mathsf{E}_p(V_p \mathsf{E}_{p+1}(V_{p+1} \cdots V_k)) = 1$$

and thus since $B \le 1$ we have

$$\mathsf{E}_p(V_p \cdots V_k B) \le 1 \ ,$$

so that (14.120) implies that, since $m_p^* < m_{p+1}^*$

$$J_{p+1}(u) \ge J_{p+1} + \frac{1}{m_p^*} \log \mathsf{E}_{p+1}(V_{p+1} \cdots V_k B) \ ,$$

and, using the definition of V_p in the second line

$$\exp m_p^* J_{p+1}(u) \ge \mathsf{E}_{p+1}(V_{p+1} \cdots V_k B) \exp m_p^* J_{p+1}$$
$$= V_p \mathsf{E}_{p+1}(V_{p+1} \cdots V_k B) \exp m_p^* J_p$$
$$= \mathsf{E}_{p+1}(V_p \cdots V_k B) \exp m_p^* J_p \ .$$

Since J_p does not depend on the r.v.s $(z_{i,n}^\ell)$ for $n \geq p$, and since $\mathsf{E}_p = \mathsf{E}_p \mathsf{E}_{p+1}$, taking expectation E_p in the last inequality entails

$$\mathsf{E}_p \exp m_p^* J_{p+1}(u) \geq (\exp m_p^* J_p) \mathsf{E}_p (V_p \cdots V_k B) .$$

Since $J_p(u) = (m_p^*)^{-1} \log \mathsf{E}_p \exp m_p^* J_{p+1}(u)$, taking logarithms completes the induction.

Using (14.120) for $p = 0$, we get

$$\log \mathsf{E}_1 (V_1 \cdots V_k B) \leq m_1^* (J_1(u) - J_1)$$

and taking expectation and using (14.119) yields

$$\mathsf{E} \log \mathsf{E}_1 (V_1 \cdots V_k B) \leq -m_1^* \varepsilon_2 N . \tag{14.121}$$

The quantity $\mathsf{E}_1(V_1 \cdots V_k B)$ depends only on the r.v.s $H_N(\boldsymbol{\sigma})$, $z_{0,i}^\ell$ and h_i. We will use concentration of measure to prove that

$$\mathsf{P}\left(|\log \mathsf{E}_1(V_1 \cdots V_k B) - \mathsf{E} \log \mathsf{E}_1(V_1 \cdots V_k B)| \geq \frac{m_1^* \varepsilon_2}{2} N \right)$$
$$\leq K_1 \exp\left(-\frac{N}{K_1} \right) . \tag{14.122}$$

The proof (14.118) is then completed as follows. From (14.121) and (14.122) we deduce

$$\mathsf{P}\left(\log \mathsf{E}_1(V_1 \cdots V_k B) > -\frac{m_1^* \varepsilon_2}{2} N \right) \leq K_1 \exp\left(-\frac{N}{K_1} \right) ,$$

so that

$$\mathsf{P}\left(\mathsf{E}_1(V_1 \cdots V_k B) > \exp\left(-\frac{m_1^* \varepsilon_2}{2} N \right) \right) \leq K_1 \exp\left(-\frac{N}{K_1} \right) ,$$

and since $\mathsf{E}_1(V_1 \cdots V_k B) \leq 1$ we obtain

$$\mathsf{E}(V_1 \cdots V_k B) = \mathsf{E}\mathsf{E}_1(V_1 \cdots V_k B) \leq K \exp(-N/K) .$$

We turn to the proof of (14.122). We will use the following representation of the r.v.s $H_N(\boldsymbol{\sigma})$: there exist M and vectors $\mathbf{x}(\boldsymbol{\sigma}) \in \mathbb{R}^M$ such that the family $(H_N(\boldsymbol{\sigma}))_{\boldsymbol{\sigma}}$ of r.v.s has the same distribution as the family $(\mathbf{x}(\boldsymbol{\sigma}) \cdot \mathbf{g})_{\boldsymbol{\sigma}}$, where $\mathbf{g} \in \mathbb{R}^M$ is a standard Gaussian vector. This is explained in Section A.2, and is obvious in the case where H_N is a sum of independent terms of the type (14.57). We note that $\|\mathbf{x}(\boldsymbol{\sigma})\|^2 = \mathsf{E} H_N(\boldsymbol{\sigma})^2 = N\xi(1)$. The basic idea is to show that, as a function of $\mathbf{g}$, the quantity $\log \mathsf{E}_1(V_1 \cdots V_k B)$ has a Lipschitz constant $\leq K\sqrt{N}$. We start the proof by using (14.54) that reads here with the usual notation as

$$\mathsf{E}_1(V_1\cdots V_k B) = \mathsf{E}'\bigg(\frac{1}{Z}\sum_{R_{1,2}=u} v_\alpha \exp\Big(-\sqrt{t}(\mathbf{x}(\boldsymbol{\sigma}^1)\cdot\mathbf{g}+\mathbf{x}(\boldsymbol{\sigma}^2)\cdot\mathbf{g}) + \sum_{i\le N}\sigma_i^1\Big(h_i+\sqrt{1-t}\Big(z_{i,0}+\sum_{1\le p\le k} z_{i,p,\alpha}\Big)\Big) + \sum_{i\le N}\sigma_i^2\Big(h_i+\sqrt{1-t}\Big(z_{i,0}+\sum_{1\le p\le k} z_{i,p,\gamma}\Big)\Big)\Big)\bigg), \quad (14.123)$$

where Z is the normalizing factor and E' denotes expectation in the r.v.s $z^\ell_{i,p,\alpha}$ for $p\ge 1$ and in the weights v_α. We realize that the right-hand side of (14.123) depends not only on $\mathbf{g}$ but also on the r.v.s $z_{i,0}$ and h_i. For simplicity we will complete the proof only in the case where h_i is not random. (When h_i is random not necessarily Gaussian, some extra work is required, using martingale difference sequences. This is left to the enterprising reader.) To understand the dependence of the right-hand side of (14.123) on the r.v.s $z_{i,0}$ we set

$$g'_i = \frac{z_{i,0}}{\sqrt{\mathsf{E}z^2_{i,0}}} = \frac{z_{i,0}}{\sqrt{\xi'(q_1)}}\ ,$$

so that the vector $\mathbf{g}' = (g'_i)_{i\le N}$ is a standard Gaussian random vector in $\mathbb{R}^N$. Then (14.123) allows one to consider the quantity $\log \mathsf{E}_1(V_1\cdots V_k B)$ as a function of the pair $(\mathbf{g},\mathbf{g}')$. One then sees by direct computation of the gradient and trivial estimates that the Lipschitz constant of this function is $\le K\sqrt{N}$. □

The center of the approach is Theorem 14.5.7. The basic tool will be a kind of "two dimensional" extension of Guerra's bound (14.88), in the spirit of Theorem 13.5.1. It is the purpose of the next section. This bound will depend on many parameters. We will exhibit a choice of these parameters that proves Theorem 14.5.1. A crucial idea is that we will not use an optimal choice of parameters, because we want to avoid the (intractable?) corresponding optimization problem. Rather (as in the proof of Theorem 13.6.2) we will observe a choice of parameters that witnesses the obvious bound $\Psi_r(t,u)\le 2\psi(t)$. We will then use a variational argument to show that a suitable small change of these parameters improves enough the bound to reach (14.112).

14.6 Bounds for Coupled Copies

In this section we develop bounds for the expression $\Psi_r(t,u)$ of (14.100), a first step towards the proof of Theorem 14.5.7. First, let us transform this expression using Theorem 14.2.1. For each p we consider independent copies $(z^1_{i,p,j_1,\dots,j_p}, z^2_{i,p,j_1,\dots,j_p})$ of the pair (z^1_p, z^2_p) of (14.96) and (14.97). For $\alpha\in\mathbb{N}^{*k}$, $\alpha=(j_1,\dots,j_k)$, we define

$$(z^1_{i,p,\alpha}, z^2_{i,p,\alpha}) = (z^1_{i,p,j_1,\ldots,j_p}, z^2_{i,p,j_1,\ldots,j_p}) \ . \tag{14.124}$$

Consider weights (v_α) forming a Poisson-Dirichlet cascade as in (14.13), but associated to the sequence (14.48) rather than to the sequence $m_1, \ldots, m_k$. Then

$$\begin{aligned}\Psi_r(t,u) = \frac{1}{N}\mathsf{E}\log \sum_{R_{1,2}=u,\alpha} v_\alpha \exp\Bigg(&-\sqrt{t}H_N(\boldsymbol{\sigma}^1) - \sqrt{t}H_N(\boldsymbol{\sigma}^2) \\ &+ \sum_{\ell=1,2}\sum_{i\le N}\sigma_i^\ell\Big(h_i + \sqrt{1-t}\sum_{p\le k} z^\ell_{i,p,\alpha}\Big)\Bigg) \ . \end{aligned} \tag{14.125}$$

The expression "coupled copies" refers to the fact that in (14.125) the summation is only over all pairs $(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2)$ with $R_{1,2} = u$. To find bounds for such a quantity we will develop a kind of two-dimensional version of the scheme of Lemma 14.4.1.

Let us first think of a rather general problem, to find bounds for a quantity

$$\frac{1}{N}\mathsf{E}\log \sum_{R_{1,2}=u,\alpha} w_\alpha \exp\big(-\sqrt{t}H_N(\boldsymbol{\sigma}^1) - \sqrt{t}H_N(\boldsymbol{\sigma}^2) - H^0(\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2,\alpha)\big) \tag{14.126}$$

where α belongs to countable set A, w_α are random weights, and where $H^0(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2, \alpha)$ is a random function, independent of the randomness of H_N. The summation is over all values of α and all values of $\boldsymbol{\sigma}^1$ and $\boldsymbol{\sigma}^2$ for which $R_{1,2} = u$. Of course $\sqrt{t}H_N$ is a general Hamiltonian of the type (14.55) where ξ is replaced by $t\xi$, so we assume $t = 1$ to simplify notation. We first present a general scheme. This scheme might look complicated at a first sight, but it is in fact a straightforward generalization of the scheme of Lemma 14.4.1. The system consists now of pairs $(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2)$ of configurations. We need a "replica" of this system, and we denote by $(\boldsymbol{\tau}^1, \boldsymbol{\tau}^2)$ another pair of configurations. For $\ell, \ell' = 1, 2$ we use the notation

$$R^{\ell,\ell'} = R(\boldsymbol{\sigma}^\ell, \boldsymbol{\tau}^{\ell'}) = \frac{1}{N}\sum_{i\le N}\sigma_i^\ell \tau_i^{\ell'} \ .$$

The interpolating Hamiltonian is a family $H(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2, \alpha)$ $(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2 \in \Sigma_N, \alpha \in A)$ of jointly Gaussian r.v.s, independent of all other randomnesses. We assume that there exists numbers $q^{\ell,\ell'}_{\alpha,\gamma}$ such that, for any $\alpha, \gamma \in A$, any $\ell, \ell' = 1, 2$ we have

$$\frac{1}{N}\mathsf{E}H(\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2,\alpha)H(\boldsymbol{\tau}^1,\boldsymbol{\tau}^2,\gamma) = \sum_{\ell,\ell'=1,2} R^{\ell,\ell'}\xi'(q^{\ell,\ell'}_{\alpha,\gamma}) \ . \tag{14.127}$$

We assume also that for each α

$$q^{1,1}_{\alpha,\alpha} = q^{2,2}_{\alpha,\alpha} = 1 \ ; \qquad q^{1,2}_{\alpha,\alpha} = q^{2,1}_{\alpha,\alpha} = u \ . \tag{14.128}$$

For $0 \le s \le 1$ we consider the Hamiltonian

$$-H_s(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2, \alpha) = -\sqrt{s}(H_N(\boldsymbol{\sigma}^1) + H_N(\boldsymbol{\sigma}^2)) - \sqrt{1-s}H(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2, \alpha) - H^0(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2, \alpha) .$$

For a function $f(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2, \alpha)$ we define

$$\langle f(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2, \alpha)\rangle_s = \frac{1}{Z_s} \sum_{R_{1,2}=u,\alpha} w_\alpha f(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2, \alpha) \exp(-H_s(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2, \alpha)) ,$$

where Z_s is the normalizing factor $\sum_{R_{1,2}=u,\alpha} w_\alpha \exp(-H_s(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2, \alpha))$.

Lemma 14.6.1. *Assume that ξ is convex, and recall the definition (14.60) of θ. Then the function*

$$\varphi^*(s) = \frac{1}{N} \mathsf{E} \log \sum_{R_{1,2}=u,\alpha} w_\alpha \exp(-H_s(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2, \alpha))$$

satisfies

$$\varphi^{*\prime}(s) \le -\theta(1) - \theta(u) + \frac{1}{2} \sum_{\ell,\ell'=1,2} \mathsf{E}\langle \theta(q^{\ell,\ell'}_{\alpha,\gamma})\rangle_s . \tag{14.129}$$

Proof. Let us define

$$U(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2, \boldsymbol{\tau}^1, \boldsymbol{\tau}^2, \alpha, \gamma) = \frac{1}{N}\Big(\mathsf{E}(H_N(\boldsymbol{\sigma}^1) + H_N(\boldsymbol{\sigma}^2))(H_N(\boldsymbol{\tau}^1) + H_N(\boldsymbol{\tau}^2)) - \mathsf{E}H(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2, \alpha)H(\boldsymbol{\tau}^1, \boldsymbol{\tau}^2, \gamma)\Big) ,$$

so that, using (14.127) and (14.55),

$$U(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2, \boldsymbol{\tau}^1, \boldsymbol{\tau}^2, \alpha, \gamma) = \sum_{\ell,\ell'=1,2} (\xi(R^{\ell,\ell'}) - R^{\ell,\ell'}\xi'(q^{\ell,\ell'}_{\alpha,\gamma})) .$$

In particular, using (14.128), when $R_{1,2} = u$ we have

$$\begin{aligned} U(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2, \boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2, \alpha, \alpha) &= 2\xi(1) + 2\xi(u) - 2\xi'(1) - 2u\xi'(u) \\ &= -2\theta(1) - 2\theta(u) . \end{aligned}$$

Also, since ξ is convex, we have $\xi(R^{\ell,\ell'}) - R^{\ell,\ell'}\xi'(q^{\ell,\ell'}_{\alpha,\gamma}) \ge -\theta(q^{\ell,\ell'}_{\alpha,\gamma})$ so that

$$U(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2, \boldsymbol{\tau}^1, \boldsymbol{\tau}^2, \alpha, \gamma) \ge -\sum_{\ell,\ell'=1,2} \theta(q^{\ell,\ell'}_{\alpha,\gamma}) ,$$

and to conclude we simply differentiate and integrate by parts to obtain that

$$\varphi^{*\prime}(s) = \frac{1}{2}\Big(\mathsf{E}\langle U(\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2,\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2,\alpha,\alpha)\rangle_s - \mathsf{E}\langle U(\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2,\boldsymbol{\tau}^1,\boldsymbol{\tau}^2,\alpha,\gamma)\rangle_s\Big)\ . \quad \square$$

We should stress the main point: the terms created by the interaction between $H_N(\boldsymbol{\sigma}^1)$ and $H_N(\boldsymbol{\sigma}^2)$ is $\langle \xi(R_{1,2}) - R_{1,2}\xi'(q_{\alpha,\alpha})\rangle_s = \langle \xi(R_{1,2}) - R_{1,2}\xi'(u)\rangle_s$. It has the wrong sign to be bounded above by the inequality $\langle \xi(R_{1,2}) - R_{1,2}\xi'(u)\rangle_s \geq -\theta(u)$. It is the restriction of the summation to the pairs with $R_{1,2} = u$ that makes this term equal to $-\theta(u)$ and saves the day.

In Proposition 14.6.3 below we describe a specific choice of the Hamiltonian $H(\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2,\alpha)$, that is sufficient to prove Theorem 14.5.7. It seems however fruitful in the long range to discuss more general schemes. This is in particular the case because, as will be explained in detail in the next chapter, these schemes are related to some of the darkest remaining mysteries.

The scheme we present now seems to be the natural (or even canonical) 2-dimensional generalization of the scheme leading to Guerra's bound (14.88). The main difference is that the sequence $0 = q_0 \leq q_1 \leq \ldots \leq q_{k+1} \leq q_{k+2} = 1$ of (14.70) is now replaced by four different sequences $(\rho_p^{\ell,\ell'})$ for $\ell,\ell' = 1,2$.

Let us assume that $A = \mathbb{N}^{*\kappa}$ for a certain integer κ and that for $\ell,\ell' = 1,2$, $0 \leq p \leq \kappa$ we are given pairs of Gaussian r.v.s (y_p^1, y_p^2) such that, for certain numbers $(\rho_p^{\ell,\ell'})_{0\leq p\leq \kappa+1}$, we have

$$\mathsf{E}y_p^\ell y_p^{\ell'} = \xi'(\rho_{p+1}^{\ell,\ell'}) - \xi'(\rho_p^{\ell,\ell'})\ . \tag{14.130}$$

The sequence $\rho_p^{\ell,\ell'}$ need not be non-decreasing. Let us also assume that

$$\rho_0^{\ell,\ell'} = 0\ ; \tag{14.131}$$

$$\rho_{\kappa+1}^{1,1} = \rho_{\kappa+1}^{2,2} = 1\ ; \qquad \rho_{\kappa+1}^{1,2} = \rho_{\kappa+1}^{2,1} = u\ . \tag{14.132}$$

Then, for $n \leq \kappa + 1$ we have

$$\sum_{0\leq p<n} \mathsf{E}y_p^\ell y_p^{\ell'} = \xi'(\rho_n^{\ell,\ell'})\ . \tag{14.133}$$

Considering independent copies $(y^1_{i,p,j_1,\ldots,j_p}, y^2_{i,p,j_1,\ldots,j_p})$ of the pair (y_p^1, y_p^2), let us write as usual, when $\alpha = (j_1,\ldots,j_\kappa)$

$$(y^1_{i,p,\alpha}, y^2_{i,p,\alpha}) = (y^1_{i,p,j_1,\ldots,j_p}, y^2_{i,p,j_1,\ldots,j_p})\ . \tag{14.134}$$

Let us choose the interpolating Hamiltonian $H(\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2,\alpha)$ as

$$H(\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2,\alpha) = \sum_{\ell=1,2}\sum_{i\leq N}\sigma_i^\ell\bigg(\sum_{0\leq p\leq\kappa} y^\ell_{i,p,\alpha}\bigg)\ . \tag{14.135}$$

We observe that $\mathsf{E}y^\ell_{i,p,\alpha}y^{\ell'}_{i,p,\gamma} = 0$ if $p \geq (\alpha,\gamma)$ and $= \mathsf{E}y_p^\ell y_p^{\ell'}$ if $p < (\alpha,\gamma)$. Using (14.133) it follows that

$$\sum_{0 \le p \le \kappa} \mathsf{E} y_{i,p,\alpha}^{\ell} y_{i,p,\gamma}^{\ell'} = \sum_{0 \le p < (\alpha,\gamma)} \mathsf{E} y_p^{\ell} y_p^{\ell'} = \xi'(\rho_{(\alpha,\gamma)}^{\ell,\ell'})$$

and hence (14.127) holds for $q_{\alpha,\gamma}^{\ell,\ell'} = \rho_{(\alpha,\gamma)}^{\ell,\ell'}$. Finally (14.128) is a consequence of (14.132).

Let us further assume that the Hamiltonian $H^0(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2, \alpha)$ is of the type

$$-H^0(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2, \alpha) = \sum_{\ell=1,2} \sum_{i \le N} \sigma_i^{\ell} \left(h_i + \sum_{0 \le p \le \kappa} Z_{i,p,\alpha}^{\ell} \right), \tag{14.136}$$

where the pairs $(Z_{i,p,\alpha}^1, Z_{i,p,\alpha}^2)$ are built as usual from pairs (Z_p^1, Z_p^2) of jointly Gaussian r.v.s. which will be specified later (although a glance back at (14.125) might provide a clue as to how this will be done). Let us finally assume that the weights (w_α) form a Poisson-Dirichlet cascade associated to a sequence $0 < n_1 < \ldots < n_\kappa < 1$ (so that according to our conventions we will denote these weights by (v_α)). As in the case of (14.76) we deduce from (14.38) that $\mathsf{E}\langle \mathbf{1}_{\{(\alpha,\gamma)=p\}} \rangle_s = n_p - n_{p-1}$, so that (14.129) implies that (defining as usual $n_0 = 0$)

$$\varphi^{*\prime}(s) \le -\theta(1) - \theta(u) + \frac{1}{2} \sum_{\ell,\ell'=1,2} \sum_{1 \le p \le \kappa} (n_p - n_{p-1}) \theta(\rho_p^{\ell,\ell'}) \, . \tag{14.137}$$

Since $\rho_{\kappa+1}^{1,2} = \rho_{\kappa+1}^{2,1} = u$ and $\rho_{\kappa+1}^{1,1} = \rho_{\kappa+1}^{2,2} = 1$, this rearranges as

$$\varphi^{*\prime}(s) \le -\frac{1}{2} \sum_{\ell,\ell'=1,2} \sum_{1 \le p \le \kappa} n_p (\theta(\rho_{p+1}^{\ell,\ell'}) - \theta(\rho_p^{\ell,\ell'})) - (\theta(1) + \theta(u))(1 - n_\kappa) \, , \tag{14.138}$$

and we have bounded the quantity (14.126) by

$$\varphi^*(0) - \frac{1}{2} \sum_{\ell,\ell'=1,2} \sum_{1 \le p \le \kappa} n_p (\theta(\rho_{p+1}^{\ell,\ell'}) - \theta(\rho_p^{\ell,\ell'})) - (\theta(1) + \theta(u))(1 - n_\kappa) \, . \tag{14.139}$$

Let us now try to find a manageable bound for $\varphi^*(0)$. We use the simple fact that for any value of λ and α we have

$$\begin{aligned} &\sum_{R_{1,2}=u} \exp(-H(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2, \alpha) - H^0(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2, \alpha)) \qquad (14.140) \\ &= \exp(-\lambda N u) \sum_{R_{1,2}=u} \exp(-H(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2, \alpha) - H^0(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2, \alpha) + \lambda N R_{1,2}) \\ &\le \exp(-\lambda N u) \sum_{\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2} \exp(-H(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2, \alpha) - H^0(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2, \alpha) + \lambda N R_{1,2}) \, , \end{aligned}$$

where the last summation is over *all* values of $\boldsymbol{\sigma}^1$ and $\boldsymbol{\sigma}^2$. The quantity

$$\frac{1}{N}\mathsf{E}\log\sum_{\alpha,\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2} v_\alpha \exp(-H(\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2,\alpha)-H^0(\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2,\alpha)+\lambda N R_{1,2})$$
$$=\frac{1}{N}\mathsf{E}\log\sum_\alpha v_\alpha \sum_{\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2}\exp\biggl(\sum_{i\le N}\biggl(\sum_{\ell=1,2}\sigma_i^\ell\biggl(h_i+\sum_{0\le p\le\kappa}(y^\ell_{i,p,\alpha}+Z^\ell_{i,p,\alpha})\biggr)$$
$$+\lambda\sigma_i^1\sigma_i^2\biggr)\biggr) \tag{14.141}$$

can be easily computed through Theorem 14.2.1 because the sites decouple exactly as in the case of the computation of (14.80). First let us observe that

$$\sum_{\varepsilon_1,\varepsilon_2=\pm1}\exp(\varepsilon_1x_1+\varepsilon_2x_2+\varepsilon_1\varepsilon_2\lambda)=4(\mathrm{ch}x_1\mathrm{ch}x_2\mathrm{ch}\lambda+\mathrm{sh}x_1\mathrm{sh}x_2\mathrm{sh}\lambda) \tag{14.142}$$

so that, setting $g^\ell_{i,p,\alpha}=y^\ell_{i,p,\alpha}+Z^\ell_{i,p,\alpha}$ the quantity (14.141) is

$$\frac{1}{N}\mathsf{E}\log\sum_\alpha v_\alpha\prod_{i\le N}\biggl(\mathrm{ch}\biggl(h_i+\sum_{0\le p\le\kappa}g^1_{i,p,\alpha}\biggr)\mathrm{ch}\biggl(h_i+\sum_{0\le p\le\kappa}g^2_{i,p,\alpha}\biggr)\mathrm{ch}\lambda$$
$$+\mathrm{sh}\biggl(h_i+\sum_{0\le p\le\kappa}g^1_{i,p,\alpha}\biggr)\mathrm{sh}\biggl(h_i+\sum_{0\le p\le\kappa}g^2_{i,p,\alpha}\biggr)\mathrm{sh}\lambda\biggr)+2\log 2\ . \tag{14.143}$$

Setting $g^\ell_p=y^\ell_p+Z^\ell_p$, let us define

$$Y_{\kappa+1}=\log\biggl(\mathrm{ch}\biggl(h+\sum_{0\le p\le\kappa}g^1_p\biggr)\mathrm{ch}\biggl(h+\sum_{0\le p\le\kappa}g^2_p\biggr)\mathrm{ch}\lambda$$
$$+\mathrm{sh}\biggl(h+\sum_{0\le p\le\kappa}g^1_p\biggr)\mathrm{sh}\biggl(h+\sum_{0\le p\le\kappa}g^2_p\biggr)\mathrm{sh}\lambda\biggr) \tag{14.144}$$

and recursively, for $p\ge1$,

$$Y_p=\frac{1}{n_p}\log\mathsf{E}_p\exp n_pY_{p+1}\ . \tag{14.145}$$

Let us further define $Y_0=\mathsf{E}Y_1$, so that the quantity (14.143) is equal to $2\log2+Y_0$. We therefore have proved the bound

$$\varphi^*(0)\le 2\log2+Y_0-2\lambda u\ .$$

In summary, when H^0 is chosen as in (14.136) and the weights (v_α) as above, we have proved the bound

$$\frac{1}{N}\mathsf{E}\log\sum_{R_{1,2}=u,\alpha}v_\alpha\exp(-H_N(\boldsymbol{\sigma}^1)-H_N(\boldsymbol{\sigma}^2)-H^0(\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2,\alpha))$$
$$\le 2\log2+Y_0-\lambda u-\frac12\sum_{\ell,\ell'=1,2}\sum_{1\le p\le\kappa}n_p(\theta(\rho^{\ell,\ell'}_{p+1})-\theta(\rho^{\ell,\ell'}_p))$$
$$-(\theta(u)+\theta(1))(1-n_\kappa)\ .$$

This bound was proved under the conditions $0 < n_1 < n_2 < \ldots < n_\kappa < 1$ but since the right-hand side is a continuous function of $(n_1, \ldots, n_\kappa)$, the inequality also holds under the less restrictive condition $0 \le n_1 \le n_2 \le \ldots \le n_\kappa \le 1$. In particular when

$$0 \le n_1 \le n_2 \le \ldots \le n_\kappa = 1 \,, \tag{14.146}$$

then

$$\begin{aligned}\frac{1}{N}\mathsf{E}\log \sum_{R_{1,2}=u,\alpha} & v_\alpha \exp(-H_N(\boldsymbol{\sigma}^1) - H_N(\boldsymbol{\sigma}^2) - H^0(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2, \alpha)) \\ \le 2\log 2 + Y_0 - \lambda u - \frac{1}{2} & \sum_{\ell,\ell'=1,2} \sum_{1\le p\le \kappa} n_p(\theta(\rho_{p+1}^{\ell,\ell'}) - \theta(\rho_p^{\ell,\ell'}))\,. \end{aligned} \tag{14.147}$$

Although (14.147) probably looks complicated to the first time reader, this bound is not hard to discover, one simply copies the scheme of proof of Theorem 14.4.3.

The bound (14.147) depends on many parameters, namely the sequence (14.146), the four sequences $(\rho_p^{\ell,\ell'})_{0\le p\le \kappa+1}$ for $\ell, \ell' = 1, 2$, and λ. It does not give a clue on how to choose these parameters in an efficient way. This should be expected, since the same problem already arises in Guerra's bound (14.88). The real problem however is that the "two-dimensional bound" (14.147) does not connect perfectly with the one dimensional bound (14.88). Let us explain this in a simple situation where this problem already arises, namely the case where

$$-H^0(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2, \alpha) = \sum_{i\le N} h_i(\sigma_i^1 + \sigma_i^2)\,. \tag{14.148}$$

In that case the left-hand side of (14.147) is simply

$$\frac{1}{N}\mathsf{E}\log \sum_{R_{1,2}=u} \exp\Big(-H_N(\boldsymbol{\sigma}^1) - H_N(\boldsymbol{\sigma}^2) + \sum_{i\le N} h_i(\sigma_i^1 + \sigma_i^2)\Big)\,. \tag{14.149}$$

There is a natural bound for this quantity, namely the sum $\sum_{R_{1,2}=u}$ is smaller than the sum for all values of $\boldsymbol{\sigma}^1$ and $\boldsymbol{\sigma}^2$, and therefore the quantity (14.149) is $\le 2p_N$, where p_N is as in (14.88). Consequently, given sequences $\mathbf{m}$ and $\mathbf{q}$ as in Section 14.4, the quantity (14.149) is $\le 2\mathcal{P}_k(\mathbf{m}, \mathbf{q})$. We would expect that the complicated bound (14.147) should be able to recover this simple bound; that is, we should be able to choose the parameters in (14.147) so that the right-hand side is $\le 2\mathcal{P}_k(\mathbf{m}, \mathbf{q})$. It turns out that this can be done when $u \ge 0$, in a way that connects extremely well with the discussion at the end of Section 14.4 (but we do not know how to do it in full generality when $u < 0$). This is somehow the center of the proof of Parisi's formula. Let us consider sequences $0 \le q_1 \le \ldots \le q_k \le q_{k+1} \le 1$ and $0 \le m_1 \le \ldots \le m_{k-1} \le m_k \le 1$ as in (14.70) and (14.91) and $\mathbf{m} = (m_1, \ldots, m_k)$, $\mathbf{q} = (q_1, \ldots, q_{k+1})$. Let us define $m_{k+1} = 1$ and $q_{k+2} = 1$. The quantity $\mathcal{P}_k(\mathbf{m}, \mathbf{q})$ is then given by (14.89), i.e.

$$\mathcal{P}_k(\mathbf{m}, \mathbf{q}) = \log 2 + X_0 - \frac{1}{2} \sum_{1 \le p \le k+1} m_p(\theta(q_{p+1}) - \theta(q_p)) \; .$$

Let us assume in a first stage that $u = q_\tau$ for some $1 \le \tau \le k+1$. We will use the bound (14.147) for $\kappa = k+1$. For $p \le \kappa = k+1$, we consider jointly Gaussian r.v.s (y_p^1, y_p^2) such that

$$\mathsf{E}((y_p^1)^2) = \mathsf{E}((y_p^2)^2) = \xi'(q_{p+1}) - \xi'(q_p) \tag{14.150}$$

$$p < \tau \quad \Rightarrow \quad y_p^1 = y_p^2 \; ; \qquad p \ge \tau \quad \Rightarrow \quad \mathsf{E} y_p^1 y_p^2 = 0 \; . \tag{14.151}$$

Then the conditions (14.130) are satisfied for

$$\rho_p^{1,1} = \rho_p^{2,2} = q_p \; ; \qquad \rho_p^{1,2} = \rho_p^{2,1} = q_{\min(p,\tau)}$$

and (14.132) holds since $\rho_{k+1}^{1,1} = \rho_{k+1}^{2,2} = q_\tau = u$. Next, we choose

$$n_p = \frac{m_p}{2} \;\; \text{if } p < \tau \; ; \qquad n_p = m_p \;\; \text{if } \tau \le p \le k+1 = \kappa \; .$$

We claim that for this choice of parameters, when $\lambda = 0$, the right-hand side of (14.147) is $2\mathcal{P}_k(\mathbf{m}, \mathbf{q})$.

We first estimate the double sum. If $p < \tau$, for any values of ℓ, ℓ' we have

$$n_p(\theta(\rho_{p+1}^{\ell,\ell'}) - \theta(\rho_p^{\ell,\ell'})) = \frac{m_p}{2}(\theta(q_{p+1}) - \theta(q_p)) \; .$$

Next, for $p \ge \tau$, if $\ell = \ell'$ we have

$$n_p(\theta(\rho_{p+1}^{\ell,\ell}) - \theta(\rho_p^{\ell,\ell})) = m_p(\theta(q_{p+1}) - \theta(q_p)) \; ,$$

while if $\ell \ne \ell'$ we have

$$n_p(\theta(\rho_{p+1}^{\ell,\ell'}) - \theta(\rho_p^{\ell,\ell'})) = m_p(\theta(q_\tau) - \theta(q_\tau)) = 0 \; .$$

Therefore

$$\sum_{\ell,\ell'=1,2} \sum_{1 \le p \le k+1} n_p(\theta(\rho_{p+1}^{\ell,\ell'}) - \theta(\rho_p^{\ell,\ell'})) = 2 \sum_{1 \le p \le k+1} m_p(\theta(q_{p+1}) - \theta(q_p)) \; . \tag{14.152}$$

Next, we check that $Y_0 = 2X_0$. The argument should come as no surprise: it has already occurred in the proof of (14.50). We define

$$Y_{k+1}^\ell = \log \operatorname{ch}\Bigl(h + \sum_{0 \le p \le k} y_p^\ell\Bigr)$$

and $Y_{k+1} = Y^1_{k+1} + Y^2_{k+1}$. We define Y_p through (14.145) and Y^ℓ_p through the relation

$$Y^\ell_p = \frac{1}{m_p} \log \mathsf{E}_p \exp m_p Y^\ell_{p+1}$$

and we prove by decreasing induction over p that $Y_p = Y^1_p + Y^2_p$. If we remember the definition of X_0 it is obvious that

$$Y_0 = \mathsf{E} Y^1_1 + \mathsf{E} Y^2_1 = 2X_0 \ ,$$

and we have completed the proof that for the previous choices, the right-hand side of (14.147) is $2\mathcal{P}_k(\mathbf{m}, \mathbf{q})$.

Let us now explain why, when $u \geq 0$, it is not a restriction to assume that $u = q_\tau$. The general fact is that if we think of $\mathbf{m}$ and $\mathbf{q}$ as defining a non-decreasing step function on $[0,1]$, the value of $\mathcal{P}_k(\mathbf{m}, \mathbf{q})$ depends only on the step function; one can merge two consecutive intervals $[q_p, q_{p+1}[$ and $[q_{p+1}, q_{p+2}[$ when $m_p = m_{p+1}$ without changing the value of $\mathcal{P}_k(\mathbf{m}, \mathbf{q})$. This is straightforward to check (see Lemma 14.7.1 below). When $q_\tau \leq u < q_{\tau+1}$ one simply splits the interval $[q_\tau, q_{\tau+1}[$ in the intervals $[q_\tau, u[$ and $[u, q_{\tau+1}[$ (keeping the same m on both) to make u appear as an endpoint.

The previous argument assumes that $u \geq 0$. We do not know how to extend it in full generality to the case $u < 0$. This is a bad omen, and the first sign that we do not really understand yet what is happening. (The full extent of how much we still do not understand will become clear later in Section 15.7.)

Research Problem 14.6.2. (Level 3?) When $u < 0$, given two sequences $\mathbf{m}$ and $\mathbf{q}$ can one find the parameters in the bound (14.147) such that this bound is $\leq 2\mathcal{P}_k(\mathbf{m}, \mathbf{q})$?

We know how to solve this problem when we assume that

$$\forall x \ , \quad \xi(x) = \xi(-x) \ , \tag{14.153}$$

a condition which occurs e.g. in (14.57) when p is even. This condition implies

$$\forall x \ , \quad \xi'(x) = -\xi'(x) \ ; \qquad \theta(x) = \theta(-x) \ .$$

We proceed as follows. By inserting $|u|$ if necessary in the list $\rho_1, \ldots, \rho_{\tau+1}$ as we just explained in the case $u > 0$ we assume without loss of generality that $u = -q_\tau$ for a certain τ. We then replace in (14.151) the condition $y^1_p = y^2_p$ for $p < \tau$ by the condition $y^1_p = -y^2_p$. Then condition (14.130) holds for

$$\rho^{1,1}_p = \rho^{2,2}_p = q_\tau \ ; \ \rho^{1,2}_p = \rho^{2,1}_p = -q_{\min(p,\tau)} \ , \tag{14.154}$$

and since $\theta(x) = \theta(-x)$, (14.152) remains true. It them turns out that with some extra work one can prove that $Y_0 \leq 2X_0$; this will be done later in Proposition 14.8.6. Consequently for the previous choice of parameters the bound (14.147) is $\leq 2\mathcal{P}_k(\mathbf{m},\mathbf{q})$. The scheme (14.154) is effective at proving this inequality, but it is not really canonical. One might get the feeling "that we are not doing the right things".

These considerations suggest a specialization of the bound (14.147). This scheme is not really canonical either, but it suffices to prove the Parisi formula. We assume (14.153) and we consider numbers ρ_p, $0 \leq p \leq \kappa+1$ such that for a certain integer $1 \leq \tau \leq \kappa$ we have

$$0 \leq \rho_0 \leq \rho_1 \leq \ldots \leq \rho_\tau = |u| \leq \rho_{\tau+1} \leq \ldots \leq \rho_{\kappa+1} = 1\ , \tag{14.155}$$

and we consider $\eta \in \{-1,1\}$ with $u = \eta|u|$.

For $0 \leq p \leq \kappa$ we then consider pairs (y_p^1, y_p^2) of jointly Gaussian r.v.s as follows

$$\mathsf{E}(y_p^1)^2 = \mathsf{E}(y_p^2)^2 = \xi'(\rho_{p+1}) - \xi'(\rho_p) \tag{14.156}$$

$$y_p^1 = \eta y_p^2 \ \text{ if } p < \tau\ ; \qquad y_p^1 \text{ and } y_p^2 \text{ independent if } p \geq \tau\ . \tag{14.157}$$

Thus (14.130) holds for

$$\rho_p^{1,1} = \rho_p^{2,2} = \rho_p\ ; \qquad \rho_p^{1,2} = \rho_p^{2,1} = \eta\rho_{\min(p,\tau)}\ , \tag{14.158}$$

and (14.132) holds because $u = \eta|u| = \eta\rho_\tau$. Also, as in the proof of (14.152) we have

$$\sum_{\ell,\ell'=1,2}\ \sum_{1\leq p\leq\kappa} n_p(\theta(\rho_{p+1}^{\ell,\ell'}) - \theta(\rho_p^{\ell,\ell'})) = 4\sum_{p<\tau} n_p(\theta(\rho_{p+1}) - \theta(\rho_p)) + 2\sum_{\tau\leq p\leq\kappa} n_p(\theta(\rho_{p+1}) - \theta(\rho_p))\ .$$

For further reference, we state the specialization of (14.147) that we will use to bound the quantity $\Psi_r(t,u)$ of (14.100).

Proposition 14.6.3. *Let us assume (14.153) and let us keep the notation (14.155), (14.156) and (14.157). Consider pairs of r.v.s* (Z_p^1, Z_p^2) *for* $0 \leq p \leq \kappa$ *and independent copies* $(Z_{i,p}^1, Z_{i,p}^2)_{i\geq 1}$ *of these. Assume they are independent of all the other forms of randomness. Let us consider*

$$n_0 = 0 \leq n_1 \leq n_2 \leq \ldots \leq n_{\kappa-1} \leq n_\kappa = 1\ . \tag{14.159}$$

Let us define $G_{\kappa+1}$ *by*

$$\exp G_{\kappa+1} = \sum_{R_{1,2}=u} \exp\Big(-\sqrt{t}H_N(\boldsymbol{\sigma}^1) - \sqrt{t}H_N(\boldsymbol{\sigma}^2) + \sum_{i\leq N}\sum_{\ell=1,2}\sigma_i^\ell\Big(h_i + \sum_{1\leq p\leq\kappa} Z_{i,p}^\ell\Big)\Big) \tag{14.160}$$

and, for $p \geq 1$, *define recursively*

$$G_p = \frac{1}{n_p} \log \mathsf{E}_p \exp n_p G_{p+1} , \tag{14.161}$$

where E_p *denotes expectation in the r.v.s* $Z^\ell_{i,n}$ *for* $n \geq p$. *Then for any* λ *we have, recalling the notation* ρ_q *of (14.155),*

$$\begin{aligned}\frac{1}{N}\mathsf{E}G_1 \;\leq\; & 2\log 2 + Y_0 - \lambda u - 2t \sum_{p<\tau} n_p(\theta(\rho_{p+1}) - \theta(\rho_p)) \\ & -t \sum_{\tau \leq p \leq \kappa} n_p(\theta(\rho_{p+1}) - \theta(\rho_p)) , \end{aligned} \tag{14.162}$$

where Y_0 *is defined as follows. Recalling the r.v.s* y^ℓ_p *as in (14.156) and (14.157), and letting* $g^\ell_p = \sqrt{t} y^\ell_p + Z^\ell_p$, *define*

$$\begin{aligned} Y_{\kappa+1} = \log\Bigg(& \mathrm{ch}\Big(h + \sum_{0\leq p\leq\kappa} g^1_p\Big)\mathrm{ch}\Big(h + \sum_{0\leq p\leq\kappa} g^2_p\Big)\mathrm{ch}\lambda \\ & + \mathrm{sh}\Big(h + \sum_{0\leq p\leq\kappa} g^1_p\Big)\mathrm{sh}\Big(h + \sum_{0\leq p\leq\kappa} g^2_p\Big)\mathrm{sh}\lambda\Bigg) \end{aligned} \tag{14.163}$$

and, recursively, for $p \geq 1$ *define*

$$Y_p = \frac{1}{n_p} \log \mathsf{E}_p \exp n_p Y_{p+1} , \tag{14.164}$$

where E_p *denotes expectation in the r.v.s* g^ℓ_n *for* $n \geq p$. *Finally define* $Y_0 = \mathsf{E}Y_1$.

It should be useful to keep in mind that the bound (14.162) depends on the sequence (ρ_p) of (14.155), on λ (and on the sequence (14.159)).

Proof. This is the bound (14.147) when one replaces H_N by $\sqrt{t}H_N$ and one uses the special choice (14.158). The left-hand side of (14.147) has been found through Theorem 14.2.1 to be equal to $N^{-1}\mathsf{E}G_1$. □

The reader certainly wonders why we have transformed the left-hand side of (14.147) using Theorem 14.2.1 to state Proposition 14.6.3. The reason is that with this formulation it is easier to understand what happens when we insert new elements in the list (14.159), as will be an essential feature of the proof.

Let us continue our outline of the proof of Theorem 14.5.7. We shall consider situations where the quantity $N^{-1}\mathsf{E}G_1$ of (14.162) satisfies

$$\Psi_r(t,u) = \frac{1}{N}\mathsf{E}G_1 - (1-t)(\xi'(1) - \xi'(q_{k+1})) , \tag{14.165}$$

so that the bound (14.162) of Proposition 14.6.3 will provide a bound for $\Psi_r(t,u)$, which will be our main tool. The sneaky point is that a certain sequence (14.159) will witness naturally (14.165), but that the sequence (14.159) that will actually be used in the bound (14.162) will be different (some new elements will have been inserted). This largely motivates the formulation of Proposition 14.6.3.

Among all the parameters in the bound (14.162) λ is especially important. The quantity $Y_0 = Y_0(\lambda)$ depends on λ, and we end this section with the study of this dependence. It clarifies matters to use a setting slightly more general than the one of Proposition 14.6.3.

Let us consider an integer κ and Gaussian r.v.s g_p, g_p^1, g_p^2 for $0 \le p \le \kappa$. The r.v.s g_p are independent, and the pairs (g_p^1, g_p^2) are independent of each other. We assume that

$$\mathsf{E}g_p^2 = \mathsf{E}(g_p^1)^2 = \mathsf{E}(g_p^2)^2 \;, \tag{14.166}$$

and that for a certain integer $\tau \ge 1$ we have

$$p \ge \tau \quad \Rightarrow \quad g_p^1 \text{ and } g_p^2 \text{ are independent.} \tag{14.167}$$

The reader observes that we do not assume that $g_p^1 = g_p^2$ for $p < \tau$, and this is why this setting is more general than that of Proposition 14.6.3 (it will sometimes happen that $g_p^1 = -g_p^2$). Consider numbers $n_0 = 0 \le n_1 \le \ldots \le n_\kappa = 1$. For $1 \le p \le \kappa + 1$, let us define

$$\zeta_p^\ell = h + \sum_{0 \le n < p} g_n^\ell \;.$$

Starting with

$$Y_{\kappa+1}(\lambda) = \log(\mathrm{ch}(\zeta_{\kappa+1}^1)\mathrm{ch}(\zeta_{\kappa+1}^2)\mathrm{ch}\lambda + \mathrm{sh}(\zeta_{\kappa+1}^1)\mathrm{sh}(\zeta_{\kappa+1}^2)\mathrm{sh}\lambda) \;, \tag{14.168}$$

we define successively for $p \ge 1$

$$Y_p(\lambda) = \frac{1}{n_p} \log \mathsf{E}_p \exp n_p Y_{p+1}(\lambda) \;, \tag{14.169}$$

where E_p denotes expectation in the r.v.s g_n, g_n^ℓ for $n \ge p$. When $n_p = 0$ this means that $Y_p(\lambda) = \mathsf{E}_p Y_{p+1}(\lambda)$. We define $Y_0(\lambda) = \mathsf{E}Y_1(\lambda)$.

Starting with

$$D_{\kappa+1}(x) = \log \mathrm{ch} x \;, \tag{14.170}$$

for $p \ge 0$ we define recursively the function D_p as follows:

$$\text{For } p \ge \tau \;,\; D_p(x) = \frac{1}{n_p} \log \mathsf{E} \exp n_p D_{p+1}(x + g_p) \tag{14.171}$$

$$\text{For } p < \tau \;,\; D_p(x) = \frac{1}{2n_p} \log \mathsf{E} \exp 2n_p D_{p+1}(x + g_p) \;. \tag{14.172}$$

We define $\zeta_p = h + \sum_{0 \le n < p} g_n$ and

$$W_p = \exp n_p(D_{p+1}(\zeta_{p+1}) - D_p(\zeta_p)) \text{ if } p \geq \tau \tag{14.173}$$
$$W_p = \exp 2n_p(D_{p+1}(\zeta_{p+1}) - D_p(\zeta_p)) \text{ if } p < \tau \;. \tag{14.174}$$

Proposition 14.6.4. *a) Let us assume that*

$$p < \tau \quad \Rightarrow \quad g_p = g_p^1 = g_p^2 \;. \tag{14.175}$$

Then we have

$$Y_0(0) = 2\mathsf{E}D_0(h) = 2\mathsf{E}D_1(h + g_0) \tag{14.176}$$
$$Y_0'(0) = \mathsf{E}(W_1 \cdots W_{\tau-1} D_\tau'(\zeta_\tau)^2) \;. \tag{14.177}$$

When $\tau = 1$ this means that $Y_0'(0) = \mathsf{E}(D_1'(\zeta_1)^2)$.
b) Let us assume that

$$p < \tau \quad \Rightarrow \quad n_p = 0 \;. \tag{14.178}$$

Then we have

$$Y_0(0) = \mathsf{E}D_\tau(\zeta_\tau^1) + \mathsf{E}D_\tau(\zeta_\tau^2) \tag{14.179}$$
$$Y_0'(0) = \mathsf{E}(D_\tau'(\zeta_\tau^1) D_\tau'(\zeta_\tau^2)) \;. \tag{14.180}$$

It should already be clear that the case a) will be extremely useful. The motivation for the less important case b) will become apparent in due time.

Proof. First for $p \geq \tau$ we rewrite (14.171) as

$$\begin{aligned} D_p(x) &= \frac{1}{n_p} \log \mathsf{E}_p \exp n_p D_{p+1}(x + g_p) \\ &= \frac{1}{n_p} \log \mathsf{E}_p \exp n_p D_{p+1}(x + g_p^\ell) \;, \end{aligned} \tag{14.181}$$

and using this for $x = \zeta_p^\ell$ (that does not depend on g_p or g_p^ℓ) we obtain

$$D_p(\zeta_p^\ell) = \frac{1}{n_p} \log \mathsf{E}_p \exp n_p D_{p+1}(\zeta_{p+1}^\ell) \;. \tag{14.182}$$

In a similar manner for $p < \tau$ we get from (14.172) that

$$D_p(\zeta_p^\ell) = \frac{1}{2n_p} \log \mathsf{E}_p \exp 2n_p D_{p+1}(\zeta_{p+1}^\ell) \;. \tag{14.183}$$

We use (14.182) and the fact that g_p^1 and g_p^2 are independent for $p \geq \tau$ to prove recursively that then we have

$$Y_p(0) = D_p(\zeta_p^1) + D_p(\zeta_p^2) \;. \tag{14.184}$$

When $n_p = 0$ for $p < \tau$, then for $p < \tau$ we have $Y_p(\lambda) = \mathsf{E}_p Y_{p+1}(\lambda)$, so by iteration we have $Y_0(\lambda) = \mathsf{E}Y_\tau(\lambda)$ and taking expectation in (14.184) for

$p = \tau$ yields (14.179). On the other hand, if we assume (14.175) we have $\zeta_\tau = \zeta_\tau^1 = \zeta_\tau^2$, so that for $p = \tau$ (14.184) yields

$$Y_\tau(0) = 2D_\tau(\zeta_\tau)$$

and we prove recursively in a straightforward fashion using (14.183) that $Y_p(0) = 2D_p(\zeta_p)$ for $1 \le p \le \tau$. (This is again exactly the argument by which we proved (14.50).) Taking expectation when $p = 1$ yields (14.176).

Next, differentiating (14.168) in λ at $\lambda = 0$ we get

$$Y'_{\kappa+1}(0) = \mathrm{th}\zeta^1_{\kappa+1}\mathrm{th}\zeta^2_{\kappa+1} = D'_{\kappa+1}(\zeta^1_{\kappa+1})D'_{\kappa+1}(\zeta^2_{\kappa+1}) ,$$

and differentiating (14.169) in λ at $\beta = 0$ we get

$$Y'_p(0) = \mathsf{E}_p(Y'_{p+1}(0)\exp n_p(Y_{p+1}(0) - Y_p(0)) . \tag{14.185}$$

For $p \ge \tau$ we can combine (14.185) with (14.184) to obtain

$$Y'_p(0) = \mathsf{E}_p(Y'_{p+1}(0)W_p^1 W_p^2) , \tag{14.186}$$

where $W_p^\ell = \exp n_p(D_{p+1}(\zeta^\ell_{p+1}) - D_p(\zeta^\ell_p))$. Differentiation of (14.181) in x shows that

$$D'_p(x) = \mathsf{E}_p(D'_{p+1}(x + g_p^\ell)\exp n_p(D_{p+1}(z + g_p) - D_p(x)) ,$$

and using this for $x = \zeta_p^\ell$ yields

$$D'_p(\zeta_p^\ell) = \mathsf{E}_p(D'_{p+1}(\zeta^\ell_{p+1})W_p^\ell) ,$$

so that using independence we see recursively from (14.186) that for $p \ge \tau$ we have

$$Y'_p(0) = D'_p(\zeta_p^1)D'_p(\zeta_p^2) . \tag{14.187}$$

When $n_p = 0$ for $p < \tau$, writing (14.187) for $p = \tau$ and taking expectation yields (14.180). On the other hand if we assume (14.175) then (14.187) yields

$$Y'_\tau(0) = D'_\tau(\zeta_\tau)^2 .$$

We then use (14.185) recursively to obtain (14.177) using (14.171) and (14.184). □

Lemma 14.6.5. *We have* $0 \le Y_0''(\lambda) \le 1$.

Proof. We use Theorem 14.2.1 to represent $Y_0(\lambda)$ as

$$Y_0(\lambda) = \mathsf{E}\log\sum_\alpha v_\alpha(\mathrm{ch}(\zeta_\alpha^1)\mathrm{ch}(\zeta_\alpha^2)\mathrm{ch}\lambda + \mathrm{sh}(\zeta_\alpha^1)\mathrm{sh}(\zeta_\alpha^2)\mathrm{sh}\lambda)$$

where v_α are random weights and ζ_α are r.v.s, so that by straightforward differentiation we have

$$Y_0''(\lambda) = 1 - \mathsf{E}\left(\frac{\sum_\alpha v_\alpha(\mathrm{ch}(\zeta_\alpha^1)\mathrm{ch}(\zeta_\alpha^2)\mathrm{sh}\lambda + \mathrm{sh}(\zeta_\alpha^1)\mathrm{sh}(\zeta_\alpha^2)\mathrm{ch}\lambda)}{\sum_\alpha v_\alpha(\mathrm{ch}(\zeta_\alpha^1)\mathrm{ch}(\zeta_\alpha^2)\mathrm{ch}\lambda + \mathrm{sh}(\zeta_\alpha^1)\mathrm{sh}(\zeta_\alpha^2)\mathrm{sh}\lambda)}\right)^2 .$$

Now, if the numbers T, T_1, T_2 satisfy $|T|, |T_1|, |T_2| \le 1$, we have $|T + T_1T_2| \le 1 + TT_1T_2$ (e.g. $-T - T_1T_2 \le 1 + TT_1T_2$ because $(1+T)(1+T_1T_2) \ge 0$) and using this for $T = \mathrm{th}\lambda$, $T_1 = \mathrm{th}x_1$ and $T_2 = \mathrm{th}x_2$ it follows that

$$|\mathrm{ch}x_1\mathrm{ch}x_2\mathrm{sh}\lambda + \mathrm{sh}x_1\mathrm{sh}x_2\mathrm{ch}\lambda| \le \mathrm{ch}x_1\mathrm{ch}x_2\mathrm{ch}\lambda + \mathrm{sh}x_1\mathrm{sh}x_2\mathrm{sh}\lambda . \qquad \square$$

14.7 Operators

We must expect that (14.105) and (14.106) will be used in the course of the proof of Theorem 14.5.7, and our first task is to understand these. This motivates the results of the present section, that do not otherwise use the ideas of Sections 14.5 or Section 14.6.

In order to use condition (14.106) we must gain understanding on how $\mathcal{P}_k(\mathbf{m}, \mathbf{q})$ depends on $\mathbf{m}$ and $\mathbf{q}$. The problem is to understand the dependence of the difficult term, i.e. X_0. Let us repeat once more how this quantity is constructed. We consider Gaussian r.v.s z_p with $\mathsf{E}z_p^2 = \xi'(q_{p+1}) - \xi'(q_p)$ for $0 \le p \le k+1$. Starting with $X_{k+2} = \log\mathrm{ch}\big(h + \sum_{0\le p\le k+1} z_p\big)$, we define recursively

$$X_p = \frac{1}{m_p}\log\mathsf{E}_p\exp m_pX_{p+1} \tag{14.188}$$

for $p \ge 1$ and $X_0 = \mathsf{E}X_1$, where E_p is expectation in the r.v.s z_n for $n \ge p$ (or, equivalently, in z_p alone).

It is convenient to look at the previous construction in terms of functions. Starting with the function

$$A_{k+2}(x) = \log\mathrm{ch}x , \tag{14.189}$$

for $p \ge 1$ we define recursively the functions

$$A_p(x) = \frac{1}{m_p}\log\mathsf{E}\exp m_pA_{p+1}(x + z_p) , \tag{14.190}$$

and we define $A_0(x) = \mathsf{E}A_1(x + z_0)$. It should be obvious recursively that

$$X_p = A_p\Big(h + \sum_{0\le n<p} z_n\Big) \tag{14.191}$$

and in particular $X_1 = A_1(h + z_0)$,

$$X_0 = \mathsf{E}A_1(h + z_0) = \mathsf{E}A_0(h) \ . \tag{14.192}$$

Relation (14.190) brings forward the following operation. Given $0 \le m \le 1$ and $v > 0$, given a function A we consider the function

$$T_{m,v}(A)(x) = \frac{1}{m} \log \mathsf{E} \exp mA(x + g\sqrt{v}) \ , \tag{14.193}$$

where g is standard Gaussian. When $m = 0$ this means that

$$T_{0,v}(A)(x) = \mathsf{E}A(x + g\sqrt{v}) \ .$$

With this definition we observe that

$$A_p(x) = T_{m_p,\xi'(q_{p+1})-\xi'(q_p)}(A_{p+1})(x) \ , \tag{14.194}$$

Let us start with a simple property (that reflects the fact that we can merge two consecutive intervals $[q_p, q_{p+1}[$ with the same value of m).

Lemma 14.7.1. *We have*

$$T_{m,a} \circ T_{m,b} = T_{m,a+b} \ . \tag{14.195}$$

Proof. If g' is a standard Gaussian r.v. we have

$$T_{m,a}(B)(x) = \frac{1}{m} \log \mathsf{E} \exp mB(x + g'\sqrt{a})$$

so if $B(x) = T_{m,b}(A)(x)$, then $\exp mB(x) = \mathsf{E} \exp mA(x + g\sqrt{b})$ and therefore

$$\begin{aligned} T_{m,a} \circ T_{m,b}(A)(x) &= \frac{1}{m} \log \mathsf{E} \exp mB(x + g'\sqrt{a} + g\sqrt{b}) \\ &= T_{m,a+b}(A)(x) \ , \end{aligned}$$

since $g'\sqrt{a} + g\sqrt{b}$ has the same distribution as $g\sqrt{a+b}$. □

We will have to understand what happens when we compose many operators $T_{m,v}$, but first of course we must collect the properties of such an operator. To lighten notation we fix an arbitrary function A and we write

$$B(x, v, m) = T_{m,v}(A)(x) = \frac{1}{m} \log \mathsf{E} \exp mA(x + g\sqrt{v}) \tag{14.196}$$

and

$$B' = \frac{\partial B}{\partial x} \ ; \quad B'' = \frac{\partial^2 B}{\partial x^2} \ .$$

To further lighten notation we do not always write the arguments x, v, m in the lemma below.

Lemma 14.7.2. *We have*

$$\exp B(x,v,m) \le \mathsf{E}\exp A(x+g\sqrt{v}) \tag{14.197}$$

$$\text{If } A'' > 0 \text{ then } B'' > 0 \tag{14.198}$$

$$\frac{\partial B}{\partial v} = \frac{1}{2}B'' + \frac{m}{2}B'^2 \,. \tag{14.199}$$

Form (14.195) $(T_{m,a})_{a\ge 0}$ is a semi-group, and (14.199) is simply the corresponding heat equation (a fact we shall not use).

Proof. We leave the easy case $m=0$ to the reader and we assume $m>0$. Hölder's inequality implies

$$\mathsf{E}\exp mA(x+g\sqrt{v}) \le (\mathsf{E}\exp A(x+g\sqrt{v}))^m \,,$$

and this proves (14.197). Differentiation of (14.196) in x yields

$$\begin{aligned} B' &= \frac{\mathsf{E}A'(x+g\sqrt{v})\exp mA(x+g\sqrt{v})}{\mathsf{E}\exp mA(x+g\sqrt{v})} \\ &= \mathsf{E}A'(x+g\sqrt{v})\exp m(A(x+g\sqrt{v})-B(x,v,m)) \,. \end{aligned} \tag{14.200}$$

To lighten notation, we write $Y = x+g\sqrt{v}$,

$$Q = \exp m(A(Y)-B(x,v,m)) \,, \tag{14.201}$$

so that $\mathsf{E}Q = 1$ and (14.200) means that

$$B' = \mathsf{E}(A'(Y)Q) \,. \tag{14.202}$$

We differentiate this formula again in x using that

$$\frac{\partial Q}{\partial x} = m(A'(Y)-B')Q$$

to get

$$\begin{aligned} B'' &= \mathsf{E}(A''(Y)Q) + m\mathsf{E}(A'(Y)^2Q) - mB'\mathsf{E}(A'(Y)Q) \\ &= \mathsf{E}(A''(Y)Q) + m\mathsf{E}(A'(Y)^2Q) - mB'^2 \,, \end{aligned} \tag{14.203}$$

using (14.202) in the second line. Since $\mathsf{E}Q = 1$, the Cauchy-Schwarz inequality implies

$$B'^2 = (\mathsf{E}A'(Y)Q)^2 \le \mathsf{E}(A'(Y)^2Q)$$

so that

$$B'' \ge \mathsf{E}(A''(Y)Q) \tag{14.204}$$

and this proves (14.198). Finally, proceeding as in (14.202) we have

$$\frac{\partial B}{\partial v} = \frac{1}{2\sqrt{v}} \mathsf{E}(gA'(Y)Q)$$

and integration by parts yields

$$\frac{\partial B}{\partial v} = \frac{1}{2} \mathsf{E}(A''(Y)Q) + \frac{m}{2} \mathsf{E}(A'(Y)^2 Q) \ .$$

Combining with (14.203) proves (14.199). □

Having gained some understanding for one operator $T_{m,v}$, we now consider $m' \le m$, a fixed number $a > 0$ and we study the operator $T_{m',a-v} \circ T_{m,v}$. The reason for this formulation is that we try to understand what happens when we "distribute a quantity a between the exponents m and m'", giving a share $a - v$ to m' and v to m. This occurs naturally in the way X_0 depends on q_r for a given r, since this dependence is through the product

$$T_{m_{r-1},\xi'(q_r)-\xi'(q_{r-1})} \circ T_{m_r,\xi'(q_{r+1})-\xi'(q_r)}$$

and this is of the type $T_{m_{r-1},a-v} \circ T_{m_r,v}$ for $a = \xi'(q_{r+1}) - \xi'(q_{r-1})$, a quantity that does not depend on q_r. (Thus, the dependence on q_r is only through $v = \xi'(q_{r+1}) - \xi'(q_r)$.)

Considering a new standard Gaussian r.v. g', we note the formula (recalling (14.196))

$$C(x,v,m) := T_{m',a-v} \circ T_{m,v}(A)(x) = \frac{1}{m'} \log \mathsf{E} \exp m' B(x + g'\sqrt{a-v}, v, m) \ . \tag{14.205}$$

This notation does not indicate the value of m', that we think as fixed once and for all. We also think of a as fixed once and for all.

We write

$$Z = x + g'\sqrt{a-v}$$

and

$$R = \exp m'(B(Z,v,m) - C(x,v,m)) \ . \tag{14.206}$$

Lemma 14.7.3. *We have*

$$\frac{\partial C}{\partial v}(x,v,m) = \frac{1}{2}(m - m')\mathsf{E}(B'(Z,v,m)^2 R) \ . \tag{14.207}$$

Proof. We differentiate (14.205), keeping in mind that there are two sources of dependence on v, to get

$$\frac{\partial C}{\partial v} = \mathrm{I} + \mathrm{II}$$

where

$$\mathrm{I} = \mathsf{E}\left(\frac{\partial B}{\partial v}(Z, v, m)R\right)$$

$$\begin{aligned}\mathrm{II} &= -\frac{1}{2\sqrt{a-v}}\mathsf{E}(g' B'(Z, v, m)R) \\ &= -\frac{1}{2}\mathsf{E}\big((B''(Z, v, m) + m' B'(Z, v, m)^2)R\big) \;,\end{aligned}$$

using integration by parts and keeping in mind the dependence of R on g'. We conclude with (14.199). □

We write

$$\Delta(x, v) = \frac{\partial}{\partial m}C(x, v, m)\Big|_{m=m'} \;. \tag{14.208}$$

Since we think of m' as fixed, we write $B(x, v)$ for $B(x, v, m')$ (etc.), and since we consider the case $m = m'$ we still denote by R the expression (14.206) when $m = m'$.

Lemma 14.7.4. *We have*

$$\frac{\partial \Delta}{\partial v}(x, v) = \frac{1}{2}\mathsf{E}(B'(Z, v)^2 R) \tag{14.209}$$

$$\frac{\partial}{\partial v}\mathsf{E}(B'(Z, v)^2 R) = -\mathsf{E}(B''(Z, v)^2 R) < 0 \;, \tag{14.210}$$

and therefore

$$\frac{\partial^2 \Delta}{\partial v^2}(x, v) = -\frac{1}{2}\mathsf{E}(B''(Z, v)^2 R) < 0 \;. \tag{14.211}$$

Proof. To prove (14.209) we simply use the fact that

$$\frac{\partial \Delta}{\partial v}(x, v) = \frac{\partial}{\partial m}\left(\frac{\partial C}{\partial v}(x, v, m)\right)\Big|_{m=m'} \;,$$

together with (14.207). To prove (14.210) for simplicity we write $B = B(Z, v)$, $B' = B'(Z, v)$ (etc.) and we denote by C the quantity (14.205), so that $R = \exp m'(B - C)$. The dependence of B and B' on v has two sources, the dependence through v as the second variable and the dependence through $Z = x + g'\sqrt{a-v}$. Thus, writing separately the contributions of these dependences to the derivatives, we get

$$\frac{\partial}{\partial v}\mathsf{E}(B'^2 R) = \frac{\partial}{\partial v}\mathsf{E}(B'^2 \exp m'(B - C)) = \mathrm{III} + \mathrm{IV} \tag{14.212}$$

$$\begin{aligned}\mathrm{III} &= \mathsf{E}\left(\left(2\frac{\partial B'}{\partial v}B' + m' B'^2 \frac{\partial B}{\partial v}\right)R\right) \\ \mathrm{IV} &= -\frac{1}{2\sqrt{a-v}}\mathsf{E}\big(g'(2B'B'' + m' B'^3)\exp m'(B - C)\big) \;.\end{aligned}$$

By (14.199) used for m' rather than m, and since $B = B(Z, v) = B(Z, v, m')$, we have

$$\frac{\partial B}{\partial v} = \frac{1}{2} B'' + \frac{m'}{2} B'^2 ,$$

and differentiating this relation in x we get

$$\frac{\partial B'}{\partial v} = \frac{1}{2} B^{(3)} + m' B' B'' .$$

Consequently, we have

$$\mathrm{III} = \mathsf{E}\left(\left(B^{(3)} B' + 2m' B'^2 B'' + \frac{1}{2} m' B'^2 B'' + \frac{m'^2}{2} B'^4\right) R\right) .$$

Integrating by parts in IV yields

$$\mathrm{IV} = -\frac{1}{2} \mathsf{E}\big((2B''^2 + 2B' B^{(3)} + 3m' B'' B'^2 + (2B' B'' + m' B'^3) m' B') R\big) ,$$

and the terms nicely cancel out in (14.212) to yield (14.210). As for (14.211) it follows by combining the other two relations. □

Of course one might (and probably should) feel that formula (14.210) is a kind of miracle. This is the first of several such miracles, each of which greatly contributes to make the proof of the Parisi formula possible. Probably this simply reflects the fact that one has not yet found the correct way to look at these objects.

We investigate now how X_0 depends on q_r for $1 \le r \le k+1$. Consider the function A_{r+1} defined through (14.190), so that from (14.194) we have

$$A_r(x) = T_{m_r, \xi'(q_{r+1}) - \xi'(q_r)}(A_{r+1})(x) \tag{14.213}$$

$$A_{r-1}(x) = T_{m_{r-1}, \xi'(q_r) - \xi'(q_{r-1})}(A_r)(x) . \tag{14.214}$$

Let us consider the quantity $A_p(x)$ for $p \le r-1$. It depends on q_r. We differentiate (14.190) to obtain for $p \le r-2$ the recursion formula

$$\frac{\partial}{\partial q_r} A_p(x) = \mathsf{E}\left(\frac{\partial}{\partial q_r} A_{p+1}(x + z_p) \exp m_p (A_{p+1}(x + z_p) - A_p(x))\right) . \tag{14.215}$$

For $0 \le p \le r-2$ let us define

$$\zeta_p = h + \sum_{0 \le n < p} z_n , \tag{14.216}$$

so that $\zeta_0 = h$ and $X_p = A_p(\zeta_p)$ by (14.191). By iterative use of (14.215) we then obtain

$$\frac{\partial}{\partial q_r} X_0 = \frac{\partial}{\partial q_r} \mathsf{E} A_1(h + z_0) = \mathsf{E}\left(W_1 \cdots W_{r-2} \frac{\partial}{\partial q_r} A_{r-1}(\zeta_{r-1})\right) , \tag{14.217}$$

where

$$W_p = \exp m_p(A_{p+1}(\zeta_{p+1}) - A_p(\zeta_p)) = \exp m_p(X_{p+1} - X_p) \, . \qquad (14.218)$$

When $r = 1$ or $r = 2$ this means simply that

$$\frac{\partial}{\partial q_r} X_0 = \mathsf{E}\left(\frac{\partial}{\partial q_r} A_{r-1}(\zeta_{r-1})\right) \, .$$

To compute $\partial A_{r-1}(x)/\partial q_r$ we use the formula (14.207) when $A = A_{r+1}$, $a = \xi'(q_{r+1}) - \xi'(q_{r-1})$, $v = \xi'(q_{r+1}) - \xi'(q_r)$, $m' = m_{r-1}$, $m = m_r$. Thus

$$A_r(x) = B(x, \xi'(q_{r+1}) - \xi'(q_r), m_r)$$
$$A_{r-1}(x) = C(x, \xi'(q_r) - \xi'(q_{r-1}), m_{r-1}) \, .$$

Since $g\sqrt{a-v}$ has the same distribution as z_{r-1}, formula (14.207) implies that

$$\frac{\partial}{\partial q_r} A_{r-1}(x) = -\frac{1}{2}\xi''(q_r)(m_r - m_{r-1}) \qquad (14.219)$$
$$\times \, \mathsf{E}\big(A_r'(x + z_{r-1})^2 \exp m_{r-1}(A_r(x + z_{r-1}) - A_{r-1}(x))\big) \, .$$

Substitution of this formula in (14.217) yields

$$\frac{\partial}{\partial q_r} X_0 = -\frac{1}{2}\xi''(q_r)(m_r - m_{r-1})\mathsf{E}(W_1 \cdots W_{r-1} A_r'(\zeta_r)^2) \, ;$$

the extra term W_{r-1} above compared to (14.217) arises from the term $\exp m_{r-1}(A_r(x + z_{r-1}) - A_{r-1}(x))$ in (14.219). Finally, recalling (14.89) we get, since $\theta'(q) = q\xi''(q)$

$$\frac{\partial}{\partial q_r} \mathcal{P}_k(\mathbf{m}, \mathbf{q}) = \frac{1}{2}(m_r - m_{r-1})\xi''(q_r)(-\mathsf{E}(W_1 \cdots W_{r-1} A_r'(\zeta_r)^2) + q_r) \, . \qquad (14.220)$$

Proposition 14.7.5. *Assume (14.103) and (14.104), and consider the quantities A_r, ζ_r, and W_r defined just above in (14.213), (14.214), (14.216) and (14.218). Then if we cannot decrease $\mathcal{P}_k(\mathbf{m}, \mathbf{q})$ by changing $\mathbf{q}$, we have*

$$0 = q_0 < q_1 < \ldots < q_k < q_{k+1} < 1 \qquad (14.221)$$

and (for $r = 1, \ldots, k+1$)

$$q_r = \mathsf{E}(W_1 \cdots W_{r-1} A_r'(\zeta_r)^2) \, . \qquad (14.222)$$

Proof. When $q_{r-1} < q_r < q_{r+1}$ we must have $\partial \mathcal{P}_k(\mathbf{m}, \mathbf{q})/\partial q_r = 0$ and then (14.220) implies (14.222), since $m_r \neq m_{r-1}$ by (14.103) and $\xi''(q_r) \neq 0$. This condition $q_{r-1} < q_r < q_{r+1}$ is true for $r \neq 1$ and $r \neq k+1$, since (14.104) states that

$$q_0 = 0 \le q_1 < q_2 < \ldots < q_{k+1} \le 1 = q_{k+2} \; .$$

When $r = 1$, we need to show we cannot have $q_1 = 0$. This is because (14.220) shows that $\partial \mathcal{P}_k(\mathbf{m}, \mathbf{q})/\partial q_1 < 0$ for every q_1 small enough. When $r = k+1$, we need to show that we cannot have $q_{k+1} = 1$. First, since

$$A'_{k+2}(x) = \mathrm{th} x$$

and $|\mathrm{th} x| < 1$, computation of A'_r as usual through (14.190) shows that $|A'_r(x)| < 1$ and since $\mathsf{E}(W_1 \cdots W_{r-1}) = 1$ we have $\mathsf{E}(W_1 \cdots W_{r-1} A'_r(\zeta_r)^2) < 1$ and (14.220) shows that $\partial \mathcal{P}_k(\mathbf{m}, \mathbf{q})/\partial q_{k+1}|_{q_{k+1}=1} > 0$. □

How will we use the fact that $\mathcal{P}_k(\mathbf{m}, \mathbf{q})$ is close to $\mathcal{P}(\xi, h)$? Simply by writing that we cannot decrease much $\mathcal{P}_k(\mathbf{m}, \mathbf{q})$ by "adding one level of replica-symmetry breaking". To add a level of replica-symmetry breaking, we consider $q_{r-1} \le u \le q_r$ and $m_{r-1} \le m \le m_r$. Consider the lists

$$0 = m_0 < m_1 < \ldots < m_{r-1} \le m \le m_r < \ldots < m_k < m_{k+1} = 1 \tag{14.223}$$

$$0 = q_0 < q_1 < \ldots < q_{r-1} \le u \le q_r < \ldots < q_{k+1} < q_{k+2} = 1 \; , \tag{14.224}$$

and the corresponding sequences $\mathbf{m}(m) = (m_1, \ldots, m_{r-1}, m, m_r, \ldots, m_k)$ and $\mathbf{q}(u) = (q_1, \ldots, q_{r-1}, u, q_r, \ldots, q_{k+1})$. We define

$$\Phi(m, u) = \mathcal{P}_{k+1}(\mathbf{m}(m), \mathbf{q}(u)) \; . \tag{14.225}$$

In words, we split the interval $[q_{r-1}, q_r[$ into two intervals $[q_{r-1}, u[$ and $[u, q_r[$. To the interval $[q_{r-1}, u[$ we attach the parameter m_{r-1}. To the interval $[u, q_r[$ we attach the parameter m.

We note right away that by definition of $\mathcal{P}(\xi, h)$ we have

$$\Phi(m, u) \ge \mathcal{P}(\xi, h) \; , \tag{14.226}$$

and that

$$\Phi(m, q_r) = \mathcal{P}_k(\mathbf{m}, \mathbf{q}) \; , \tag{14.227}$$

since when $u = q_r$ the interval $[u, q_r[$ is empty.

Lemma 14.7.6. *Assume (14.105) and (14.106). Then for $m_{r-1} \le m \le m_r$ and $q_{r-1} \le u \le q_r$ we have*

$$\Phi(m, u) \ge \Phi(m_{r-1}, u) - \varepsilon \; . \tag{14.228}$$

Moreover for $q_{r-1} \le u \le q_r$ we have

$$\Phi(m_r, u) \ge \Phi(m_{r-1}, u) \; . \tag{14.229}$$

Proof. We have $\Phi(m_{r-1}, u) = \mathcal{P}_k(\mathbf{m}, \mathbf{q})$ because when $m = m_{r-1}$ the parameter m_{r-1} is attached to both the intervals $[q_{r-1}, u[$ and $[u, q_r[$ and we can merge them. (A formal proof of this fact is given in (14.237) below.) Thus, using (14.226) in the first inequality and (14.105) in the second inequality, we get

$$\Phi(m, u) \geq \mathcal{P}(\xi, h) \geq \mathcal{P}_k(\mathbf{m}, \mathbf{q}) - \varepsilon = \Phi(m_{r-1}, u) - \varepsilon \; .$$

Moreover, when $m = m_r$, the parameter m_r is attached to both the intervals $[u, q_r[$ and $[q_r, q_{r+1}[$, and we can merge these to see that $\Phi(m_r, u) = \mathcal{P}_{k+1}(\mathbf{m}(m_r), \mathbf{q}(u)) = \mathcal{P}_k(\mathbf{m}', \mathbf{q}')$ for certain sequences $\mathbf{m}'$ and $\mathbf{q}'$. But (14.106) implies

$$\mathcal{P}_k(\mathbf{m}', \mathbf{q}') \geq \mathcal{P}_k(\mathbf{m}, \mathbf{q}) = \Phi(m_{r-1}, u) \; . \qquad \square$$

A basic idea is that we will get considerable information from the fact that we cannot violate condition (14.228) by slightly increasing m starting from the value m_{r-1}. Therefore the function

$$f(u) := \frac{\partial}{\partial m} \Phi(m, u) \bigg|_{m = m_{r-1}} \tag{14.230}$$

will play a important role. Of course, before we can implement this basic idea, we must learn how to calculate $\Phi(m, u)$.

To compute $\mathcal{P}_{k+1}(\mathbf{m}(m), \mathbf{q}(u))$ the difficult part is to understand what happens to the term X_0 when we split the interval $[q_{r-1}, q_r[$. We are going to construct first functions $(C_p)_{p \leq k+3}$ that, after this interval has been split, play the rôle that the functions $(A_p)_{p \leq k+2}$ (defined before through (14.190)) play before this interval $[q_{r-1}, q_r[$ has been split. For $p \leq r$ these functions also depend on m and a parameter v. For $p \geq r$ we have $C_{p+1} = A_p$, so in particular $C_{r+1} = A_r$. We define $a = \xi'(q_r) - \xi'(q_{r-1})$,

$$C_r(x, v, m) = T_{m,v}(C_{r+1})(x) = T_{m,v}(A_r)(x) \tag{14.231}$$

$$C_{r-1}(x, v, m) = T_{m_{r-1}, a-v}(C_r(\cdot, v, m))(x) \; . \tag{14.232}$$

These two different operations reflect what happens on the interval $[q_{r-1}, q_r[$, that is now split in two different intervals. Let us note that by Lemma 14.7.1 we have

$$C_{r-1}(x, v, m_{r-1}) = A_{r-1}(x) \; . \tag{14.233}$$

For $p \leq r-2$ we define

$$C_p(x, v, m) = \frac{1}{m_p} \log \mathsf{E} \exp m_p C_{p+1}(x + z_p, v, m) \; , \tag{14.234}$$

and we set

$$S(v, m) = \mathsf{E} C_1(h + z_0, v, m) \; . \tag{14.235}$$

If we compare with (14.192) we see that when $v = \xi'(q_r) - \xi'(u)$, this is the value corresponding to X_0 when we have replaced $\mathbf{m}$ and $\mathbf{q}$ by $\mathbf{m}(m)$ and $\mathbf{q}(u)$ respectively. Therefore by definition of $\mathcal{P}_{k+1}$ we get

$$\begin{aligned}\varPhi(m,u) &= \mathcal{P}_{k+1}(\mathbf{m}(m),\mathbf{q}(u)) \\ &= \log 2 + S(\xi'(q_r) - \xi'(u), m) \\ &- \frac{1}{2}\sum_{0\le p\le k+1} m_p(\theta(q_{p+1}) - \theta(q_p)) + \frac{1}{2}(m_{r-1} - m)(\theta(q_r) - \theta(u)) \ .\end{aligned} \tag{14.236}$$

The last term occurs since we have replaced m_{r-1} by m on the interval $[u, q_r[$.

Let us observe that when $m = m_{r-1}$, by (14.233), for any value of v and each $1 \le p \le r-1$ we have $C_p(x, v, m_{r-1}) = A_p(x)$ and therefore

$$S(v, m_{r-1}) = X_0 \ , \tag{14.237}$$

so that

$$\varPhi(m_{r-1}, u) = \mathcal{P}_k(\mathbf{m}, \mathbf{q}) \ .$$

Let us define

$$U(v) = 2\frac{\partial S}{\partial m}(v, m)\bigg|_{m=m_{r-1}} \ . \tag{14.238}$$

This quantity will also play a important role. Recalling the definition (14.230), we note that (14.238) and (14.236) imply

$$f(u) = \frac{1}{2}U(\xi'(q_r) - \xi'(u)) - \frac{1}{2}(\theta(q_r) - \theta(u)) \ . \tag{14.239}$$

We note that since $S(0, m)$ does not depend on m we have $U(0) = 0$. The next step is to compute the derivatives of $U(v)$. For simplicity of notation let us write

$$\varDelta_{r-1}(x, v) = \frac{\partial}{\partial m}C_{r-1}(x, v, m)\bigg|_{m=m_{r-1}} \ .$$

We recall that when $m = m_{r-1}$ we have $C_p = A_p$ for $p \le r-1$. Therefore, recalling the quantities W_p of (14.218), differentiation of the relation (14.235) and recursive differentiation of the relations (14.234) yields as usual (e.g. as in (14.217)) the relation

$$U(v) = 2\mathsf{E}(W_1 \cdots W_{r-2}\varDelta_{r-1}(\zeta_r, v)) \ . \tag{14.240}$$

The function $\varDelta_{r-1}(x, v)$ is the function $\varDelta(x, v)$ given by (14.208) in the case $m' = m_{r-1}$ and $A = A_r$. In that case, keeping in mind that

$$C_r(x, v, m) = B(x, v, m) \ ,$$

formula (14.209) means

$$\frac{\partial \Delta_{r-1}}{\partial v}(x,v) = \frac{1}{2}\mathsf{E}\big(C_r'(Z,v,m_{r-1})^2 \exp m_{r-1}(C_r(Z,v,m_{r-1}) - A_{r-1}(x)\big) \tag{14.241}$$

where

$$Z = x + g'\sqrt{a-v} = x + g'\sqrt{\xi'(q_r) - \xi'(q_{r-1}) - v}\ .$$

Let us set

$$\zeta_r(v) = \zeta_{r-1} + g'\sqrt{a-v}$$

and

$$W_{r-1}(v) = \exp m_{r-1}(C_r(\zeta_r(v), v, m_{r-1}) - A_{r-1}(\zeta_{r-1}))\ .$$

Since the quantity $W_1 \cdots W_{r-2}$ does not depend on v, we deduce from (14.240) that

$$U'(v) = 2\mathsf{E}\left(W_1 \cdots W_{r-2}\frac{\partial \Delta_{r-1}}{\partial v}(\zeta_{r-1}, v)\right)\ , \tag{14.242}$$

and using (14.241) with $x = \zeta_{r-1}$ yields

$$U'(v) = \mathsf{E}\big(W_1 \cdots W_{r-2} W_{r-1}(v) C_r'(\zeta_r(v), v, m_{r-1})^2\big)\ . \tag{14.243}$$

When $r = 1$ this means that

$$U'(v) = \mathsf{E}(C_r'(\zeta_r(v), v, m_{r-1})^2)\ . \tag{14.244}$$

Then since $m_{r-1} = m_0 = 0$ we also have $C_r(x, v, m_{r-1}) = \mathsf{E}A_r(x + \gamma\sqrt{v})$ and therefore

$$C_r'(x, v, m_{r-1}) = \mathsf{E}A_r'(x + g\sqrt{v}) = \mathsf{E}A_1'(x + g\sqrt{v})\ . \tag{14.245}$$

When $v = 0$ we have $C_r(x, 0, m_{r-1}) = T_{m,0}(A_r)(x) = A_r(x)$ so that $C_r'(x, 0, m_{r-1}) = A_r'(x)$, and since and (in distribution) $\zeta_r(0) = h + \sum_{n<r} z_n = \zeta_r$ we deduce from (14.243) that

$$U'(0) = \mathsf{E}(W_1 \cdots W_{r-1} A_r'(\zeta_r)^2)\ . \tag{14.246}$$

We now repeat the previous work for second derivatives. From (14.211), as in (14.241) we get

$$\frac{\partial^2 \Delta_{r-1}}{\partial v^2}(x,v) = -\frac{1}{2}\mathsf{E}\big(C_r''(Z,v,m_{r-1})^2 \exp m_{r-1}(C_r(Z,v,m_{r-1}) - A_{r-1}(x))\big).$$

Differentiating (14.242) and combining with the above for $x = \zeta_{r-1}$ we get as in (14.243)

$$U''(v) = -\mathsf{E}\big(W_1 \cdots W_{r-2} W_{r-1}(v) C_r''(\zeta_r(v), v, m_{r-1})^2\big)\ , \tag{14.247}$$

and in particular, as in (14.246),

$$U''(0) = -\mathsf{E}(W_1 \cdots W_{r-1} A_r''(\zeta_r)^2)\ . \tag{14.248}$$

Proposition 14.7.7. *Under (14.106), for* $v < \xi'(q_r) - \xi'(q_{r-1})$ *we have*

$$U''(v) < 0 \; . \tag{14.249}$$

Proof. Obvious from (14.247). □

The expression (14.246) coincides with the expression (14.222). We have proved the following.

Proposition 14.7.8. *Under (14.106) we have*

$$U'(0) = q_r \; . \tag{14.250}$$

We now have gathered important information about the function f of (14.230).

Proposition 14.7.9. *If* $k, \mathbf{q}, \mathbf{m}$ *satisfy condition* $\mathrm{MIN}(\varepsilon)$*, we have*

$$f(q_r) = f'(q_r) = 0 \quad ; \quad r \geq 2 \Rightarrow f(q_{r-1}) = 0 \; . \tag{14.251}$$

Proof. For $u = q_r$ the quantity $\Phi(m, u)$ does not depend on m so that $f(q_r) = 0$. By (14.239) we have

$$f'(u) = -\frac{\xi''(u)}{2}(U'(\xi'(q_r) - \xi'(u)) - u) \tag{14.252}$$

since $\theta'(u) = u\xi''(u)$, and so $f'(q_r) = 0$ by (14.250). Let us now assume that $r \geq 2$. When $u = q_{r-1}$, to compute $\mathcal{P}_{k+1}(\mathbf{m}(m), \mathbf{q}(u))$ we may ignore the empty interval $[q_{r-1}, u[$, that is we may remove u from the sequence $\mathbf{q}(u)$ and we may remove m_{r-1} from the sequence $\mathbf{m}$, and then $\mathcal{P}_{k+1}(\mathbf{m}(m), \mathbf{q}(u)) = \mathcal{P}_k(\mathbf{m}^*(m), \mathbf{q})$ where the sequence $\mathbf{m}^*(m)$ is obtained from the sequence $\mathbf{m}$ by replacing the value m_{r-1} by m. Therefore we have $f(q_{r-1}) = 0$, for otherwise we could decrease $\mathcal{P}_k(\mathbf{m}, \mathbf{q})$ by a small variation of $m_{r-1} > 0$. □

Proposition 14.7.10. *For a number* M *depending on* ξ *and* h *only, we have*

$$\left|\frac{\partial \Phi}{\partial m}(m, u)\right| \leq M \; , \qquad \left|\frac{\partial^2 \Phi}{\partial m^2}(m, u)\right| \leq M \; . \tag{14.253}$$

The essential point is that the number M does not depend on k. The proof relies on the following result.

Lemma 14.7.11. *There exists a number* L *with the following property. Consider a r.v.* Y *and for* $0 < m \leq 1$ *define*

$$c(m) = \frac{1}{m} \log \mathsf{E} \exp mY \; .$$

Then

$$|c'(m)| \leq L\mathsf{E} \exp L|Y| \tag{14.254}$$

$$|c''(m)| \leq L\mathsf{E} \exp L|Y| \; . \tag{14.255}$$

Proof. This is elementary calculus, and it should be possible to find a more elegant proof than the rather repulsive argument we give. Let us first define

$$W = W(m) = \frac{\exp mY}{\mathsf{E}\exp mY} = \exp m(Y - c(m))$$

so that $\mathsf{E}W = 1$. We claim that

$$\mathsf{E}(W-1)^4 \le Lm^4 \mathsf{E}\exp L|Y| \,. \tag{14.256}$$

To see this we use that $|e^x - 1| \le |x|e^{|x|}$ and hence $(e^x - 1)^4 \le x^4 e^{4|x|}$ for $x = m(Y - c(m))$ to get

$$\mathsf{E}(W-1)^4 \le m^4 \mathsf{E}\big((Y - c(m))^4 \exp 4m|Y - c(m)|\big) \,. \tag{14.257}$$

We now use rough bounds to control the last term, such as

$$(Y - c(m))^4 \le (|Y| + |c(m)|)^4 \le \exp 4(|Y| + c(m))$$

because $x \le e^x$ for $x \ge 0$; and moreover

$$\exp 4m|Y - c(m)| \le \exp 4(|Y| + |c(m)|)$$

so that

$$\begin{aligned}\mathsf{E}((Y - c(m))^4 \exp 4(|Y| + |c(m)|) &\le \mathsf{E}\exp 8(|Y| + |c(m)|)\\ &= \exp 8|c(m)|\mathsf{E}\exp 8|Y| \,.\end{aligned}$$

Now Hölder's inequality yields

$$c(m) \le \frac{1}{m}\log \mathsf{E}\exp m|Y| \le \log \mathsf{E}\exp|Y| \,,$$

and, Jensen's inequality used for the convex function exp implies

$$c(m) \ge \frac{1}{m}\log \mathsf{E}\exp(-m|Y|) \ge \frac{1}{m}\log\exp(-m\mathsf{E}|Y|)) = -\mathsf{E}|Y| \,,$$

so that

$$|c(m)| \le \max(\mathsf{E}|Y|, \log \mathsf{E}\exp|Y|) = \log \mathsf{E}\exp|Y|$$

and

$$\exp 8|c(m)| \le (\mathsf{E}\exp|Y|)^8 \le \mathsf{E}\exp 8|Y| \,.$$

Finally, $(\mathsf{E}\exp 8|Y|)^2 \le \mathsf{E}\exp 16|Y|$, and this proves (14.256). By Hölder's inequality we obtain that for $1 \le p \le 4$

$$\mathsf{E}|W-1|^p \le Lm^p \mathsf{E}\exp L|Y| \,. \tag{14.258}$$

We compute

$$\begin{aligned} c'(m) &= -\frac{1}{m^2}\log \mathsf{E}\exp mY + \frac{1}{m}\frac{\mathsf{E}Y\exp mY}{\mathsf{E}\exp mY} \\ &= \frac{1}{m^2}(-mc(m) + m\mathsf{E}YW) \\ &= \frac{1}{m^2}\mathsf{E}(m(Y-c(m))W) = \frac{1}{m^2}\mathsf{E}W\log W \ . \end{aligned} \tag{14.259}$$

Next, we have

$$W - 1 \le W\log W \le L(W-1)^2 \tag{14.260}$$

and since $\mathsf{E}W = 1$, taking expectation shows that

$$0 \le c'(m) \le \frac{L}{m^2}\mathsf{E}(W-1)^2$$

and (14.258) proves (14.254). The proof of (14.255) is similar but requires more tedious computation. To compute $c''(m)$, we first use the definition of $W = W(m)$ and (14.259) to get

$$\begin{aligned} W'(m) &= (Y - c(m))\exp m(Y-c(m)) - mc'(m)\exp m(Y-c(m)) \\ &= (Y-c(m))W - \frac{1}{m}W\mathsf{E}(W\log W) \\ &= \frac{1}{m}W\log W - \frac{1}{m}W\mathsf{E}(W\log W) \end{aligned}$$

and therefore

$$\frac{\mathrm{d}}{\mathrm{d}m}(W\log W) = \frac{1}{m}((1+\log W)(W\log W - W\mathsf{E}(W\log W))) \ ,$$

so that, since $\mathsf{E}W = 1$,

$$\frac{\mathrm{d}}{\mathrm{d}m}\mathsf{E}W\log W = \frac{1}{m}(\mathsf{E}W\log^2 W - (\mathsf{E}W\log W)^2) \ ,$$

and thus, using (14.259),

$$c''(m) = \frac{1}{m^3}(-2\mathsf{E}W\log W + \mathsf{E}W\log^2 W - (\mathsf{E}W\log W)^2) \ .$$

We consider the function $\psi(x) = -2x\log x + x\log^2 x + 2(x-1)$, so that $\psi'(x) = \log^2 x$ and $\psi(1) = \psi'(1) = \psi''(1) = 0$ and therefore

$$|\psi(x)| \le L|x-1|^3 \ , \tag{14.261}$$

since this is true when x is either $\le 1/2$ or ≥ 2.

By (14.260) and (14.258) we have $(\mathsf{E}W\log W)^2 \le Lm^4\exp L|Y|$ and we conclude the proof of (14.255) by using that (14.261) implies

$$|-2W\log W + W\log^2 W + 2(W-1)| \le L|W-1|^3$$

so that

$$\begin{aligned}|-2\mathsf{E}W\log W+\mathsf{E}W\log^2 W| &= |\mathsf{E}(-2W\log W+W\log^2 W+2(W-1))| \\ &\le \mathsf{E}|-2W\log W+W\log^2 W+2(W-1)| \\ &\le L\mathsf{E}|W-1|^3 ,\end{aligned}$$

and we conclude with (14.258). □

Proof of Proposition 14.7.10. We recall that

$$\Phi(m,u)=\mathcal{P}_{k+1}(\mathbf{m}(m),\mathbf{q}(u)) .$$

Thinking of $\mathbf{m}(m)$ as a sequence $(m'_1,\dots,m'_{k+1})$ it suffices to prove that the first two derivatives of the quantity $\mathcal{P}_k(\mathbf{m},\mathbf{q})$ with respect to m_r (where r is any number $1\le r\le k$) are bounded by a number M depending only on ξ and h (and then to change k into $k+1$). This allows for simpler notation. We consider as usual the Gaussian variables z_p with $\mathsf{E}z_p^2=\xi'(q_{p+1})-\xi'(q_p)$ for $0\le p\le k+1$, the function $A_{k+2}(x)=\log \mathrm{ch} x$, and we define recursively for $p\ge r+1$

$$A_p(x)=\frac{1}{m_p}\log\mathsf{E}\exp m_p A_{p+1}(x+z_p) .$$

We notice that (recursively) $A_p(x)\ge 0$. To insist on the dependence on m_r, we write

$$A_r(x,m_r)=\frac{1}{m_r}\log\mathsf{E}\exp m_r A_{r+1}(x+z_r) ,$$

and for $p<r$ we define

$$A_p(x,m_r)=\frac{1}{m_p}\log\mathsf{E}\exp m_p A_{p+1}(x+z_p,m_r) .$$

so that again $A_p(x,m_r)\ge 0$. When $r=1$ Proposition 14.7.10 is a direct consequence of Lemma 14.7.11. When $r>1$, we consider for $\alpha\in\mathbb{N}^{*(r-1)}$ random weights (v_α) forming a Poisson-Dirichlet cascade associated with the sequence $m_1,\dots,m_{r-1}$. We denote by $z_{p,j_1,\dots,j_p}$ independent copies of the r.v.s z_p; and for $p\le r-1$ we set $z_{p,\alpha}=z_{p,j_1,\dots,j_p}$ when $\alpha=(j_1,\dots,j_{r-1})$. We set $y_\alpha=\sum_{1\le p\le r}z_{p,\alpha}$. It follows from Theorem 14.2.1 that

$$A_1(x,m_r)=\mathsf{E}\log\sum_\alpha v_\alpha\exp A_r(x+y_\alpha,m_r) ,$$

and consequently

$$\mathsf{E}A_1(h+z_0,m_r)=\mathsf{E}\log\sum_\alpha v_\alpha\exp A_r(\zeta_\alpha,m_r) , \tag{14.262}$$

where $\zeta_\alpha=h+z_0+y_\alpha$. To control the derivatives of $\mathcal{P}_k(\mathbf{m},\mathbf{q})$ with respect to m_r, it suffices to control the derivatives of $a(m_r):=\mathsf{E}A_1(h+z_0,m_r)$ with respect to m_r. The first derivative is

$$a'(m_r) = \mathsf{E}\frac{\sum_\alpha v_\alpha \frac{\partial A_r}{\partial m_r}(\zeta_\alpha, m_r)\exp A_r(\zeta_\alpha, m_r)}{\sum_\alpha v_\alpha \exp A_r(\zeta_\alpha, m_r)} .$$

Since $A_r \geq 0$ as we already observed, using independence in the second line and that $\sum_\alpha \mathsf{E} v_a = 1$ in the third line we obtain

$$\begin{aligned}
|a'(m_r)| &\leq \mathsf{E}\sum_\alpha v_\alpha \left|\frac{\partial A_r}{\partial m_r}(\zeta_\alpha, m_r)\right| \exp A_r(\zeta_\alpha, m_r) \\
&\leq \sum_\alpha \mathsf{E} v_\alpha \mathsf{E}\left|\frac{\partial A_r}{\partial m_r}(\zeta_\alpha, m_r)\right| \exp A_r(\zeta_\alpha, m_r) \\
&\leq \max_\alpha \left(\mathsf{E}\left(\frac{\partial A_r}{\partial m_r}(\zeta_\alpha, m_r)\right)^2\right)^{1/2} (\mathsf{E}\exp 2A_r(\zeta_\alpha, m_r))^{1/2} .
\end{aligned}$$

It follows from (14.254) that, denoting E_z expectation in z_r only,

$$\left|\frac{\partial A_r}{\partial m_r}(\zeta_\alpha, m_r)\right| \leq L\mathsf{E}_z \exp LA_{r+1}(\zeta_\alpha + z_r) ,$$

so that

$$\begin{aligned}
\mathsf{E}\left(\frac{\partial A_r}{\partial m_r}(\zeta_\alpha, m_r)\right)^2 &\leq L\mathsf{E}(\mathsf{E}_z \exp LA_{r+1}(\zeta_\alpha + z_r))^2 \\
&\leq L\mathsf{E}\exp 2LA_{r+1}(\zeta_\alpha + z_r) . \qquad (14.263)
\end{aligned}$$

Now, by iteration of (14.197) if g is a standard normal r.v. we have

$$\exp A_{r+1}(x) \leq \mathsf{E}\mathrm{ch}(x + g\sqrt{\xi'(1) - \xi'(q_{r+1})}) ,$$

and therefore

$$\mathsf{E}\exp 2LA_{r+1}(\zeta_\alpha + z_r) \leq \mathsf{E}\mathrm{ch}(h + g\sqrt{\xi'(1)})^{2L} \leq M .$$

We proceed in the same way to bound $\mathsf{E}\exp 2A_r(\zeta_\alpha, m_r)$. The proof for the second derivative is entirely similar, using now (14.255). □

Exercise 14.7.12. Find a proof that does not use Theorem 14.2.1.

In the remainder of this chapter, we denote by M a number depending only on ξ and h, that need not be the same at each occurrence. Of course, M_1, etc. denote a specific number.

Proposition 14.7.13. *The function $f(u)$ of (14.230) satisfies*

$$f(u) \geq -M\sqrt{\varepsilon} . \qquad (14.264)$$

Proof. Proposition 14.7.10 implies

$$\left|\frac{\partial^2 \varPhi}{\partial m^2}(m,u)\right| \le M_1 \;,$$

so that by Taylor's formula and since $f(u) = \partial\varPhi(m,q)/\partial m|_{m=m_{r-1}}$, whenever $m_{r-1} \le m \le m_r$ we have

$$\varPhi(m,u) \le \varPhi(m_{r-1},u) + (m - m_{r-1})f(u) + M_1(m-m_{r-1})^2 \;. \tag{14.265}$$

Combining with (14.229) yields

$$(m_r - m_{r-1})f(u) + M_1(m_r - m_{r-1})^2 \ge 0 \;, \tag{14.266}$$

and combining with (14.228) we obtain

$$(m - m_{r-1})f(u) + M_1(m - m_{r-1})^2 \ge -\varepsilon \;. \tag{14.267}$$

It follows from (14.266) that

$$-f(u) \le M_1(m_r - m_{r-1}) \;.$$

Without loss of generality we may assume that $f(u) \le 0$, and thus

$$m_{r-1} \le m := m_{r-1} - \frac{f(u)}{2M_1} \le m_r \;.$$

Using (14.267) for this value of m yields that $-f(u)^2/4M_1 \ge -\varepsilon$. □

Lemma 14.7.14. *We have* $|f^{(3)}(u)| \le M$.

Proof. Using (14.239) it suffices to show that

$$|U^{(3)}(v)| \le M \;. \tag{14.268}$$

Using (14.240) and since $\mathsf{E}(W_1 \cdots W_{r-2}) = 1$, it suffices to show that

$$\left|\frac{\partial^3 \varDelta_{r-1}}{\partial v^3}(t,u)\right| \le M \;, \tag{14.269}$$

which should be obvious. □

Lemma 14.7.15. *When* $h \ne 0$*, there exists a constant* M *depending on* h *and* ξ *only such that*

$$q_1 \ge \frac{1}{M} \;. \tag{14.270}$$

Before we prove this, we gather some properties of the functions A_p. Some of these properties will be used only later. It is (14.274) which will be used to prove Lemma 14.7.15.

Lemma 14.7.16. *We have*

$$A_p(x) = A_p(-x)\ ;\ |A_p'(x)| \le 1\ ;\ 0 \le A_p''(x) \le 1 \tag{14.271}$$

$$|A_p^{(3)}| \le 4 \tag{14.272}$$

$$A_{p+1}(x) \le A_p(x) \le A_{p+1}(x) + \xi'(q_{p+1}) - \xi'(q_p) \tag{14.273}$$

$$A_p''(x) \ge \frac{1}{M\mathrm{ch}^2 x}\ , \tag{14.274}$$

where M depends on ξ only.

Proof. We use Theorem 14.2.1 to obtain a representation

$$A_p(x) = \mathsf{E}\log\sum_\alpha v_\alpha \mathrm{ch}(x+z_\alpha)\ ,$$

where (v_α) are certain weights and (z_α) are symmetric r.v.s. Then $A_p(x) = A_p(-x)$ and

$$A_p'(x) = \mathsf{E}\frac{\sum_\alpha v_\alpha \mathrm{sh}(x+z_\alpha)}{\sum_\alpha v_\alpha \mathrm{ch}(x+z_\alpha)}\ . \tag{14.275}$$

Now

$$\left|\frac{\sum_\alpha v_\alpha \mathrm{sh}(x+z_\alpha)}{\sum_\alpha v_\alpha \mathrm{ch}(x+z_\alpha)}\right| \le \frac{\sum_\alpha v_\alpha |\mathrm{sh}(x+z_\alpha)|}{\sum_\alpha v_\alpha \mathrm{ch}(x+z_\alpha)} \le 1 \tag{14.276}$$

since $|\mathrm{sh}x| \le \mathrm{ch}x$. In particular (14.275) shows that $|A_p'(x)| \le 1$. We differentiate (14.275) to obtain

$$A_p''(x) = 1 - \mathsf{E}\left(\frac{\sum_\alpha v_\alpha \mathrm{sh}(x+z_\alpha)}{\sum_\alpha v_\alpha \mathrm{ch}(x+z_\alpha)}\right)^2$$

and combining with (14.276) this proves (14.271). Furthermore we have

$$A_p^{(3)}(x) = -2\mathsf{E}\left(\frac{\sum_\alpha v_\alpha \mathrm{sh}(x+z_\alpha)}{\sum_\alpha v_\alpha \mathrm{ch}(x+z_\alpha)}\right) + 2\mathsf{E}\left(\frac{\sum_\alpha v_\alpha \mathrm{sh}(x+z_\alpha)}{\sum_\alpha v_\alpha \mathrm{ch}(x+z_\alpha)}\right)^3\ ,$$

and this proves (14.272). To prove (14.273) we think of A_{p+1} and q_{p+1} as fixed. Then (14.190) "defines A_p as a function $G(x, q_p)$". More specifically, consider the function $B(x, v, m_p)$ of (14.196) in the case $A = A_{p+1}$, and for $y \le q_{p+1}$ define

$$G(x,y) = B(x, \xi'(q_{p+1}) - \xi'(y), m_p)\ , \tag{14.277}$$

so that

$$\frac{\partial G}{\partial y} = -\xi''(y)\frac{\partial B}{\partial v}\ ;\ \frac{\partial G}{\partial x} = \frac{\partial B}{\partial x}\ ;\ \frac{\partial^2 G}{\partial x^2} = \frac{\partial^2 B}{\partial x^2}\ .$$

Consequently (14.199) implies

$$\frac{\partial G}{\partial y}(x,y) = -\xi''(y)\left(\frac{1}{2}\frac{\partial^2 G}{\partial x^2}(x,y) + \frac{m_p}{2}\left(\frac{\partial G}{\partial x}(x,y)\right)^2\right) . \tag{14.278}$$

Now, when $q_p = y$ we have $G(x,y) = A_p(x)$ so that (14.271) entails

$$-\xi''(y) \le \frac{\partial G}{\partial y}(x,y) \le 0 .$$

Finally we observe that $G(x,q_{p+1}) = A_{p+1}(x)$ and that $G(x,q_p) = A_p(x)$, so that

$$A_{p+1}(x) - A_p(x) = \int_{q_p}^{q_{p+1}} \frac{\partial G}{\partial y}(x,z)\mathrm{d}z ,$$

and this proves (14.273).

To prove (14.274) we show by decreasing induction over p that

$$A_p''(x) \ge \frac{\exp\bigl(-2(\xi'(1) - \xi'(q_p))\bigr)}{\mathrm{ch}^2 x} . \tag{14.279}$$

For the induction from $p+1$ to p, we use (14.277), recalling that $A_p(x) = G(x,q_p)$ and (14.204) to get

$$A_p''(x) \ge \mathsf{E}\bigl(A_{p+1}''(x+z_p)\exp m_p(A_{p+1}(x+z_p) - A_p(x))\bigr)$$

so that, by the induction hypothesis,

$$A_p''(x) \ge \alpha_{p+1}\mathsf{E}\left(\frac{1}{\mathrm{ch}^2(x+z_p)}\exp m_p(A_{p+1}(x+z_p) - A_p(x))\right) , \tag{14.280}$$

where

$$\alpha_{p+1} = \exp(-2(\xi'(1) - \xi'(q_{p+1})) .$$

Now, since $(\log \mathrm{ch} x)'' \le 1$,

$$\log \mathrm{ch}(x+y) \le \log \mathrm{ch} x + y\,\mathrm{th} x + \frac{y^2}{2} ,$$

and thus

$$\frac{1}{\mathrm{ch}^2(x+y)} \ge \frac{1}{\mathrm{ch}^2 x}\exp(-2y\,\mathrm{th} x - y^2) . \tag{14.281}$$

Since $A_{p+1}(x+z_p) \ge A_{p+1}(x) + z_p A_{p+1}'(x)$ by convexity, we deduce from (14.280) and (14.281) that if $c = m_p A_{p+1}'(x) - 2\mathrm{th} x$,

$$A_p''(x) \ge \frac{\alpha_{p+1}}{\mathrm{ch}^2 x}\exp m_p(A_{p+1}(x) - A_p(x))\mathsf{E}\exp(cz_p - z_p^2) . \tag{14.282}$$

Moreover, using (A.6) in the first inequality and that $1 + x \le e^x$ in the second inequality, we get

$$\mathsf{E}\exp(cz_p - z_p^2) \geq \frac{1}{\sqrt{1+2\mathsf{E}z_p^2}} \geq \exp(-\mathsf{E}z_p^2) = \exp\bigl(-(\xi'(q_{p+1}) - \xi'(q_p))\bigr) \; .$$

Combining with (14.282) we deduce from (14.273) that

$$A_p''(x) \geq \frac{1}{\mathrm{ch}^2 x}\alpha_{p+1}\exp\bigl(-2(\xi'(q_{p+1}) - \xi'(q_p))\bigr)$$

and this completes the proof of (14.279). □

Proof of Lemma 14.7.15. First, since $A_0(x) = A_0(-x)$ we have $A_0'(0) = 0$. It then follows from (14.274) that $|A_0'(x)| \geq |\mathrm{th}(x)|/M$. Now (14.222) implies

$$q_1 = \mathsf{E}A_1'(\zeta_1)^2 = \mathsf{E}A_1'(h+z_0)^2 \; .$$

Denoting by E_0 integration in z_0 only (but not in h) Jensen's inequality implies

$$q_1 \geq \mathsf{E}(\mathsf{E}_0 A_1'(h+z_0))^2 = \mathsf{E}A_0'(h)^2 \geq \mathsf{E}(\mathrm{th}h)^2 \geq \frac{1}{M} \; . \qquad □$$

Proposition 14.7.17. *Assume that $h \neq 0$. Then we have*

$$-f''(q_r) = \frac{1}{2}\xi''(q_r)(-\xi''(q_r)U''(0) - 1) \leq M\varepsilon^{1/6} \; . \tag{14.283}$$

Proof. The expression for $f''(q_r)$ is straightforward from (14.252), since $U'(0) = q_r$. From (14.251) and Lemma 14.7.14, Taylor's formula yields

$$f(u) \leq \frac{1}{2}(u-q_r)^2 f''(q_r) + M|u-q_r|^3 \; . \tag{14.284}$$

Assume first

$$q := q_r - \varepsilon^{1/6} \geq q_{r-1} \; . \tag{14.285}$$

Using (14.284) and that $f(q) \geq -M\sqrt{\varepsilon}$ (by (14.264)) yields

$$-M\sqrt{\varepsilon} \leq \frac{1}{2}\varepsilon^{1/3} f''(q_r) + M\sqrt{\varepsilon} \; ,$$

and this finishes the proof in that case. Assume now that (14.285) fails. Then, since $q_0 = 0$ (and provided that ε is small enough), Lemma 14.7.15 implies that $r \geq 2$. Then $f(q_{r-1}) = 0$ by (14.251), and the right-hand side of (14.284) is ≥ 0 for $u = q_{r-1}$ and thus

$$-f''(q_r) \leq 2M(q_r - q_{r-1}) \; ,$$

and proof is finished since $q_r - q_{r-1} \leq \varepsilon^{1/6}$ because (14.285) fails. □

14.8 Main Estimate: Methodology

In this section we describe in more detail how we shall prove Theorem 14.5.7, therefore completing the proof of the Parisi Formula. From this point to the end of Section 14.10 we assume that $h \neq 0$. We consider $k, \mathbf{m}, \mathbf{q}$ that satisfy the condition MIN(ε) of Definition 14.5.3. That is, $\mathbf{q}$ and $\mathbf{m}$ are as in (14.104) and (14.103) respectively,

$$0 < m_1 < \ldots < m_k < 1 \tag{14.286}$$

$$0 = q_0 \leq q_1 < q_2 < \ldots < q_k < q_{k+1} \leq 1 = q_{k+2} \,, \tag{14.287}$$

we assume that $\mathcal{P}_k(\mathbf{m}, \mathbf{q})$ cannot be decreased by a change of $\mathbf{m}$ or $\mathbf{q}$, and that $\mathcal{P}_k(\mathbf{m}, \mathbf{q}) < \mathcal{P}(\xi, h) + \varepsilon$. We intend to prove that if $t \leq t_0$ and if ε is small enough then for each $1 \leq r \leq k+1$ we have

$$\Psi_r(t, u) \leq 2\psi(t) - \frac{(u - q_r)^2}{K} \,, \tag{14.288}$$

where K does not depend on t, u or r. Since we think of r as fixed once and for all, we will omit the subscript r.

All the bounds we will obtain for $\Psi(t, u)$ will rely on Proposition 14.6.3. We will choose the sequence $n_0, \ldots, n_\kappa$ as in (14.159), i.e.

$$n_0 = 0 \leq n_1 \leq n_2 \leq \ldots \leq n_{\kappa-1} \leq n_\kappa = 1 \,, \tag{14.289}$$

and the pairs (Z_p^1, Z_p^2) in a way that (14.165) holds, i.e.

$$\Psi(t, u) = \frac{1}{N}\mathsf{E}G_1 - (1-t)(\xi'(1) - \xi'(q_{k+1})) \,. \tag{14.290}$$

Therefore the right-hand side of (14.162) will provide a bound for $\Psi(t, u)$. A detailed study of this bound will yield (14.288).

Our first concern is to ensure (14.290). There is a canonical method to ensure this equality, and we describe it now. When $Z_p^1 = Z_p^2 = 0$, it follows from (14.161) that $G_p = G_{p+1}$. “Nothing happens for that p”. Suppose that we can remove from the list $n_0, \ldots, n_\kappa$ some terms for which “nothing happens” and then get exactly the list

$$0, m_1^* = \frac{m_1}{2}, \ldots, m_{r-1}^* = \frac{m_{r-1}}{2}, m_r^* = m_r, \ldots, m_k^* = m_k, m_{k+1}^* = 1 \,. \tag{14.291}$$

Suppose moreover that for $0 \leq p \leq k+1$ the p^{th}-term of that list occurs as a term $n_{a(p)}$ of the list (14.289) in such a way that $(Z_{a(p)}^1, Z_{a(p)}^2) = (\sqrt{1-t}z_p^1, \sqrt{1-t}z_p^2)$, where the pairs $(z_p^1, z_p^2)_{1 \leq p \leq k+1}$ are as follows:

$$0 \leq p \leq k+1 \quad \Rightarrow \quad \mathsf{E}(z_p^1)^2 = \mathsf{E}(z_p^2)^2 = \xi'(q_{p+1}) - \xi'(q_p) \tag{14.292}$$

$$p < r \quad \Rightarrow \quad z_p^1 = z_p^2 \; ; \qquad p \ge r \quad \Rightarrow \quad z_p^1 \text{ and } z_p^2 \text{ are independent} \; . \tag{14.293}$$

(The difference with (14.96) and (14.97) is that now p goes up to $k+1$ rather than k.) We claim that then (14.290) holds. This is because to compute G_1 we need to "take only in account the values of n_a for which something happens". That is, starting with

$$G_{k+2}^* = \log \sum_{R_{1,2}=u} \exp\Bigg(-\sqrt{t} H_N(\boldsymbol{\sigma}^1) - \sqrt{t} H_N(\boldsymbol{\sigma}^2) + \sum_{\ell=1,2} \sum_{i \le N} \sigma_i^\ell \Big(h_i + \sqrt{1-t} \sum_{0 \le p \le k+1} z_{i,p}^\ell \Big) \Bigg) \; , \tag{14.294}$$

we define recursively for $1 \le p \le k+1$

$$G_p^* = \frac{1}{m_p^*} \log \mathsf{E}_p \exp m_p^* G_{p+1}^* \; , \tag{14.295}$$

where E_p denotes expectation in the r.v.s z_n^ℓ for $n \ge p$. Then $G_1 = G_1^*$. Moreover since $m_{p+1}^* = 1$, and since $\mathsf{E}(z_{k+1}^\ell)^2 = \xi'(q_{k+2}) - \xi'(q_{k+1}) = \xi'(1) - \xi'(q_{k+1})$ and $\mathsf{E} z_{k+1}^1 z_{k+1}^2 = 0$ we have

$$\begin{aligned} G_{k+1}^* &= (1-t)(\xi'(1) - \xi'(q_{k+1})) \\ &\quad + \log \sum_{R_{1,2}=u} \exp\Bigg(-\sqrt{t} H_N(\boldsymbol{\sigma}^1) - \sqrt{t} H_N(\boldsymbol{\sigma}^2) \\ &\quad + \sum_{\ell=1,2} \sum_{i \le N} \sigma_i^\ell \Big(h_i + \sqrt{1-t} \sum_{0 \le p \le k} z_{i,p}^\ell \Big) \Bigg) \; , \end{aligned} \tag{14.296}$$

and recalling the definition (14.98) of $J_{k+1}(u)$ this means

$$G_{k+1}^* = (1-t)(\xi'(1) - \xi'(q_{k+1})) + J_{k+1}(u) \; .$$

Since the recursions (14.99) and (14.295) are identical, for each $1 \le p \le k+1$ we have

$$G_p^* = (1-t)(\xi'(1) - \xi'(q_{k+1})) + J_p(u) \; .$$

Taking $p = 1$ and expectation, this proves (14.290) since $\Psi(t,u) = \mathsf{E} J_1(u)$ and $\mathsf{E} G_1^* = \mathsf{E} G_1$.

We will call this situation the *canonical situation*. Thus we have shown that equality (14.290) occurs whenever the canonical situation occurs. The right-hand side of (14.162) then provides the following bound for $\Psi(t,u)$:

$$\begin{aligned} \Psi(t,u) &\le 2\log 2 + Y_0 - \lambda u - (1-t)(\xi'(1) - \xi'(q_{k+1})) \\ &\quad - t\Bigg(2 \sum_{p<\tau} n_p(\theta(\rho_{p+1}) - \theta(\rho_p)) + \sum_{\tau \le p \le \kappa} n_p(\theta(\rho_{p+1}) - \theta(\rho_p)) \Bigg) \; . \end{aligned} \tag{14.297}$$

Besides the sequence (14.289) this bounds depends on λ and on the sequence (14.155). Let us denote by $\Psi^*(t,u)$ the infimum of the right-hand side of (14.297) over all choices of these parameters for which the canonical situation occurs. Then

$$\Psi(t,u) \le \Psi^*(t,u)\ , \tag{14.298}$$

and $\Psi^*(t,u)$ does not depend on N. Also, given any choice of parameters, the right-hand side of (14.297) is a continuous function of t and u. Therefore, $\Psi^*(t,u)$ is an upper semi-continuous function of t and u. We now completely forget about Ψ and work only with Ψ^*. That is, we will prove that

$$\Psi^*(t,u) \le 2\psi(t) - \frac{(u-q_r)^2}{K}\ . \tag{14.299}$$

The basic tool is the following obvious consequence of the definition of $\Psi^*(t,u)$. Whenever the canonical situation occurs,

$$\begin{aligned}\Psi^*(t,u) \le\ & 2\log 2 + Y_0 - \lambda u - (1-t)(\xi'(1) - \xi'(q_{k+1})) \\ & - t\Bigg(2\sum_{p<\tau} n_p(\theta(\rho_{p+1}) - \theta(\rho_p)) + \sum_{\tau\le p\le\kappa} n_p(\theta(\rho_{p+1}) - \theta(\rho_p))\Bigg)\ .\end{aligned} \tag{14.300}$$

A useful observation is that to prove (14.299), it suffices essentially to prove this estimate for u close to q_r, and to prove that $\Psi^*(t,u) < 2\psi(t)$ for $u \ne q_r$. We will have to distinguish a number of cases. The next proposition addresses the case where $u \le q_r$ is close to q_r. We remind the reader that M denotes a quantity depending only on ξ and h, that need not be the same at each occurrence, and that $M_1, M_2, M^* \ldots$ denote specific such quantities.

Proposition 14.8.1. *There exists M_1 with the following property. Assume that $k, \mathbf{m}, \mathbf{q}$ satisfy the condition* $\mathrm{MIN}(\varepsilon)$ *of Definition 14.5.3, where ε is such that $M_1\varepsilon^{1/6} \le 1 - t_0$. Then*

$$q_{r-1} \le u \le q_r\ ,\ M_1(q_r - u) \le 1 - t_0 \ \ \Rightarrow\ \ \Psi^*(t,u) \le 2\psi(t) - \frac{(1-t_0)^2}{M_1}(u-q_r)^2\ . \tag{14.301}$$

We consider the function

$$\gamma(c) = \inf\big\{|\xi(y) - \xi(x) + (x-y)\xi'(y)|\ ;\ 0 \le x,y \le 1\ ,\ |x-y| > c\big\}\ . \tag{14.302}$$

Since we assume that $\xi''(x) > 0$ for $x > 0$, we have $\gamma(c) > 0$ for $c > 0$. In (14.304) below, the constant M_1 is the same that as in (14.310).

Proposition 14.8.2. *There exists a number M_3 with the following property. Assume that $k, \mathbf{m}, \mathbf{q}$ satisfy the condition* $\mathrm{MIN}(\varepsilon)$ *of Definition 14.5.3, where ε is such that*

$$M_3\varepsilon^{1/2} \le (1-t_0)\gamma\left(\frac{1-t_0}{M_1}\right)\ . \tag{14.303}$$

Then for $r \geq 2$ we have

$$q_{r-1} \leq u \leq q_r \;, \quad M_1(q_r - u) \geq 1 - t_0 \quad \Rightarrow \quad \Psi^*(t,u) < 2\psi(t) \;. \qquad (14.304)$$

In (14.304), the constant M_1 is the same that as in (14.310), so that when $r \geq 2$, Propositions 14.8.2 and 14.8.1 together cover all values $q_{r-1} \leq u \leq q_r$. The case $r = 1$ requires a specific argument, and is the object of the next result, that also covers some of the cases where $u < 0$.

Proposition 14.8.3. *Assume that $k, \mathbf{m}, \mathbf{q}$ satisfy the condition* MIN(ε) *of Definition 14.5.3, where $M_1 \varepsilon^{1/6} \leq 1 - t_0$. Then if $r = 1$ and $|u| \leq q_1$ we have*

$$\Psi^*(t,u) \leq 2\psi(t) - \frac{(1-t_0)^2}{M}(u - q_1)^2 \;.$$

Of course, we also need the results corresponding to Propositions 14.8.1 and 14.8.2 when u is to right of q_r rather than to the left. These are valid for each value of $1 \leq r \leq k+1$.

Proposition 14.8.4. *There exists M_2 such that whenever $k, \mathbf{m}, \mathbf{q}$ satisfy the condition* MIN(ε) *of Definition 14.5.3 where $M_2 \varepsilon^{1/6} \leq 1 - t_0$, then*

$$q_r \leq u \leq q_{r+1} \;, \; M_1(u - q_r) \leq 1 - t_0 \;\Rightarrow\; \Psi^*(t,u) \leq 2\psi(t) - \frac{(1-t_0)^2}{M_2}(u - q_r)^2 \;. \qquad (14.305)$$

Proposition 14.8.5. *There exists a quantity M_4 such that whenever $k, \mathbf{m}, \mathbf{q}$ satisfy the condition* MIN(ε) *of Definition 14.5.3 where*

$$M_4 \varepsilon^{1/2} \leq (1 - t_0)\gamma\left(\frac{1-t_0}{M_4}\right) \;, \qquad (14.306)$$

then we have

$$q_r \leq u \leq q_{r+1} \;, \; M_1(u - q_r) \geq 1 - t_0 \quad \Rightarrow \quad \Psi^*(t,u) < 2\psi(t) \;. \qquad (14.307)$$

Finally, a last effort is needed to cover the missing cases:

Proposition 14.8.6. *Assume that $t > 0$, $u < q_{r-1}$ or $u > q_{r+1}$. Then we have*

$$\Psi^*(t,u) < 2\psi(t) \;. \qquad (14.308)$$

Proposition 14.8.7. *If $t = 0$ and $u \neq q_r$ then $\Psi^*(t,u) < 2\psi(t)$.*

Proof of Theorem 14.5.7 (The Main Estimate). The previous results show that if ε is small enough (depending only on ξ, h and t_0) then for $|u - q_r| \geq (1-t_0)/M_1$ and $0 \leq t \leq t_0$ we have $\Psi^*(t,u) < 2\psi(t)$. More specifically, if $r \geq 2$, this follows from Proposition 14.8.2 if $q_{r-1} \leq u \leq q_r$ and

$M_1(q_r - u) \geq 1 - t_0$; from Proposition 14.8.5 if $q_r \leq u \leq q_{r+1}$, $M_1(u - q_r) \geq 1 - t_0$; from Proposition 14.8.1 if $q_{r-1} \leq u \leq q_r$ and $M_1(q_r - u) \leq 1 - t_0$; from Proposition 14.8.4 if $q_r \leq u \leq q_{r+1}$ and $M_1(u - q_r) \leq 1 - t_0$; and from Propositions 14.8.6 and 14.8.7 in the other cases. If $r = 1$, one also needs Proposition 14.8.3 to cover the case $|u| \leq q_1$.

Since the set of values of u, t given by $|u - q_r| \geq (1 - t_0)/M_1$ and $0 \leq t \leq t_0$ is compact and $\Psi^*(t, u)$ is upper semi-continuous, so there exists $\varepsilon_0 >$ with $\Psi^*(t, u) \leq 2\psi(t) - \varepsilon_0$ for these values of u, t. On the other hand, we have also proved that

$$t \leq t_0 \ , \quad |u - q_r| \leq \frac{1 - t_0}{M_1} \quad \Rightarrow \quad \Psi^*(t, u) \leq 2\psi(t) - \frac{(u - q_r)^2}{M}$$

and if M is large enough this holds for any value of u. □

14.9 Main Estimate: The Critical Cases

In this section we prove Propositions 14.8.1 to 14.8.5. The case of Proposition 14.8.6 (and of the much easier Proposition 14.8.7) is a bit different and will be covered in the next section. The title of the section is a bit misleading. *Each* of the Propositions of the preceding section is critical, but somehow "there is more room in Proposition 14.8.6".

We chose $\kappa = k + 2$ and

$$\begin{aligned} &n_0 = 0 \ , \quad n_1 = \frac{m_1}{2} \ , \quad \ldots \ , \quad n_{r-1} = \frac{m_{r-1}}{2} \ , \quad n_r = m \ , \\ &n_{r+1} = m_r \ , \quad \ldots \ , \quad n_{k+1} = m_k \ , \quad n_{k+2} = 1 \ , \end{aligned} \tag{14.309}$$

where m is any number $m_{r-1}/2 \leq m \leq m_r$.

Let us recall the pairs (z_p^1, z_p^2) as in (14.292) and (14.293) and define

$$\begin{aligned} &Z_p^\ell = \sqrt{1 - t} z_p^\ell \ \text{ if } p < r \\ &Z_r^\ell = 0 \\ &Z_p^\ell = \sqrt{1 - t} z_{p-1}^\ell \ \text{ if } r + 1 \leq p \leq k + 2 \ . \end{aligned} \tag{14.310}$$

If we remove from the list (14.309) the term n_r "for which nothing happens", we get the list (14.291). Thus the canonical situation occurs and (14.300) holds.

In order to prove Proposition 14.8.1 and 14.8.2, we consider $q_{r-1} \leq u \leq q_r$ and the sequence

$$\begin{aligned} &\rho_0 = 0 \ , \quad \rho_1 = q_1 \ , \quad \ldots \ , \quad \rho_{r-1} = q_{r-1} \ , \quad \rho_r = u \ , \\ &\rho_{r+1} = q_r \ , \quad \ldots \ , \quad \rho_{k+2} = q_{k+1} \ , \quad \rho_{k+3} = q_{k+2} = 1 \ . \end{aligned} \tag{14.311}$$

We compute the terms $n_p(\theta(\rho_{p+1}) - \theta(\rho_p))$ as follows.

$$\begin{aligned}0 \le p < r-1 \Rightarrow n_p(\theta(\rho_{p+1}) - \theta(\rho_p)) &= \frac{m_p}{2}(\theta(q_{p+1}) - \theta(q_p)) \\ n_{r-1}(\theta(\rho_r) - \theta(\rho_{r-1})) &= \frac{m_{r-1}}{2}(\theta(u) - \theta(q_{r-1})) \\ n_r(\theta(\rho_{r+1}) - \theta(\rho_r)) &= m(\theta(q_r) - \theta(u)) \\ r+1 \le p \le k+2 \Rightarrow n_p(\theta(\rho_{p+1}) - \theta(\rho_p)) &= m_p(\theta(q_p) - \theta(q_{p-1})) ,\end{aligned}$$

so that, collecting the terms in the bound (14.300) with $\tau = r$ gives

$$\begin{aligned}\Psi^*(t,u) \le 2\log 2 + Y_0 - \lambda u - (1-t)(\xi'(1) - \xi'(q_{k+1})) \\ - t\biggl(\sum_{p \le k+1} m_k(\theta(q_{k+1}) - \theta(q_k)) + (m - m_{r-1})(\theta(q_r) - \theta(u))\biggr) .\end{aligned} \tag{14.312}$$

This bound now depends only on u and the parameters m and λ. The basic idea is that (as we shall soon see) the choice $m = m_{r-1}$ and $\lambda = 0$ witnesses the inequality $\Psi^*(t,u) \le 2\psi(t)$, so that we can improve on that inequality by a small variation of either m or λ.

We think of Y_0 in (14.312) as a function $Y_0(\lambda, m, u)$,

$$Y_0 = Y_0(\lambda, m, u) ,$$

and we recall the definitions (14.235) of $S(v,m)$ and (14.238) of $U(v)$.

Lemma 14.9.1. *If $m \ge m_{r-1}$ we have*

$$Y_0(0,m,u) = 2S(v,m) \tag{14.313}$$

where

$$v = t(\xi'(q_r) - \xi'(u)) . \tag{14.314}$$

Moreover

$$\frac{\partial}{\partial\lambda} Y_0(\lambda, m_{r-1}, u)\bigg|_{\lambda=0} = U'(v) . \tag{14.315}$$

Proof. We recall that $y_p^1 = y_p^2$ for $0 \le p < r$ and are independent for $r \le p \le \kappa = k+2$, and that $\mathsf{E}(y_p^1)^2 = \mathsf{E}(y_p^2)^2 = \xi'(\rho_{p+1}) - \xi'(\rho_p)$. Consider the r.v.s

$$g_p^\ell = \sqrt{t} y_p^\ell + Z_p^\ell .$$

Thus $g_p^1 = g_p^2$ for $p < r$, while these variables are independent for $p \ge r$. Also,

$$\mathsf{E}(g_p^\ell)^2 = t\mathsf{E}(y_p^\ell)^2 + \mathsf{E}(Z_p^\ell)^2 ,$$

which we compute now for the various values of p. For $r+1 \le p \le \kappa$, and since then $\rho_{p+1} = q_p$ and $\rho_p = q_{p-1}$

$$\begin{aligned}\mathsf{E}(g_p^\ell)^2 &= t(\xi'(q_p) - \xi'(q_{p-1})) + (1-t)(\xi'(q_p) - \xi'(q_{p-1})) \\ &= \xi'(q_p) - \xi'(q_{p-1}) .\end{aligned} \tag{14.316}$$

For $p = r$,

$$\mathsf{E}(g_r^\ell)^2 = v = t(\xi'(q_r) - \xi'(u)) \,. \tag{14.317}$$

For $p = r - 1$,

$$\begin{aligned}\mathsf{E}(g_{r-1}^\ell)^2 &= t(\xi'(u) - \xi'(q_{r-1})) + (1-t)(\xi'(q_r) - \xi'(q_{r-1})) \\ &= a - v \,,\end{aligned} \tag{14.318}$$

where $a = \xi'(q_r) - \xi'(q_{r-1})$. And for $0 \le p \le r-2$ we have

$$\mathsf{E}(g_p^\ell)^2 = t(\xi'(q_{p+1}) - \xi'(q_p)) + (1-t)(\xi'(q_{p+1}) - \xi'(q_p)) = \xi'(q_{p+1}) - \xi'(q_p) \,. \tag{14.319}$$

Recall the sequence $(D_p)_{p \le \kappa+1}$ of functions constructed through (14.170), (14.171) and (14.172) with $g_p = g_p^1$, and that the sequence (A_p) is constructed through (14.190). It follows from (14.316) that

$$D_p = A_{p-1} \tag{14.320}$$

for $r+1 \le p \le \kappa+1 = k+3$. This is proved by decreasing induction over p. Continuing the decreasing induction on p we prove that for $p \le r$ we have

$$D_p(x) = C_p(x, v, m) \,, \tag{14.321}$$

where the functions $C_p(x, v, m)$ are constructed recursively through (14.231) and (14.232). To see this, for $p = r$ we use (14.231) and (14.317). For $p = r-1$ we use (14.232) and (14.318); and for $p \le r-2$ we use (14.234) and (14.316). It follows from (14.176) that if $\mathsf{E}z_0^2 = \xi'(q_1)$ then

$$Y_0(0, m, u) = 2\mathsf{E}D_1(h + z_0) = 2\mathsf{E}C_1(h + z_0, v, m) \,.$$

Comparing with (14.235) proves (14.313).

It remains to prove (14.315). Using (14.177) for $\tau = r$ and using (14.321) in the second line yields

$$\begin{aligned}\frac{\partial}{\partial\lambda} Y_0(\lambda, m_{r-1}, u)\Big|_{\lambda=0} &= \mathsf{E}(W_1 \cdots W_{r-1} D_r'(\zeta_r)^2) \\ &= \mathsf{E}(W_1 \cdots W_{r-1} C_r'(\zeta_r, v, m_{r-1})^2) \,,\end{aligned} \tag{14.322}$$

where $\zeta_p = h + \sum_{0 \le n < p} g_p^1$ and

$$W_p = \exp 2n_p(D_{p+1}(\zeta_{p+1}) - D_p(\zeta_p)) = \exp m_p(D_{p+1}(\zeta_{p+1}) - D_p(\zeta_p)) \,.$$

Now, (14.321), (14.233) and (14.234) prove that for $p \le r-1$ it holds that

$$D_p(x) = C_p(x, v, m_{r-1}) = A_p(x) \,.$$

On then checks that the terms denoted by W_{r-1} and ζ_r the right-hand-side of (14.322) coincide respectively with the terms denoted by $W_{r-1}(v)$ and $\zeta_r(v)$

in the right-hand side of (14.243). Therefore the right-hand sides of (14.322) and (14.243) coincide, and this finishes the proof. □

As a consequence of (14.313), recalling the expression (14.86) of $\psi(t)$ and (14.237), the choice $\lambda = 0$, $m = m_{r-1}$ in (14.312) witnesses the inequality $\Psi^*(t,u) \le 2\psi(t)$.

We now set

$$\alpha(\lambda) = Y_0(\lambda, m_{r-1}, u) - \lambda u \tag{14.323}$$

and

$$\beta(u) := \alpha'(0) = \frac{\partial}{\partial\lambda} Y_0(\lambda, m_{r-1}, u)\Big|_{\lambda=0} - u \ , \tag{14.324}$$

so that, using (14.315),

$$\beta(u) = U'(t(\xi'(q_r) - \xi'(u))) - u \ , \tag{14.325}$$

and, by straightforward differentiation,

$$\beta'(q_r) = -t\xi''(q_r)U''(0) - 1 \ . \tag{14.326}$$

Let us note also that (14.250) implies

$$\beta(q_r) = 0 \ . \tag{14.327}$$

Lemma 14.9.2. *There exists M^* with the following property. Assume that $k, \mathbf{m}, \mathbf{q}$ satisfy the condition* MIN(ε) *of Definition 14.5.3, where ε is such that $M^*\varepsilon^{1/6} \le 1 - t_0$. Then*

$$\beta'(q_r) \le -\frac{1-t_0}{2} \ . \tag{14.328}$$

Proof. We recall that $U''(v) < 0$ by (14.249), and that, by (14.240) we have $|U''| \le M$. Also, (14.283) yields

$$\xi''(q_r)(-\xi''(q_r)U''(0) - 1) \le M\varepsilon^{1/6} \ . \tag{14.329}$$

Now, if $\xi''(q_r)(-U''(0)) \le 1/2$, then $-t\xi''(q_r)U''(0) \le 1/2$ since $-U''(0) \ge 0$ and (14.326) implies that $\beta'(q_r) \le -1/2$ and we are done. So we may assume

$$\xi''(q_r)(-U''(0)) \ge \frac{1}{2} \ ,$$

and since $|U''(0)| \le M$ by (14.240), this implies that $\xi''(q_r) \ge 1/M'$. Consequently (14.329) implies

$$-\xi''(q_r)U''(0) - 1 \le \frac{M\varepsilon^{1/6}}{\xi''(q_r)} \le MM'\varepsilon^{1/6} = M''\varepsilon^{1/6} \ ,$$

so that

$$-\xi''(q_r)U''(0) \le 1 + M''\varepsilon^{1/6}$$

and, recalling (14.326),

$$\beta'(q_r) = -t\xi''(q_r)U''(0) - 1 \le t_0(1 + M''\varepsilon^{1/6}) - 1 \le t_0 + M''\varepsilon^{1/6} - 1 .$$

Consequently $\beta'(q_r) \le -(t_0 - 1)/2$ whenever $2M''\varepsilon^{1/6} \le 1 - t_0$. □

Proof of Proposition 14.8.1 We use the bound (14.312) for $m = m_{r-1}$. When $\lambda = 0$, this bound is $2\psi(t)$, and we will show that we may improve it by a small variation of λ. By Lemma 14.6.5, the function $\alpha(\lambda) = Y_0(\lambda, m_{r-1}, u) - \lambda u$ of (14.323) satisfies $|\alpha''(\lambda)| \le 1$, so that

$$\inf_\lambda \alpha(\lambda) \le \alpha(0) - \frac{\alpha'(0)^2}{2} , \tag{14.330}$$

and, recalling the definition (14.324) of $\beta(u)$,

$$\Psi^*(t, u) \le 2\psi(t) - \frac{\beta(u)^2}{2} . \tag{14.331}$$

To bound $|\beta(u)|$ from below, we may write, since $|\beta''(u)| \le M^\sim$ and since $\beta(q_r) = 0$ by (14.327)

$$\beta(u) \ge (u - q_r)\beta'(q_r) - M^\sim(u - q_r)^2 \ge \frac{1}{4}(q_r - u)(1 - t_0) \tag{14.332}$$

whenever $\beta'(q_r) \le -(1 - t_0)/2$ and $q_r - u \le (1 - t_0)/4M^\sim$. Combining with (14.331) this finishes the proof. □

Without loss of generality we assume

$$M^* \le M_1 . \tag{14.333}$$

Proof of Proposition 14.8.2. We already noticed that the choice $m = m_{r-1}$ and $\lambda = 0$ in (14.300) witnesses that $\Psi^*(t, u) \le 2\psi(t)$, and that we can improve on this inequality by making either a small change of λ or m. Since $r \ge 2$ we have $m_r > m_{r-1} > 0$. we may use any value m with $m_{r-1}/2 \le m \le m_r$. In particular we may either increase or decrease m from the value m_{r-1} (which is not true when $r = 1$, where we may only increase m). Therefore it suffices to show that the partial derivatives of the right-hand side of (14.312) at $\lambda = 0$ and $m = m_{r-1}$ with respect to λ and m cannot be both zero. The derivative with respect to λ was computed in the proof of Proposition 14.8.1. It is

$$U'\big(t(\xi'(q_r) - \xi'(q(u)))\big) - u . \tag{14.334}$$

To compute the derivative with respect to m of the right-hand side of (14.312) when $\lambda = 0$ we observe that then $Y_0 = Y_0(0, m, u) = 2S(v, m)$ by (14.313). Using (14.312) and (14.238), this derivative with respect to m is given by

$$d := U(t(\xi'(q_r) - \xi'(u))) - t(\theta(q_r) - \theta(u)) . \tag{14.335}$$

We assume that the quantity (14.334) is 0, and we prove that $d \neq 0$ (so that the partial derivatives with respect to λ and m cannot both be zero and the proof is finished). The function

$$t \mapsto d(t) = U\big(t(\xi'(q_r) - \xi'(u))\big) - t(\theta(q_r) - \theta(u))$$

is concave because $U'' < 0$ by (14.249), so that

$$d(1) \le d(t) + (1-t)d'(t) \ , \tag{14.336}$$

and

$$d'(t) = (\xi'(q_r) - \xi'(u))U'\big(t(\xi'(q_r) - \xi'(u))\big) - (\theta(q_r) - \theta(u)) \ .$$

We assume that the quantity (14.334) is 0, so that

$$d'(t) = u(\xi'(q_r) - \xi'(u)) - \theta(q_r) + \theta(u) \ ,$$

and, replacing $\theta(x)$ by its value $x\xi'(x) - \xi(x)$, we obtain

$$d'(t) = \xi(q_r) - \xi(u) + (u - q_r)\xi'(q_r) \ .$$

Since we assume $q_r - u \ge (1-t_0)/M_1$, the definition of the function $\gamma(c)$ of (14.302) implies $|d'(t)| \ge \gamma((1-t_0)/M_1)$ and since $d'(t) \le 0$ because ξ is convex,

$$d'(t) \le -\gamma\left(\frac{1-t_0}{M_1}\right)$$

and (14.336) yields

$$\begin{aligned} d = d(t) &\ge d(1) - (1-t)d'(t) \ge (1-t)\gamma\left(\frac{1-t_0}{M_1}\right) + d(1) \\ &\ge (1-t_0)\gamma\left(\frac{1-t_0}{M_1}\right) + d(1) \ . \end{aligned}$$

Now, using (14.264) and (14.239) we have

$$d(1) = U(\xi'(q_r) - \xi'(u)) - (\theta(q_r) - \theta(u)) = f(u) \ge -M\sqrt{\varepsilon}$$

and thus

$$d \ge (1-t_0)\gamma\left(\frac{1-t_0}{M_1}\right) - M\sqrt{\varepsilon} \ ,$$

and the right-hand side is > 0 under (14.303). □

Of course the preceding proof, while simple, is a kind of miracle. Is there any real reason why both partial derivatives cannot be zero at the same time, or is it just a strike of luck? Proposition 14.8.2 is absolutely critical, and it is uncomfortable to prove it using arguments that succeed for mysterious reasons.

Propositions 14.8.1 and 14.8.2 are based on the technique "of splitting the interval $[q_{r-1}, q_r[$ into two intervals $[q_{r-1}, u[$ and $[u, q_r[$ and assigning the value m_{r-1} to the interval $[q_{r-1}, u[$ and the value $m_{r-1} \le m \le m_r$ to the interval $[u, q_r[$". Similarly we can split the interval $[q_r, q_{r+1}[$ into two intervals $[q_r, u[$ and $[u, q_{r+1}[$, assigning the value $m_{r-1} \le m \le m_r$ to the interval $[q_r, u[$ and the value m_r to the interval $[u, q_r[$. In this manner we obtain the "mirror images" Propositions 14.8.4 and 14.8.5. The proof of both of them use the choice (14.310).

In the remainder of this section we prove Proposition 14.8.3, so that now $r = 1$. We still use the sequence (14.309), but with $m = 0$, so this sequence becomes

$$n_0 = 0 \;, \quad n_1 = 0 \;, \quad n_2 = m_1 \;, \quad \dots \;, \quad n_{k+1} = m_k \;, \quad n_{k+2} = 1 \;, \tag{14.337}$$

and we still use the r.v Z_p^ℓ defined in (14.310).

Rather than (14.311) we now use the sequence

$$\rho_0 = 0 \;, \; \rho_1 = |u| \;, \; \rho_2 = q_1 \;, \;\; \dots \;, \; \rho_{k+3} = 1 \;. \tag{14.338}$$

If $\eta \in \{-1, 1\}$ is such that $\eta u = |u|$, recalling (14.156) and (14.157), we now have

$$y_0^2 = \eta y_0^1 \;,$$

and y_p^1, y_p^2 are independent for $p \ge 1$. We observe that in the case where $u \ge 0$, the construction is simply the case $r = 1$ and $m = 0$ of the previous construction.

In the bound (14.312), the quantity Y_0 is now a function $Y_0(\lambda, u)$, which we study first.

Lemma 14.9.3. *We have*

$$Y_0(0, u) = 2X_0 \tag{14.339}$$

$$\frac{\partial}{\partial \lambda} Y_0(\lambda, u)\bigg|_{\lambda=0} = \mathsf{E}(A_1'(h + g_0^1 + g_1^1) A_1'(h + g_0^2 + g_1^2)) \;. \tag{14.340}$$

Proof. This relies on the second part of Proposition 14.6.4 used for $\tau = 2$ and $g_p^\ell = \sqrt{t} y_p^\ell + Z_p^\ell$. Since (14.316) holds for $p \ge 2$, so does (14.320), and hence $D_2 = A_1$. Now we observe that

$$\begin{aligned} \mathsf{E}(g_0^1 + g_1^1)^2 &= t\mathsf{E}(y_0^1 + y_0^2)^2 + \mathsf{E}(Z_0^1)^2 \\ &= t\big(\xi'(q_1) - \xi'(|u|) + \xi'(|u|) - \xi'(0)\big) + (1-t)\xi'(q_1) = \xi'(q_1) \;, \end{aligned}$$

and similarly $\mathsf{E}(g_0^2 + g_1^2)^2 = \xi'(q_1)$. Thus (14.179) implies

$$Y_0(0) = \mathsf{E} A_1(h + g_0^1 + g_1^1) + \mathsf{E} A_1(h + g_0^2 + g_1^2) = 2X_0 \;.$$

This proves (14.339) and (14.180) proves (14.340). □

As a consequence of the expression (14.86) for $\psi(t)$, when $\lambda = 0$, the bound (14.312) witnesses that $\Psi^*(t,u) \le 2\psi(t)$ so that if we set

$$\beta(u) = \frac{\partial}{\partial \lambda} Y_0(\lambda, u)\Big|_{\lambda=0} - u = \mathsf{E}(A_1'(h+g_0^1+g_1^1)A_1'(h+g_0^2+g_1^2)) - u \ , \quad (14.341)$$

the inequality (14.331) holds true.

Lemma 14.9.4. *If* $0 \le u \le q_1$ *we have*

$$\beta'(u) \le \beta'(q_1) \ . \quad (14.342)$$

If $-q_1 \le u \le 0$ *we have*

$$\beta(u) \ge \beta(0) \ . \quad (14.343)$$

Proof of Proposition 14.8.3. We recall that for $u \ge 0$ the construction is the same as that of Proposition 14.8.1, so that we can use Lemma 14.9.2 to obtain that if $M^* \varepsilon^{1/6} \le 1 - t_0$, then

$$\beta'(q_1) \le -\frac{1-t_0}{2} \ .$$

Since $\beta(q_1) = 0$ by (14.327), (14.342) implies that for $0 \le u \le q_1$

$$\beta(u) \ge \frac{1-t_0}{2}(q_1 - u) \ ,$$

while for $-q_1 \le u \le 0$, (14.343) yields

$$\beta(u) \ge \beta(0) \ge \frac{1-t_0}{2} q_1 \ge \frac{1-t_0}{4}(q_1 - u) \ .$$

Using (14.331) then concludes the proof. □

Lemma 14.9.5. *Consider two independent pairs of jointly Gaussian r.v.s* $\chi_1, \chi_2, \chi_1', \chi_2'$*, all of variance* a*, and a standard Gaussian r.v.* χ*. Then*

$$\begin{aligned} &|\mathsf{E}A'(h+\chi_1)A'(h+\chi_2) - \mathsf{E}A'(h+\chi_1')A'(h+\chi_2')| \\ &\le |\mathsf{E}\chi_1\chi_2 - \mathsf{E}\chi_1'\chi_2'| \mathsf{E}A''(h+\chi\sqrt{a})^2 \ . \end{aligned} \quad (14.344)$$

Proof. We consider the function

$$\varphi(s) = \mathsf{E}\big(A'(h+\sqrt{s}\chi_1 + \sqrt{1-s}\chi_1')A'(h+\sqrt{s}\chi_2+\sqrt{1-s}\chi_2')\big) \ ,$$

we compute $\varphi'(s)$ (using Lemma 1.3.1) to find

$$\varphi'(s) = (\mathsf{E}\chi_1\chi_2 - \mathsf{E}\chi_1'\chi_2')\mathsf{E}\big(A''(h+\sqrt{s}\chi_1+\sqrt{1-s}\chi_1')A''(h+\sqrt{s}\chi_2+\sqrt{1-s}\chi_2')\big)$$

and we use the Cauchy-Schwarz inequality to find that $|\varphi'(s)|$ is bounded by the right-hand side of (14.344). □

Proof of Lemma 14.9.4. Let us first assume that $u \geq 0$. Setting $a = \xi'(q_1)$ and $v = t(\xi'(q_1) - \xi'(u))$, when $u \geq 0$ we have $g_0^1 = g_0^2$ and

$$\mathsf{E}(g_0^\ell)^2 = t\xi'(u) + (1-t)\xi'(q_1) = a - v \ ,$$

while g_1^1 and g_1^2 are independent with

$$\mathsf{E}(g_1^1)^2 = \mathsf{E}(g_1^2)^2 = v \ .$$

Thus if χ, χ_1, χ_2 are standard Gaussian r.v.s we get

$$\begin{aligned} &\mathsf{E}(A_1'(h + g_0^1 + g_1^2)A_1'(h + g_0^2 + g_1^2)) \qquad (14.345) \\ &= \mathsf{E}(A_1'(h + \chi\sqrt{a-v} + \chi_1\sqrt{v})A_1'(h + \chi\sqrt{a-v} + \chi_2\sqrt{v})) =: -\Phi(v) \ , \end{aligned}$$

so that (14.341) implies

$$\beta(u) = -\Phi(v) - u = -\Phi\big(t(\xi'(q_1) - \xi'(u))\big) - u \ . \tag{14.346}$$

It is not difficult to prove using (14.245) that $-\Phi(v) = U'(v)$, so that the formula (14.346) coincides with the formula (14.325) (as it should, since we are making the same construction), but this fact is irrelevant to our proof.

We compute $\Phi'(v)$ by differentiating the formula (14.345) and integrating by parts (or by using Lemma 1.3.1). It is straightforward to find

$$\Phi'(v) = \mathsf{E}(A_1''(h + \chi\sqrt{a-v} + \chi_1\sqrt{v})A_1''(h + \chi\sqrt{a-v} + \chi_2\sqrt{v})) \ .$$

The Cauchy-Schwarz inequality (together with the fact that for $j = 1, 2$, $h + \chi\sqrt{a-v} + \chi_j\sqrt{v}$ has the same distribution as $h + \chi\sqrt{a}$) implies the magic fact:

$$\Phi'(v) \leq \mathsf{E}A_1''(h + \chi\sqrt{a})^2 = \Phi'(0). \tag{14.347}$$

Therefore by (14.346)

$$\beta'(u) = t\xi''(u)\Phi'(v) - 1 \leq t\xi''(q_1)\Phi'(0) - 1 = \beta'(q_1) \ . \tag{14.348}$$

So, we have proved (14.342), and we turn to the proof of (14.343). We assume now $u \leq 0$, so that

$$\begin{aligned} g_0^1 &= \sqrt{t}y_0^1 + \sqrt{1-t}z_0^1 \ ; \qquad & g_0^2 &= -\sqrt{t}y_0^1 + \sqrt{1-t}z_0^1 \\ g_1^1 &= \sqrt{t}y_1^1 \ ; & g_1^2 &= \sqrt{t}y_1^2 \ , \end{aligned}$$

where $y_0^1, z_0^1, y_1^1, y_0^2$ are independent Gaussian r.v.s and where

$$\mathsf{E}(y_0^1)^2 = \xi'(|u|) \ ; \ \mathsf{E}(y_1^1)^2 = \mathsf{E}(y_1^2)^2 = \xi'(q_1) - \xi'(|u|) \ ; \ \mathsf{E}(z_0^1)^2 = \xi'(q_1) \ .$$

Thus

$$\mathsf{E}(g_0^1 + g_1^1)^2 = \mathsf{E}(g_0^2 + g_1^2)^2 = \xi'(q_1) \ ,$$

while

$$\mathsf{E}((g_0^1+g_1^1)(g_0^2+g_1^2)) = (1-t)\xi'(q_1) - t\xi'(|u|) \; . \tag{14.349}$$

Consider $\chi_1 = g_0^1 + g_1^1$ and $\chi_2 = g_0^2 + g_1^2$ so that (14.341) implies

$$\beta(u) = \mathsf{E}A_1'(h+\chi_1)A_1'(h+\chi_2) - u \; . \tag{14.350}$$

Let us also consider χ_1' and χ_2' defined as χ_1 and χ_2 but for the value $u = 0$. Thus

$$\beta(0) = \mathsf{E}A_1'(h+\chi_1')A_1'(h+\chi_2') \; . \tag{14.351}$$

Consequently to prove (14.343) it suffices, using (14.350) and (14.351) to prove that

$$|\mathsf{E}A_1'(h+\chi_1)A_1'(h+\chi_2) - \mathsf{E}A_1'(h+\chi_1')A_1'(h+\chi_2')| \le |u| \; , \tag{14.352}$$

or even, using (14.344), that

$$|\mathsf{E}\chi_1\chi_2 - \mathsf{E}\chi_1'\chi_2'|\mathsf{E}A''(h+\chi\sqrt{a})^2 \le |u| \; . \tag{14.353}$$

Now (14.349) implies

$$\mathsf{E}\chi_1\chi_2 = (1-t)\xi'(q_1) - t\xi'(|u|) \; ; \; \mathsf{E}\chi_1'\chi_2' = (1-t)\xi'(q_1) \; ,$$

so that

$$|\mathsf{E}\chi_1\chi_2 - \mathsf{E}\chi_1'\chi_2'| = t\xi'(|u|) \; .$$

Since we assume in (14.101) that $\xi^{(3)}(x) \ge 0$ for $x > 0$, we have $\xi''(|u|) \le \xi''(q_1)$ for $|u| < q_1$, so that $\xi'(|u|) \le |u|\xi''(q_1)$, and therefore

$$|\mathsf{E}\chi_1\chi_2 - \mathsf{E}\chi_1'\chi_2'| \le t|u|\xi''(q_1) \; . \tag{14.354}$$

Also, $\beta'(q_1) = t\xi''(q_1)\Phi'(0) - 1$ by the last equality of (14.348), so that, since $M^*\varepsilon^{1/6} \le M_1\varepsilon^{1/6} \le 1 - t_0$ and therefore $\beta'(q_1) < 0$ by Lemma 14.9.2 we get

$$t\xi''(q_1)\Phi'(0) \le 1 \; .$$

Recalling (14.347) and (14.354) this implies (14.353). □

14.10 Main Estimate: Proof of Proposition 14.8.6

This proposition does not require the work of Section 14.7. We use again the bound (14.300) with $\lambda = 0$. We remind the reader that, for this bound to hold, we must be in the canonical situation as defined in the beginning of Section 14.8. That is, it is possible to remove from the list (14.289) some terms for which (Z_p^1, Z_p^2) is zero and find the sublist

$$\begin{aligned} 0\,,\; m_1^* = \frac{m_1}{2}\,,\; \dots\,,\; m_{r-1}^* = \frac{m_{r-1}}{2}\,, \\ m_r^* = m_r\,,\; \dots\,,\; m_k^* = m_k\,,\; m_{k+1}^* = 1\,, \end{aligned} \tag{14.355}$$

in such a way that if the p^{th}-term of this list occurs as $n_{a(p)}$ in the list (14.289), then $(Z^1_{a(p)}, Z^2_{a(p)}) = (\sqrt{1-t}z^1_p, \sqrt{1-t}z^2_p)$.

Our task is to find a choice of parameters so that the right-hand side (14.300) is $< 2\psi(t)$. Basically our strategy is to (cross our fingers and) make a simple choice ensuring

$$2\sum_{p<\tau} n_p(\theta(\rho_{p+1}) - \theta(\rho_p)) + \sum_{\tau \le p \le \kappa} n_p(\theta(\rho_{p+1}) - \theta(\rho_p))$$
$$= \sum_{p \le k+1} m_p(\theta(q_{p+1}) - \theta(q_p)) \, . \tag{14.356}$$

Our good luck will be that $Y_0 < 2X_0$, completing the proof.

A convenient way to achieve (14.356) is to arrange "that the same terms occur in the left and the right-hand sides". To make this possible we must first rework that right-hand side. Let us consider the unique integer $1 \le s \le k+1$ such that

$$q_{s-1} < |u| \le q_s \tag{14.357}$$

and the list

$$0 \, , \quad \frac{m_1}{2} \, , \, \dots \, , \frac{m_{s-1}}{2} \, , \quad m_{s-1} \, , \quad m_s \, , \quad \dots \, , \quad m_k \, , \quad 1 \, . \tag{14.358}$$

Let us denote these numbers by $(m'_j)_{0 \le j \le k+2}$, and let us define the numbers $(\rho'_j)_{0 \le j \le k+3}$ by

$$\rho'_0 = 0 \, , \;\; \rho'_1 = q_1 \, , \;\; \dots \, , \;\; \rho'_{s-1} = q_{s-1} \, , \;\; \rho'_s = |u| \, , \;\; \rho'_{s+1} = q_s \, , \dots \, , \;\; \rho'_{k+3} = 1 \, . \tag{14.359}$$

Thus

$$\sum_{p \le k+1} m_p(\theta(q_{p+1}) - \theta(q_p)) \tag{14.360}$$
$$= \sum_{1 \le j \le s-1} 2m'_j(\theta(\rho'_{j+1}) - \theta(\rho'_j)) + \sum_{s \le j \le k+2} m'_j(\theta(\rho'_{j+1}) - \theta(\rho'_j)) \, .$$

A convenient way to ensure (14.356) is now to arrange that the *non-zero* terms occurring in the left-hand side of (14.356) are exactly the terms occurring in the right-hand side of (14.360).

The list (14.289) we will construct is in a sense the smallest one with the previous properties; it will turn out then that $Y_0 < 2X_0$, proving the result. The list (14.355) contains $k+2$ elements, and the list (14.358) contains $k+3$ elements. We merge the two lists, repeating the elements that occur in both lists and keeping track from which list the element arise. Formally, we obtain a list $0 = n_0 = n_1 \le \dots \le n_{2k+3} = n_{2k+4} = n_\kappa = 1$ and two disjoint subsets I and J of $\{0, \dots, 2k+4\}$ such that

$$\text{the numbers } (n_p)_{p \in I} \text{ are the numbers (14.355),} \tag{14.361}$$

while

$$\text{the numbers } (n_p)_{p\in J} \text{ are the numbers (14.358).} \tag{14.362}$$

Let us recall that we denote by $(m^*_j)_{0\le j\le k+1}$ the numbers (14.355) and by $(m'_j)_{0\le j\le k+2}$ the numbers (14.358). Thus if $p \in I$, there is a unique $0 \le j \le k+1$ such that $n_p = m^*_j$, while if $p \in J$ there is a unique $0 \le j \le k+2$ such that $n_p = m'_j$. We define $Z^\ell_p = 0$ for $p \in J$; for $p \in I$, if $n_p = m^*_j$, we define $Z^\ell_p = \sqrt{1-t}z^\ell_j$. In this manner we ensure that we are in the canonical situation.

Let us now define the list $(\rho_p)_{0\le p\le 2k+5}$. Each member of this list is a member of the list (14.359) and conversely, but some elements of the list (14.358) occur several times in the list $(\rho_p)_{0\le p\le 2k+5}$. Starting with $\rho_0 = 0$, we construct the numbers (ρ_p) as follows. If $p \in I$, then $\rho_{p+1} = \rho_p$, while, if $p \in J$, ρ_{p+1} is the element of the list (14.359) just right to ρ_p. Equivalently, if $n_p = m'_j$ for some $j \le k+1$ and $m'_{j+1} = n_{p'}$ then $\rho_p = \rho'_j$ and $\rho_s = \rho'_{j+1}$ for $p < s \le p'$, whereas if $n_p = m'_{k+2} = 1$, then $\rho_s = 1$ for $s > p$. The list $(\rho_j)_{j\in J}$ is exactly the list (14.359). We denote by

$$\tau \text{ the unique integer in } J \text{ such that } \rho_\tau = |u| = \rho'_s \ , \tag{14.363}$$

so that

$$n_\tau = m_{s-1} = m'_s \ .$$

We define the pairs of r.v.s (y^1_p, y^2_p) as we must, by (14.156) and (14.157), i.e.

$$\mathsf{E}(y^1_p)^2 = \mathsf{E}(y^2_p)^2 = \xi'(\rho_{p+1}) - \xi'(\rho_p) \tag{14.156}$$

and

$$y^1_p = \eta y^2_p \ \text{ if } p < \tau\ ; \qquad y^1_p \text{ and } y^2_p \text{ independent if } p \ge \tau\ . \tag{14.157}$$

This construction ensures that $y^\ell_p = 0$ for $p \in I$ since then $\rho_{p+1} = \rho_p$. Since $Z^\ell_p = 0$ for $p \in J$, we see that for each p we have either $Z^\ell_p = 0$ or $y^\ell_p = 0$. This greatly simplifies matters.

From the construction we have

$$\begin{aligned}&\sum_{1\le p<\tau} 2n_p(\theta(\rho_{p+1}) - \theta(\rho_p)) + \sum_{\tau\le p\le 2k+4} n_p(\theta(\rho_{p+1}) - \theta(\rho_p))\\ &= \sum_{1\le j\le s-1} 2m'_j(\theta(\rho'_{j+1}) - \theta(\rho'_j)) + \sum_{s\le j\le k+2} m'_j(\theta(\rho'_{j+1}) - \theta(\rho'_j))\ ,\end{aligned}$$

since exactly the same terms occur on the left-hand side and on the right-hand side. Therefore (14.356) holds.

Consequently (14.312) implies that to prove Proposition 14.8.6 it suffices to show that

$$Y_0 < 2X_0\ . \tag{14.364}$$

The basic principle is to prove first that $Y_0 \le 2X_0$, through many uses of the Cauchy-Schwarz inequality and Hölder's inequality. The strict inequality is then obtained by showing that one cannot have equality in each of these inequalities.

Let us define $g_p^\ell = Z_p^\ell + \sqrt{t} y_p^\ell$, and let us consider the sets

$$F_I = \{p \in I \; ; \; n_p = m_j^* \; , \; r \le j \le k+1\} \tag{14.365}$$

$$F_J = \{p \in J \; ; \; n_p = m_j' \; , \; s \le j \le k+2\} \; . \tag{14.366}$$

These sets simply consist respectively of the last $k+2-r$ elements of I and of the last $k+3-s$ elements of J. We further set

$$F = F_1 \cup F_2 \; , \tag{14.367}$$

and we observe that $F_I = F \cap I$ and $F_J = F \cap J$. The set F will be used throughout the proof.

Next, we show that

$$p \in F \Rightarrow g_p^1 \text{ and } g_p^2 \text{ are independent.} \tag{14.368}$$

Indeed, if $p \in I$ with $n_p = m_j^*$ for $j \ge r$, then $g_p^\ell = Z_j^\ell$ for $\ell = 1, 2$ and these are independent. On the other hand, if $p \in J$ with $n_p = m_j'$ for $j \ge s$, then $p \ge \tau$ and since $p \in J$, $g_p^\ell = \sqrt{t} y_p^\ell$ and these are independent.

Let us observe that

$$p \notin F \; , \; p \in I \Rightarrow g_p^1 = g_p^2 \; . \tag{14.369}$$

This is simply because then $n_p = m_j^*$ with $j \le r-1$, and consequently $g_p^1 = g_p^2 = Z_j^1 = Z_j^2$. Also,

$$p \notin F \; , \; p \in J \Rightarrow g_p^1 = \eta g_p^2 \; . \tag{14.370}$$

This is simply because then $n_p = m_j'$ with $j \le s-1$, so that $p < \tau$ and then $g_p^1 = \sqrt{t} y_p^1 = \eta \sqrt{t} y_p^2 = \eta g_p^2$.

In particular, combining (14.369) and (14.370) yields

$$p \notin F \Rightarrow g_p^1 = \pm g_p^2 \; . \tag{14.371}$$

Let us review the definition of Y_0. Starting with

$$Y_{\kappa+1} = \log \mathrm{ch}\Big(h + \sum_{0 \le p \le \kappa} g_p^1\Big) + \log \mathrm{ch}\Big(h + \sum_{0 \le p \le \kappa} g_p^2\Big)$$

we define recursively

$$Y_p = \frac{1}{n_p} \log \mathsf{E}_p \exp n_p Y_{p+1} \; ,$$

where E_p is expectation in g_p^1 and g_p^2. Let us define

$$U_{\kappa+1}^\ell = \log \operatorname{ch}\Bigl(h + \sum_{0\le p\le\kappa} g_p^\ell\Bigr)$$

and, recursively

$$U_p^\ell = \frac{1}{n_p^*} \log \mathsf{E}_p \exp n_p^* U_{p+1}^\ell$$

where

$$n_p^* = 2n_p \text{ if } p \notin F \; ; \qquad n_p^* = n_p \text{ if } p \in F \; .$$

The notation n_p^* will also be used throughout the proof. Thus, we have $Y_{\kappa+1} = U_{\kappa+1}^1 + U_{\kappa+1}^2$, and we prove recursively that

$$Y_p \le U_p^1 + U_p^2 \; . \tag{14.372}$$

When $p \in F$, this is simply independence,

$$\mathsf{E}_p \exp n_p(U_{p+1}^1 + U_{p+1}^2) = \mathsf{E}_p \exp n_p U_{p+1}^1 \mathsf{E}_p \exp n_p U_{p+1}^2 \; , \tag{14.373}$$

while, if $p \notin F$, we use that $n_p^* = 2n_p$ and the Cauchy-Schwarz inequality to write

$$\begin{aligned}\mathsf{E}_p \exp n_p(U_{p+1}^1 + U_{p+1}^2) &\le (\mathsf{E}_p \exp n_p^* U_{p+1}^1)^{1/2} (\mathsf{E}_p \exp n_p^* U_{p+1}^2)^{1/2} \\ &= \exp n_p(U_p^1 + U_p^2) \; . \end{aligned} \tag{14.374}$$

Thus

$$Y_0 \le \mathsf{E}U_1^1 + \mathsf{E}U_1^2 = 2\mathsf{E}U_1^1 \; . \tag{14.375}$$

To analyze the sequence (U_p^ℓ), it is convenient to think in term of the operators $T_{m,v}$ of (14.193). Starting with the function $\log\operatorname{ch}$, one applies successively the operators $T_{n_p^*, a_p}$ starting with the largest value of p, where $a_p = \mathsf{E}(g_p^1)^2 = \mathsf{E}(g_p^2)^2$.

An important property of the sequence (n_p^*, a_p) is as follows.

Lemma 14.10.1. *Each number n_p^* is one of the numbers $0, m_1, \ldots, m_k, 1$. Moreover*

For each integer $0 \le b \le k$, the sum of the values of a_p for which $n_p^* = m_b$ is exactly $\xi'(q_{b+1}) - \xi'(q_b)$. (14.376)

Proof. By construction, when $p \in F$, n_p is one of the numbers m_k, and then $n_p^* = n_p$, while when $p \notin F$, n_p is one of the numbers $m_k/2$ and then $n_p^* = 2n_p$. Therefore each number n_p^* is one of the numbers $0, m_1, \ldots, m_k, 1$.

The proof of (14.376) is done by inspection. For example (the most delicate case) there are 3 values of p such that $n_p^* = m_{s-1}$. There is one value in J for which $n_p^* = 2n_p$, and $a_p = t(\xi'(|u|) - \xi'(q_{s-1}))$. This corresponds to the term $m_{s-1}/2$ of the list (14.358). There is another value in J, the value $p = \tau$, this time corresponding to the term m_{s-1} of that list, for which $a_p = t(\xi'(q_s) - \xi'(|u|))$. Finally there is one value from the list (14.355), and for this value $a_p = (1-t)(\xi'(q_s) - \xi'(q_{s-1}))$. □

Let us now sketch the structure of the proof. The sequence (n_p^*) need not be non-decreasing. We will show (through Hölder's inequality) that if we rearrange the operators $T_{n_p^*, a_p}$, applying first the one with the largest value of n_p^* etc., we can only increase the resulting function. Gathering the operators $T_{m,v}$ with the same value of m through (14.195) and (14.376) this amounts to using the sequence of operators $T_{m_p, \xi'(q_{p+1}) - \xi'(q_p)}$, the ones that define X_0. In this manner we prove that $\mathsf{E} U_0^1 \le X_0$, and combining with (14.375) that $Y_0 \le 2X_0$. Analysis of the cases of equality shows that there can be equality only when $s = 1$ and $-q_1 \le u \le 0$, a case that has been covered by another method in Proposition 14.8.3.

Let us now complete the details, and first look at the cases of equality (14.374). we may think of U_{p+1}^ℓ as

$$U_{p+1}^\ell = A_{p+1}(h + g_0^1 + \cdots + g_p^\ell) \, , \tag{14.377}$$

where the function A_{p+1} is obtained from the function $\log \mathrm{ch}$ by application of operations of the type $T_{m,v}$, so is strictly convex by (14.198). Also $A_{p+1}(x) = A_{p+1}(-x)$. To lighten notation we will write

$$A = A_{p+1} \, .$$

Lemma 14.10.2. *When $Y_0 = \mathsf{E} U_1^1 + \mathsf{E} U_1^2$, then for each p we have*

$$Y_p = U_p^1 + U_p^2 \tag{14.378}$$

and when $p \notin F$ we have

$$\mathsf{E}_p \exp n_p (U_{p+1}^1 + U_{p+1}^2) = (\mathsf{E}_p \exp n_p^* U_{p+1}^1)^{1/2} (\mathsf{E}_p \exp n_p^* U_{p+1}^2)^{1/2} \, . \tag{14.379}$$

The equalities are everywhere.

Proof. It goes by (increasing) induction in p. Assuming (14.378) for p, we prove it for $p+1$. Suppose first that $p \in F$. We assume $n_p \ne 0$, leaving the easier case $n_p = 0$ to the reader. From (14.372), we have

$$Y_{p+1} \le U_{p+1}^1 + U_{p+1}^2 \, ,$$

and (14.373) implies

$$\begin{aligned} Y_p = \frac{1}{n_p} \log \mathsf{E}_p \exp n_p Y_{p+1} &\le \frac{1}{n_p} \log \mathsf{E}_p \exp n_p (U_{p+1}^1 + U_{p+1}^2) \\ &= U_p^1 + U_p^2 = Y_p \, , \end{aligned}$$

so that

$$\mathsf{E}_p \exp n_p Y_{p+1} = \mathsf{E}_p \exp n_p (U^1_{p+1} + U^2_{p+1}) \, .$$

Since $Y_{p+1} \le U^1_{p+1} + U^2_{p+1}$, and hence $\exp n_p Y_{p+1} \le \exp n_p (U^1_{p+1} + U^2_{p+1})$, this proves that $\exp n_p Y_{p+1} = \exp n_p (U^1_{p+1} + U^2_{p+1})$ "almost surely for E_p"; but since both sides are continuous functions of $(g^\ell_k)_{k \le p}$, there is true equality, so that $Y_{p+1} = U^1_{p+1} + U^2_{p+1}$.

Assume that now $p \notin F$. Then, using (14.372) in the first inequality and the first line of (14.374) in the second one,

$$\begin{aligned} Y_p &= \frac{1}{n_p} \log \mathsf{E}_p \exp n_p Y_{p+1} \\ &\le \frac{1}{n_p} \log \mathsf{E}_p \exp n_p (U^1_{p+1} + U^2_{p+1}) \\ &\le \frac{1}{n_p} \log (\mathsf{E}_p \exp n^*_p U^1_{p+1})^{1/2} + \frac{1}{n_p} \log (\mathsf{E}_p \exp n^*_p U^2_{p+1})^{1/2} \\ &= U^1_p + U^2_p = Y_p \, , \end{aligned}$$

so the second inequality must be an equality, which proves (14.379); and we proceed similarly as before to obtain $Y_{p+1} = U^1_{p+1} + U^2_{p+1}$, which completes the proof of (14.378). □

Let us make a simple observation. If a sum $w_0 + \cdots + w_p$ of independent r.v.s is a.s. constant, each of these r.v.s must be a.s. constant. Since the pairs (g^1_p, g^2_p) are independent as p varies, when we have a relation such as

$$g^1_0 + \cdots + g^1_p = g^2_0 + \cdots + g^2_p \quad \text{a.s}$$

then $\sum_{p' \le p} (g^1_{p'} - g^2_{p'}) = 0$ a.s., so that for each $p' \le p$ we have $g^1_{p'} = g^2_{p'}$ a.s. (To lighten notation, we will not indicate any more that the equalities between r.v.s g^ℓ_p are always understood a.s.) When we have a relation such as

$$h + g^1_0 + \cdots + g^1_p = -h - g^2_0 - \cdots - g^2_p$$

then we must have $g^1_{p'} = -g^2_{p'}$ for each $p' \le p$ (and this case is possible only when h=0).

Lemma 14.10.3. *If $n_p > 0$, $g^1_p \ne 0$, $p \notin F$ and (14.379) holds we have $g^1_p = g^2_p$ and*

$$\forall p' \le p \, , \; g^1_{p'} = g^2_{p'} \, . \tag{14.380}$$

Proof. When (14.379) holds, we have (since $n^* = 2n_p$ because $p \notin F$)

$$\begin{aligned} &\mathsf{E} \exp n_p (A(x_1 + g^1_p) + A(x_2 + g^2_p)) \\ &= (\mathsf{E} \exp 2n_p A(x_1 + g^1_p))^{1/2} (\mathsf{E} \exp 2n_p A(x_2 + g^2_p))^{1/2} \, , \end{aligned} \tag{14.381}$$

for almost every (and hence for every) value of $x_1 = h + g_0^1 + \cdots + g_{p-1}^1$ and $x_2 = h + g_0^2 + \cdots + g_{p-1}^2$. When (14.381) holds, then the two random variables $\exp n_p A(x_1 + g_p^1)$ and $\exp n_p A(x_2 + g_p^2)$ must be proportional, that is, since $n_p \neq 0$, the r.v.s $A(x_1 + g_p^1) - A(x_2 + g_p^2)$ must be constant. We remember from (14.371) that there are 2 possible cases, $g_p^2 = \pm g_p^1$. If $g_p^2 = g_p^1 \neq 0$, then the function $y \mapsto A(x_1 + y) - A(x_2 + y)$ must be constant, so that its derivative $A'(x_1 + y) - A'(x_2 + y)$ must be 0. But since A is strictly convex A' is strictly increasing, so we must have $x_1 = x_2$, and (14.380) is proved in that case. If $g_p^2 = -g_p^1 \neq 0$, then the function

$$y \mapsto A(x_1 + y) - A(x_2 - y) = A(x_1 + y) - A(y - x_2)$$

must be constant, so since A is strictly convex we must have $x_1 = -x_2$ i.e.

$$h + g_0^1 + \cdots + g_{p-1}^1 = -h - g_0^2 - \cdots - g_{p-1}^2 \,.$$

As already explained this implies that $g_{p'}^1 = -g_{p'}^2$ for each $p' \leq p$. This however is impossible because one of the pairs (g_n^1, g_n^2) for $n \leq p-1$ is the pair $(Z_0^1, Z_0^2) = (\sqrt{1-t}z_0^1, \sqrt{1-t}z_0^2)$, and in that case we have $g_n^1 = g_n^2 \neq 0$. □

Corollary 14.10.4. *In the equality case $Y_0 = 2X_0$ we have*

$$n_p > 0\,, \quad g_p^1 \neq 0\,, \quad p \notin F \quad \Rightarrow \quad \forall p' \leq p\,, \quad g_{p'}^1 = g_{p'}^2\,. \tag{14.382}$$

Proof. Combine Lemmas 14.10.2 and 14.10.3. □

Lemma 14.10.5. *If $a, a' \geq 0$ and $m \geq m'$, for any function A and any x we have*

$$T_{m,a} \circ T_{m',a'}(A)(x) \leq T_{m',a'} \circ T_{m,a}(A)(x)\,. \tag{14.383}$$

If A is strictly convex, if $a, a' > 0$ and $m > m'$, we cannot have equality in (14.383). (14.384)

To express this in words, when we transpose two operators $T_{m,a}$ and $T_{m',a'}$ to apply first the operator with the largest value of m we can only increase the result of the application of these operators. The increase is strict when we operate on strictly convex functions and when the operators are not the identity.

Proof. Consider independent standard Gaussian r.v.s g and g', and denote by E and E' expectation in g and g' respectively. Then

$$T_{m,a} \circ T_{m',a'}(A)(x) = \frac{1}{m} \log \mathsf{E} \exp \frac{m}{m'} \log \mathsf{E}' \exp m' A(x + g\sqrt{a} + g'\sqrt{a'})\,,$$

$$T_{m',a'} \circ T_{m,a}(A)(x) = \frac{1}{m'} \log \mathsf{E}' \exp \frac{m'}{m} \log \mathsf{E} \exp m A(x + g\sqrt{a} + g'\sqrt{a'})\,.$$

Let $\alpha = m/m' \geq 1$ and $X = \exp m' A(x + g\sqrt{a} + g'\sqrt{a'})$, so

$$T_{m,a} \circ T_{m',a'}(A)(x) = \frac{1}{m} \log \mathsf{E} \exp \alpha \log \mathsf{E}' X = \frac{1}{m} \log \mathsf{E}(\mathsf{E}' X)^\alpha ,$$
$$T_{m',a'} \circ T_{m,a}(A)(x) = \frac{1}{m'} \log \mathsf{E}' \exp \frac{1}{\alpha} \log \mathsf{E} X^\alpha = \frac{1}{m'} \log \mathsf{E}' (\mathsf{E} X^\alpha)^{1/\alpha}$$

and the required inequality is

$$(\mathsf{E}(\mathsf{E}'X)^\alpha)^{1/\alpha} \le \mathsf{E}'(\mathsf{E}X^\alpha)^{1/\alpha}$$

or, equivalently,

$$\|\mathsf{E}'X\|_\alpha \le \mathsf{E}'\|X\|_\alpha , \tag{14.385}$$

if one sets $\|Y\|_\alpha = (\mathsf{E}Y^\alpha)^{1/\alpha}$. This holds true by convexity. Let us now analyze the case where there is equality in (14.385). First we find Y, depending on g only, with

$$\mathsf{E}(Y\mathsf{E}'X) = \|\mathsf{E}'X\|_\alpha \,; \quad \|Y\|_{\alpha'} = 1$$

where $1/\alpha + 1/\alpha' = 1$. Then Hölder's inequality implies

$$\mathsf{E}YX \le \|Y\|_{\alpha'}\|X\|_\alpha = \|X\|_\alpha , \tag{14.386}$$

and thus, using in the last equality that we assume equality in (14.385),

$$\|\mathsf{E}'X\|_\alpha = \mathsf{E}(Y\mathsf{E}'X) = \mathsf{E}'\mathsf{E}YX \le \mathsf{E}'\|X\|_\alpha = \|\mathsf{E}'X\|_\alpha ,$$

so that we must have $\mathsf{E}YX = \|X\|_\alpha$ for almost all g', and hence for all g'. That is, we are in the equality case of Hölder's inequality (14.386).This is the case only if, as a function of g, X is proportional to Y, i.e. $X = \lambda(g')Y$. Since Y depends on g only, taking logarithms and recalling the value of X, we see that for any given value of x we must have a decomposition

$$A(x + g\sqrt{a} + g'\sqrt{a'}) = B_1(g) + B_2(g') .$$

This implies that the difference

$$A(x + g\sqrt{a} + g'\sqrt{a'}) - A(x + g\sqrt{a})$$

does not depend on g, and this is impossible since A is strictly convex. □

Lemma 14.10.6. *We have* $\mathsf{E}U_1^1 \le X_0$*, and there can be equality only if the following occurs*

$$0 \le p < p' \le 2k+4 , \quad g^1_{p'}, g^1_p \ne 0 \quad \Rightarrow \quad n^*_p \le n^*_{p'} . \tag{14.387}$$

Proof. Let us denote by T_p the operator $T_{n^*_p,a_p}$. Starting with the function $A_{2k+4} = \log \mathrm{ch} x$, we apply recursively the operators $T_{2k+4}, \dots, T_1$ in that order to obtain a function A_1 such that $U_1^1 = A_1(h + g_0^1)$. Let us now consider a permutation π of $\{1, \dots, 2k+4\}$ with the property that the sequence $(n^*_{\pi(p)})$ is non-decreasing. As already explained, (14.376) and (14.195) imply that if we apply to the function $B_{2k+4} = \log \mathrm{ch}$ the operators $T_{\pi(2k+4)}, \ldots, T_{\pi(1)}$ we obtain a function $B_1(x)$ such that $X_0 = \mathsf{E} B_1(h + g_0^1)$.

We can obtain the permutation π as the composition of a sequence of transpositions of two consecutive elements, each transposition tending, as in (14.383), "to apply first the operator with the largest value of n_p^*." If we use (14.383) for each such transposition we see that $B_1(x) \le A_1(x)$ for each x, and this proves that $\mathsf{E} U_1^1 \le X_0$. There can be equality only if we are in the equality case of (14.383) each time we perform a transposition. The operators are applied only to strictly convex functions (which obtained by applying operators of the type $T_{m,a}$ to the function $\log \mathrm{ch}$) so we can have equality only if we never have to perform a "non-trivial" transposition, that is, if the original list does not contain integers p, p' with $p < p'$, $g_p^1 \ne 0$, $g_{p'}^1 \ne 0$ and $n_p^* > n_{p'}^*$. □

In both (14.382) and (14.387) occurs the condition $g_p^1 \ne 0$, and figuring out when this is the case helps to use these conditions.

Lemma 14.10.7. *We have $g_p^1 \ne 0$ unless $p = \tau$ and $|u| = q_s$.*

Proof. Indeed, if $p \in I$, then $n_p = m_j^*$ with $j \le k+1$ and $g_p^1 = Z_p^1 = \sqrt{1-t} z_j^1$ with $\mathsf{E}(z_j^1)^2 = \xi'(q_{j+1}) - \xi'(q_j) > 0$. When $p \in J$, then $n_p = m'_j$ where $0 \le j \le k+2$ and $g_p^1 = \sqrt{t} y_p^1$, with $\mathsf{E}(y_p^1)^2 = \xi'(\rho'_{j+1}) - \xi'(\rho'_j)$. This can be 0 only if $\rho'_j = \rho'_{j+1}$. By (14.287), (14.357) and (14.359) this can happen only when $j = s$ i.e. $p = \tau$, and then only if $|u| = q_s$. □

Before we prove Proposition 14.8.6 we make two observations that will be used many times during its proof. We recall that $r \ge 1$ and $s \ge 1$.

Lemma 14.10.8. *There exists $p(J) \in J$ with*

$$n_{p(J)} = m_{s-1}/2 \ , \ p(J) \notin F \ , \ g_{p(J)}^1 = \eta g_{p(J)}^2 \ne 0 \ . \tag{14.388}$$

There exists $p(I) \in I$ such that

$$n_{p(I)} = m_{r-1}/2 \ , \ p(I) \notin F \ , \ g_{p(I)}^1 = g_{p(I)}^2 \ne 0 \ . \tag{14.389}$$

Proof. The existence of $p(J)$ follows from (14.358). Since $n_{p(J)} = m_{s-1}/2 = m'_{s-1}$, the definition (14.366) of F_J shows that $p(J) \notin F_J = F \cap F \cap J$ so that $p(J) \notin F$. Then (14.369) implies $g_{p(J)}^1 = \eta g_{p(J)}^2$ and these variables are not zero by Lemma 14.10.7 since $n_{p(J)} \ne \tau$ by definition of τ. The other case is similar. □

Proof of Proposition 14.8.6. We have seen that (14.308) holds when $Y_0 < 2X_0$. Let us investigate the possibility of the equality case $Y_0 = 2X_0$, with the aim of showing that it cannot occur when $t > 0, u < q_{r-1}$ or $u > q_{r+1}$ unless $r = 1$ and $|u| \le q_1$.

The method of proof is to combine conditions (14.382) and (14.387) to obtain our conclusions.

First, we prove that we cannot have $u < -q_1$. In that case we have $\eta = -1$ and $|u| > q_1$ so (14.357) shows that $s \geq 2$, so $s-1 \geq 1$ and therefore $m_{s-1} > 0$ and $n_{p(J)} = m_{s-1}/2 > 0$. Using (14.388) we observe that the choice $p = p' = p(J)$ contradicts (14.382). Thus we cannot have $u < -q_1$.

Next, we consider the case where $-q_1 \leq u < 0$, so $\eta = -1$, $0 < |u| \leq q_1$ and $s = 1$. We prove that then $r = 1$. Assume for contradiction that $r \geq 2$, so that then $m_{r-1} > 0$. Consider the smallest element p' of J, so that $n_{p'} = 0 < m_{r-1}/2 = n_{p(I)}$ and thus $p' < p(I)$. Also, $g^2_{p'} = -g^1_{p'}$ by (14.370) and $g^1_{p'} \neq 0$ since $p' \neq \tau$. The choice of p' and $p = p(I)$ then contradicts (14.382). We have ruled out this case.

In the remainder of the proof we assume $u \geq 0$. Then we always have that $g^1_p = g^2_p$ if $p \notin F$. If $p \in F$ then g^1_p and g^2_p are independent.

Let us recall the element τ of J such that $n_\tau = m_{s-1}$, so that $\tau \in F_J \subset F$ and consequently $n^*_\tau = n_\tau$.

Let us examine the case $u < q_{r-1}$, so since $u > q_{s-1}$ by (14.357) we have $s \leq r-1$. Since $p = p(I) \notin F$ by (14.389) we have $n^*_p = 2n_p = m_{r-1}$. Therefore

$$n^*_\tau = n_\tau = m_{s-1} < m_s \leq m_{r-1} = n^*_p \,.$$

Let us assume first that $u < q_s$, so that $g^1_\tau \neq 0$. Then (14.387) (used for $p' = \tau$) shows that $\tau \leq p$. But since g^1_τ and g^2_τ are independent and $g^1_\tau \neq 0$ they are not equal, and (14.382) cannot hold. Next, let us assume that $u = q_s$. Since we assume $u < q_{r-1}$ then $s < r-1$. We consider $p' \in J$ with $n_{p'} = m_s < m_{r-1}$, so again $g^1_{p'} \neq 0$ and

$$n^*_{p'} = n_{p'} = m_s < m_{r-1} = n^*_p \,,$$

and we argue as before using p' instead of τ.

Finally we examine the case where $u > q_{r+1}$, so since $u \leq q_s$ by (14.357) we have $s > r+1$ and thus $s-1 \geq r+1$. We consider $p = p(J)$, so that $p \notin F$ and consequently $n^*_p = 2n_p = m_{s-1}$. Consider $p' \in I$ with $n_{p'} = m_r$. Then

$$n^*_p = m_{s-1} \geq m_{r+1} > m_r = n_{p'} = n^*_{p'} \,.$$

Since g^1_p and $g^1_{p'}$ are $\neq 0$ (because $p \neq \tau$ and $p' \neq \tau$), (14.387) implies $p' < p$. Since $p \notin F$ this contradicts (14.382) because $g^1_{p'}$ and $g^2_{p'}$ are independent and not 0.

In conclusion (14.308) holds unless $Y_0 = 2X_0$. Provided that $t > 0$, $u < q_{r-1}$ or $u > q_{r+1}$, we have shown that then we must have $-q_1 \leq u < 0$ and $r = 1$. But in that case we have proved (14.308) in Proposition 14.8.3. □

Proof of Proposition 14.8.7. We use again the bound (14.300) when $\kappa = k+1$ and $n_p = m^*_p$, where m^*_p is the sequence (14.291). In that case $g^\ell_p = Z^\ell_p = z^\ell_p$. The bound (14.300) then reads $\Psi^*(0,u) \leq 2\log 2 + Y_0 - \lambda u$. The case $\lambda = 0$ witnesses the inequality $\Psi^*(0,u) \leq 2\psi(0)$. To improve on this, it suffices to prove that

$$\left.\frac{\partial Y_0}{\partial \lambda}\right|_{\lambda=0} \neq u \tag{14.390}$$

and this is true because (14.177) used for $\tau = r$ and (14.222) prove that the left-hand side of (14.390) is q_r. □

14.11 Parisi Measures

Let us consider a function ξ that satisfies (14.101) (although some of the results we will prove do not use these conditions). Given a sequence $\mathbf{m}$

$$m_0 = 0 < m_1 < \ldots < m_k < m_{k+1} = 1 \ , \tag{14.391}$$

and a sequence $\mathbf{q}$

$$q_0 = 0 \le q_1 \le \ldots \le q_{k+1} \le 1 = q_{k+2} \ , \tag{14.392}$$

consider the probability measure μ that is sum of the point masses $m_p - m_{p-1}$ put at the points q_p, $1 \le p \le k+1$. As we already explained, the quantity $\mathcal{P}_k(\mathbf{m}, \mathbf{q})$ of (14.89) depends only on the function $x(q)$ such that $x(q) = m_p$ for $q_p \le q < q_{p+1}$, so that it depends only on μ. Thinking of h as fixed once and for all, we denote this quantity by $\mathcal{P}(\xi, \mu)$.

We shall prove that this definition of $\mathcal{P}(\xi, \mu)$ can be "smoothly" extended to any probability measure μ on $[0,1]$, and that, quite remarkably, the measure μ that minimizes $\mathcal{P}(\xi, \mu)$ is (in some sense) "the law of the overlaps $R_{1,2}$".

The results of this section are simple consequences of a technical lemma. We fix the sequence (14.391) and numbers $w_1 \le w_2 \le \ldots \le w_{k+2}$, and we define the random function

$$F_{k+2}(w_1, \ldots, w_{k+2}) = \log \operatorname{ch}\left(h + \sum_{0 \le p \le k+1} z_p \sqrt{w_{p+1} - w_p}\right) ,$$

where $w_0 = 0$, and where the z_p are independent standard Gaussian r.v.s. For $1 \le p \le k+1$ we then define recursively the random functions

$$F_p(w_1, \ldots, w_{k+2}) = \frac{1}{m_p} \log \mathsf{E}_p \exp m_p F_{p+1}(w_1, \ldots, w_{k+2}) ,$$

where E_p denotes expectation in the r.v.s $(z_\ell)_{\ell \ge p}$. We then define

$$F_0(w_1, \ldots, w_{k+2}) = \mathsf{E} F_1(w_1, \ldots, w_{k+2}) \ . \tag{14.393}$$

Lemma 14.11.1. *We have*

$$\frac{\partial F_0}{\partial w_{k+2}} = \frac{1}{2} \ , \tag{14.394}$$

and for each $1 \le r \le k+1$ *we have*

$$-\frac{1}{2}(m_r - m_{r-1}) \le \frac{\partial F_0}{\partial w_r} \le 0 \ . \tag{14.395}$$

Moreover

$$\sum_{1 \le p,r \le k+1} \left| \frac{\partial^2 F_0}{\partial w_p \partial w_r} \right| \le L \ . \tag{14.396}$$

The notable fact is that the bound (14.396) holds independently of the value of k.

Proof. We observe that, since $m_{k+1} = 1$,

$$F_{k+1}(w_1, \ldots, w_{k+2}) = \frac{1}{2}(w_{k+2} - w_{k+1}) + \log \mathrm{ch}\left(h + \sum_{0 \le p \le k} z_p \sqrt{w_{p+1} - w_p} \right) ,$$

so that (14.394) is obvious.

For $\alpha = (j_1, \ldots, j_k) \in \mathbb{N}^{*k}$ and $0 \le p \le k$ consider independent standard normal r.v.s $z_{p,\alpha} = z_{p,j_1,\ldots,j_p}$. Let us set

$$\zeta_\alpha = h + \sum_{0 \le p \le k} z_{p,\alpha} \sqrt{w_{p+1} - w_p} \ ,$$

and use Theorem 14.2.1 to see that when the weights v_α form a Poisson-Dirichlet cascade associated with the sequence $m_1, \ldots, m_k$, we have

$$F_0(w_1, \ldots, w_{p+2}) = \frac{1}{2}(w_{k+2} - w_{k+1}) + \mathsf{E} \log \sum_\alpha v_\alpha \mathrm{ch} \zeta_\alpha \ . \tag{14.397}$$

To prove (14.395), let us first assume that $1 \le r \le k$. Let us define $g_\alpha = \sum_{0 \le p \le k} z_{p,\alpha} \sqrt{w_{p+1} - w_p}$, so that $\zeta_\alpha = h + g_\alpha$ and let us further define

$$g'_\alpha := \frac{\partial g_\alpha}{\partial w_r} = -\frac{z_{r,\alpha}}{2\sqrt{w_{r+1} - w_r}} + \frac{z_{r-1,\alpha}}{2\sqrt{w_r - w_{r-1}}} \ ,$$

so that from (14.397) and the definition of ζ_α we get

$$\frac{\partial F_0}{\partial w_r} = \mathsf{E} \frac{\sum_\alpha v_\alpha g'_\alpha \mathrm{sh}(h + g_\alpha)}{\sum_\alpha v_\alpha \mathrm{ch}(h + g_\alpha)} \ . \tag{14.398}$$

Since $\mathsf{E} z_{p,\alpha} z_{r,\gamma} = 0$ unless $r = p$ we have

$$\mathsf{E}g'_\alpha g_\gamma = \frac{1}{2}(-\mathsf{E}z_{r,\alpha}z_{r,\gamma} + \mathsf{E}z_{r-1,\alpha}z_{r-1,\gamma}) ,$$

and since $\mathsf{E}z_{r,\alpha}z_{r,\gamma} = 0$ if $(\alpha,\gamma) \le r$ and $= 1$ if $(\alpha,\gamma) > r$ we have

$$\mathsf{E}g'_\alpha g_\gamma = \frac{1}{2} \text{ if } (\alpha,\gamma) = r \; ; \quad \mathsf{E}g'_\alpha g_\gamma = 0 \text{ otherwise} \tag{14.399}$$

and integration by parts in (14.398) yields

$$\frac{\partial F_0}{\partial w_r} = -\frac{1}{2}\mathsf{E}\frac{\sum_{(\alpha,\gamma)=r} v_\alpha v_\gamma \mathrm{sh}\zeta_\alpha \mathrm{sh}\zeta_\gamma}{\left(\sum_\alpha v_\alpha \mathrm{ch}\zeta_\alpha\right)^2} . \tag{14.400}$$

When $r = k+1$, we have

$$\frac{\partial F_0}{\partial w_r} = -\frac{1}{2} + \mathsf{E}\frac{\sum_\alpha v_\alpha g'_\alpha \mathrm{sh}(h+g_\alpha)}{\sum_\alpha v_\alpha \mathrm{ch}(h+g_\alpha)} , \tag{14.401}$$

and

$$g'_\alpha = \frac{\partial g_\alpha}{\partial w_r} = \frac{z_{k,\alpha}}{2\sqrt{w_r - w_{r-1}}} .$$

Then (14.399) still holds true, and since $(\alpha,\alpha) = r+1$ and therefore $\mathsf{E}g_\alpha g'_\alpha = 1/2$ we see as previously that (14.400) still holds. Taking $U(\alpha) = \mathrm{th}\zeta_\alpha$ and $F(\alpha) = \log \mathrm{ch}\zeta_\alpha$ this yields, using the notation (14.17)

$$\mathsf{E}\frac{\sum_{(\alpha,\gamma)=r} v_\alpha v_\gamma \mathrm{sh}\zeta_\alpha \mathrm{sh}\zeta_\gamma}{\left(\sum_\alpha v_\alpha \mathrm{ch}\zeta_\alpha\right)^2} = \mathsf{E}\langle \mathbf{1}_{\{(\alpha,\gamma)=r\}} U(\alpha)U(\gamma)\rangle ,$$

and (14.400) and (14.37) imply (14.395) (since $|U_{k+1}| \le 1$).

To prove (14.396) we differentiate (14.400) again. Assuming $p \ne r$, we find, after integration by parts,

$$\begin{aligned}\frac{\partial^2 F_0}{\partial w_p \partial w_r} &= 2\mathsf{E}\frac{\sum_{(\alpha,\gamma)=r,\,(\alpha,\delta)=p} v_\alpha v_\gamma v_\delta \mathrm{ch}\zeta_\alpha \mathrm{sh}\zeta_\gamma \mathrm{sh}\zeta_\delta}{\left(\sum_\alpha v_\alpha \mathrm{ch}\zeta_\alpha\right)^3} \\ &\quad - \frac{3}{2}\mathsf{E}\frac{\left(\sum_{(\alpha,\gamma)=r} v_\alpha v_\gamma \mathrm{sh}\zeta_\alpha \mathrm{sh}\zeta_\gamma\right)\left(\sum_{(\alpha,\gamma)=p} v_\alpha v_\gamma \mathrm{sh}\zeta_\alpha \mathrm{sh}\zeta_\gamma\right)}{\left(\sum_\alpha v_\alpha \mathrm{ch}\zeta_\alpha\right)^4}\end{aligned}$$

and, defining $c_\alpha = v_\alpha \mathrm{ch}\zeta_\alpha$ (so that $v_\alpha|\mathrm{sh}\zeta_\alpha| \le c_\alpha$) we get

$$\begin{aligned}\left|\frac{\partial^2 F_0}{\partial w_p \partial w_r}\right| &\le 2\mathsf{E}\frac{\sum_{(\alpha,\gamma)=r,\,(\alpha,\delta)=p} c_\alpha c_\gamma c_\delta}{\left(\sum_\alpha c_\alpha\right)^3} \\ &\quad + \frac{3}{2}\frac{\left(\sum_{(\alpha,\gamma)=r} c_\alpha c_\gamma\right)\left(\sum_{(\alpha,\gamma)=p} c_\alpha c_\gamma\right)}{\left(\sum_\alpha c_\alpha\right)^4} .\end{aligned}$$

We claim that the sum over all the values of p and r with $r \ne p$ is $\le 7/2$. For example, we have

$$\sum_{r,p}\Bigl(\sum_{(\alpha,\gamma)=r} c_\alpha c_\gamma\Bigr)\Bigl(\sum_{(\alpha',\gamma')=p} c_{\alpha'} c_{\gamma'}\Bigr) \le \Bigl(\sum_{\alpha,\gamma} c_\alpha c_\gamma\Bigr)\Bigl(\sum_{\alpha',\gamma'} c_{\alpha'} c_{\gamma'}\Bigr) = \Bigl(\sum_\alpha c_\alpha\Bigr)^4 .$$

We then show by a similar calculation that

$$\left|\frac{\partial^2 F_0}{\partial w_p^2}\right| \le L(m_p - m_{p-1}) . \qquad \square$$

Given the sequences (14.391) and (14.392), consider the function $x(q)$ such that $x(q) = m_p$ for $q_p \le q < q_{p+1}$, and $x(1) = 1$. We consider two other sequences $\mathbf{m}^\sim, \mathbf{q}^\sim$ and the corresponding function $x^\sim(q)$.

Theorem 14.11.2. (F. Guerra) *We have*

$$|\mathcal{P}_k(\mathbf{m},\mathbf{q}) - \mathcal{P}_{k'}(\mathbf{m}^\sim,\mathbf{q}^\sim)| \le \frac{1}{2}\xi''(1)\int_0^1 |x(q) - x^\sim(q)|\mathrm{d}q . \tag{14.402}$$

Proof. We have already noticed that in the list $\mathbf{m}$ we can insert values "for which nothing happens" without changing the value of $\mathcal{P}_k(\mathbf{m},\mathbf{q})$. More specifically, if we insert a new element $m_p < m < m_{p+1}$ in the list $\mathbf{m}$, we simply repeat the element q_{p+1} in the list $\mathbf{q}$. Thus we can insert in the sequence $\mathbf{m}$ the elements of $\mathbf{m}^\sim$ that it does not contain, and in the sequence $\mathbf{m}^\sim$ the elements of $\mathbf{m}$ it does not contain. Therefore we can assume that $\mathbf{m} = \mathbf{m}^\sim$. Recalling the function F_0 of (14.393) and the definition (14.89) of $\mathcal{P}_k(\mathbf{m},\mathbf{q})$ we observe that

$$\begin{aligned}\mathcal{P}_k(\mathbf{m},\mathbf{q}) &= \log 2 + F_0(\xi'(q_1),\dots,\xi'(q_{k+1}),\xi'(1)) \\ &\quad + \frac{1}{2}\sum_{1\le p\le k+1}\theta(q_p)(m_p - m_{p-1}) - \frac{1}{2}\theta(1) .\end{aligned} \tag{14.403}$$

For $1 \le p \le k+1$, let $q_{p,t} = tq_p + (1-t)q_p^\sim$, and denote by $\mathbf{q}_t$ the corresponding list. Let

$$\begin{aligned}\varphi(t) = \mathcal{P}_k(\mathbf{m},\mathbf{q}_t) &= F_0(\xi'(q_{1,t}),\dots,\xi'(q_{k+1,t}),\xi'(1)) \\ &\quad + \frac{1}{2}\sum_{1\le p\le k+1}\theta(q_{p,t})(m_p - m_{p-1}) - \frac{1}{2}\theta(1) .\end{aligned}$$

Differentiating and recalling that $\theta'(q) = q\xi''(q)$ we obtain

$$\varphi'(t) = \sum_{1\le p\le k+1}\xi''(q_{p,t})(q_p - q_p^\sim)\left(a_{p,t} + \frac{1}{2}q_{p,t}(m_p - m_{p-1})\right)$$

where

$$a_{p,t} = \frac{\partial F_0}{\partial w_p}\bigg|_{w_1=\xi'(q_{1,t}),\ldots,w_{k+1}=\xi'(q_{k+1,t})} .$$

Now, by (14.395), we have

$$-\frac{1}{2}(m_p - m_{p-1}) \le a_{p,t} \le 0 ,$$

and since $\xi''(q_{p,t}) \le \xi''(1)$ (because we assume $\xi^{(3)}(x) \ge 0$ for $x > 0$) it follows that

$$|\varphi'(t)| \le \frac{\xi''(1)}{2} \sum_{1\le p\le k+1} |q_p - q_p^\sim|(m_p - m_{p-1}) .$$

So it suffices to prove that

$$\sum_{1\le p\le k+1} |q_p - q_p^\sim|(m_p - m_{p-1}) \le \int_0^1 |x(q) - x^\sim(q)| \mathrm{d}q .$$

The right-hand side is the area between the graphs of the functions x and $x^\sim$. For each $1 \le p \le k+1$ the rectangle

$$\{(x,y) \; ; \; \min(q_p, q_p^\sim) < x < \max(q_p, q_p^\sim) \; ; \; m_{p-1} < y < m_p\}$$

is entirely contained between these graphs, its area is $(m_p - m_{p-1})|q_p - q_p^\sim|$ and these rectangles are disjoint. □

When the probability measure μ has all its mass concentrated at a finite number of points $q_1, \ldots, q_{k+1}$, we have defined $\mathcal{P}(\xi,\mu)$ as $\mathcal{P}_k(\mathbf{m},\mathbf{q})$, where for $1 \le p \le k+1$, $m_p = \mu([0,q_p])$. The function $x(q)$ is given in that case by $x(q) = \mu([0,q])$ for $0 \le q \le 1$. It follows from (14.403) that we can define $\mathcal{P}(\xi,\mu)$ for any probability measure on $[0,1]$ by

$$\mathcal{P}(\xi,\mu) = \lim\{\mathcal{P}(\xi,\mu') \; ; \; \mu' \to \mu\}$$

where μ' has its mass concentrated on a finite number of points and the convergence is for the weak topology on the space of probability measures on $[0,1]$. The map $\mu \mapsto \mathcal{P}(\xi,\mu)$ is then continuous for this weak topology.

We define

$$\mathcal{P}(\xi) = \inf_\mu \mathcal{P}(\xi,\mu) = \inf_{k,\mathbf{m},\mathbf{q}} \mathcal{P}_k(\mathbf{m},\mathbf{q}) .$$

It seems obvious that the probability measures μ for which $\mathcal{P}(\xi,\mu) = \mathcal{P}(\xi)$ will be of special interest.

Research Problem 14.11.3. (Level 2) Find useful conditions under which one can prove that there is a unique measure μ for which $\mathcal{P}(\xi,\mu) = \mathcal{P}(\xi)$.

Not knowing how to solve this nagging question, we find technically useful to consider only those μ for which $\mathcal{P}(\xi,\mu) = \mathcal{P}(\xi)$ that have an extra regularity condition.

Definition 14.11.4. *We say that a probability measure μ is* stationary *if it has all its mass concentrated on a finite number of points and if we cannot decrease $\mathcal{P}(\xi,\mu)$ by changing the location of these points (and keeping their masses constant).*

Definition 14.11.5. *We say that a probability measure μ is a* Parisi measure *(corresponding to the function ξ) if the following two conditions hold:*

$$\mathcal{P}(\xi,\mu) = \mathcal{P}(\xi) \tag{14.404}$$

$$\mu \textit{ is the limit of a sequence } (\mu_n) \textit{ of stationary measures.} \tag{14.405}$$

It is an interesting technical question whether (14.405) is a consequence of (14.404). On the other hand, it is rather obvious that there exists at least one Parisi measure. Indeed, given $n \geq 0$, we find $k, \mathbf{m}, \mathbf{q}$ with $\mathcal{P}_k(\mathbf{m},\mathbf{q}) \leq \mathcal{P}(\xi) + 2^{-n}$. Fixing $\mathbf{m}$ we then choose $\mathbf{q}$ to make $\mathcal{P}_k(\mathbf{m},\mathbf{q})$ as small as possible, and obtain in this way a stationary measure μ_n with $\mathcal{P}(\xi,\mu_n) \leq \mathcal{P}(\xi) + 2^{-n}$. The sequence (μ_n) has a converging subsequence, the limit of which is a Parisi measure.

We get now interested in Hamiltonians of the type

$$-H_{N,\boldsymbol{\beta}}(\boldsymbol{\sigma}) = \sum_{i \leq N} h_i \sigma_i + \sum_{p \geq 1} \frac{\beta_p}{N^{p-1/2}} \sum_{i_1,\ldots,i_{2p}} g_{i_1,\ldots,i_{2p}} \sigma_{i_1} \cdots \sigma_{i_{2p}} \ , \tag{14.406}$$

where $\boldsymbol{\beta} = (\beta_p)_{p \geq 1}$, the summation is over $1 \leq i_1, \ldots, i_{2p} \leq N$, and $g_{i_1,\ldots,i_{2p}}$ are independent standard Gaussian r.v.s. In that case the function $\xi = \xi_{\boldsymbol{\beta}}$ is given by $\xi(x) = \sum_{p \geq 1} \beta_p^2 x^{2p}$. Consider $r \geq 1$ and define

$$\beta_p(t) = \beta_p \ \text{ if } p \neq r \ \ ; \ \ \beta_r(t) = \beta_r + t \ .$$

The corresponding function ξ_t is given by

$$\xi_t(x) = \xi(x) + (2\beta_r t + t^2) x^{2r} \tag{14.407}$$

and the function θ_t by

$$\theta_t(x) = \theta(x) + (2r-1)(2\beta_r t + t^2) x^{2r} \ . \tag{14.408}$$

Theorem 14.11.6. *The map $t \mapsto \mathcal{P}(\xi_t)$ is differentiable at $t = 0$. Its derivative at $t = 0$ is given by*

$$\beta_r \left(1 - \int q^{2r} \mathrm{d}\mu(q)\right) ,$$

where μ is any Parisi measure. Moreover, if μ is a Parisi measure, we have

$$\beta_r \neq 0 \quad \Rightarrow \quad \lim_{N\to\infty} \nu(R_{1,2}^{2r}) = \int q^{2r} \mathrm{d}\mu(q) . \tag{14.409}$$

In this statement, "μ a Parisi measure" refers to the case $t = 0$, and $\nu(R_{1,2}^{2r})$ to the Hamiltonian (14.406).

It follows from (14.409) that if $\beta_r \neq 0$ whenever $r \geq 1$ (or at least sufficiently many values of r) then "μ is the limiting distribution of $|R_{1,2}|$", in the sense that for any continuous function f on $[0, 1]$ we have

$$\lim_{N\to\infty} \mathsf{E}\langle f(|R_{1,2}|)\rangle = \int f(q) \mathrm{d}\mu(q) . \tag{14.410}$$

When the external field in (14.406) is 0, the distribution of $R_{1,2}$ is symmetric around 0. In Section 14.12 we will show that when the external field is not zero, the overlap $R_{1,2}$ is essentially > 0, so that $\lim_{N\to\infty} \mathsf{E}\langle f(|R_{1,2}|)\rangle = \lim_{N\to\infty} \mathsf{E}\langle f(R_{1,2})\rangle$ and (14.410) proves that under the previous conditions, μ is indeed the limiting law of the overlap.

The proof of Theorem 14.11.6 is based on three simple lemmas.

Lemma 14.11.7. *The function $t \mapsto \mathcal{P}(\xi_t)$ is convex.*

Proof. Let

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

Theorem 14.5.1 implies

$$\mathcal{P}(\xi_t) = \lim_{N\to\infty} p_N(t). \tag{14.411}$$

The left-hand side is of the type

$$p_N(t) = \frac{1}{N} \mathsf{E} \log \sum_{\boldsymbol{\sigma}} \exp(-H_1(\boldsymbol{\sigma}) + tH_2(\boldsymbol{\sigma})) ,$$

where H_1 and H_2 are independent of t. It is thus a convex function of t by Hölder's inequality (as we used so many times in Chapter 12). □

The (excruciatingly painful) frustration is that this argument is very indirect. Moreover it applies only to the case where the function ξ is of the type $\xi(x) = \sum_{p\geq 1} \beta_p^2 x^{2p}$, while one should certain expect that it holds true even if there are terms with odd exponent.

Research Problem 14.11.8. (Level $\geq$ 2) Find a direct proof of Lemma 14.11.7 (i.e. not using Theorem 14.5.1).

Maybe we should emphasize here how little is known about the functional $\mathcal{P}(\xi)$. In particular, the author does not know the solution to the following.

Exercise 14.11.9. When $\xi(x) = \beta x^2/2$, the case of the ordinary SK model (and in presence of an external field) find an analytical proof that when β is small enough the Parisi measure is concentrated at one single point.

Lemma 14.11.10. *There exists a number A depending on $\boldsymbol{\beta}$ only with the following property. If μ is a probability measure that has all its mass concentrated on a finite set, the function*

$$f_\mu(t) = \mathcal{P}(\xi_t, \mu) \tag{14.412}$$

satisfies

$$\forall t\ ,\ |t| \leq 1\ , \quad |f''_\mu(t)| \leq A\ . \tag{14.413}$$

Proof. The quantity $f_\mu(t)$ is given (for a certain $\mathbf{m}$ and $\mathbf{q}$) by the left-hand side of (14.403) for the function ξ_t of (14.407) and the function θ_t of (14.408), so the result is an immediate consequence of (14.396). □

We recall Definition 14.11.4.

Lemma 14.11.11. *If μ is stationary, the function $f_\mu(t)$ of (14.412) satisfies*

$$f'_\mu(0) = \beta_r \left(1 - \int q^{2r} \mathrm{d}\mu(q)\right)\ . \tag{14.414}$$

Proof. As in the previous lemma $f_\mu(t)$ is given by the right-hand side of (14.403) for the function ξ_t of (14.407) and the function θ_t of (14.408). For $1 \leq p \leq k+1$ let

$$a_p = \left.\frac{\partial F_0}{\partial w_p}\right|_{w_1=\xi'(q_1),\ldots,w_{k+1}=\xi'(q_{k+1})}\ .$$

Recalling (14.407) and (14.408) we obtain

$$\left.\frac{\partial}{\partial t}(\xi'_t(q_p))\right|_{t=0} = 2\beta_r(2rq_p^{2r-1})\ \ ;\ \ \left.\frac{\partial}{\partial t}(\theta_t(q_p))\right|_{t=0} = 2\beta_r(2r-1)q_p^{2r}\ .$$

Thus, recalling (14.394) it is straightforward that

$$\begin{aligned} f'_\mu(0) &= 2\beta_r \sum_{1\le p\le k+1} q_p^{2r-1}\left(2ra_p + (2r-1)\frac{q_p}{2}(m_p - m_{p-1})\right) \\ &\quad + \beta_r(2r-(2r-1)) \\ &= \beta_r\left(1 + \sum_{1\le p\le k+1} q_p^{2r-1}(4ra_p + (2r-1)q_p(m_p - m_{p-1}))\right) . \end{aligned} \tag{14.415}$$

On the other hand, since μ is stationary, we cannot decrease the right-hand side of (14.403) by a small variation of q_p, so that

$$\xi''(q_p)\left(a_p + \frac{1}{2}q_p(m_p - m_{p-1})\right) = 0$$

i.e.

$$a_p = -\frac{1}{2}q_p(m_p - m_{p-1})$$

and substitution in (14.415) yields

$$f'_\mu(0) = \beta_r\left(1 - \sum_{1\le p\le k+1} q_p^{2r}(m_p - m_{p-1})\right) = \beta_r\left(1 - \int q^{2r}\mathrm{d}\mu(q)\right) . \quad \square$$

The properly amazing part of the previous argument is that the only information it uses about the function F_0 is in (14.394)!

Proof of Theorem 14.11.6. Consider a Parisi measure μ. By definition (see (14.405)) it is the limit of a sequence (μ_n) of stationary measures. Now (14.413) implies that for each t,

$$\mathcal{P}(\xi_t) \le \mathcal{P}(\xi_t, \mu_n) = f_{\mu_n}(t) \le f_{\mu_n}(0) + t f'_{\mu_n}(0) + \frac{t^2 A}{2} . \tag{14.416}$$

Since μ_n is stationary, the value of $f'_{\mu_n}(0)$ is given by (14.414) for μ_n rather then μ. Letting $n \to \infty$ in (14.416), and recalling that $\lim_{n\to\infty} \mathcal{P}(\xi, \mu_n) = \lim_{n\to\infty} f_{\mu_n}(0) = \mathcal{P}(\xi)$, we obtain

$$\mathcal{P}(\xi_t) \le \mathcal{P}(\xi) + t\beta_r\left(1 - \int q^{2r}\mathrm{d}\mu(q)\right) + \frac{t^2 A}{2} ,$$

and since the function $f(t) = \mathcal{P}(\xi_t)$ is convex by Lemma 14.11.7 this proves that f is differentiable at $t = 0$ and that $f'(0) = \beta_r\left(1 - \int q^{2r}\mathrm{d}\mu(q)\right)$. Moreover (14.411) and Griffiths' lemma imply

$$f'(0) = \lim_{N\to\infty} p'_N(0) ,$$

and

$$p'_N(0) = \frac{1}{N}\nu\left(\frac{1}{N^{r-1/2}}\sum_{i_1,\ldots,i_{2r}} g_{i_1,\ldots,i_{2r}}\sigma_{i_1}\cdots\sigma_{i_{2r}}\right)$$
$$= \beta_r(1-\nu(R_{1,2}^{2r}))$$

by integration by parts as in Lemma 1.3.11. □

Let us discuss the application of Theorem 14.11.6 to the ordinary SK model beyond the AT line. From Theorem 13.3.1 and the Parisi formula (14.102) we know that the Parisi measure μ cannot be supported by one point. (Indeed, when the Parisi measure is supported at the point q, the value of $\mathcal{P}(\xi,\mu)$ is the right-hand side of (1.72) and is then $\geq \mathrm{SK}(\beta,h)$.) By (14.409) we know that

$$\lim_{N\to\infty}\nu(R_{1,2}^2) = \int x^2 \mathrm{d}\mu(x)\ . \tag{14.417}$$

Unfortunately, we do not see how to prove even that $\nu((R_{1,2}-q)^2)\geq\delta>0$ for some δ independent of N. The situation improves if we assume that the external field h is a Gaussian r.v. with positive variance. In that case we can extend Theorem 13.3.1 with the same proof "to the case $r=1/2$" to obtain

$$\lim_{N\to\infty}\nu(R_{1,2}) = \int x \mathrm{d}\mu(x)\ ,$$

which, when combined with (14.417) and the fact that μ is not concentrated at one point gives for large N that $\nu(R_{1,2}^2)-\nu(R_{1,2})^2\geq\delta>0$.

14.12 Positivity of the Overlap

When $h=0$, the Hamiltonian $H_{N,\boldsymbol{\beta}}$ of (14.405) is invariant under the symmetry $\boldsymbol{\sigma}\mapsto-\boldsymbol{\sigma}$; therefore (at given disorder) the law of the overlap $R_{1,2}$ in $[-1,1]$ is symmetric around 0. In this section we prove that the situation changes dramatically when $\mathsf{E}h^2>0$. The overlap becomes >0.

Theorem 14.12.1. *Consider the Gibbs measure G_N relative to the Hamiltonian (14.406), and assume that $\mathsf{E}h^2>0$. Consider a Parisi measure μ relative to this Hamiltonian, and the smallest point c in the support of μ. Then $c>0$ and for any $c'<c$, for some number K independent of N we have*

$$\mathsf{E}G_N^2(\{R_{1,2}\leq c'\})\leq K\exp\left(-\frac{N}{K}\right)\ . \tag{14.418}$$

An important feature of this result is that it holds in complete generality. There is no need to make assumptions on the values of the coefficients β_p.

Theorem 14.12.1 is completely expected "on physical grounds" (i.e. by analogy with simpler models) but the proof is very far from being trivial.

A consequence of Theorem 14.12.1 is that, even though we do not know that the Parisi measure μ is unique, the smallest point c of its support does not depend on μ. It would be nice if in the left-hand side of (14.13) one could replace the set $\{R_{1,2} < c'\}$ by the set $\{R_{1,2} \in A\}$ where A is any compact set that does not intersect the support of the Parisi measure. Unfortunately, this seems to be wrong, see Problem 15.7.7.

Another consequence of Theorem 14.12.1 is the following.

Theorem 14.12.2. *Consider the Hamiltonian (14.406). Assume that* $\mathsf{E}h^2 > 0$ *and that* $\beta_p \neq 0$ *for all* $p \geq 1$*. Then this Hamiltonian satisfies the extended Ghirlanda-Guerra identities. That is, given a continuous function* ψ *on* $\mathbb{R}$ *we have*

$$\lim_{N\to\infty} \sup_{|f|\leq 1} \left| n\nu(\psi(R_{1,n+1})f) - \nu(\psi(R_{1,2}))\nu(f) - \sum_{2\leq \ell\leq n} \nu(\psi(R_{1\,\ell})f) \right| = 0\,, \tag{14.419}$$

where the supremum is taken over all the functions f *on* Σ_N^n *with* $|f| \leq 1$.

Proof. When $\psi(x) = x^{2p}$ for $p \geq 1$, (14.419) follows from Theorem 14.11.6 (and more precisely the assertion about the differentiability of the map $t \mapsto \mathcal{P}(\xi_t)$ at $t = 0$) and Theorem 12.1.3 by integration by parts as in Theorem 12.1.10. By linearity the result also holds if ψ is a polynomial in x^2. If c is any positive number, any continuous function ψ can be approximated uniformly well by polynomials in x^2 on the interval $[c, 1]$, and the result then follows from Theorem 14.12.1. □

Theorem 14.12.1 does not play a big part in the picture that we try to outline in the next chapter. Its only purpose is to make certain statements more natural. These statements would remain striking even if we assumed that the overlap is positive (instead of deducing this from Theorem 14.12.1). The proof of this theorem does not involve any radically new idea, but rather a careful use of the tools we have already built. The reader who finds the remainder of this section too technical should simply move on to the next chapter.

We set $\xi(x) = \sum_{p\geq 1} \beta_p x^{2p}$, $\theta(x) = x\xi'(x) - \xi(x)$. To prove Theorem 14.12.1 we will show the following.

Proposition 14.12.3. *If* c *is as in Theorem 14.12.1, given* $c' < c$ *there exists a number* $\varepsilon^* > 0$*, depending only on* $\boldsymbol{\beta}$*,* h *and* c*, such that if* $u \leq c'$ *we have*

$$\frac{1}{N} \mathsf{E} \log \sum_{R_{1,2}=u} \exp(-H_{N,\boldsymbol{\beta}}(\boldsymbol{\sigma}^1) - H_{N,\boldsymbol{\beta}}(\boldsymbol{\sigma}^2)) \leq 2\mathcal{P}(\xi, h) - \varepsilon^*\,. \tag{14.420}$$

The reader observes that here there is no Poisson-Dirichlet component and no α's.

Proof of Theorem 14.12.1. By Theorem 14.5.1, for large N we have $p_N \geq \mathcal{P}(\xi, h) - \varepsilon^*/4$, so that (14.420) implies

$$\frac{1}{N}\mathsf{E}\log \sum_{R_{1,2}=u} \exp(-H_{N,\boldsymbol{\beta}}(\boldsymbol{\sigma}^1) - H_{N,\boldsymbol{\beta}}(\boldsymbol{\sigma}^2)) \leq 2p_N - \frac{\varepsilon^*}{2} .$$

We then use concentration of measure as in Proposition 13.4.3 to deduce that with overwhelming probability

$$\frac{1}{N}\log \sum_{R_{1,2}=u} \exp(-H_{N,\boldsymbol{\beta}}(\boldsymbol{\sigma}^1) - H_{N,\boldsymbol{\beta}}(\boldsymbol{\sigma}^2)) \leq \frac{2}{N}\log\sum_{\boldsymbol{\sigma}} \exp(-H_{N,\boldsymbol{\beta}}(\boldsymbol{\sigma})) - \frac{\varepsilon^*}{4} ,$$

and thus

$$G_N^{\otimes 2}(\{R_{1,2} = u\}) \leq \exp\left(-\frac{\varepsilon^* N}{4}\right) .$$

Consequently, for $u \leq c'$ we have

$$\mathsf{E}(G_N^{\otimes 2}(\{R_{1,2} = u\})) \leq K \exp\left(-\frac{N}{K}\right) ,$$

where K depends only on $\boldsymbol{\beta}$, h and c'. This implies (14.418). □

The main difficulty in the proof of Proposition 14.12.3 is the case $u \leq 0$. Interestingly, in this case the arguments use no information about the Parisi measure μ. It is for the case $0 \leq u < c' < c$ that this information will be useful. The case $u \leq 0$ relies on the following.

Proposition 14.12.4. *Consider a number* $0 \leq v \leq 1$*, integers* $\tau \leq \kappa$*, and numbers*

$$0 = \rho_0 \leq \rho_1 \leq \ldots \leq \rho_\tau = v \leq \rho_{\tau+1} \leq \ldots \leq \rho_{\kappa+1} = 1 \tag{14.421}$$

$$0 = n_0 \leq n_1 \leq \ldots \leq n_\kappa = 1 . \tag{14.422}$$

For $0 \leq p \leq \kappa$*, consider independent pairs of Gaussian r.v.s* (y_p^1, y_p^2) *with*

$$\mathsf{E}(y_p^1)^2 = \mathsf{E}(y_p^2)^2 = \xi'(\rho_{p+1}) - \xi'(\rho_p) \tag{14.423}$$

and

$$y_p^2 = -y_p^1 \text{ if } p < \tau ; \quad y_p^1 \text{ and } y_p^2 \text{ independent if } p \geq \tau .$$

Then we have

$$\frac{1}{N}\mathsf{E}\log\sum_{R_{1,2}=-v}\exp(-H_{N,\boldsymbol{\beta}}(\boldsymbol{\sigma}^1)-H_{N,\boldsymbol{\beta}}(\boldsymbol{\sigma}^2))$$
$$\le 2\log 2+Y_0+\lambda v-2\sum_{p<\tau}n_p(\theta(\rho_{p+1})-\theta(\rho_p))$$
$$-\sum_{\tau\le p\le\kappa}n_p(\theta(\rho_{p+1})-\theta(\rho_p))\ , \tag{14.424}$$

where Y_0 is computed as follows: starting with

$$Y_{\kappa+1}=\log\Bigg(\mathrm{ch}\Big(h+\sum_{0\le p\le\kappa}y_p^1\Big)\mathrm{ch}\Big(h+\sum_{0\le p\le\kappa}y_p^2\Big)\mathrm{ch}\lambda$$
$$+\ \mathrm{sh}\Big(h+\sum_{0\le p\le\kappa}y_p^1\Big)\mathrm{sh}\Big(h+\sum_{0\le p\le\kappa}y_p^2\Big)\mathrm{sh}\lambda\Bigg)\ ,$$

we define recursively

$$Y_p=\frac{1}{n_p}\log\mathsf{E}_p\exp n_pY_{p+1}\ , \tag{14.425}$$

where E_p denotes expectation in the r.v.s (y_n^1,y_n^2) for $n\ge p$. (Of course (14.425) means that $Y_p=\mathsf{E}_pY_{p+1}$ if $n_p=0$).

Proof. This is a special case of Proposition 14.6.3 when $Z_p^1=Z_p^2=0$ for each p and $t=1$, with $v=-u$. □

The reader observes that the meaning of the notation v has changed compared with the beginning of this chapter (e.g. in (14.314)).

The strategy is to find sequences (14.421) and (14.422) such that the bound (14.424) proves (14.420). For this we will use sequences $\mathbf{m}$ and $\mathbf{q}$ that satisfy condition MIN(ε) of Definition 14.5.3 page 373 for ε small enough.

There are two main steps in the proof, which are described in the next two results.

Proposition 14.12.5. *There exists $\delta>0$ and $\varepsilon_0>0$, depending on $\boldsymbol{\beta}$ and h only with the following property. Whenever we can find $k,\mathbf{m},\mathbf{q}$ that satisfy condition* MIN(ε_0) *and that for a certain integer s with $1\le s\le k+1$, we have*

$$m_{s-1}\le\delta\ ;\ q_s\ge v-\delta\ , \tag{14.426}$$

then we can find the parameters in (14.424) so that the right-hand side is $\le\mathcal{P}(\xi,h)-1/M$.

Proposition 14.12.6. *Consider δ as in Proposition 14.12.5. Then we can find $\varepsilon_1>0$ with the following property. Whenever we can find $k,\mathbf{m},\mathbf{q}$ such* *that* $\mathcal{P}_k(\mathbf{m},\mathbf{q}) \leq \mathcal{P}(\xi,h)+\varepsilon_1$*, and such that for some integer* s *with* $1 \leq s \leq k+1$ *we have*

$$q_s \leq v-\delta \; ; \qquad m_s \geq \delta \; , \tag{14.427}$$

then we can find the parameters in (14.424) such that the right-hand side is $\leq \mathcal{P}(\xi,h)-1/M$.

Proof of (14.420) for $u \leq 0$**.** Consider $v \geq 0$, and δ and ε_0 as in Proposition 14.12.5. Consider $k,\mathbf{m},\mathbf{q}$ that satisfy MIN$(\min(\varepsilon_0,\varepsilon_1))$. Let s be the largest integer $\leq k+1$ such that $m_{s-1} \leq \delta$. If $q_s \geq v-\delta$, we conclude with Proposition 14.12.5. Otherwise we have $q_s \leq v-\delta$. If $s=k+1$, then $m_s=m_{k+1}=1\geq\delta$. If $s<k+1$ the definition of s then shows that we have $m_s \geq \delta$, and we conclude with Proposition 14.12.6. □

The basic idea of Proposition 14.12.5 is that the case $\delta>0$ very small should be a (very small!) perturbation of the case $\delta=0$, so that a suitably quantitative solution of the case $\delta=0$ should suffice. Of course, it will require work to prove estimates that do not depend on k, but we should be confident that this is possible since similar tasks has already been performed.

We first present the argument in the case where $\delta=0$, i.e. when

$$0 \leq v \leq q_1 \; . \tag{14.428}$$

In that case we use Proposition 14.12.4 with $\tau=1$, $\kappa=k+2$ and

$$0=\rho_0 \leq \rho_1 = v \leq \rho_2 = q_1 \leq \rho_3 = q_2 \leq \ldots \leq \rho_{k+3} = q_{k+2} = 1 \tag{14.429}$$

$$n_0 = 0 = n_1 \; ; \quad n_p = m_{p-1} \;\; \text{for } 2 \leq p \leq \kappa = k+2 \; . \tag{14.430}$$

Let us denote by

$$\alpha(\lambda) \text{ the right-hand side of (14.424)} \; ,$$

keeping implicit the dependence on the other parameters, and on v in particular. As we have used many times, we have $\alpha(0)=2\mathcal{P}_k(\mathbf{m},\mathbf{q})$. Also Lemma 14.6.5 shows that $0 \leq \alpha''(\lambda) \leq 1$, and thus

$$\inf_\lambda \alpha(\lambda) \leq \alpha(0) - \frac{\alpha'(0)^2}{2} \; . \tag{14.431}$$

The strategy now is to try to bound from below $\beta(v) := \alpha'(0)$. Let us recall the following construction (that is again essential throughout this section). Starting with the function $A_{k+2}(x)=\log \mathrm{ch} x$, we define recursively the functions $A_p(x)$ by (14.190) i.e.

$$A_p(x) = \frac{1}{m_p} \log \mathsf{E} \exp m_p A_{p+1}(x+z_p) \; , \tag{14.432}$$

where z_p is a Gaussian r.v. with $\mathsf{E} z_p^2 = \xi'(q_{p+1})-\xi'(q_p)$.

We think of Y_0 as a function $Y_0(\lambda)$. This is exactly the function of Proposition 14.6.4 in the case $g_p^\ell = y_p^\ell$. Recalling the sequence D_p of that Proposition, we then have $D_{k+3}(x) = \log \operatorname{ch} x = A_{k+2}(x)$, and we see from (14.430) by decreasing induction over p that $D_p = A_{p-1}$ for $p \geq 2$. We then use (14.180) for $\tau = 2$ to compute $Y_0'(0)$, and since $D_2 = A_1$ we get

$$\beta(v) = \mathsf{E}\big(A_1'(h + y_0^1 + y_1^1)A_1'(h + y_0^2 + y_1^2)\big) + v \, . \tag{14.433}$$

We note that by construction the following relations hold:

$$\mathsf{E}(y_0^1 + y_1^1)^2 = \mathsf{E}(y_0^2 + y_1^2)^2 = \xi'(q_1) \; ; \; \mathsf{E}(y_0^1 + y_1^1)(y_0^2 + y_1^2) = -\mathsf{E}(y_0^1)^2 = -\xi'(v) \, . \tag{14.434}$$

Let us first consider the case $\rho_1 = v = 0$. Then $y_0^1 = y_0^2 = 0$ and y_1^1 and y_1^2 are independent with $\mathsf{E}(y_1^1)^2 = \mathsf{E}(y_1^2)^2 = \xi'(q_1) = \mathsf{E}z_1^2$. Therefore

$$A_0(x) = \mathsf{E}A_1(x + z_1) = \mathsf{E}A_1(x + y_1^1) = \mathsf{E}A_1(x + y_1^2) \, ,$$

so that

$$A_0'(x) = \mathsf{E}A_1'(x + y_1^1) = \mathsf{E}A_1'(x + y_1^2) \, ,$$

and taking first expectation in y_1^1 and y_1^2 we deduce from (14.433) that

$$\beta(0) = \mathsf{E}A_0'(h)^2 \, . \tag{14.435}$$

Throughout the section, we denote by M a number that depends only on $\boldsymbol{\beta}$ and h, and thus need not be the same at each occurrence. Since $\mathsf{E}h^2 > 0$, we know from (14.274) and (14.435) that

$$\beta(0) \geq \frac{1}{M} \, . \tag{14.436}$$

Combining with (14.431) this implies

$$\inf_\lambda \alpha(\lambda) \leq 2\mathcal{P}_k(\mathbf{m}, \mathbf{k}) - \frac{1}{M'} \, . \tag{14.437}$$

Thus, in the case $v = 0$ we are in very good shape. To use the same idea when $v \neq 0$ we would like actually to prove that for all $0 \leq v \leq q_1$ we have

$$\beta(v) \geq \frac{1}{M} \, . \tag{14.438}$$

We will deduce this from (14.436) and comparison of $\beta(v)$ and $\beta(0)$ using Lemma 14.9.5. From (14.433) and (14.434) we obtain (as we have just used)

$$\beta(0) = \mathsf{E}A_1'(h + \chi_1)A_1'(h + \chi_2) \, ,$$

where χ_1 and χ_2 are independent Gaussian r.v.s with $\mathsf{E}\chi_1^2 = \mathsf{E}\chi_2^2 = \xi'(q_1) = \mathsf{E}z_0^2$, while

$$\beta(v) = \mathsf{E}A_1'(h + \chi_1')A_1'(h + \chi_2') + v \, ,$$

where χ'_1 and χ'_2 are jointly Gaussian r.v.s with $\mathsf{E}\chi_1'^2 = \mathsf{E}\chi_2'^2 = \xi'(q_1)$ and $\mathsf{E}\chi'_1\chi'_2 = -\xi'(v)$. Thus Lemma 14.9.5 implies

$$\beta(v) \geq \beta(0) + v - \xi'(v)\mathsf{E}A_1''(h+z_0)^2 , \tag{14.439}$$

and since $\xi'(v) \leq v\xi''(v) \leq v\xi''(q_1)$,

$$\beta(v) \geq \beta(0) + v\big(1 - \xi''(q_1)\mathsf{E}A_1''(h+z_0)^2\big) . \tag{14.440}$$

Let us define

$$\varepsilon = \mathcal{P}_k(\mathbf{m},\mathbf{q}) - \mathcal{P}(\xi, h) . \tag{14.441}$$

We appeal to (14.283) for $r = 1$ to obtain

$$\xi''(q_1)(-\xi''(q_1)U''(0) - 1) \leq M\varepsilon^{1/6} , \tag{14.442}$$

where, using (14.248),

$$U''(0) = -\mathsf{E}A_1''(h+z_0)^2 . \tag{14.443}$$

Now we deduce from Lemma 14.7.15 page 411 that $q_1 \geq 1/M$ and thus $\xi''(q_1) \geq 1/M$. Therefore (14.442) implies

$$-\xi''(q_1)U''(0) - 1 \leq M\varepsilon^{1/6}$$

and, combining with (14.443),

$$\xi''(q_1)\mathsf{E}A_1''(h+z_0)^2 \leq 1 + M\varepsilon^{1/6} ,$$

so that (14.440) yields

$$\beta(v) \geq \beta(0) - M\varepsilon^{1/6} \geq \frac{1}{M} - M\varepsilon^{1/6} .$$

So, if ε is small enough, we have $\beta(v) \geq 1/M$ for all $v \leq q_1$ and we conclude as in the case $v = 0$ with (14.437). This finishes the proof when (14.428) holds, i.e. $0 \leq v \leq q_1$.

Before we start the proof of Proposition 14.12.5 let us make a general observation. To keep the notation manageable it is very desirable to assume that $v = q_a$ for a certain integer a. We have explained why this is possible, as we can insert any element we wish in the list $\mathbf{q}$. Of course, when using consequences of relation (14.424) as in (14.222) and (14.283) we will not be able to use the value $r = a$.

Let us now start the proof of Proposition 14.12.5. It is unfortunately necessary in the argument to distinguish whether $q_s \geq v$ or $q_s \leq v$. We consider first the case $q_s \leq v$ (which is the most difficult). So we assume that for a certain number $\delta > 0$ (that will not depend on N and will be very small) and a certain integer s with $1 \leq s \leq k+1$ we have

$$m_{s-1} \le \delta \; ; \; v - \delta \le q_s \le v \; , \tag{14.444}$$

As previously explained, we may assume without loss of generality that $v = q_a$ for a certain integer a, so that

$$v - \delta \le q_s \le \ldots \le q_a = v \le q_{a+1} \le \ldots \le q_{k+2} = 1 \; . \tag{14.445}$$

We define $\tau = 1$,

$$\rho_0 = 0 \; , \quad \rho_1 = v = q_a \; , \quad \rho_2 = q_{a+1} \; , \quad \ldots \; , \quad \rho_{\kappa+1} = q_{k+2} = 1 \; , \tag{14.446}$$

where $\kappa = k + 2 - a$. We define

$$n_0 = 0 \; , \quad n_1 = m_a \; , \quad n_2 = m_{a+1} \; , \quad \ldots \; , \; n_\kappa = m_{k+1} = 1 \; . \tag{14.447}$$

We denote by $\alpha(\lambda)$ the right-hand side of (14.424), keeping the dependence on N implicit. The first goal is the following.

Lemma 14.12.7. *We have*

$$\alpha(0) \le 2\mathcal{P}_k(\mathbf{m}, \mathbf{q}) + M\delta \; . \tag{14.448}$$

Proof. First, we observe that

$$\begin{aligned}\sum_{1 \le p \le \kappa} n_p(\theta(\rho_{p+1}) - \theta(\rho_p)) &= \sum_{a \le p \le k+1} m_p(\theta(q_{p+1}) - \theta(q_p)) \\ &= \sum_{1 \le p \le k+1} m_p(\theta(q_{p+1}) - \theta(q_p)) - C \; ,\end{aligned} \tag{14.449}$$

where

$$\begin{aligned} C &= \sum_{1 \le p < a} m_p(\theta(q_{p+1}) - \theta(q_p)) \\ &= \sum_{1 \le p \le s-1} m_p(\theta(q_{p+1}) - \theta(q_p)) + \sum_{s \le p < a} m_p(\theta(q_{p+1}) - \theta(q_p)) \; . \end{aligned}$$

Since $m_p \le \delta$ for $p \le s-1$ and $m_p \le 1$ for $p \ge s$ we get

$$C \le \delta\theta(q_s) + \theta(q_a) - \theta(q_s) \le M\delta \; ,$$

using that $q_a - q_s = v - q_s \le \delta$ by (14.444). Combining with (14.449) we obtain

$$\sum_{1 \le p \le \kappa} n_p(\theta(q_{p+1}) - \theta(q_p)) \ge \sum_{1 \le p \le k+1} m_p(\theta(\rho_{p+1}) - \theta(\rho_p)) - M\delta \; . \tag{14.450}$$

Next, when $\lambda = 0$, copying the proof of (14.184) we get that

$$Y_0 = \mathsf{E} A_a(h + y_0^1) + \mathsf{E} A_a(h + y_0^2) = 2\mathsf{E} A_a(h + y_0^1) \; ,$$

because $y_0^2 = -y_0^1$. Now $\mathsf{E}(y_0^1)^2 = \xi'(\rho_1) - \xi'(0) = \xi'(q_a)$, so that y_1^1 and $\sum_{0\le p<a} z_p$ have the same distribution, and therefore

$$Y_0 = 2\mathsf{E}A_a\Big(h + \sum_{0\le p<a} z_p\Big)\ . \tag{14.451}$$

On the other hand, Jensen's inequality implies

$$\mathsf{E}\exp mX \ge \exp m\mathsf{E}X\ ,$$

and consequently (14.432) entails that $A_p(x) \ge \mathsf{E}A_{p+1}(x+z_p)$. Iteration of this relation yields

$$\mathsf{E}A_0(h) \ge \mathsf{E}A_a\Big(h + \sum_{0\le p<a} z_p\Big)$$

and (14.451) implies that $Y_0 \le 2\mathsf{E}A_0(h)$. Using the definition (14.89) of $\mathcal{P}_k(\mathbf{m},\mathbf{q})$ completes the proof. □

The next goal is to bound $\alpha'(0)$ from below. Proposition 14.6.4 (and specifically (14.180)) implies

$$\alpha'(0) = \mathsf{E}(A'_a(h+y_0^1)A'_a(h+y_0^2)) + v \tag{14.452}$$

where $y_0^2 = -y_0^1$, $\mathsf{E}(y_0^1)^2 = \xi'(v)$. Consider two independent Gaussian r.v.s χ'_1 and χ'_2 with $\mathsf{E}(\chi'_1)^2 = \mathsf{E}(\chi'_2)^2 = \xi'(v)$. It then follows from Lemma 14.9.5 (used for $\chi_1 = y_0^1$, $\chi_2 = y_0^2$), and since $|\mathsf{E}\chi_1\chi_2 - \mathsf{E}\chi'_1\chi'_2| = \xi'(v)$ that

$$\mathsf{E}(A'_a(h+y_0^1)A'_a(h+y_0^2)) \ge \mathsf{E}(A'_a(h+\chi'_1)A'_a(h+\chi'_2)) - \xi'(v)\mathsf{E}A''_a\big(h+\chi\sqrt{\xi'(v)}\big)^2\ ,$$

where χ is standard Gaussian. Since $\xi'(v) \le v\xi''(v)$, combining with (14.452) yields

$$\alpha'(0) \ge \mathsf{E}(A'_a(h+\chi'_1)A'_a(h+\chi'_2)) + v\Big(1 - \xi''(v)\mathsf{E}A''_a\big(h+\chi\sqrt{\xi'(v)}\big)^2\Big)\ . \tag{14.453}$$

Lemma 14.12.8. *There exists a number M depending on $\boldsymbol{\beta}$ and h only such that*

$$\mathsf{E}(A'_a(h+\chi'_1)A'_a(h+\chi'_2)) \ge \frac{1}{M}\ . \tag{14.454}$$

Proof. It follows from (14.274) that

$$A''_a(x) \ge \frac{1}{M\mathrm{ch}^2 x} \tag{14.455}$$

where M depends on $\boldsymbol{\beta}$ only. Defining the function

$$A_*(x) = \mathsf{E}A_a(x+\chi'_1)\ ,$$

we then have, using (14.281) in the second inequality,

$$A_*''(x) = \mathsf{E}A_a''(x+\chi_1') \geq \mathsf{E}\frac{1}{M\mathrm{ch}^2(x+\chi_1')} \geq \frac{1}{M'\mathrm{ch}^2 x} . \tag{14.456}$$

Since $A_a(x) = A_a(-x)$, we have $A_*'(0) = 0$, and (14.456) implies

$$|A_*'(x)| \geq \frac{1}{M}|\mathrm{th}x| , \tag{14.457}$$

so that

$$\mathsf{E}A_*'(h)^2 \geq \frac{1}{M} .$$

But

$$\mathsf{E}A_*'(h)^2 = \mathsf{E}(A_a'(h+\chi_1')A_a'(h+\chi_2')) . \qquad \square$$

Let us remark for further use that arguing as in (14.457) one obtains that for each p one has

$$|A_p'(x)| \geq \frac{1}{M}|\mathrm{th}x| . \tag{14.458}$$

To use (14.453) the next task is to bound from above the quantity

$$\xi''(v)\mathsf{E}A_a''(h+\chi\sqrt{\xi'(v)})^2 .$$

The starting point of the proof is that (14.283) yields

$$\xi''(q_s)(-\xi''(q_s)U''(0)-1) \leq M\varepsilon^{1/6} , \tag{14.459}$$

where, using (14.248) and setting $\zeta_p = h + \sum_{0\leq n<p} z_n$,

$$U''(0) = -\mathsf{E}(W_1\cdots W_{s-1}A_s''(\zeta_s)^2) \tag{14.460}$$

for

$$W_p = \exp m_p(A_{p+1}(\zeta_{p+1}) - A_p(\zeta_p)) .$$

By (14.222) we have $q_1 = \mathsf{E}A_1'(\zeta_1)^2$, and (14.458) implies that $q_1 \geq 1/M$, so that $q_s \geq q_1 \geq 1/M$ and $\xi''(q_s) \geq 1/M$. Therefore (14.459) and (14.460) imply

$$\xi''(q_s)\mathsf{E}(W_1\cdots W_{s-1}A_s''(\zeta_s)^2) \leq 1 + M\varepsilon^{1/6} . \tag{14.461}$$

Lemma 14.12.9. *Assuming (14.444) there exists $\delta_0 > 0$ depending only on $\boldsymbol{\beta}$ and h such that when $\delta \leq \delta_0$ we have*

$$\xi''(v)\mathsf{E}A_a''\big(h+\chi\sqrt{\xi'(v)}\big)^2 \leq \xi''(q_s)\mathsf{E}(W_1\cdots W_{s-1}A_s''(\zeta_s)^2) + M\sqrt{\delta} . \tag{14.462}$$

Accepting this for the moment we see that combining with (14.454), (14.461) and (14.453) implies

$$\alpha'(0) \geq \frac{1}{M} - M\varepsilon^{1/6} - M\sqrt{\delta}\,. \tag{14.463}$$

Proof of Proposition 14.12.5. First we complete the proof in the case where $q_s \leq v$. Let us denote by M_1 the constant of (14.448) and (14.463), and without loss of generality assume that $M_1 \geq 1$. Let us set $\delta = 1/16M_1^4$. Then if $\varepsilon \leq \varepsilon_1 = (1/4M_1^2)^6$, (14.463) implies

$$\alpha'(0) \geq \frac{1}{M_1} - \frac{M_1}{4M_1^2} - \frac{M_1}{4M_1^2} = \frac{1}{2M_1}\,,$$

and combining with (14.431) and (14.448) yields

$$\inf_\lambda \alpha(\lambda) \leq 2\mathcal{P}_k(\mathbf{m},\mathbf{q}) + M_1\delta - \frac{1}{8M_1^2} \leq 2\mathcal{P}(\xi,h) - \frac{1}{16M_1^2}\,,$$

and this completes the proof in the case where $q_s \leq v$.

It remains to prove the result when $q_s > v = q_a$. Then (14.444) holds for $s' = a+1$, so that without loss of generality we may assume $s = a+1$. We define

$$0 = \rho_0 \leq \rho_1 = v = q_a \leq \rho_2 = q_{a+1} \leq \ldots \leq \rho_{\kappa+1} = q_{k+2}$$

where $\kappa = k+1-a$, and

$$n_0 = 0\,,\quad n_1 = 0\,,\quad n_2 = m_{a+1}\,,\quad \ldots\,,\quad n_\kappa = m_{k+1} = 1\,. \tag{14.464}$$

The difference between (14.464) and (14.447) is that now $n_1 = 0$. Rather than (14.452) we find

$$\alpha'(0) = \mathsf{E}\big(A'_{a+1}(h + y_0^1 + y_1^1)A'_{a+1}(h + y_0^2 + y_1^2)\big) + v\,. \tag{14.465}$$

Consider two independent Gaussian r.v.s χ'_1 and χ'_2 with $\mathsf{E}(\chi'_1)^2 = \mathsf{E}(\chi'_2)^2 = \xi'(q_{a+1})$. We use Lemma 14.9.5 with $\chi_1 = y_0^1 + y_1^1$, $\chi_2 = y_0^2 + y_1^2$, so that $\mathsf{E}\chi_1\chi_2 = \mathsf{E}y_0^1y_0^2 = -\mathsf{E}(y_0^1)^2 = -\xi'(v)$ and $\mathsf{E}\chi'_1\chi'_2 = 0$. We get

$$\begin{aligned}\mathsf{E}\big(A'_{a+1}(h + y_0^1 + y_1^1)A'_{a+1}(h + y_0^2 + y_1^2)\big) &\geq \mathsf{E}\big(A'_{a+1}(h + \chi'_1)A'_{a+1}(h + \chi'_2)\big)\\ &\quad - \xi'(v)\mathsf{E}A''_{a+1}\big(h + \chi\sqrt{\xi'(q_{a+1})}\big)^2\,,\end{aligned}$$

where χ is standard Gaussian. Then to finish the argument as previously it suffices to show that

$$\xi''(v)\mathsf{E}A''_{a+1}\big(h + \chi\sqrt{\xi'(q_{a+1})}\big)^2 \leq 1 + M\delta + M\varepsilon^{1/6}\,. \tag{14.466}$$

Using (14.459) for $a+1$ rather than s, instead of (14.461) we reach

$$\xi''(q_{a+1})\mathsf{E}\big(W_1\cdots W_a A''_{a+1}(\zeta_{a+1})^2\big) \leq 1 + M\varepsilon^{1/6}\,.$$

This implies (14.466) since $\xi''(q_{a+1}) \geq \xi''(v)$ because $q_{a+1} \geq v$, since (as in (14.469)) $\mathsf{E}|W_1 \cdots W_a - 1| \leq M\delta$ and since $\mathsf{E}A''_{a+1}(\zeta_{a+1})^2 = \mathsf{E}A''_{a+1}(h + \chi\sqrt{\xi'(q_{a+1})})^2$. □

Proof of Lemma 14.12.9. Using (14.204) in the case $A = A_{p+1}$ and $m = m_p$ we get

$$A''_p(x) \geq \mathsf{E}_p\big(\exp m_p(A_{p+1}(x+z_p) - A_p(x))A''_{p+1}(x+z_p)\big) ,$$

where as usual E_p denotes expectation in the r.v.s z_n for $n \geq p$. This implies

$$A''_p(\zeta_p) \geq \mathsf{E}_p(W_p A''_{p+1}(\zeta_{p+1})) ,$$

and since W_p does not depend on the r.v.s z_n for $n \geq p$ this inequality can be iterated to give

$$A''_s(\zeta_s) \geq \mathsf{E}_s(W_s \cdots W_{a-1} A''_a(\zeta_a)) . \tag{14.467}$$

We shall prove that

$$\mathsf{E}|W_s \cdots W_{a-1} - 1| \leq M\sqrt{\delta} \tag{14.468}$$

and

$$\mathsf{E}|W_1 \cdots W_{s-1} - 1| \leq M\delta . \tag{14.469}$$

Now, since $0 \leq A''_a \leq 1$,

$$\begin{aligned} |\mathsf{E}_s(W_s \cdots W_{a-1}A''_a(\zeta_a)) - A''_a(\zeta_a)| &\leq \mathsf{E}_s(|W_s \cdots W_{a-1}A''_a(\zeta_a) - A''_a(\zeta_a)|) \\ &\leq \mathsf{E}_s(|W_s \cdots W_{a-1} - 1|) . \end{aligned} \tag{14.470}$$

Combining with (14.467), and letting $Y = \mathsf{E}_s(|W_s \cdots W_{a-1} - 1|)$ yields

$$A''_s(\zeta_s) \geq \mathsf{E}_s A''_a(\zeta_a) - Y ,$$

so that

$$A''_s(\zeta_s)^2 \geq (\mathsf{E}_s A''_a(\zeta_a))^2 - 2Y$$

and, taking expectation and using (14.468),

$$\mathsf{E}A''_s(\zeta_s)^2 \geq \mathsf{E}(\mathsf{E}_s A''_a(\zeta_a))^2 - M\sqrt{\delta} . \tag{14.471}$$

Now, by (14.272) we get

$$|A''_a(\zeta_a) - A''_a(\zeta_s)| \leq 5|\zeta_a - \zeta_s| , \tag{14.472}$$

and

$$\zeta_a - \zeta_s = \sum_{s \leq n < a} z_n ,$$

so that $\mathsf{E}_s(\zeta_a-\zeta_s)^2=\xi'(q_a)-\xi'(q_s)$. Now, since $q_a\ge q_s\ge q_a-\delta$, we have $\xi'(q_a)-\xi'(q_s)\le\xi''(1)(q_a-q_s)\le M\delta$, and finally

$$\mathsf{E}_s(\zeta_a-\zeta_s)^2\le M\delta\ . \tag{14.473}$$

Therefore (14.472) implies that

$$\mathsf{E}_s A''_a(\zeta_s)\ge A''_a(\zeta_a)-M\sqrt{\delta}$$

and, taking square and expectation, (14.471) entails

$$\mathsf{E}A''_s(\zeta_s)^2\ge\mathsf{E}A''_a(\zeta_s)^2-M\sqrt{\delta}\ .$$

Using again (14.472) and (14.473) we get

$$\mathsf{E}A''_a(\zeta_s)^2\ge\mathsf{E}A''_a(\zeta_a)^2-M\sqrt{\delta}\ ,$$

and finally

$$\mathsf{E}A''_s(\zeta_s)^2\ge\mathsf{E}A''_a(\zeta_a)^2-M\sqrt{\delta}\ . \tag{14.474}$$

Moreover by the same argument as in (14.470), (14.469) implies

$$\mathsf{E}(W_1\cdots W_{s-1}A''_s(\zeta_s)^2)\ge\mathsf{E}A''_s(\zeta_s)^2-M\delta\ ,$$

and combining with (14.474),

$$\mathsf{E}(W_1\cdots W_{s-1}A''_s(\zeta_s)^2)\ge\mathsf{E}A''_a(\zeta_a)^2-M\sqrt{\delta}=\mathsf{E}A''_a(h+\chi\sqrt{\xi'(v)})^2-M\sqrt{\delta}\ .$$

Since $\xi''(v)\le\xi''(q_s)+M\delta$ (because $q_s\ge v-\delta$) this proves (14.462).

We turn to the proof of (14.469). The basic reason why this is true is that $m_p\le\delta$ for $p\le s-1$ so that each factor W_p is very close to 1. We have

$$W_1\cdots W_{s-1}=\exp S \tag{14.475}$$

where

$$S:=m_{s-1}A_s(\zeta_s)+\sum_{1\le p\le s-1}(m_{p-1}-m_p)A_p(\zeta_p)\ .$$

Now, (14.197) and Hölder's inequality imply

$$\exp 4B(x,v,m)\le(\mathsf{E}\exp A(x+g\sqrt{v}))^4\le\mathsf{E}\exp 4A(x+g\sqrt{v})\ ,$$

and iteration of this inequality shows that

$$\mathsf{E}\exp 4A_p(\zeta_p)\le M\ .$$

Since $m_{s-1}+\sum_{1\le p\le s-1}|m_{p-1}-m_p|\le 2m_{s-1}\le 2\delta$ and $\mathsf{E}A_p(\zeta_p)^4\le M$ we get

$$\mathsf{E}S^2\le\delta M\ .$$

We also have $A_p \geq 0$. Using simply that

$$\frac{m_{s-1}}{2} + \sum_{1 \leq p \leq s-1} \frac{|m_{p-1} - m_p|}{2} \leq m_{s-1} \leq 1$$

and Hölder's inequality

$$\mathsf{E} \exp 2|S| \leq (\mathsf{E} \exp 4A_s(\zeta_s))^{m_{s-1}/2} \prod_{1 \leq p \leq s-1} (\mathsf{E} \exp 4A_s(\zeta_s))^{|m_{p-1} - m_p|/2} , \tag{14.476}$$

we obtain that $\mathsf{E} \exp 2|S| \leq M$. Using the inequality

$$\mathsf{E}|\exp S - 1| \leq \mathsf{E}|S| \exp |S| \leq (\mathsf{E}S^2)^{1/2} (\mathsf{E} \exp 2|S|)^{1/2} \tag{14.477}$$

we have proved (14.469), and we turn to the proof of (14.468). We write

$$W_s \cdots W_{a-1} = \exp S^* \tag{14.478}$$

where

$$\begin{aligned} S^* &:= m_{a-1} A_a(\zeta_a) - m_s A_s(\zeta_s) + \sum_{s+1 \leq p \leq a-1} (m_{p-1} - m_p) A_p(\zeta_p) \qquad (14.479) \\ &= -m_s(A_s(\zeta_s) - A_a(\zeta_a)) + \sum_{s+1 \leq p \leq a-1} (m_{p-1} - m_p)(A_p(\zeta_p) - A_a(\zeta_a)) . \end{aligned}$$

The reason why S^* is close to 0 is that the terms $A_p(\zeta_p) - A_a(\zeta_a)$ are small. Indeed, iteration of (14.273) shows that for $s \leq p \leq a$ we have

$$A_a(x) \leq A_p(x) \leq A_a(x) + \xi'(q_a) - \xi'(q_p) .$$

Since $v = q_a \geq q_p \geq q_a - \delta = v - \delta$ for $s \leq p \leq a$, we have $0 \leq q_a - q_p \leq \delta$ and $\xi'(q_a) - \xi'(q_p) \leq M\delta$, so that

$$|A_p(x) - A_a(x)| \leq M\delta .$$

Since $|A'_a| \leq 1$, we have

$$\begin{aligned} |A_p(\zeta_p) - A_a(\zeta_a)| &\leq |A_p(\zeta_p) - A_a(\zeta_p)| + |A_a(\zeta_p) - A_a(\zeta_a)| \\ &\leq M\delta + |\zeta_p - \zeta_a| . \end{aligned} \tag{14.480}$$

Now,

$$\zeta_a - \zeta_p = \sum_{p \leq n < a} z_n$$

so that

$$\mathsf{E}(\zeta_a - \zeta_p)^2 = \xi'(q_a) - \xi'(q_p) \leq M\delta . \tag{14.481}$$

Therefore $\mathsf{E}(A_p(\zeta_p) - A_a(\zeta_a))^4 \leq M\delta^2$ and since

$$|m_s| + \sum_{s+1 \le p \le a-1} |m_{p-1} - m_p| \le 2 \tag{14.482}$$

we get that $\mathsf{E}S^{*2} \le M\delta$. Denoting by M_3 the constant in (14.480) and (14.481) we get

$$\mathsf{E}\exp 4|A_p(\zeta_p) - A_p(\zeta_a)| \le \exp(4M_3\delta + 4|g|) \ ,$$

where g is a Gaussian r.v. with

$$\mathsf{E}g^2 \le M_3\delta \ .$$

Consequently, if $M_3\delta \le 1$ we get $\mathsf{E}\exp 4|A_p(\zeta_p) - A_p(\zeta_a)| \le L$. It then follows from (14.479), (14.482) and Hölder's inequality as in (14.476) that $\mathsf{E}\exp 2|S^*| \le L$. The conclusion then follows from (14.477). □

Let us now describe the approach to Proposition 14.12.6. We choose the parameters in Proposition 14.12.4 as follows. We take $\lambda = 0$. As we just explained, without loss of generality we may assume that $v = q_a$ for a certain integer a. We then choose $\kappa = k+1$, $\rho_p = q_p$, $\tau = a$, $n_p = m_p$ if $p \ge a$ and $n_p = m_p/2$ if $p < a$. The key point is as follows.

Lemma 14.12.10. *Under the conditions of Proposition 14.12.6 the quantity Y_0 in the right-hand side of (14.424) satisfies*

$$Y_0 \le 2\mathsf{E}A_0(h) - \frac{1}{M} \ , \tag{14.483}$$

where M depends only on δ, $\boldsymbol{\beta}$ and h.

Proof of Proposition 14.12.6. Recalling the identity

$$\begin{aligned}&-2\sum_{p<\tau} n_p(\theta(\rho_{p+1}) - \theta(\rho_p)) - \sum_{\tau \le p \le \kappa} n_p(\theta(\rho_{p+1}) - \theta(\rho_p))\\ &= -\sum_{p \le k+1} m_p(\theta(q_{p+1}) - \theta(q_p))\end{aligned}$$

Proposition 14.12.6 then follows from (14.483). □

We will obtain (14.483) by improving upon the argument by which we have proved in the previous chapter that $Y_0 \le 2A_0(h)$. For convenience we reproduce this argument now.

Proceeding e.g as in (14.184) independence yields

$$Y_a = A_a\Big(h + \sum_{0 \le p < a} y_p^1\Big) + A_a\Big(h + \sum_{0 \le p < a} y_p^2\Big) \ .$$

Now we prove that if $p \le a$ we have

$$Y_p \le A_p\Big(h + \sum_{0\le r<p} y_r^1\Big) + A_p\Big(h + \sum_{0\le r<p} y_r^2\Big) .$$

The proof is by decreasing induction over p. For the induction step from $p+1$ to p we write, using the induction hypothesis in the first line and then the Cauchy-Schwarz inequality:

$$\begin{aligned} &\mathsf{E}_p \exp n_p Y_{p+1} \\ &\le \mathsf{E}_p \exp \frac{m_p}{2}\Big(A_{p+1}\Big(h + \sum_{0\le r\le p} y_r^1\Big) + A_{p+1}\Big(h + \sum_{0\le r\le p} y_r^2\Big)\Big) \\ &\le \prod_{j=1,2}\Big(\mathsf{E}_p \exp m_p A_{p+1}\Big(h + \sum_{0\le r\le p} y_p^j\Big)\Big)^{1/2} . \end{aligned} \tag{14.484}$$

Now, since y_p^1 and y_p^2 are distributed like z_p for $j = 1, 2$, by definition of A_p we have

$$\mathsf{E}_p \exp m_p A_{p+1}\Big(h + \sum_{0\le r\le p} y_r^j\Big) = \exp m_p A_p\Big(h + \sum_{0\le r<p} y_r^j\Big) ,$$

and taking logarithms in (14.484) completes the induction step. In this manner, we prove that

$$Y_0 \le 2\mathsf{E}A_0(h) . \tag{14.485}$$

Since $y_2^2 = -y_2^1$, we expect however that when $h \ne 0$, there cannot be equality in (14.484). We also expect that the small differences between the two sides of (14.484) keep accumulating over the values of p. To prove this we will require non-trivial estimates that do not depend on how many times we use the Cauchy-Schwarz inequality. A very important point is that we want that the accumulated small differences add to a quantity that is bounded below, depending only on $\boldsymbol{\beta}$ and h. There is equality in (14.484) when $m_p = 0$, so that to bound from below these differences we need to bound m_p from below. There is also equality in (14.484) when $y_p^1 = 0$, so that we can expect the difference between the two sides of (14.484) to be somewhat proportional to $\mathsf{E}(y_p^1)^2$. To have a chance to succeed we then need to control from below the sum of these quantiles for values of m_p (with $p \le a$) that are not too small. This control is provided by condition (14.427). This concludes our scheme of proof of Lemma 14.12.10, and we start the real proof.

Finding estimates that properly quantify the difference between the two sides of (14.484) is a nice exercise in elementary analysis. We state the main estimate first.

Proposition 14.12.11. *Consider a function $A(x)$, and assume that*

$$A(x) = A(-x) \ ; \qquad |A'| \le 1 \ ; \qquad A''(x) \le 1 \tag{14.486}$$

and that, for a certain number C

$$A''(x) \geq \frac{1}{C\mathrm{ch}^2 x} \,. \tag{14.487}$$

Then, if $0 < m \leq 1$, $0 \leq w \leq 10^{-2}$, $0 \leq \alpha \leq 1$ *and* z *is standard Gaussian, we have, for any numbers* x_1, x_2

$$\begin{aligned}
&\frac{2}{m}\log \mathsf{E}\exp\frac{m}{2}\Bigg(A(x_1+z\sqrt{w})+A(x_2+z\sqrt{w}) \\
&\quad -\frac{\alpha(x_1-x_2)^2}{\mathrm{ch}^4(x_1+z\sqrt{w})\mathrm{ch}^4(x_2+z\sqrt{w})}\Bigg) \\
&\leq \frac{1}{m}\log\mathsf{E}\exp mA(x_1+z\sqrt{w})+\frac{1}{m}\log\mathsf{E}\exp mA(x_2+z\sqrt{w}) \\
&\quad -\Big(\alpha(1-Lw)+\frac{mw}{LC^2}\Big)\frac{(x_1-x_2)^2}{\mathrm{ch}^4 x_1\mathrm{ch}^4 x_2} \,.
\end{aligned} \tag{14.488}$$

To understand this statement one should first consider the case $\alpha = 0$. In that case we get a quantitative improvement in the use of the Cauchy-Schwarz inequality. This improvement is quantified by the term

$$\frac{mw}{LC^2}\frac{(x_1-x_2)^2}{\mathrm{ch}^4 x_1\mathrm{ch}^4 x_2} \,.$$

When $\alpha \neq 0$, the term

$$\alpha\frac{(x_1-x_2)^2}{\mathrm{ch}^4 x_1\mathrm{ch}^4 x_2}$$

represents how much we have improved on the Cauchy-Schwarz inequality in the previous steps. Most of this improvement is preserved in the new step (as is shown by the term $\alpha(1-Lw)$), while some new improvement is gained (as is shown by the term mw/LC^2). That the improvements on the Cauchy-Schwarz inequality keep accumulating over the steps (at least for small α) is testified by the fact that

$$\alpha \leq \frac{m}{2L^2C^2} \quad \Rightarrow \quad \alpha(1-Lw)+\frac{mw}{LC^2} \geq \alpha + \frac{mw}{2LC^2} \,. \tag{14.489}$$

Proof of Lemma 14.12.10. Let us denote by L_0 the number of (14.488). We may and do assume that $L_0 \geq 1$. Since we can always insert new elements in the list $\mathbf{q}$, we may assume without loss of generality that

$$\forall p \,, \qquad \xi'(q_{p+1}) - \xi'(q_p) \leq \frac{1}{10^2 L_0} \,. \tag{14.490}$$

Let us define the function

$$Y_a(x_1, x_2) = A_a(x_1) + A_a(x_2)$$

and, recursively, for $1 \le p < a$

$$Y_p(x_1, x_2) = \frac{2}{m_p} \log \mathsf{E} \exp \frac{m_p}{2} Y_{p+1}(x_1 + y_p^1, x_2 + y_p^2) ,$$

and finally $Y_0(x_1, x_2) = \mathsf{E} Y_1(x_1 + y_0^1, x_2 + y_0^2)$. It should be obvious that

$$Y_0 = \mathsf{E} Y_0(h, h) . \tag{14.491}$$

We proved in Lemma 14.7.16 that there exists a constant C, depending on $\boldsymbol{\beta}$ only, such that

$$\forall p , \quad A_p''(x) \ge \frac{1}{C \mathrm{ch}^2 x} . \tag{14.492}$$

We prove that for $s \le p \le a$ we have

$$Y_p(x_1, x_2) \le A_p(x_1) + A_p(x_2) - \alpha_p \frac{(x_1 + x_2)^2}{\mathrm{ch}^4 x_1 \mathrm{ch}^4 x_2} \tag{14.493}$$

where

$$\alpha_p = \delta \min \left(\frac{1}{2 L_0^2 C^2}, \frac{\xi'(q_a) - \xi'(q_p)}{2 L_0 C^2} \right) . \tag{14.494}$$

The proof is by decreasing induction over p. This is true for $p = a$. For the induction from $p + 1$ to p, we observe that

$$A_{p+1}(x_2 + y_p^2) = A_{p+1}(x_2 - y_p^1) = A_{p+1}(-x_2 + y_p^1) .$$

We use (14.488) with $-x_2$ instead of x_2, and with

$$A = A_{p+1} , \quad \alpha = \alpha_{p+1} , \quad m = m_p , \quad w = \xi'(q_{p+1}) - \xi'(q_p) . \tag{14.495}$$

Since $m_p \ge m_s \ge \delta$ we obtain

$$Y_p(x_1, x_2) \le A_p(x_1) + A_p(x_2) - \alpha' \frac{(x_1 + x_2)^2}{\mathrm{ch}^4 x_1 \mathrm{ch}^4 x_2} \tag{14.496}$$

for

$$\alpha' = \alpha_{p+1}(1 - L_0 w) + \frac{\delta w}{L_0 C^2} = \alpha_{p+1} + w \left(\frac{\delta}{L_0 C^2} - \alpha_{p+1} L_0 \right) . \tag{14.497}$$

Since $\alpha_{p+1} \le \delta / 2 L_0^2 C^2$, we have

$$\alpha' \ge \alpha_{p+1} + \frac{w \delta}{2 L_0 C^2}$$

and thus, recalling the definition (14.495) of w we obtain

$$\begin{aligned}\alpha' &\geq \alpha_{p+1} + \delta\frac{(\xi'(q_{p+1}) - \xi'(q_p))}{2L_0C^2} \\ &= \delta\min\left(\frac{1}{2L_0C^2}, \frac{\xi'(q_a) - \xi'(q_{p+1})}{2L_0C^2}\right) + \delta\frac{(\xi'(q_{p+1}) - \xi'(q_p))}{2L_0C^2} \\ &\geq \delta\min\left(\frac{1}{2L_0C^2}, \frac{\xi'(q_a) - \xi'(q_p)}{2L_0C^2}\right) = \alpha_p \ .\end{aligned}$$

This proves that (14.493) holds for $p \geq s$ for the value of α_p given by (14.494). Since $\delta > 0$ depends only on $\boldsymbol{\beta}$ and h, and since $q_s \leq q_a - \delta$, we have $\alpha_s \geq 1/M$ (where M depends only on $\boldsymbol{\beta}$ and h). Next, we prove that for $p \leq s$, (14.493) holds for

$$\alpha_p = \alpha_s \exp\bigl(-2L_0(\xi'(q_s) - \xi'(q_p))\bigr) \ . \tag{14.498}$$

This is true for $p = s$, and the proof is again by decreasing induction over p. We use (14.488) as before, but since we no longer control m_p from below, instead of (14.497) we may only assert that

$$\alpha' \geq \alpha_{p+1}(1 - L_0 w) \geq \alpha_{p+1}\exp(-2L_0 w) \tag{14.499}$$

because $w = \xi'(q_{p+1}) - \xi'(q_p) \leq 1/2L_0$ by (14.490) and since $1 - x \geq \exp(-2x)$ for $x \leq 1/2$. Using again the value of w, (14.499) completes the proof that (14.493) holds for the value (14.498). Taking $p = 0$, we have shown that

$$Y_0(x_1, x_2) \leq A_0(x_1) + A_0(x_2) - \frac{1}{M}\frac{(x_1 + x_2)^2}{\mathrm{ch}^4 x_1 \mathrm{ch}^4 x_2} \ ,$$

so that if $\mathsf{E}h^2 > 0$

$$Y_0 = \mathsf{E}Y_0(h, h) \leq 2\mathsf{E}A_0(h) - \frac{1}{M} \ . \qquad \square$$

Let us observe the following simple fact. Since $(\log \mathrm{ch} x)'' = 1/\mathrm{ch}^2 x \leq 1$ we have

$$\log \mathrm{ch}(x_1 + x) \leq \log \mathrm{ch} x_1 + x \mathrm{th} x_1 + \frac{x^2}{2}$$

and thus

$$\mathrm{ch}(x_1 + x) \leq \mathrm{ch} x_1 \exp\Bigl(x \mathrm{th} x_1 + \frac{x^2}{2}\Bigr) \ . \tag{14.500}$$

We turn to the proof of Proposition 14.12.11. We first consider the case $\alpha = 0$, where this proposition improves upon the Cauchy-Schwarz inequality.

For $j = 1, 2$ let us set

$$X_j = \exp\frac{m}{2}A(x_j + z\sqrt{w}) \ ; \qquad b_j = (\mathsf{E}X_j^2)^{1/2} \ ,$$

and let us further set

$$b = \mathsf{E}X_1X_2 \ .$$

We assume without loss of generality that $x_1 \geq x_2$. The starting point of the proof is the identity

$$b_1^2 b_2^2 - b^2 = b_2^2 \mathsf{E}\left(X_1 - \frac{b}{b_2^2} X_2\right)^2 \tag{14.501}$$

and we try to find a lower bound for the right-hand side. Consider the function

$$f(t) = \frac{1}{2}(A(x_1 + t) - A(x_2 + t)) ,$$

so that

$$2f'(t) = A'(x_1 + t) - A'(x_2 + t) = \int_{x_2+t}^{x_1+t} A''(s)\mathrm{d}s \geq \frac{1}{C}\int_{x_2+t}^{x_1+t} \frac{\mathrm{d}s}{\mathrm{ch}^2 s} .$$

Now, for $|t| \leq 1$ we have, using (14.500) in the last inequality,

$$\begin{aligned} x_2 + t \leq s \leq x_1 + t \quad \Rightarrow \quad \mathrm{ch}s &\leq \max(\mathrm{ch}(x_1 + t), \mathrm{ch}(x_2 + t)) \\ &\leq \mathrm{ch}(x_1 + t)\mathrm{ch}(x_2 + t) \\ &\leq L\mathrm{ch}x_1\mathrm{ch}x_2 , \end{aligned}$$

and thus

$$|t| \leq 1 \quad \Rightarrow \quad f'(t) \geq \frac{x_1 - x_2}{L C \mathrm{ch}^2 x_1 \mathrm{ch}^2 x_2} , \tag{14.502}$$

so that if $0 \leq t_0 \leq 1$,

$$\begin{aligned} t_0 \leq t \leq 1 \quad &\Rightarrow \quad f(t) \geq f(t_0) \geq f(0) + \frac{t_0(x_1 - x_2)}{L C \mathrm{ch}^2 x_1 \mathrm{ch}^2 x_2} \\ -1 \leq t \leq -t_0 \quad &\Rightarrow \quad f(t) \leq f(-t_0) \leq f(0) - \frac{t_0(x_1 - x_2)}{L C \mathrm{ch}^2 x_1 \mathrm{ch}^2 x_2} . \end{aligned} \tag{14.503}$$

Let us assume for definiteness that $\exp mf(0) \geq b/b_2^2$. (The case where the reverse inequality holds is similar.) We observe that $X_1 = X_2 \exp mf(z\sqrt{w})$, and that since $w \leq 1$ we have

$$\frac{1}{2} \leq z \leq 1 \Rightarrow \frac{\sqrt{w}}{2} \leq z\sqrt{w} \leq 1 .$$

So it follows from (14.503) used for $t_0 = \sqrt{w}/2$ and $t = z\sqrt{w}$ that, using the inequality $\exp x \geq 1 + x$ in the second line and that $\exp mf(0) \geq b/b_2^2$ in the third line,

$$\begin{aligned} \frac{1}{2} \leq z \leq 1 \quad \Rightarrow \quad X_1 &\geq X_2 \exp\left(mf(0) + m\sqrt{w}\frac{x_1 - x_2}{L C \mathrm{ch}^2 x_1 \mathrm{ch}^2 x_2}\right) \\ &\geq X_2 \left(1 + \frac{m\sqrt{w}(x_1 - x_2)}{L C \mathrm{ch}^2 x_1 \mathrm{ch}^2 x_2}\right) \exp mf(0) \\ &\geq X_2 \left(1 + \frac{m\sqrt{w}(x_1 - x_2)}{L C \mathrm{ch}^2 x_1 \mathrm{ch}^2 x_2}\right) \frac{b}{b_2^2} . \end{aligned}$$

Therefore

$$\frac{1}{2} \le z \le 1 \quad \Rightarrow \quad X_1 - \frac{b}{b_2^2} X_2 \ge \frac{m\sqrt{w}(x_1 - x_2)}{LC\mathrm{ch}^2 x_1 \mathrm{ch}^2 x_2} \frac{b}{b_2^2} X_2 \ ,$$

and thus

$$\mathsf{E}\left(X_1 - \frac{b}{b_2^2} X_2\right)^2 \ge \frac{b^2}{b_2^4} \frac{m^2 w (x_2 - x_1)^2}{LC^2 \mathrm{ch}^4 x_1 \mathrm{ch}^4 x_2} \mathsf{E}(\mathbf{1}_{\{1/2 \le z \le 1\}} X_2^2) \ . \tag{14.504}$$

Let us assume for the moment that we know that

$$\mathsf{E}(\mathbf{1}_{\{1/2 \le z \le 1\}} X_2^2) \ge \frac{1}{L} \mathsf{E} X_2^2 = \frac{1}{L} b_2^2 \ . \tag{14.505}$$

Then, combining (14.501) and (14.504) yields

$$b_1^2 b_2^2 \ge b^2 \left(1 + \frac{m^2 w (x_2 - x_1)^2}{LC^2 \mathrm{ch}^4 x_1 \mathrm{ch}^4 x_2}\right) \ .$$

Since $\log(1+y) \ge y/2$ for $0 \le y \le 1$, taking logarithms and dividing by m implies (14.488) when $\alpha = 0$.

To prove (14.505) we simply use that, since $|A'| \le 1$ and $|A''| \le 1$, we have

$$|A(x_2 + z\sqrt{w}) - A(x_2)| \le |z|\sqrt{w} + \frac{z^2 w}{2} \ ,$$

so that

$$X_2^2 = \exp mA(x_2 + z\sqrt{w}) \le \exp mA(x_2) \exp(2|z|\sqrt{w} + z^2 w) \ ,$$

and

$$\mathbf{1}_{\{1/2 \le z \le 1\}} X_2^2 \ge \frac{1}{L} \exp mA(x_2) \mathbf{1}_{\{1/2 \le z \le 1\}} \ .$$

Taking expectation, and since $w \le 10^{-2}$, we obtain respectively

$$\mathsf{E} X_2^2 \le L \exp mA(x_2)$$

and

$$\mathsf{E}(\mathbf{1}_{\{1/2 \le z \le 1\}} X_2^2) \ge \frac{1}{L} \exp mA(x_2) \ .$$

To complete the proof of Proposition 14.12.11, the major step is as follows.

Lemma 14.12.12. *Consider a function $B(x)$ with $|B'| \le 1$, $|B''| \le 1$. Consider a number $0 \le w \le 10^{-2}$. Then if $0 < \gamma \le 5$ and $|d| \le 8$ we have, for $m \le 1$*

$$\frac{1}{m} \log \mathsf{E} \exp m\big(B(z\sqrt{w}) - \gamma \exp(dz\sqrt{w} - 4z^2 w)\big)$$
$$\le \frac{1}{m} \log \mathsf{E} \exp mB(z\sqrt{w}) - \gamma(1 - Lw) \ . \tag{14.506}$$

Proof of Proposition 14.12.11. The inequality (14.500) implies

$$\mathrm{ch}^4(x_1+x)\mathrm{ch}^4(x_2+x) \le \mathrm{ch}^4 x_1 \mathrm{ch}^4 x_2 \exp(4x(\mathrm{th}x_1+\mathrm{th}x_2)+4x^2)$$

so that if set

$$d = -4(\mathrm{th}x_1+\mathrm{th}x_2)\ ; \qquad \gamma = \frac{\alpha(x_1-x_2)^2}{\mathrm{ch}^4 x_1 \mathrm{ch}^4 x_2}$$

then

$$-\frac{\alpha(x_1-x_2)^2}{\mathrm{ch}^4(x_1+z\sqrt{w})\mathrm{ch}^4(x_2+z\sqrt{w})} \le -\gamma \exp(dz\sqrt{w}-4z^2w)\ ,$$

and therefore setting

$$B(x) = \frac{1}{2}(A(x_1+x)+A(x_2+x))$$

we get

$$\begin{aligned}&\frac{1}{m}\log \mathsf{E}\exp\frac{m}{2}\Bigg(A(x_1+z\sqrt{w})+A(x_2+z\sqrt{w})\\ &-\frac{\alpha(x_1-x_2)^2}{\mathrm{ch}^4(x_1+z\sqrt{w})\mathrm{ch}^4(x_2+z\sqrt{w})}\Bigg)\\ &\le \frac{1}{m}\log\mathsf{E}\exp m\left(B(z\sqrt{w})-\frac{\gamma}{2}\exp(dz\sqrt{w}-4z^2w)\right)\ .\end{aligned} \tag{14.507}$$

Trivial bounds show that $\gamma \le 10$. Using (14.506) for $\gamma/2$ rather than γ, we obtain

$$\begin{aligned}&\frac{2}{m}\log\mathsf{E}\exp m\left(B(z\sqrt{w})-\frac{\gamma}{2}\exp(dz\sqrt{w}-4z^2w)\right)\\ &\le \frac{2}{m}\log\mathsf{E}\exp mB(z\sqrt{w})-\gamma(1-Lw)\ ,\end{aligned}$$

and combining with (14.507) and recalling the definition of B, we get

$$\begin{aligned}&\frac{2}{m}\log\mathsf{E}\exp\frac{m}{2}\Bigg(A(x_1+z\sqrt{w})+A(x_2+z\sqrt{w})\\ &-\frac{\alpha(x_1-x_2)^2}{\mathrm{ch}^4(x_1+z\sqrt{w})\mathrm{ch}^4(x_2+z\sqrt{w})}\Bigg)\\ &\le \frac{2}{m}\log\mathsf{E}\exp\frac{m}{2}(A(x_1+z\sqrt{w})+A(x_2+z\sqrt{w}))-\gamma(1-Lw)\ .\end{aligned}$$

Combining with the case $\alpha = 0$ of (14.488) (that we previously proved) we have finished the proof of Proposition 14.12.11. □

Proof of Lemma 14.12.12. We first reduce to the case $d = 0$ by considering for $0 \le t \le d$ the function

$$\varphi(t) = \frac{1}{m} \log \mathsf{E} V(t)$$

where

$$V(t) = \exp m\big(B(z\sqrt{w}) - \gamma \exp(tz\sqrt{w} - 4z^2 w)\big) .$$

The goal is to prove that

$$|\varphi'(t)| \le Lw\gamma . \tag{14.508}$$

We compute $\varphi'(t)$,

$$\varphi'(t) = \frac{\gamma}{\mathsf{E} V(t)} \mathsf{E}(-z\sqrt{w} \exp(tz\sqrt{w} - 4z^2 w)V(t)) , \tag{14.509}$$

and we integrate by parts in z,

$$\mathsf{E}(-z\sqrt{w} \exp(tz\sqrt{w} - 4z^2 w)V(t)) = \mathsf{E}(DV(t)) \tag{14.510}$$

where

$$\begin{aligned} &D = \exp(tz\sqrt{w} - 4z^2 w) \\ &\times \Big(-tw + 8zw^{3/2} + m\gamma(tw - 8zw^{3/2}) \exp(tz\sqrt{w} - 4z^2 w) - mwB'(z\sqrt{w})\Big) . \end{aligned}$$

This looks intractable, but we notice that for $z^2 w \le 1$, since $0 \le t \le d \le 8$ we have

$$|tw - 8zw^{3/2}| \le Lw \; ; \; \exp(tz\sqrt{w} - 4z^2 w) \le L ,$$

and therefore

$$z^2 w \le 1 \quad \Rightarrow \quad |D| \le Lw .$$

Since, recalling (14.509) and (14.510)

$$\varphi'(t) = \frac{\gamma}{\mathsf{E} V(t)} \mathsf{E}(DV(t))$$

we then get

$$|\varphi'(t)| \le L\gamma w + \frac{\gamma}{\mathsf{E} V(t)} \mathsf{E}(\mathbf{1}_{\{z^2 w \ge 1\}} DV(t)) . \tag{14.511}$$

We use Hölder's inequality to obtain

$$\mathsf{E}(\mathbf{1}_{\{z^2 w \ge 1\}} DV(t)) \le \mathsf{P}\left(|z| \ge \frac{1}{\sqrt{w}}\right)^{1/2} (\mathsf{E} D^4)^{1/4} (\mathsf{E} V(t)^4)^{1/4} . \tag{14.512}$$

Using that $|B'| \le 1$ in the second inequality we deduce

$$V(t) \le \exp mB(z\sqrt{w}) \le \exp m(B(0) + |z|\sqrt{w})$$

and since $w \le 1$ we obtain

$$(\mathsf{E}V(t)^4)^{1/4} \le L \exp mB(0) \; .$$

We note also that, when $|z| \le 1$, and since $w \le 1$ we have $z^2 w \le 1$, so that since $|B'| \le 1$ we have $B(z\sqrt{w}) \ge B(0) - L$ and therefore since $\gamma \le 5$ we have $V(t) \ge (\exp mB(0))/L$. Consequently,

$$\mathsf{E}V(t) \ge \mathsf{E}(\mathbf{1}_{\{|z|\le 1\}} V(t)) \ge \frac{1}{L} \exp mB(0) \; .$$

We observe that

$$D = D^* \Big(-tw + 8zw^{3/2} + m\gamma(tw - 8zw^{3/2})D^* - mwB'(z\sqrt{w}) \Big) \; ,$$

where $D^* = \exp(tz\sqrt{w} - 4z^2 w)$. Since $w \le 10^{-2}$ we have $\mathsf{E}D^{*10} \le L$. Moreover

$$|D| \le LD^*(1 + |z|)(1 + D^*) \; ,$$

so that using Hölder's inequality again we get $\mathsf{E}D^4 \le L$. Since $\mathsf{P}(|z| \ge 1/\sqrt{w}) \le 2\exp(-1/2w) \le Lw^2$, combining these estimates with (14.511) we have proved (14.508). Thus

$$\varphi(d) \le \varphi(0) + L\gamma w \tag{14.513}$$

i.e.

$$\begin{aligned} &\frac{1}{m} \log \mathsf{E} \exp m(B(z\sqrt{w}) - \gamma \exp(dz\sqrt{w} - 4z^2 w)) \\ &\le \frac{1}{m} \log \mathsf{E} \exp m(B(z\sqrt{w}) - \gamma \exp(-4z^2 w)) + L\gamma w \; . \end{aligned} \tag{14.514}$$

We now consider the function

$$\psi(t) = \frac{1}{m} \log \mathsf{E} \exp m(B(z\sqrt{w}) - \gamma \exp(-tz^2 w)) \; ,$$

and we proceed very much as above to prove that

$$\psi(4) \le \psi(0) + L\gamma w = -\gamma + L\gamma w + \frac{1}{m} \log \mathsf{E} \exp mB(z\sqrt{w}) \; ,$$

which, when combined with (14.512) completes the proof. □

It remains now (in the unlikely event that any reader has resisted up to this point) to prove (14.420) in the case where $0 \le u \le c'$. Rather than Proposition 14.12.4 we use another special case of Proposition 14.6.3. Consider

$$0 = \rho_0 \le \rho_1 = u \le \rho_2 \le \ldots \le \rho_{\kappa+1} = 1$$

and $n_0, \ldots, n_\kappa$ as in (14.422). Consider independent pairs (y_p^1, y_p^2) of Gaussian r.v.s as in (14.423), but now with

$$y_0^2 = y_0^1 \; ; \; y_p^1 \text{ and } y_p^2 \text{ are independent if } p \geq 1 .$$

Then the bound (14.424) holds for $v = -u$.

The point c is the smallest point of the support of the Parisi measure μ, and by definition μ is the limit of a sequence of stationary measures μ_n such that $\mathcal{P}(\xi, \mu_n) \to \mathcal{P}(\xi)$. If c_1 is a given point with $c' < c_1 < c$, given any $\varepsilon > 0$, for large n we have $\mu_n([-1, c_1]) < \varepsilon$ and $\mathcal{P}(\xi, \mu_n) \leq \mathcal{P}(\xi) + \varepsilon$. By definition of stationarity, the measure μ_n corresponds to k, $\mathbf{m}$ and $\mathbf{q}$ that satisfy condition MIN(ε). Consider the smallest $s \leq k+2$ for which $c_1 < q_s$. Then $\mu_n([-1, c_1]) < \varepsilon = m_{s-1}$, so that

$$m_{s-1} \leq \varepsilon \, .$$

This condition implies in turn as in (14.469) that

$$\mathsf{E}|W_1 \cdots W_{s-1} - 1| \leq M\varepsilon \, . \tag{14.515}$$

We define $\kappa = k + 3 - s$ and

$$0 = \rho_0 \leq \rho_1 = u \leq \rho_2 = q_s \leq \ldots \leq \rho_{\kappa+1} = q_{k+2} = 1 \, .$$

With these choices consider the right-hand side of (14.424) as a function $\alpha(\lambda)$. As in Lemma 14.12.7 this implies

$$\alpha(0) \leq 2\mathcal{P}(\xi, h) + M\varepsilon \, . \tag{14.516}$$

Lemma 14.12.13. *If χ denotes a Gaussian r.v. with $\mathsf{E}\chi^2 = \xi'(q_s)$, then we have*

$$|q_s - \mathsf{E}A_s'(h+\chi)^2| \leq M\varepsilon \tag{14.517}$$

$$\xi''(q_s)\mathsf{E}A_s''(h+\chi)^2 \leq 1 + M\varepsilon^{1/6} \, . \tag{14.518}$$

Proof. To prove (14.517) we combine (14.515) with (14.222), and to prove (14.518) we combine it with (14.461) (that remains true with the same proof). □

Proposition 14.6.4 (and specifically (14.180)) implies

$$\alpha'(0) = \mathsf{E}A_a'(h+\chi_1)A_a'(h+\chi_2) - u \, , \tag{14.519}$$

where $\mathsf{E}\chi_1^2 = \mathsf{E}\chi_2^2 = \xi'(q_s)$ and $\mathsf{E}\chi_1\chi_2 = \xi'(u)$. The key point is the following:

Lemma 14.12.14. *There exists a number $\gamma > 0$ depending only on ξ, h, c and c' such that*

$$\mathsf{E}A_s'(h+\chi)^2 \leq \mathsf{E}A_s'(h+\chi_1)A_s'(h+\chi_2) + (\xi'(q_s) - \xi'(u))\mathsf{E}A_s''(h+\chi)^2 - \gamma \, . \tag{14.520}$$

Let us accept this for a moment and finish the proof of Theorem 14.12.1. Since $\xi'(q_s) - \xi'(u) \le \xi''(q_s)(q_s - u)$, combining (14.520) with (14.518) yields

$$\mathsf{E}A_s'(h+\chi)^2 \le \mathsf{E}A_s'(h+\chi_1)A_s'(h+\chi_2) + q_s - u - \gamma + M\varepsilon^{1/6} .$$

It then follows from (14.519) and (14.517) that

$$\alpha'(0) \ge \gamma + \mathsf{E}A_s'(h+\chi)^2 - q_s - M\varepsilon^{1/6} \ge \gamma - M\varepsilon^{1/6} - M\varepsilon ,$$

so that when ε is small enough we have $\alpha'(0) \ge \gamma/2$, and (14.420) then follows from (14.431) and (14.516) when ε is small enough.

Proof of Lemma 14.12.14. We revisit the proof of Lemma 14.9.5. This proof used the Cauchy-Schwarz inequality, in a case where there cannot be equality because the function A_s'' is not constant, which follows from (14.274) and the fact that

$$\int_0^\infty A_s''(x)\mathrm{d}x \le 1$$

since $A_s'(x) \le 1$. The fact that the "non equality" is uniform over all parameters involved requires only straightforward but tedious estimates, starting with the identity (14.501). □

The proof of Lemma 14.12.14 raises the question of how one could control the higher derivatives of the functions A_p. For example, is it true that $A_p^{(3)}(x) > 0$ for $x > 0$?

14.13 Notes and Comments

The possible relevance of Poisson-Dirichlet cascades to the SK model was put forward in particular in [16] and in [11]. However Guerra's original proof of Theorem 14.4.3 [49] does not use Poisson-Dirichlet cascades, and is entirely written using the recursive construction of the r.v.s F_p. This is also the case of the author's original proof of (14.93). We have chosen here to use Poisson-Dirichlet cascades, despite the fact that the proof is then quite longer, because they allow to bring forward simple principles such as Lemma 14.4.1, and since they seem to be intrinsically relevant. It is however remarkable that Guerra's original scheme is somewhat more powerful than the use of Poisson-Dirichlet cascades. This is apparent when one studies the setting of Section 12.5, and is explained in detail in [111], but much remains to be understood.

Research Problem 14.13.1. Given $a < 0$, compute (at least for p even)

$$\lim_{N\to\infty} \frac{1}{Na} \log \mathsf{E}Z_N^a , \tag{14.521}$$

where Z_N is the partition function corresponding to the Hamiltonian (14.57).

When $a \geq 0$, the limit (14.521) is computed in [111]. When $a \geq 1$, this limit is given by a "replica-symmetric formula" similar to (1.395). The key argument towards this result is given in Section 12.5.

The results of Section 14.3 have been discovered by analyzing what makes the calculations of Guerra work when they are written in terms of Poisson-Dirichlet cascades. Of course direct probabilistic proofs can be found, but it is interesting to note that these results do not seem to have been identified earlier. Section 14.11 follows the presentation of D. Panchenko [69] of the results of [109].

15. The Parisi Solution

15.1 Introduction

The Parisi formula that we proved in the previous chapter is only the tip of the iceberg. Underneath lies a very beautiful structure, that we call the Parisi Solution. This structure is only partially understood at the rigorous level. In this chapter we describe it at the heuristic level, and we rigorously prove that significant parts of it indeed hold true. As we pointed out, it is not necessary to have mastered all the material of the previous chapter to enjoy the present one. It probably suffices to have a fair understanding of Sections 14.2 to 14.6.

15.2 Ghirlanda-Guerra Identities and Poisson Dirichlet Cascades

In this section we explore a fundamental relationship between the Ghirlanda-Guerra identities and Poisson-Dirichlet cascades of Section 14.2. We consider a Poisson-Dirichlet cascade associated to a sequence $0 < m_1 < \ldots < m_k < 1$. Let us think of the weights (v_α) as defining a random probability measure G on $A = \mathbb{N}^{*k}$; for a function h of "replicas" $\alpha^1, \ldots, \alpha^n \in A$ we define

$$\langle h \rangle = \sum_{\alpha^1, \ldots, \alpha^n} v_{\alpha^1} \cdots v_{\alpha^n} h(\alpha^1, \ldots, \alpha^n) \ . \tag{15.1}$$

That is, each variable α^ℓ is integrated independently according to G. We write $\nu(h) = \mathsf{E}\langle h \rangle$. For $\alpha, \gamma \in A$ we recall that we denote (α, γ) "the first coordinate on which the sequences α and γ differ", see (14.36).

Theorem 15.2.1. (Ghirlanda-Guerra identities for Poisson-Dirichlet cascades) *Consider a function h of $\alpha^1, \ldots, \alpha^n$ that depends on $\alpha^1, \ldots, \alpha^n$ only through the integers $(\alpha^\ell, \alpha^{\ell'})$ for $1 \le \ell < \ell' \le n$. Then for each $1 \le r \le k+1$ we have*

$$\nu(\mathbf{1}_{\{(\alpha^1, \alpha^{n+1}) = r\}} h) = \frac{1}{n} \nu(\mathbf{1}_{\{(\alpha^1, \alpha^2) = r\}}) \nu(h) + \frac{1}{n} \sum_{2 \le \ell \le n} \nu(\mathbf{1}_{\{(\alpha^1, \alpha^\ell) = r\}} h) \ . \tag{15.2}$$

M. Talagrand, *Mean Field Models for Spin Glasses*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys in Mathematics 55, DOI 10.1007/978-3-642-22253-5_8,

For the proof, we fix $1 \le r \le k+1$ once and for all, and we consider independent standard Gaussian r.v.s $(g_{j_1,\ldots,j_r})$ for $j_1,\ldots,j_r \in \mathbb{N}^*$, and for $\alpha = (j_1,\ldots,j_k) \in A$ we write $g_\alpha = g_{j_1,\ldots,j_r}$. Thus g_α depends only on $\alpha|r = (j_1,\ldots,j_r)$.

Lemma 15.2.2. *For each t we have*

$$\mathsf{E}\frac{\sum_\alpha v_\alpha g_\alpha^2 \exp t g_\alpha}{\sum_\alpha v_\alpha \exp t g_\alpha} - \left(\mathsf{E}\frac{\sum_\alpha v_\alpha g_\alpha \exp t g_\alpha}{\sum_\alpha v_\alpha \exp t g_\alpha}\right)^2 = 1\ . \tag{15.3}$$

Proof. Let us recall the following elementary fact (A.6): if g is a standard Gaussian r.v. and $a < 1/2$ we have

$$\mathsf{E}\exp(ag^2 + bg) = \frac{1}{\sqrt{1-2a}} \exp \frac{b^2}{2(1-2a)}\ . \tag{15.4}$$

Consider $s < 1/2$. Using (14.8) for the function $F(x_1,\ldots,x_k) = \exp(sx_r^2 + tx_r)$ entails

$$\begin{aligned}\mathsf{E}\log\sum_\alpha v_\alpha \exp(sg_\alpha^2 + tg_\alpha) &= \frac{1}{m_r}\log\mathsf{E}\exp(m_r s g^2 + m_r t g) \qquad (15.5)\\ &= \frac{m_r t^2}{2(1-2m_r s)} - \frac{1}{2m_r}\log(1-2m_r s)\ ,\end{aligned}$$

using (15.4) in the second line. Taking $s=0$ and differentiating in t gives

$$\mathsf{E}\frac{\sum_\alpha v_\alpha g_\alpha \exp t g_\alpha}{\sum_\alpha v_\alpha \exp t g_\alpha} = m_r t\ . \tag{15.6}$$

Differentiating in s at $s = 0$ gives

$$\mathsf{E}\frac{\sum_\alpha v_\alpha g_\alpha^2 \exp t g_\alpha}{\sum_\alpha v_\alpha \exp t g_\alpha} = m_r^2 t^2 + 1\ .$$

This proves the result. □

Let us define

$$v_{\alpha,t} = \frac{v_\alpha \exp t g_\alpha}{\sum_{\gamma\in A} v_\gamma \exp t g_\gamma}$$

and define $\langle\cdot\rangle_t$ as $\langle\cdot\rangle$ in (15.1), but with $v_{\alpha,t}$ instead of v_α. If h' is a (possibly random) function on A^n, we then write $\nu_t(h') = \mathsf{E}\langle h'\rangle_t$, where the expectation E is as always over all sources of randomness. Then (15.3) means that if the (random) function $u : A \to \mathbb{R}$ is given by $u(\alpha) = g_\alpha$, we have

$$\nu_t(u^2) - \nu_t(u)^2 = \nu_t((u - \nu_t(u))^2) = \nu_t((u - m_r t)^2) = 1\ ,$$

using that $\nu_t(u) = m_r t$ by (15.6).

Therefore, if $h : A^n \to \mathbb{R}$ is such that $|h| \le 1$, and if $u^1(\alpha^1, \dots, \alpha^n) = u(\alpha^1) = g_{\alpha^1}$, then

$$|\nu_t(u^1 h) - m_r t \nu_t(h)| = |\nu_t((u^1 - m_r t)h)| \le \nu_t(|u - m_r t|) \\ \le \nu_t((u - m_r t)^2)^{1/2} \le 1 \; . \tag{15.7}$$

On the other hand, we have

$$\nu_t(u^1 h) = \mathsf{E} \frac{\sum_{\alpha^1,\dots,\alpha^n} g_{\alpha^1} h(\alpha^1, \dots, \alpha^n) v_{\alpha^1} \cdots v_{\alpha^n} \exp t(g_{\alpha^1} + \cdots + g_{\alpha^n})}{\left(\sum_\gamma v_\gamma \exp t g_\gamma\right)^n} \; . \tag{15.8}$$

Now

$$\mathsf{E} g_{\alpha^1} g_{\alpha^\ell} = 1 \quad \text{if } (\alpha^1, \alpha^\ell) > r$$

and is 0 otherwise. Integration by parts of (15.8) in the r.v.s g_{α^1} gives

$$\nu_t(u^1 h) = t \Big(\nu_t(h) + \sum_{2 \le \ell \le n} \nu_t(\mathbf{1}_{\{(\alpha^1,\alpha^\ell) > r\}} h) - n \nu_t(\mathbf{1}_{\{(\alpha^1,\alpha^{n+1}) > r\}} h) \Big) \; ,$$

and combining with (15.7) yields

$$t \Big((1 - m_r) \nu_t(h) + \sum_{2 \le \ell \le n} \nu_t(\mathbf{1}_{\{(\alpha^1,\alpha^\ell) > r\}} h) - n \nu_t(\mathbf{1}_{\{(\alpha^1,\alpha^{n+1}) > r\}} h) \Big) = \mathcal{R} \tag{15.9}$$

where $|\mathcal{R}| \le 1$.

Lemma 15.2.3. *If h depends on $\alpha^1, \dots, \alpha^n$ only through the quantities $(\alpha^\ell, \alpha^{\ell'})$ for $1 \le \ell < \ell' \le n$, then $\nu_t(h) = \nu(h)$.*

Proof of Theorem 15.2.1. Combining (15.9) with Lemma 15.2.3 we obtain

$$t \Big((1 - m_r) \nu(h) + \sum_{2 \le \ell \le n} \nu(\mathbf{1}_{\{(\alpha^1,\alpha^\ell) > r\}} h) - n \nu(\mathbf{1}_{\{(\alpha^1,\alpha^{n+1}) > r\}} h) \Big) = \mathcal{R}$$

where $|\mathcal{R}| \le 1$. Letting $t \to \infty$ yields

$$n \nu(\mathbf{1}_{\{(\alpha^1,\alpha^{n+1}) > r\}} h) = (1 - m_r) \nu(h) + \sum_{2 \le \ell \le n} \nu(\mathbf{1}_{\{(\alpha^1,\alpha^\ell) > r\}} h) \; .$$

We write the same relation for $r - 1$ instead of r. Subtraction and the fact that from (14.39) we have $m_r - m_{r-1} = \nu(\mathbf{1}_{\{(\alpha^1,\alpha^2) = r\}})$ complete the proof of (15.2). □

Proof of Lemma 15.2.3. The principle of the proof is to show that there is a (random) permutation σ of A with the following properties:

The sequence $(v_{\sigma(\alpha),t})_\alpha$ is distributed like the sequence $(v_\alpha)_\alpha$. (15.10)

$$\forall \alpha, \gamma \in A\,, \qquad (\sigma(\alpha), \sigma(\gamma)) = (\alpha, \gamma)\,. \tag{15.11}$$

It then follows from (15.11) that, since h depends on $\alpha^1, \ldots, \alpha^n$ only through the quantities $(\alpha^\ell, \alpha^{\ell'})$, we have

$$h(\alpha^1, \ldots, \alpha^n) = h(\sigma(\alpha^1), \ldots, \sigma(\alpha^n))\,. \tag{15.12}$$

Therefore, using (15.12) in the third line and (15.10) in the last line,

$$\begin{aligned}\nu_t(h) &= \mathsf{E} \sum_{\alpha^1, \ldots, \alpha^n} v_{\alpha^1,t} \cdots v_{\alpha^n,t} h(\alpha^1, \ldots, \alpha^n) \\ &= \mathsf{E} \sum_{\alpha^1, \ldots, \alpha^n} v_{\sigma(\alpha^1),t} \cdots v_{\sigma(\alpha^n),t} h(\sigma(\alpha^1), \ldots, \sigma(\alpha^n)) \\ &= \mathsf{E} \sum_{\alpha^1, \ldots, \alpha^n} v_{\sigma(\alpha^1),t} \cdots v_{\sigma(\alpha^n),t} h(\alpha^1, \ldots, \alpha^n) \\ &= \nu(h)\,.\end{aligned}$$

We turn to the construction of σ. Recalling the notation

$$u^*_\alpha = u_{\alpha|1} u_{\alpha|2} \cdots u_{\alpha|k-1} u_\alpha = u_{j_1} u_{j_1 j_2} \cdots u_{j_1 \ldots j_k}$$

of (14.1) we first notice that

$$v_{\alpha,t} = \frac{u^*_\alpha \exp t g_\alpha}{\sum u^*_\gamma \exp t g_\gamma}\,. \tag{15.13}$$

Next, it follows from Corollary 13.1.2 that for each $j_1, \ldots, j_{r-1}$ there is a permutation $\theta_{j_1,\ldots,j_{r-1}}$ of $\mathbb{N}^*$ such that the sequence

$$\big(u_{j_1 \ldots j_{r-1} \theta_{j_1,\ldots,j_{r-1}}(j)} \exp t g_{j_1,\ldots,j_{r-1},\theta_{j_1,\ldots,j_{r-1}}(j)}\big)_{j \geq 1}$$

has the same distribution as the sequence

$$(c u_{j_1 \ldots j_{r-1} j})_{j \geq 1}$$

where $c = (\mathsf{E} \exp t m_r g)^{1/m_r}$, for a standard Gaussian r.v. g. Moreover the randomness of $\theta_{j_1,\ldots,j_{r-1}}$ depends only on the randomness of the variables $u_{j_1,\ldots,j_{r-1} j}$. Let us then define

$$\sigma(\alpha) = \sigma((j_1, \ldots, j_k)) = (j_1, j_2, \ldots, j_{r-1}, \theta_{j_1,\ldots,j_{r-1}}(j_r), j_{r+1}, \ldots, j_k)\,.$$

We first show that the family $(u^*_{\sigma(\alpha)} \exp t g_{\sigma(\alpha)})_\alpha$ is distributed like the family $(c u^*_\alpha)_\alpha$. First we note that by definition of $g_\alpha = g_{j_1,\ldots,j_r}$ we have

$$g_{\sigma(\alpha)} = g_{j_1,\ldots,j_{r-1},\theta_{j_1,\ldots,j_{r-1}}(j_r)}\,.$$

Now, since

$$u^*_\alpha = u_{j_1} u_{j_1 j_2} \cdots u_{j_1 j_2 \ldots j_k}$$

we have

$$c^{-1} u^*_{\sigma(\alpha)} \exp t g_{\sigma(\alpha)} = u'_{j_1} u'_{j_1 j_2} \cdots u'_{j_1 j_2 \ldots j_k} \ ,$$

where, for $p < r$,

$$u'_{j_1 \ldots j_p} = u_{j_1 \ldots j_p} \ ,$$

where

$$u'_{j_1 \ldots j_r} = c^{-1} u_{j_1 \ldots j_{r-1} \theta_{j_1,\ldots,j_{r-1}}(j_r)} \exp t g_{j_1,\ldots,j_{r-1},\theta_{j_1,\ldots,j_{r-1}}(j_r)} \ ,$$

and where, for $p > r$,

$$u'_{j_1 \ldots j_p} = u_{j_1 \ldots j_{r-1} \theta_{j_1,\ldots,j_{r-1}}(j_r) j_{r+1} \ldots j_p} \ .$$

The collection of sequences $u'_{j_1 \ldots j_p}$ is distributed like the collection of sequences $u_{j_1 \ldots j_p}$, and therefore the family $(c^{-1} u^*_{\sigma(\alpha)} \exp t g_{\sigma(\alpha)})_\alpha$ is distributed like the family $(u^*_\alpha)_\alpha$. Consequently by (15.13) the family $(v_{\sigma(\alpha),t})_\alpha$ is distributed like the family $(v_\alpha)_\alpha$. It remains to prove (15.11); this follows from the fact that for any s, the first s components of $\sigma(\alpha)$ are determined by the first s components $j_1, \ldots, j_s$ of α (and conversely). □

We end up this section by a kind of converse to Theorem 15.2.1 in the case $k = 1$ which is of fundamental importance. In this setting, quantities $\mathbf{1}_{\{\alpha^\ell = \alpha^{\ell'}\}}$ play the rôle of the overlaps. So, to lighten notation we write

$$R_{\ell,\ell'} = \mathbf{1}_{\{\alpha^\ell = \alpha^{\ell'}\}} = \mathbf{1}_{\{(\alpha^\ell, \alpha^{\ell'}) = 2\}} \ ,$$

so that (15.2) implies that

$$\nu(R_{1,n+1} h) = \frac{1}{n} \nu(R_{1,2}) \nu(h) + \frac{1}{n} \sum_{2 \le \ell \le n} \nu(R_{1,\ell} h) \ . \tag{15.14}$$

Of special interest is the function h defined as follows. Consider integers $n_1, \ldots, n_k$ and $n = \sum_{s \le k} n_s$. Consider a partition of $\{1, \ldots, n\}$ into sets $I_1, \ldots, I_k$ with $\mathrm{card} I_s = n_s$. We define $h = h(\alpha^1, \ldots, \alpha^n) = 1$ if

$$\forall s \le k \ , \quad \ell, \ell' \in I_s \quad \Rightarrow \quad \alpha^\ell = \alpha^{\ell'} \ ,$$

and we define $h = 0$ otherwise. If we define h_s by

$$h_s = 1 \text{ if } \alpha^\ell = \alpha^{\ell'} \text{ for all } \ell, \ell' \in I_s, \text{ and } h_s = 0 \text{ otherwise,} \tag{15.15}$$

then $\langle h_s \rangle = \sum_\alpha v_\alpha^{n_s}$, and $h = \prod_{s \le k} h_s$, so that by independence between replicas,

$$\langle h \rangle = \prod_{s \le k} \langle h_s \rangle = \prod_{s \le k} \sum_\alpha v_\alpha^{n_s}$$

and

$$\nu(h) = \mathsf{E} \prod_{s \le k} \sum_{\alpha} v_\alpha^{n_s} \ .$$

The sequence (v_α) has a Poisson-Dirichlet distribution; let us call m its parameter. We define

$$S^{(m)}(n_1, \ldots, n_k) = \mathsf{E} \prod_{s \le k} \sum_{\alpha} v_\alpha^{n_s} \ , \tag{15.16}$$

so that

$$\nu(h) = S^{(m)}(n_1, \ldots, n_k) \ . \tag{15.17}$$

We observe that by (13.17) we have

$$\nu(R_{1,2}) = \mathsf{E} \sum_{\alpha} v_\alpha^2 = 1 - m \ , \tag{15.18}$$

and our goal is now to apply (15.14) to the function h. We compute the various terms. We recall (15.15). Assuming without loss of generality that $1 \in I_1$, then

$$R_{1,n+1} h_1 = h_1'$$

where $h_1' = 1$ if $\alpha^\ell = \alpha^{\ell'}$ for each $\ell, \ell' \in I_1 \cup \{n+1\}$ and $h_1' = 0$ otherwise. That is, the function $R_{1,n+1} h_1$ has the same structure as the function h_1 except that we have replaced n_1 by $n_1 + 1$. Thus by (15.17) we get

$$\nu(R_{1,n+1} h) = S^{(m)}(n_1 + 1, n_2, \ldots, n_k) \ .$$

Assuming that $\ell \in I_s$ with $s \ge 2$, we have $R_{1,\ell} h_1 h_s = 1$ if $\alpha^\ell = \alpha^{\ell'}$ for $\ell, \ell' \in I_1 \cup I_s$ and $R_{1,\ell} h_1 h_s = 0$ otherwise. In words, multiplying h by $R_{1,\ell}$ amounts to merge the sets I_1 and I_s. Therefore by (15.17) we get

$$\nu(R_{1,\ell} h) = S^{(m)}(n_2, \ldots, n_{s-1}, n_s + n_1, n_{s+1}, \ldots, n_k) \ .$$

Of course if $\ell \in I_1$ we have $R_{1,\ell} h = h$, so that $\nu(R_{1,\ell} h) = S^{(m)}(n_1, \ldots, n_k)$. Consequently (15.14) implies that

$$\begin{aligned} & S^{(m)}(n_1 + 1, n_2, \ldots, n_k) \\ &= \frac{n_1 - m}{n} S^{(m)}(n_1, \ldots, n_k) \\ &+ \sum_{2 \le s \le k} \frac{n_s}{n} S^{(m)}(n_2, \ldots, n_{s-1}, n_s + n_1, n_{s+1}, \ldots, n_k) \ . \end{aligned} \tag{15.19}$$

The terms in the right-hand side are in a sense simpler than the term in the left-hand side. We have either decreased k or decreased $\sum_{s \le k} n_s$. We observe that since $\sum_\alpha v_\alpha = 1$, we have

$$S^{(m)}(1, n_2, \ldots, n_k) = S^{(m)}(n_2, \ldots, n_k) \ , \tag{15.20}$$

so that when $n_1 = 1$, the relation (15.19) is simply

$$S^{(m)}(2, n_2, \ldots, n_k) = \frac{1-m}{n} S^{(m)}(n_2, \ldots, n_k) + \sum_{2 \le s \le k} \frac{n_s}{n} S^{(m)}(n_2, \ldots, n_{s-1}, n_s + 1, n_{s+1}, \ldots, n_k) . \tag{15.21}$$

In this manner, the relations (15.19) and (15.21) completely determine the numbers $S^{(m)}(n_1, \ldots, n_k)$ for $n_1, \ldots, n_k \ge 2$.

Let us now consider for each N a non-increasing random sequence $(w_{\alpha,N})_{\alpha \ge 1}$. We assume $w_{\alpha,N} \ge 0$, and that $\sum_\alpha w_{\alpha,N} \le 1$. We do NOT assume that $\sum_\alpha w_{\alpha,N} = 1$. Given integers $n_1, \ldots, n_k \ge 2$ we define

$$S_N(n_1, \ldots, n_k) = \mathsf{E} \prod_{s \le k} \sum_\alpha w_{\alpha,N}^{n_s} .$$

The reader has observed that we consider the numbers $S_N(n_1, \ldots, n_k)$ only for $n_1, \ldots, n_k \ge 2$. We further define

$$\mathcal{S} = \left\{ (x_\alpha)_{\alpha \ge 1} \ , \ x_\alpha \ge 0 \ , \ \sum_\alpha x_\alpha \le 1 \ , \ (x_\alpha) \text{ non-increasing} \right\} .$$

Proposition 15.2.4. *Let us assume that*

$$\lim_{N \to \infty} S_N(2) = \lim_{N \to \infty} \mathsf{E} \sum_\alpha w_{\alpha,N}^2 = 1 - m , \tag{15.22}$$

that whenever $n_1, n_2, \ldots, n_k \ge 2$,

$$\lim_{N \to \infty} \Big| S_N(n_1 + 1, n_2, \ldots, n_k) - \frac{n_1 - m}{n} S_N(n_1, \ldots, n_k) - \sum_{2 \le s \le k} \frac{n_s}{n} S_N(n_2, \ldots, n_{s-1}, n_s + n_1, n_{s+1}, \ldots, n_k) \Big| = 0 , \tag{15.23}$$

and whenever $n_2, \ldots, n_k \ge 2$,

$$\lim_{N \to \infty} \Big| S_N(2, n_2, \ldots, n_k) - \frac{1-m}{n} S_N(n_2, \ldots, n_k) - \sum_{2 \le s \le k} \frac{n_s}{n} S_N(n_2, \ldots, n_{s-1}, n_s + 1, n_{s+1}, \ldots, n_k) \Big| = 0 . \tag{15.24}$$

Then the limiting law of the sequence $(w_{\alpha,N})_{\alpha \ge 1}$ *in* $\mathcal{S}$ *is the Poisson-Dirichlet distribution* Λ_m. *In particular we have*

$$\lim_{N \to \infty} \sum_\alpha w_{\alpha,N} = 1 . \tag{15.25}$$

Proof. For each integers $n_1 \geq 2, \ldots, n_k \geq 2$ we have

$$\lim_{N\to\infty} S_N(n_1, \ldots, n_k) = S^{(m)}(n_1, \ldots, n_k) \,. \tag{15.26}$$

This is because asymptotically the numbers $S_N(n_1, \ldots, n_k)$ satisfy the relations (15.19) and (15.21), (as is shown by (15.23) and (15.24)), and that these relations determine these numbers. Formally, it is straightforward to prove (15.26) by induction over $k + n_1 + \cdots + n_k$. The use of (15.22) is to start the induction by proving (15.26) when $k = 1$ and $n_1 = 2$. Given $\mathbf{n} = (n_1, \ldots, n_k)$, with $n_1, \ldots, n_k \geq 2$, let us consider the function $f_{\mathbf{n}} : \mathcal{S} \to \mathbb{R}$ given by

$$f_{\mathbf{n}}(\mathbf{x}) = \prod_{s\leq k} \sum_{\alpha} x_\alpha^{n_s} \,,$$

where $\mathbf{x} = (x_\alpha)_{\alpha\geq 1}$ denotes the generic point of $\mathcal{S}$. First we prove that the function $f_{\mathbf{n}}$ is continuous on $\mathcal{S}$ when $\mathcal{S}$ is considered with its natural topology (the weakest that makes all the maps $\mathbf{x} \mapsto x_\gamma$ continuous). It suffices to consider the case where $f_{\mathbf{n}}(\mathbf{x}) = \sum_{\alpha\geq 1} x_\alpha^n$ for some $n \geq 2$. We note that the sequence $\mathbf{x} = (x_\alpha)_{\alpha\geq 1}$ is non-increasing and that $\sum_{\alpha\geq 1} x_\alpha \leq 1$. Thus, for each p we have $p x_p \leq \sum_{\alpha\leq p} x_\alpha \leq 1$, so that $x_p \leq 1/p$. Therefore

$$\sum_{\alpha\geq p} x_\alpha^n \leq p^{-(n-1)} \sum_{\alpha\geq p} x_\alpha \leq p^{-n+1} \,.$$

Thus

$$\left| \sum_{\alpha} x_\alpha^n - \sum_{\alpha<p} x_\alpha^n \right| \leq p^{-n+1}$$

and $f_{\mathbf{n}}$ is the uniform limit as $p \to \infty$ of the functions $\mathbf{x} \mapsto \sum_{\alpha<p} x_\alpha^n$ which are continuous, so it is continuous.

Taking a subsequence if necessary, let us denote by μ the limiting law in $\mathcal{S}$ of $(w_{\alpha,N})_{\alpha\geq 1}$, so that, since the function $f_{\mathbf{n}}$ is continuous,

$$\int f_{\mathbf{n}}(\mathbf{x}) \mathrm{d}\mu(\mathbf{x}) = \lim_{N\to\infty} S_N(n_1, \ldots, n_k) \,.$$

Thus it follows from (15.26) that

$$\int f_{\mathbf{n}}(\mathbf{x}) \mathrm{d}\mu(\mathbf{x}) = \int f_{\mathbf{n}}(\mathbf{x}) \mathrm{d}\Lambda_m(\mathbf{x}) \,, \tag{15.27}$$

and all that is left to prove is that this implies that $\mu = \Lambda_m$. As a first step we prove the following

$$\forall \mathbf{x}, \mathbf{y} \in \mathcal{S} \,, \quad \mathbf{x} \neq \mathbf{y} \quad \Rightarrow \quad \exists n \geq 2 \,, \quad \sum_{\alpha\geq 1} x_\alpha^n \neq \sum_{\alpha\geq 1} y_\alpha^n \,. \tag{15.28}$$

We proceed by contradiction. Assuming

$$\forall n \geq 2 \,, \qquad \sum_{\alpha \geq 1} x_\alpha^n = \sum_{\alpha \geq 1} y_\alpha^n \,, \tag{15.29}$$

we first see that if φ is a polynomial

$$\sum_{\alpha \geq 1} x_\alpha^2 \varphi(x_\alpha) = \sum_{\alpha \geq 1} y_\alpha^2 \varphi(y_\alpha) \,. \tag{15.30}$$

This remains true by approximation if φ is any continuous function and therefore if φ is the pointwise limit of a sequence of continuous functions. In particular

$$\forall t \,, \qquad \sum_{\alpha \geq 1} x_\alpha^2 \mathbf{1}_{\{x_\alpha \geq t\}} = \sum_{\alpha \geq 1} y_\alpha^2 \mathbf{1}_{\{y_\alpha \geq t\}} \,. \tag{15.31}$$

Assume if possible that $\mathbf{x} \neq \mathbf{y}$. Let $\gamma \geq 1$ be the smallest value of α for which $x_\alpha \neq y_\alpha$, and assume without loss of generality that $x_\gamma > y_\gamma$. Then

$$\sum_{\alpha \geq 1} x_\alpha^2 \mathbf{1}_{\{x_\alpha \geq x_\gamma\}} \geq \sum_{\alpha \leq \gamma} x_\alpha^2 > \sum_{\alpha < \gamma} x_\alpha^2 = \sum_{\alpha < \gamma} y_\alpha^2 = \sum_{\alpha \geq 1} y_\alpha^2 \mathbf{1}_{\{y_\alpha \geq x_\gamma\}}$$

and thus (15.31) fails. This proves (15.28).

In particular the family $f_{\mathbf{n}}$ "separates $\mathcal{S}$" in the sense that if $\mathbf{x} \neq \mathbf{y}$ then $f_{\mathbf{n}}(\mathbf{x}) \neq f_{\mathbf{n}}(\mathbf{y})$ for some $\mathbf{n}$. Since the product of two functions of the type $f_{\mathbf{n}}$ is still of the same type the collection of functions of the form

$$c_0 + \sum_{\ell} c_\ell f_{\mathbf{n}_\ell} \tag{15.32}$$

(where $\sum_\ell$ is a finite sum of such terms) is an algebra $\mathcal{A}$ of continuous functions on $\mathcal{S}$ that separates $\mathcal{S}$. By the Stone-Weierstrass theorem, this algebra is uniformly dense in the set of continuous functions on $\mathcal{S}$. By (15.27) $\int \psi(\mathbf{x}) \mathrm{d}\mu(\mathbf{x}) = \int \psi(\mathbf{x}) \mathrm{d}\Lambda_m(\mathbf{x})$ whenever $\psi \in \mathcal{A}$ so that $\mu = \Lambda_m$. □

15.3 The Baffioni-Rosati Theorem

Consider the set $\mathcal{C}$ of all sequences $\mathbf{x} = (x_{\ell,\ell'})_{1 \leq \ell < \ell'}$ with $|x_{\ell,\ell'}| \leq 1$, provided with the natural (product) topology. Assuming all limits exist, given a spin system, a fundamental object is the probability measure μ^* on $\mathcal{C}$ such that given a continuous function f on $\mathcal{C}$, which depends only on finitely many variables $x_{\ell,\ell'}$, we have

$$\int f(\mathbf{x}) \mathrm{d}\mu^*(\mathbf{x}) = \lim_{N \to \infty} \nu(f((R_{\ell,\ell'}))) \,. \tag{15.33}$$

Equivalently, the measure μ^* is defined by the fact that for each n and each continuous function f^* on $\mathbb{R}^{n(n-1)/2}$, writing $\mathbf{x}_n = (x_{\ell,\ell'})_{1 \leq \ell < \ell' \leq n}$, we have

$$\int f^*(\mathbf{x}_n)\mathrm{d}\mu^*(\mathbf{x}) = \lim_{N\to\infty} \nu(f^*((R_{\ell,\ell'})_{1\le\ell<\ell'\le n})) \ .$$

Thus, μ^* is the "limiting law in $\mathcal{C}$ of the family $(R_{\ell,\ell'})_{1\le\ell<\ell'}$". We will first explain the importance of μ^*. The point is (even though this is absolutely not intuitive the first time one thinks about it) that μ^* carries considerable information (at least "in the limit") about Gibbs' measure.

A first fundamental fact is that (at least for Hamiltonians of the type (14.406)),

$$\mu^* \text{ determines } \lim_{N\to\infty} \frac{1}{N} \mathsf{E} \log Z_N \ . \tag{15.34}$$

This, of course, would be really useful if we knew how to compute μ^*, a question related to Research Problem 15.3.7 below. Until this is the case, there is little point to give a formal version of (15.34) in a book, so we will simply sketch that main ideas.

Consider a family $(A(\boldsymbol{\sigma}))$ of Gaussian r.v.s, that is independent of the disorder of the spin system. Let us assume that for $\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2 \in \Sigma_N$, we have, for a certain function η

$$\mathsf{E}A(\boldsymbol{\sigma}^1)A(\boldsymbol{\sigma}^2) = \eta(R_{1,2}) \ . \tag{15.35}$$

Consider also a bounded and continuous function V, and the r.v. $U = \langle V(A(\boldsymbol{\sigma}))\rangle$. We claim that, in the limit $N \to \infty$, μ^* determines the law of U. To see this, it suffices to show that, in the limit, μ^* determines the moments of U. Denoting by E_0 expectation in the r.v.s $A(\boldsymbol{\sigma})$, we have

$$\mathsf{E}U^k = \nu(W(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^k)) \tag{15.36}$$

where

$$W(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^k) := \mathsf{E}_0 V(A(\boldsymbol{\sigma}^1)) \cdots V(A(\boldsymbol{\sigma}^k)) \ .$$

The quantity depends only on the correlations of the r.v.s $A(\boldsymbol{\sigma}^1), \ldots, A(\boldsymbol{\sigma}^k)$, and by (15.35) it is a continuous function of the overlaps $R_{\ell,\ell'}$ for $1 \le \ell < \ell' \le k$, i.e.

$$W(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n) = f((R_{\ell,\ell'})_{1\le\ell<\ell'\le n}) \ .$$

so that μ^* determines $\lim_{N\to\infty} \mathsf{E}U^k$. Assuming we consider only reasonable functions V (i.e. that do not grow too fast), a truncation argument will show that μ^* determines

$$\lim_{N\to\infty} \mathsf{E} \log\langle V(A(\boldsymbol{\sigma}))\rangle \ . \tag{15.37}$$

The approach of Proposition 1.6.8, and in particular (1.172) and (1.174), then allows one to compute $\lim_{N\to\infty} N^{-1} \log Z_N$ by estimating two quantities of the type (15.37), and this concludes our scheme of proof of (15.34).

To explain the importance of μ i n a different direction, consider at a given disorder n replicas $\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n$, that is, n configurations taken at random according to the Gibbs measure. If we understand for each $n$the joint law under Gibbs' measure of the overlaps $(R_{\ell,\ell'})_{1\leq\ell<\ell'\leq n}$, we might argue that we understand very well Gibbs' measure "in bulk" (remember that $1-R_{\ell,\ell'}$ is twice the Hamming distance of $\boldsymbol{\sigma}^\ell, \boldsymbol{\sigma}^{\ell'}$). We understand this joint law as soon as we understand the Gibbs averages of all the polynomials in the overlaps $R_{\ell,\ell'}$. These are random quantities. We understand their randomness as soon as we understand their joint law, i.e. as soon as we can compute the expectation of any polynomial in these random Gibbs' averages. But using replicas, such an expectation is the expectation of the Gibbs' average of a (bigger) polynomial in the overlaps, and in the limit is determined by μ^*.

Definition 15.3.1. *A probability measure μ^* on $\mathcal{C}$ is* symmetric *if, given any continuous function f on $\mathcal{C}$, and any permutation τ of $\mathbb{N}^*$, we have*

$$\int f(\mathbf{x})\mathrm{d}\mu^*(\mathbf{x}) = \int f(\tau(\mathbf{x}))\mathrm{d}\mu^*(\mathbf{x})$$

where $\tau(\mathbf{x}) = (x_{\tau(\ell),\tau(\ell')})_{1\leq\ell<\ell'}$.

It should be obvious that the measure μ^* given by (15.33) is symmetric. In the literature about infinite random arrays, "symmetric" is sometimes called "weakly exchangeable".

Definition 15.3.2. *We say that the symmetric measure μ^* on $\mathcal{C}$ is* ultrametric *if*

$$\mu^*(\{x_{1,3} \geq \min(x_{1,2}, x_{2,3})\}) = 1 \ . \tag{15.38}$$

An equivalent definition is to require that for each a we have

$$\mu^*(\{x_{1,2} \geq a \ , \ x_{2,3} \geq a \ , \ x_{1,3} < a\}) = 0 \ . \tag{15.39}$$

Conjecture 15.3.3. Under very general conditions the measure μ^* given by (15.33) (exists and) is ultrametric.

We don't know what the general conditions should be; it would be mighty nice to prove this conjecture for the Hamiltonian (14.55) (or even for the SK model).

Definition 15.3.4. *A probability measure μ^* on $\mathcal{C}$ satisfies the Ghirlanda-Guerra identities if given any number n, any continuous function f on $\mathcal{C}$ that depends only on $x_{\ell,\ell'}$ for $1\leq\ell<\ell'\leq n$, any continuous function φ on $\mathbb{R}$, we have*

$$\int \varphi(x_{1,n+1}) f(\mathbf{x})\mathrm{d}\mu^*(\mathbf{x}) = \frac{1}{n}\int \varphi(x_{1,n+1})\mathrm{d}\mu^*(\mathbf{x}) \int f(\mathbf{x})\mathrm{d}\mu^*(\mathbf{x}) + \frac{1}{n}\sum_{2\leq\ell\leq n}\int \varphi(x_{1,\ell}) f(\mathbf{x})\mathrm{d}\mu^*(\mathbf{x}) \ . \tag{15.40}$$

Exercise 15.3.5. Consider a continuous map ψ from $[-1, 1]$ to itself, and the map Ψ form $\mathcal{C}$ to itself given by $\Psi(\mathbf{x}) = (\psi(x_{\ell,\ell'}))_{\ell<\ell'}$ for $\mathbf{x} = (x_{\ell,\ell'})_{\ell<\ell'}$. When μ^* is a probability measure on $\mathcal{C}$ that satisfies the Ghirlanda-Guerra identities, prove that the image of μ^* under Ψ also satisfies these inequalities.

Theorem 15.3.6. *Consider a probability measure μ on $[0, 1]$. Then there exists a unique probability measure μ^* on $\mathcal{C}$ with the following properties: μ^* is symmetric, ultrametric, satisfies the Ghirlanda-Guerra identities, and for each continuous function φ on $\mathbb{R}$ we have*

$$\int \varphi(x_{1,2})\mathrm{d}\mu^*(\mathbf{x}) = \int \varphi(x)\mathrm{d}\mu(x) \; . \tag{15.41}$$

This theorem gives its name to the present section, although F. Baffioni and F. Rosati [15] prove only the uniqueness part.

The vision of the physicists is as follows. The measure μ^* given by (15.33) exists and satisfies the properties of Theorem 15.3.6: it is symmetric, ultrametric and satisfies the Ghirlanda-Guerra identities. By construction it satisfies (15.41) where μ is the "limiting law of the overlap", i.e. $\int \varphi(x)\mathrm{d}\mu(x) = \lim_{N\to\infty} \nu(\varphi(R_{1,2}))$. Therefore the measure μ (which is the Parisi measure) completely determines μ^* (which itself encompasses "all" the information about Gibbs' measure). That is, the entire system is parameterized by μ (or equivalently, since physicists like to think of the function $\mu([0, t])$ rather than of μ the entire system is parameterized by a single function.)

The work of Chapter 12 makes it reasonable to believe that (at least in some sense) the measure μ^* of (15.33) satisfies the Ghirlanda-Guerra identities, a point that we will explore further in Section 15.4. The really missing part in the picture is the fact that μ^* is ultrametric. G. Parisi pointed out that it might well be that the Ghirlanda-Guerra identities themselves imply ultrametricity. Before formally stating the corresponding question, let us observe that since $R_{\ell,\ell'} = \boldsymbol{\sigma}_\ell \cdot \boldsymbol{\sigma}_{\ell'}/N$, the matrix $(R_{\ell,\ell'})_{\ell,\ell'\geq 1}$ is positive definite. Let us define the subset $\mathcal{C}^+$ of $\mathcal{C}$ consisting of the sequences $(x_{\ell,\ell'})$ with the following property: the symmetric matrix $(q_{\ell,\ell'})$ given by $q_{\ell,\ell} = 1$ and $q_{\ell,\ell'} = x_{\ell,\ell'}$ for $\ell < \ell'$ is positive definite. Then we know that $\mu^*(\mathcal{C}^+) = 1$.

Research Problem 15.3.7. (Level 3) Consider a probability measure μ^* on $\mathcal{C}^+$. Assume that μ^* is symmetric and satisfies the Ghirlanda-Guerra identities. Does it follow that μ^* is ultrametric?

Arguably, this is one of the most important problems left today in the theory of spin glasses. A positive solution, combined with (15.34) would let us extend the Parisi formula (14.102) to the case where the Hamiltonian (14.55) contains terms "with odd values of p". (The arguments to prove the formal version of (15.34) yielding this result are not trivial, but there seems to be no point to write them down now, since the expected difficulty of Problem 15.3.7 is several orders of magnitude higher.) More importantly, a positive solution would provide very solid ground to assert that "the Parisi solution is universal". Let us also observe that the construction of Exercise 15.3.5 implies that one cannot completely dispense in Problem 15.3.7 with the hypothesis that $\mu^*(\mathcal{C}^+) = 1$.

We turn to the proof of Theorem 15.3.6. We will first prove the existence of μ^*. We will assume that μ is carried by a finite set. Since any measure on $[0,1]$ is a limit of measures carried by a finite set, the general case is then obtained through a limiting argument. This argument is absolutely straightforward, and better left to the reader. We assume that there exists $q_1 < \ldots < q_{k+1}$ such that for $1 \le r \le k+1$ we have

$$\mu(\{q_r\}) = m_r - m_{r-1} \ ,$$

where $0 = m_0 < m_1 < \ldots < m_k < 1 = m_{k+1}$. We will use a Poisson-Dirichlet cascade (v_α) of parameters $m_1, \ldots, m_k$, so $\alpha \in A = \mathbb{N}^{*k}$. For $\alpha^1, \alpha^2 \in A$, let us recall the notation (α^1, α^2) of (14.36), which denotes "the first coordinate on which α^1 and α^2 differ" so $q_{(\alpha^1,\alpha^2)}$ is the number q_r for $r = (\alpha^1, \alpha^2)$. It helps to think of the points of A as being configurations. The (random) weights v_α define "Gibbs' measure G" on A. We then define "the overlap of two configurations α and γ" by

$$q_{\alpha,\gamma} = q_{(\alpha,\gamma)} \ . \tag{15.42}$$

Consider an i.i.d. sequence $(\alpha^\ell)_{\ell \ge 1}$ on A, that is distributed like G. We define

$$\text{“}\mu^* \text{ is the law of the sequence } (R_{\ell,\ell'}) = (q_{\alpha^\ell,\alpha^{\ell'}})_{1 \le \ell < \ell'}\text{”} \ . \tag{15.43}$$

In other words, we define μ^* by the formula

$$\int f(\mathbf{x}) \mathrm{d}\mu^*(\mathbf{x}) = \mathsf{E} \sum_{\alpha^1, \ldots, \alpha^n \in A} v_{\alpha^1} \cdots v_{\alpha^n} f((q_{\alpha^\ell,\alpha^{\ell'}})_{\ell,\ell'}) \ , \tag{15.44}$$

whenever the continuous function f on $\mathcal{C}$ depends only upon the variables $(x_{\ell,\ell'})_{1 \le \ell < \ell' \le n}$.

We note that μ^* is obviously symmetric by construction. It is also true by construction that μ^* is ultrametric. To see this we observe that

$$(\alpha^1, \alpha^3) \ge \min((\alpha^1, \alpha^2), (\alpha^2, \alpha^3)) \ .$$

This is because for $j < \min((\alpha^1, \alpha^2), (\alpha^2, \alpha^3))$ the j-th coordinate of α^1 is the same as the j-th coordinate of α^2, which is the same as the j-th coordinate of α^3; so the j-th coordinates of α^1 and α^3 are the same and thus $j < (\alpha^1, \alpha^3)$. Consequently, since the sequence (q_r) is increasing we have

$$q_{\alpha^1,\alpha^3} \ge \min(q_{\alpha^1,\alpha^2}, q_{\alpha^2,\alpha^3}) \ , \tag{15.45}$$

and (15.43) shows that μ^* is supported by the set $\{x_{1,3} \geq \min(x_{1,2}, x_{2,3})\}$, which proves (15.38).

Next we prove (15.41). Using (15.44) for $n = 2$ shows that when φ is a continuous function on $\mathbb{R}$ then

$$\begin{aligned}\int \varphi(x_{1,2}) \mathrm{d}\mu^*(\mathbf{x}) &= \mathsf{E} \sum_{\alpha^1, \alpha^2 \in A} v_{\alpha^1} v_{\alpha^2} \varphi(q_{\alpha^1,\alpha^2}) \\ &= \sum_{1 \leq r \leq k+1} \varphi(q_r) \mathsf{E} \sum_{(\alpha^1,\alpha^2)=r} v_{\alpha^1} v_{\alpha^2} \,. \qquad (15.46)\end{aligned}$$

Now we use (14.38) for $F = 0$ to see that

$$\mathsf{E} \sum_{(\alpha^1,\alpha^2)=r} v_{\alpha^1} v_{\alpha^2} = m_r - m_{r-1} \qquad (15.47)$$

and thus (15.46) implies

$$\int \varphi(x_{1,2}) \mathrm{d}\mu^*(\mathbf{x}) = \sum_{1 \leq r \leq k+1} (m_r - m_{r-1}) \varphi(q_r) = \int \varphi(x) \mathrm{d}\mu(x) \,.$$

It remains to prove that μ^* satisfies the Ghirlanda-Guerra identities, which we will deduce from (15.2). Let us note that from (15.47) we have $\nu(\mathbf{1}_{\{(\alpha^1,\alpha^2)=r\}}) = m_r - m_{r-1}$. To prove that (15.2) implies (15.40) we observe that (15.44) means

$$\int f(\mathbf{x}) \mathrm{d}\mu^*(\mathbf{x}) = \nu(h) \,,$$

where the function h is given by

$$h(\alpha^1, \ldots, \alpha^n) = f((q_{\alpha^\ell,\alpha^{\ell'}})_{1 \leq \ell < \ell' \leq n}) \,.$$

This function depends on $\alpha^1, \ldots, \alpha^n$ only through the numbers $(\alpha^\ell, \alpha^{\ell'})$. In a similar manner, when φ is a continuous function on $\mathbb{R}$, then

$$\int \varphi(x_{1,\ell}) f(\mathbf{x}) \mathrm{d}\mu^*(\mathbf{x}) = \nu(\varphi(q_{\alpha^1,\alpha^\ell}) h) \,,$$

where h is as above, so that (15.40) is equivalent to

$$\nu(\varphi(q_{\alpha^1,\alpha^{n+1}}) h) = \frac{1}{n} \nu(\varphi(q_{\alpha^1,\alpha^2})) \nu(h) + \frac{1}{n} \sum_{2 \leq \ell \leq n} \nu(\varphi(q_{\alpha^1,\alpha^\ell}) h) \,,$$

an obvious consequence of (15.2).

Exercise 15.3.8. Prove that the measure μ^* constructed in the proof of Theorem 15.3.6 satisfies $\mu^*(\mathcal{C}^+) = 1$, where $\mathcal{C}^+$ is defined in page 486.

We have proved the "existence part" of Theorem 15.3.6, and we turn to the "uniqueness part".

Consider a number $0 \le a \le 1$, and the subset S_a of $[0,1]^2$ given by

$$S_a = U_a \cup V_a^1 \cup V_a^2$$

where

$$\begin{aligned} U_a &= \{(x,y) \in [0,1]^2 \; ; \; x = y \le a\} \\ V_a^1 &= \{(x,y) \in [0,1]^2 \; ; \; a = x < y\} \\ V_a^2 &= \{(x,y) \in [0,1]^2 \; ; \; a = y < x\} \; . \end{aligned}$$

An alternate definition of S_a is

$$S_a = \{(x,y) \in [0,1]^2 \; ; \; a \ge \min(x,y) \; ; \; x \ge \min(a,y) \; ; \; y \ge \min(a,x)\} \; . \tag{15.48}$$

The equivalence of these two definitions is elementary and left to the reader. The idea of the set S_a is as follows. Suppose that $x_{1,2}, x_{1,3}, x_{2,3}$ satisfy the "ultrametricity condition"

$$x_{\ell_1,\ell_3} \ge \min(x_{\ell_1,\ell_2}, x_{\ell_2,\ell_3})$$

whenever $\{\ell_1,\ell_2,\ell_3\} = \{1,2,3\}$ and the numbers ℓ_1,ℓ_2,ℓ_3 are different of each other (and using the convention that $x_{\ell_1,\ell_2} = x_{\ell_2,\ell_1}$ when $\ell_1 > \ell_2$). Then it is straightforward to see that if we know that $x_{2,3} = a$, we have $(x_{1,2}, x_{1,3}) \in S_a$. Indeed if $x_{1,2} \le a$ then $a \ge x_{1,2} \ge \min(x_{1,3}, x_{2,3}) = \min(x_{1,3}, a)$ so that $x_{1,3} \le x_{1,2}$ and by symmetry $x_{1,2} = x_{1,3}$; while if $x_{1,2} \ge a$ then $a = x_{2,3} \ge \min(x_{1,2}, x_{1,3})$ so that $x_{1,3} \le a$, and since $x_{1,3} \ge \min(x_{1,2}, x_{2,3}) = a$ we have $x_{1,3} = a$.

Lemma 15.3.9. *Consider a probability measure η on $[0,1]^2$ and denote its marginals η_1 and η_2. Then, if η is supported by S_a, it is the only probability measure supported by S_a with marginals η_1 and η_2.*

Proof. We prove that this result holds for positive measures that do not need to be probabilities. First we observe that the result is obvious when η is supported by U_a. Next, the map $(x,y) \mapsto y$ from $[0,1] \times]a,1] \cap S_a = V_a^1$ to $]a,1]$ is one to one, so that η_2 determines the restriction η_2^* of η to V_a^1. Similarly, η_1 determines the restriction η_1^* of η to V_a^2. Then $\eta - \eta_1^* - \eta_2^*$ is supported by U_a, so it is determined by its marginals, and hence by η_1 and η_2. □

Consider a finite set B of pairs (ℓ,ℓ'), $1 \le \ell < \ell'$. We write

$$\mathcal{C}_B = \{\mathbf{y} = (y_{\ell,\ell'})_{(\ell,\ell') \in B}\} \; .$$

There is a canonical map $\mathcal{C} \to \mathcal{C}_B$ that associates $\mathbf{y} = (x_{\ell,\ell'})_{(\ell,\ell')\in B}$ to $\mathbf{x} = (x_{\ell,\ell'})_{1\le\ell<\ell'}$. We denote by μ^*_B the image of μ^* under this map. Let

$$B(n) = \{(\ell,\ell') \; ; \; 1 \le \ell \le \ell' \le n\} \; .$$

We will prove by induction over n that $\mu^*_{B(n)}$ is completely determined by the conditions of Theorem 15.3.6 (this proves the uniqueness of μ^*). For $n = 2$, this follows from (15.41).

Assuming that $\mu^*_{B(n)}$ is determined by the conditions of Theorem 15.3.6, we show that this is also the case of $\mu^*_{B(n+1)}$. For $\ell \le n$ we consider

$$B(n,\ell) = B(n) \cup \{(1,n+1),\ldots,(\ell,n+1)\} \; ,$$

so that $B(n,n) = B(n+1)$. We will prove by induction on ℓ that $\mu^*_{B(n,\ell)}$ is determined by the conditions of Theorem 15.3.6. For $\ell = 1$, we will use the Ghirlanda-Guerra identities (and this is the only step where these will be used). We simply notice that writing (15.40) for μ^* implies that whenever f is a function of $(x_{\ell,\ell'})_{1\le\ell<\ell'\le n}$ we have

$$\begin{aligned}\int \varphi(x_{1,n+1}) f(\mathbf{x}) \mathrm{d}\mu^*_{B(n,1)}(\mathbf{x}) &= \frac{1}{n} \int \varphi(x_{1,n+1}) \mathrm{d}\mu^*(\mathbf{x}) \int f(\mathbf{x}) \mathrm{d}\mu^*_{B(n)}(\mathbf{x}) \\ &+ \frac{1}{n} \sum_{2\le\ell\le n} \int \varphi(x_{1,\ell}) f(\mathbf{x}) \mathrm{d}\mu^*_{B(n)}(\mathbf{x}) \; . \end{aligned} \tag{15.49}$$

Using symmetry and (15.41) we obtain

$$\int \varphi(x_{1,n+1}) \mathrm{d}\mu^*(\mathbf{x}) = \int \varphi(x) \mathrm{d}\mu(x) \; ,$$

so that since by the induction hypothesis $\mu^*_{B(n)}$ is uniquely determined by the conditions of Theorem 15.3.6, so is the right-hand side of (15.49) and hence, so is the left-hand side. As f is any function of $(x_{\ell,\ell'})_{1\le\ell,\ell'\le n}$ this shows that $\mu^*_{B(n,1)}$ is determined.

Assuming that $\ell < n$ and $\mu^*_{B(n,\ell)}$ is determined, we will show that symmetry and ultrametricity imply that $\mu^*_{B(n,\ell+1)}$ is determined. Let

$$\begin{aligned} B_1 &= B(n,\ell) = B(n,\ell+1) \setminus \{(\ell+1,n+1)\} \\ B_2 &= B(n,\ell+1) \setminus \{(\ell,n+1)\} \\ B &= B(n,\ell-1) = B_1 \cap B_2 \; . \end{aligned}$$

Let

$$S = \{\mathbf{x} \in \mathcal{C}_{B(n,\ell+1)} \; ; \; (x_{\ell,n+1}, x_{\ell+1,n+1}) \in S_{x_{\ell,\ell+1}}\} \; ,$$

where the set S_a is given by (15.48). It follows from symmetry and ultrametricity that

$$\mu^*_{B(n,\ell+1)}(S) = 1 \; . \tag{15.50}$$

Consider the projection π_1 of $\mathcal{C}_{B(n,\ell+1)}$ on $\mathcal{C}_{B_1}$ (resp. π_2 on $\mathcal{C}_{B_2}$), which is obtained by forgetting the coordinate $x_{\ell+1,n+1}$ (resp. $x_{\ell,n+1}$). The image μ_1 of $\mu^*_{B(n,\ell+1)}$ under π_1 is $\mu^*_{B(n,\ell)}$, and by the induction hypothesis it is uniquely determined. The image μ_2 of $\mu^*_{B(n,\ell+1)}$ under π_2 is obtained from $\mu_1 = \mu^*_{B(n,\ell)}$ by exchanging the labels ℓ and $\ell+1$, so, by symmetry, it is uniquely determined. The remainder of the proof consists in showing that (15.50) together with the fact that μ_1 and μ_2 are determined implies that $\mu^*_{B(n,\ell+1)}$ is determined. (This completes the proof that $\mu^*_{B(n,\ell+1)}$ is determined when $\mu^*_{B(n,\ell)}$ is determined, therefore $\mu^*_{B(n,n)} = \mu^*_{B(n+1)}$ is determined).

Let us write $\mathbf{y}$ the generic point of $\mathcal{C}_B$, and

$$\mathbf{x} = (\mathbf{y}, x_{\ell,n+1}, x_{\ell+1,n+1})$$

the generic point of $\mathcal{C}_{B(n,\ell+1)}$. Since the projection of $\mu^*_{B(n,\ell+1)}$ on $\mathcal{C}_B$ is μ^*_B, there exists a family $\mu_{\mathbf{y}}$ of probability measures on $[0,1]^2$ such that for any continuous function, and hence any Borel function f we have

$$\begin{aligned}&\int f(\mathbf{x})\mathrm{d}\mu^*_{B(n,\ell+1)}(\mathbf{x})\\ &= \int \left(\int f(\mathbf{y}, x_{\ell,n+1}, x_{\ell+1,n+1})\mathrm{d}\mu_{\mathbf{y}}(x_{\ell,n+1}, x_{\ell+1,n+1}) \right) \mathrm{d}\mu^*_B(\mathbf{y}). \quad (15.51)\end{aligned}$$

Using this for $f = \mathbf{1}_S$, (15.50) implies that $\mu_{\mathbf{y}}(S_{x_{\ell,\ell+1}}) = 1$ (μ^*_B a.e.). Using (15.51) when $f(\mathbf{x})$ does not depend on $x_{1,\ell+1}$, and since μ_1 is determined shows that (μ^*_B a.e.) the first marginal of $\mu_{\mathbf{y}}$ is determined and similarly for the second marginal. Therefore by Lemma 15.3.9 the probability $\mu_{\mathbf{y}}$ is (μ^*_B a.e.) determined. This finishes the proof of the uniqueness part of Theorem 15.3.6.

It is worth to point out that (at least when μ has no atoms) Theorem 15.3.6 can be deduced from the case where μ is uniform (by applying a suitable transformation to each coordinate of $\mathcal{C}$). In some sense, "the measure μ^* is unique", and can be considered as a fundamental object worth studying in detail.

15.4 Generic Sequences and Pure States

In this section we prove that in certain rather general situations "the system decomposes in pure states". That is, the configuration space can be written, in a somewhat canonical manner, as the union of pieces "without structure". It somewhat clarifies the situation to spell the conditions under which we can prove this. Consider, for each N a random probability measure G_N on Σ_N (that will be the Gibbs measure associated to a certain Hamiltonian). We denote by $\langle\cdot\rangle$ an average for G_N. In this notation, the choice of the sequence (G_N) remains implicit.

Definition 15.4.1. *We say that the sequence* (G_N) *satisfies the* extended Ghirlanda-Guerra identities *if for each* n *and each continuous function* ψ *on* $\mathbb{R}$ *we have*

$$\lim_{N\to\infty} \sup_f \left| n\mathsf{E}\langle\psi(R_{1,n+1})f\rangle - \mathsf{E}\langle\psi(R_{1,2})\rangle\mathsf{E}\langle f\rangle - \sum_{2\le\ell\le n} \mathsf{E}\langle\psi(R_{1,\ell})f\rangle \right| = 0 \;, \tag{15.52}$$

where the supremum is taken over all (non random) functions f *on* Σ_N^n *with* $|f| \le 1$.

The only consequence of (15.52) we will ever use is as follows. Given a continuous function f^* on $\mathbb{R}^{n(n-1)/2}$, if we define

$$f(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n) = f^*((R_{\ell,\ell'})_{1\le\ell<\ell'\le n}) \;,$$

then for any continuous function ψ on $\mathbb{R}$ we have

$$\lim_{N\to\infty} \left| n\mathsf{E}\langle\psi(R_{1,n+1})f\rangle - \mathsf{E}\langle\psi(R_{1,2})\rangle\mathsf{E}\langle f\rangle - \sum_{2\le\ell\le n} \mathsf{E}\langle\psi(R_{1,\ell})f\rangle \right| = 0 \;.$$

The reason for the formally stronger formulation (15.52) is simply that it follows naturally from our arguments.

Definition 15.4.2. *We say that a sequence* (G_N) *of random probability measures on* Σ_N *has a Parisi measure* μ *if* μ *is a probability measure on* $[0,1]$ *such that for any continuous function* φ *on* $[-1,1]$ *we have*

$$\lim_{N\to\infty} \mathsf{E}\langle\varphi(R_{1,2})\rangle = \int \varphi(x)\mathrm{d}\mu(x) \;. \tag{15.53}$$

Proposition 15.4.3. *Consider a sequence* $\boldsymbol{\beta}$ *with* $\beta_p \neq 0$ *for each* $p \ge 1$, *and the corresponding Hamiltonian* $H_{N,\boldsymbol{\beta}}$ *of (14.406). Assume that* $\mathsf{E}h^2 > 0$. *Then the sequence* (G_N) *of Gibbs' measures corresponding to this Hamiltonian satisfies the extended Ghirlanda-Guerra identities and has a Parisi measure.*

Proof. The validity of the extended Guirlanda-Guerra identities is proved in 14.12.2. The existence of the Parisi measure follows from Theorem 14.11.6 (using Theorem 14.12.1 as in the proof of Theorem 14.12.2). □

It is very difficult to explicitly compute "the" Parisi measure of a given Hamiltonian $H_{N,\boldsymbol{\beta}}$. According to the work of physicists (see the diagrams in [62], pp. 42-43) it seems that usually the Parisi measure has a positive mass at the largest point q^* of its support. This, combined with Proposition 15.4.3 shows that the hypotheses of our next theorem are very reasonable.

Theorem 15.4.4. *Consider a sequence (G_N) of random measures on Σ_N. Assume that it satisfies the extended Ghirlanda-Guerra identities and that it has a Parisi measure μ. Assume that for a certain number q^* we have $\mu([0,q^*]) = 1$ and $\mu(\{q^*\}) = a > 0$. Then we can find disjoint (random) sets $(A_\alpha)_{\alpha\geq 1}$ of Σ_N with the following properties*

$$\text{The sequence } (G_N(A_\alpha))_{\alpha\geq 1} \text{ has in the limit } N\to\infty \text{ a Poisson-Dirichlet distribution } \Lambda_{1-a}. \tag{15.54}$$

Given any $\delta > 0$, there exists N_δ such that if $N \geq N_\delta$ the following holds true

$$\forall \alpha \geq 1\ , \qquad \int_{A_\alpha^2} |R_{1,2} - q^*| \mathrm{d}G_N(\boldsymbol{\sigma}^1)\mathrm{d}G_N(\boldsymbol{\sigma}^2) \leq \delta G_N(A_\alpha)^2\ . \tag{15.55}$$

Although this might not be immediately obvious, the heuristic picture is as follows. With probability close to 1, the configuration space Σ_N is nearly the union of sets $(A_\alpha)_{\alpha\geq 1}$ on each of which the overlap is nearly q^*. Moreover, the overlap of two configurations is nearly q^* essentially only when these two configurations belong to the same A_α. The sets (A_α) are the "pure states". They have, in a sense, no structure. The overlap is constant $(= q^*)$ on them. The overlap of two configurations in two different pure states is $< q^*$.

Let us detail this still at the heuristic level. The theorem does not say that $\Sigma_N = \bigcup_{\alpha\geq 1} A_\alpha$. However it follows from (15.54) that

$$\lim_{N\to\infty} \mathsf{E} G_N\left(\bigcup_{\alpha\geq 1} A_\alpha\right) = \lim_{N\to\infty} \mathsf{E} \sum_{\alpha\geq 1} G_N(A_\alpha) = 1 \tag{15.56}$$

$$\lim_{N\to\infty} \mathsf{E} \sum_{\alpha\geq 1} G_N(A_\alpha)^2 = a\ . \tag{15.57}$$

For large N, (15.56) shows that with probability close to 1, the sets A_α nearly exhaust Σ_N.

By (15.55) the overlap of two configurations in the same set A_α is nearly q^*. So it is approximately true that

$$\bigcup_\alpha A_\alpha^2 \subset \{R_{1,2} \simeq q^*\} \tag{15.58}$$

so that

$$\mathsf{E} \sum_\alpha G_N(A_\alpha)^2 \leq \mathsf{E} G_N^{\otimes 2}(\{R_{1,2} \simeq q^*\})\ . \tag{15.59}$$

The left-hand side is nearly a by (15.57), and the right-hand side is about $\mu(\{q^*\}) = a$, so the sets (15.58) are nearly equal.

It is not difficult to prove rigorously that condition (15.55) implies that

$$\limsup_{N\to\infty} \mathsf{E}\sum_{\alpha\geq 1} G_N(A_\alpha)^2 \leq \mu(\{q^*\}) , \tag{15.60}$$

so that the sets A_α cannot be "macroscopic" (i.e. of size remaining roughly independent of N) unless $\mu(\{q^*\}) = a > 0$. This condition is necessary in order for the "pure states picture" to hold. An example is constructed in [110] (for the spherical model) of a specific sequence $\boldsymbol{\beta}$ for which $\mu(\{q^*\}) = 0$. This proves that the physicist's "pure states picture" is at best true in a "generic" way (but not for every single $\boldsymbol{\beta}$).

Theorem 15.4.4 relies on the following deterministic result.

Theorem 15.4.5. *Given $\delta > 0$, there exists $\varepsilon > 0$ with the following property. Consider a probability measure π on the unit ball B of a Hilbert space. Assume that for a certain number $q^* \geq 0$ we have*

$$\pi^{\otimes 2}(\{(x,y) \in B\times B\ ;\ x\cdot y \geq q^* + \varepsilon\}) \leq \varepsilon . \tag{15.61}$$

Then we can find disjoint sets $A_1, \dots, A_r$ with the following property

$$\forall \alpha \leq r , \qquad \int_{x,y\in A_\alpha} |x\cdot y - q^*| \mathrm{d}\pi(x)\mathrm{d}\pi(y) \leq \delta \pi(A_\alpha)^2 \tag{15.62}$$

$$\pi^{\otimes 2}\left(\{(x,y)\ ;\ |x\cdot y - q^*| \leq \varepsilon\} \setminus \bigcup_{\alpha\leq r} A_\alpha^2\right) \leq \delta . \tag{15.63}$$

Two points in the same set A_α typically have nearly a dot product q^* by (15.62), while (15.63) asserts that when two points have a dot product nearly q^*, it is essentially because they belong to the same set A_α. The theorem does not say in any sense that the union of the sets A_α nearly exhausts π. This certainly does not follow from the hypotheses (e.g. if π is concentrated at 0). we shall prove Theorem 15.4.5 in Section 15.9.

We start with the proof of Theorem 15.4.4. The overall idea is to use Theorem 15.4.5 at a given disorder and then to use Proposition 15.2.4 to prove (15.54). (A remarkable feature of the proof is the rather indirect way we obtain (15.56).)

Since μ is a Parisi measure for the sequence (G_N), by definition, for a continuous function φ on $[-1,1]$ we have

$$\lim_{N\to\infty} \mathsf{E}\langle \varphi(R_{1,2})\rangle = \int \varphi(x)\mathrm{d}\mu(x) .$$

When I is a closed subset of $\mathbb{R}$, and $\varphi \geq 0$, $\varphi = 1$ on I, then

$$\limsup_{N\to\infty} \mathsf{E} G_N^{\otimes 2}(\{R_{1,2}\in I\}) \leq \lim_{N\to\infty} \mathsf{E}\langle \varphi(R_{1,2})\rangle = \int \varphi(x)\mathrm{d}\mu(x)$$

and consequently

$$\limsup_{N\to\infty} \mathsf{E}G_N^{\otimes 2}(\{R_{1,2} \in I\}) \le \mu(I) . \tag{15.64}$$

Similarly, if J is an open subset of $\mathbb{R}$, then

$$\liminf_{N\to\infty} \mathsf{E}G_N^{\otimes 2}(\{R_{1,2} \in J\}) \ge \mu(J) . \tag{15.65}$$

For $\delta = 2^{-n}$, consider ε_n as provided by Theorem 15.4.5. Using (15.64) for $I = [q^* + \varepsilon_n, \infty[$ and (15.65) for $J =]q - \varepsilon_n, \infty[$ we find an integer N_n such that

$$N \ge N_n \Rightarrow \mathsf{E}G_N^{\otimes 2}(\{R_{1,2} \ge q^* + \varepsilon_n\}) \le 2^{-n}\varepsilon_n \tag{15.66}$$

$$N \ge N_n \Rightarrow \mathsf{E}G_N^{\otimes 2}(\{R_{1,2} \ge q^* - \varepsilon_n\}) \ge a - 2^{-n} . \tag{15.67}$$

Without loss of generality we may assume that the sequence (N_n) increases.

For $N \ge N_1$, we consider the event

$$\Omega_N \quad : \qquad G_N^{\otimes 2}(\{R_{1,2} \ge q^* + \varepsilon_n\}) \le \varepsilon_n ,$$

where n is the unique integer for which $N_n \le N < N_{n+1}$. From (15.66) we see that $\mathsf{P}(\Omega_N) \ge 1 - 2^{-n}$.

When Ω_N does not occur, we define $A_\alpha = \emptyset$ for each α. When Ω_N occurs, we apply Theorem 15.4.5 to G_N (after rescaling by a factor $\sqrt{N}$) to obtain disjoint subsets $A_1, \dots, A_r$ of Σ_N such that

$$\forall \alpha \le r , \qquad \int_{A_\alpha^2} |R_{1,2} - q^*| \mathrm{d}G_N(\boldsymbol{\sigma}^1)\mathrm{d}G_N(\boldsymbol{\sigma}^2) \le 2^{-n} G_N(A_\alpha)^2 \tag{15.68}$$

$$G_N^{\otimes 2}\left(\{|R_{1,2} - q^*| \le \varepsilon_n\} \setminus \bigcup_{\alpha \le r} A_\alpha^2\right) \le 2^{-n} . \tag{15.69}$$

We define $A_\alpha = \emptyset$ for $\alpha > r$. We then relabel if necessary the sets A_α to ensure that the sequence $(G_N(A_\alpha))_{\alpha \ge 1}$ is non-increasing. This completes the construction. The number r of sets that have been constructed is random (depends on the realization of the Gibbs measure) and depends on N.

Given $\delta > 0$, consider n with $\delta > 2^{-n}$. Then (15.55) holds for $N \ge N_n$. So, to complete the proof it remains to establish (15.54).

Consider the function $U_{1,2} = U(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2)$ defined as

$$U_{1,2} = \mathbf{1}_{\{(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) \in \bigcup_\alpha A_\alpha^2\}} , \tag{15.70}$$

i.e. $U_{1,2} = 1$ if $\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2$ both belong to one of the sets A_α, and $U_{1,2} = 0$ otherwise. The next lemma provides a quantitative version of the fact that $U_{1,2} \simeq \mathbf{1}_{\{R_{1,2} = q^*\}}$.

Lemma 15.4.6. *Consider numbers $\gamma > \gamma' > 0$ and a continuous function $\varphi : [-1,1] \to [0,1]$ such that*

$$x < q^* - \gamma \quad \Rightarrow \quad \varphi(x) = 0 \; ; \qquad x > q^* - \gamma' \quad \Rightarrow \quad \varphi(x) = 1 \; .$$

Then

$$\limsup_{N\to\infty} \mathsf{E}\langle |U_{1,2} - \varphi(R_{1,2})|\rangle \le \mu([q^* - \gamma, q^*[) \; . \tag{15.71}$$

Proof. If $U_{1,2} - \varphi(R_{1,2}) \neq 0$, there are two possibilities. Either $U_{1,2} = 1$ and $\varphi(R_{1,2}) < 1$ (so that $R_{1,2} < q^* - \gamma'$) or else $U_{1,2} = 0$ and $\varphi(R_{1,2}) > 0$ (so that $R_{1,2} \ge q^* - \gamma$). Thus

$$\{U_{1,2} \neq \varphi(R_{1,2})\} \subset V_1 \cup V_2 \tag{15.72}$$

where

$$V_1 = \{U_{1,2} = 1 \; , \; R_{1,2} \le q^* - \gamma'\}$$
$$V_2 = \{U_{1,2} = 0 \; , \; R_{1,2} \ge q^* - \gamma\} \; .$$

Now, consider an integer $n \ge 1$ and observe that

$$V_2 = \{R_{1,2} \ge q^* - \gamma\} \setminus \bigcup_\alpha A_\alpha^2 \subset V_3 \cup V_4 \tag{15.73}$$

where

$$V_3 = \{R_{1,2} \ge q^* - \gamma\} \setminus \{R_{1,2} \ge q^* - \varepsilon_n\}$$
$$V_4 = \{R_{1,2} \ge q^* - \varepsilon_n\} \setminus \bigcup_\alpha A_\alpha^2 \; .$$

Since $|U_{1,2} - \varphi(R_{1,2})| \le 1$, we deduce from (15.72) and (15.73) that

$$\mathsf{E}\langle |U_{1,2} - \varphi(R_{1,2})|\rangle \le \mathsf{E}G_N^{\otimes 2}(V_1) + \mathsf{E}G_N^{\otimes 2}(V_3) + \mathsf{E}G_N^{\otimes 2}(V_4) \; . \tag{15.74}$$

By (15.69) we have $G_N^{\otimes 2}(V_4) \le 2^{-n}$ when $N \ge N_n$ and Ω_N occurs. Moreover $\mathsf{P}(\Omega_N) \ge 1 - 2^{-n}$ so that $\mathsf{E}G_N^{\otimes 2}(V_4) \le 2^{-n+1}$ and

$$\limsup_N \mathsf{E}G_N^{\otimes 2}(V_4) \le 2^{-n+1} \; . \tag{15.75}$$

By (15.68), for $N \ge N_n$ and any α, Markov's inequality implies

$$G_N^{\otimes 2}(A_\alpha^2 \cap \{R_{1,2} \le q^* - \gamma'\}) \le \frac{2^{-n}}{\gamma'} G_N^{\otimes 2}(A_\alpha^2)$$

and thus

$$G_N^{\otimes 2}(V_1) = G_N^{\otimes 2}\left(\bigcup_{\alpha\ge 1} A_\alpha^2 \cap \{R_{1,2} \le q^* - \gamma'\}\right) \le \frac{2^{-n}}{\gamma'} \sum_{\alpha\ge 1} G_N^{\otimes 2}(A_\alpha^2) \le \frac{2^{-n}}{\gamma'} \; ,$$

which shows that

$$\limsup_N \mathsf{E} G_N^{\otimes 2}(V_1) \le \frac{2^{-n}}{\gamma'} \ . \tag{15.76}$$

Finally, for $N \ge N_n$, using (15.67) in the second line,

$$\begin{aligned}\mathsf{E} G_N^{\otimes 2}(V_3) &= \mathsf{E} G_N^{\otimes 2}(\{R_{1,2} \ge q^* - \gamma\}) - \mathsf{E} G_N^{\otimes 2}(\{R_{1,2} \ge q^* - \varepsilon_n\}) \\ &\le \mathsf{E} G_N^{\otimes 2}(\{R_{1,2} \ge q^* - \gamma\}) - a + 2^{-n}\end{aligned}$$

and using (15.64),

$$\limsup_{N\to\infty} \mathsf{E} G_N^{\otimes 2}(V_3) \le \mu([q^* - \gamma, \infty[) - a + 2^{-n} \ .$$

Combining with (15.76), (15.74) and (15.66) we obtain

$$\limsup_N \mathsf{E}\langle |U_{1,2} - \varphi(R_{1,2})|\rangle \le 2^{-n+1} + \frac{2^{-n}}{\gamma'} + 2^{-n} + \mu([q^* - \gamma, \infty[) - a \ .$$

Since $\mu([q^* - \gamma, \infty[) - a = \mu([q^* - \gamma, q^*[)$ and n is arbitrary the result follows. □

Corollary 15.4.7. *We have*

$$\lim_{N\to\infty} \mathsf{E} \sum_\alpha G_N(A_\alpha)^2 = a \ . \tag{15.77}$$

Proof. We have

$$|\langle U_{1,2}\rangle - \langle \varphi(R_{1,2})\rangle| \le \langle |U_{1,2} - \varphi(R_{1,2})|\rangle$$

and $\langle U_{1,2}\rangle = \sum_\alpha G_N(A_\alpha)^2$. Thus

$$\sum_\alpha G_N(A_\alpha)^2 \ge \langle \varphi(R_{1,2})\rangle - \langle |U_{1,2} - \varphi(R_{1,2})|\rangle \ .$$

Taking expectation and using (15.53) and (15.71) yields

$$\begin{aligned}\liminf_{N\to\infty} \mathsf{E} \sum_\alpha G_N(A_\alpha)^2 &\ge \int \varphi(x) \mathrm{d}\mu(x) - \mu([q^* - \gamma, q^*[) \\ &\ge a - \mu([q^* - \gamma, q^*[) \ .\end{aligned} \tag{15.78}$$

In a similar manner, and since $\varphi(x) = 0$ for $x \le q^* - \gamma$,

$$\begin{aligned}\limsup_{N\to\infty} \mathsf{E} \sum_\alpha G_N(A_\alpha)^2 &\le \int \varphi(x) \mathrm{d}\mu(x) + \mu([q^* - \gamma, q^*[) \\ &\le a + 2\mu([q^* - \gamma, q^*[) \ .\end{aligned} \tag{15.79}$$

Since γ is arbitrary, we have

$$\limsup_{N\to\infty} \mathsf{E}\sum_{\alpha} G_N(A_\alpha)^2 \le a \le \liminf_{N\to\infty} \mathsf{E}\sum_{\alpha} G_N(A_\alpha)^2 \;,$$

which proves (15.77). □

The sequence $w_{\alpha,N} = G_N(A_\alpha)$ satisfies (15.22) for $m = 1-a$. In order to be able to use Proposition 15.2.4, we now try to prove the relations (15.23) and (15.24) for

$$S_N(n_1,\ldots,n_k) = \mathsf{E}\prod_{s\le k}\sum_{\alpha} G_N(A_\alpha)^{n_s} \;. \tag{15.80}$$

We recall the function $U_{1,2} = U(\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2)$ of (15.70), and for $\ell < \ell'$ we write $U_{\ell,\ell'} = U(\boldsymbol{\sigma}^\ell,\boldsymbol{\sigma}^{\ell'})$.

Lemma 15.4.8. *Consider a function f^* on $[-1,1]^{n(n-1)/2}$ and assume that for a certain number B we have*

$$|f^*((x_{\ell,\ell'})_{1\le\ell<\ell'\le n}) - f^*((y_{\ell,\ell'})_{1\le\ell<\ell'\le n})| \le B\sum_{1\le\ell<\ell'\le n}|x_{\ell,\ell'}-y_{\ell,\ell'}| \;. \tag{15.81}$$

Define

$$f = f(\boldsymbol{\sigma}^1,\ldots,\boldsymbol{\sigma}^n) = f^*((U_{\ell,\ell'})_{1\le\ell<\ell'\le n}) \;. \tag{15.82}$$

Then we have

$$\lim_{N\to\infty}\left| n\mathsf{E}\langle U_{1,n+1}f\rangle - \mathsf{E}\langle U_{1,2}\rangle\mathsf{E}\langle f\rangle - \sum_{2\le\ell\le n}\mathsf{E}\langle U_{1,\ell}f\rangle\right| = 0 \;. \tag{15.83}$$

Proof. The proof relies on the extended Ghirlanda-Guerra identities, but we cannot use them for the function f of (15.82) because this function is random. We consider a continuous function φ on $[0,1]$ and we define

$$f^\sim = f^\sim(\boldsymbol{\sigma}^1,\ldots,\boldsymbol{\sigma}^n) = f^*((\varphi(R_{\ell,\ell'}))_{1\le\ell<\ell'\le n}) \;.$$

The extended Ghirlanda-Guerra identities imply

$$\lim_{N\to\infty}\left| n\mathsf{E}\langle \varphi(R_{1,n+1})f^\sim\rangle - \mathsf{E}\langle\varphi(R_{1,2}\rangle\mathsf{E}\langle f^\sim\rangle - \sum_{2\le\ell\le n}\mathsf{E}\langle \varphi(R_{1,\ell})f^\sim\rangle\right| = 0 \;. \tag{15.84}$$

By (15.81) we have

$$|f^\sim - f| \le B\sum_{1\le\ell<\ell'\le n}|\varphi(R_{\ell,\ell'}) - U_{\ell,\ell'}| \;,$$

so that

$$\mathsf{E}\langle|f^{\sim} - f|\rangle \le n^2 B\mathsf{E}\langle|\varphi(R_{1,2}) - U_{1,2}|\rangle \ .$$

Thus, if we choose φ as in Lemma 15.4.6, (15.84) together with (15.71) imply (15.83) since γ' is arbitrary. □

We now finish the proof of Theorem 15.4.4 by proving (15.54). Let us consider integers $n_1, \ldots, n_k \ge 2$ and $n = n_1 + \cdots + n_k$ and decompose the set $\{1, \ldots, n\}$ as a union of k disjoint sets $I_1, \ldots, I_k$ with $1 \in I_1$, and card$I_s = n_s$. We write $\prod_{\ell,\ell' \in I_s}$ for the product over all choices of $\ell, \ell' \in I_s$ with $\ell < \ell'$. We define

$$f = f(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n) = \prod_{1 \le s \le k} \prod_{\ell,\ell' \in I_s} U_{\ell,\ell'} \tag{15.85}$$

so this is of the type (15.82) (for $f^*((x_{\ell,\ell'})_{1 \le \ell < \ell' \le n}) = \prod_{s \le k} \prod_{\ell,\ell' \in I_s} x_{\ell,\ell'}$).

We observe that

$$\prod_{\ell,\ell' \in I_s} U_{\ell,\ell'} = 1 \quad \Leftrightarrow \quad \exists \alpha \ , \ \forall \ell \in I_s \ , \boldsymbol{\sigma}^\ell \in A_\alpha \ ,$$

and thus, since the sets A_α are disjoint,

$$\left\langle \prod_{\ell,\ell' \in I_s} U_{\ell,\ell'} \right\rangle = \sum_\alpha G_N(A_\alpha)^{n_s} \ . \tag{15.86}$$

By independence of the replicas this gives

$$\langle f \rangle = \prod_{s \le k} \left\langle \prod_{\ell,\ell' \in I_s} U_{\ell,\ell'} \right\rangle = \prod_{s \le k} \sum_\alpha G_N(A_\alpha)^{n_s} \ . \tag{15.87}$$

Next, we define $I'_1 = I_1 \cup \{n+1\}$ and $I'_s = I_s$ for $2 \le s \le k$. Then

$$U_{1,n+1} f = 1 \quad \Leftrightarrow \quad \forall s \le k \ , \ \exists \alpha \ , \ \forall \ell \in I'_s \ , \ \boldsymbol{\sigma}^\ell \in A_\alpha,$$

and thus

$$\langle U_{1,n+1} f \rangle = \sum_\alpha G_N(A_\alpha)^{n_1+1} \prod_{2 \le s \le k} \sum_\alpha G_N(A_\alpha)^{n_s} \ . \tag{15.88}$$

Finally we study $U_{1,\ell} f$. If $\ell \in I_1$, then $U_{1,\ell} f = f$. If $\ell \in I_r$ for $r \ge 2$, let us define $I'_s = I_s$ for $2 \le s \le k$, $s \ne r$ and $I'_r = I_1 \cup I_s$. Then

$$U_{1,\ell} f = 1 \quad \Leftrightarrow \quad \forall s \ , \ 2 \le s \le k \ , \ \exists \alpha \ , \ \forall \ell' \in I'_s \ , \ \boldsymbol{\sigma}^{\ell'} \in A_\alpha \ ,$$

and thus

$$\langle U_{1,\ell} f \rangle = \prod_{2 \le s \le k} \sum_\alpha G_N(A_\alpha)^{n'_s} \tag{15.89}$$

where $n'_s = n_s$ if $s \ne r$ and $n'_r = n_r + n_1$.

Now we remember that (15.77) means that $\lim_{N\to\infty} \mathsf{E}\langle U_{1,2}\rangle = a$ and using (15.87) to (15.89) we see that (15.83) implies (15.23) for $m = 1 - a$.

Let us finally consider integers $n_2, \ldots, n_k \geq 2$ and define (instead of (15.85))

$$f = \prod_{2\leq s\leq k} \prod_{\ell,\ell'\in I_s} U_{\ell,\ell'}$$

so that

$$\langle f\rangle = \prod_{2\leq s\leq k} \sum_{\alpha} G_N(A_\alpha)^{n_s} \ .$$

It is obvious that (15.88) and (15.89) still hold true, so (15.83) implies now (15.24). We can then appeal to Proposition 15.2.4 to obtain (15.54) and conclude the proof of Theorem 15.4.4.

15.5 Determinators; Panchenko's Invariance Theorem

Let us consider a sequence (G_N) of random measures on Σ_N. Let us recall the set $\mathcal{C}$ of Section 15.3. Let us denote by μ_N^* "the law in $\mathcal{C}$ of the sequence $(R_{\ell,\ell'})$", that is, the probability measure on $\mathcal{C}$ such that if f^* is a continuous function on $\mathcal{C}$ that depend only on finitely many of the coordinates $x_{\ell,\ell'}$, then

$$\int f^*(\mathbf{x})\mathrm{d}\mu_N^*(\mathbf{x}) = \mathsf{E}\langle f^*((R_{\ell,\ell'})_{\ell<\ell'})\rangle \ , \tag{15.90}$$

where $\langle\cdot\rangle$ denotes an average for G_N. In this section we return to the problem of studying the limit $\mu^* = \lim_{N\to\infty} \mu_N^*$. (Of course, the limit might not exist, in which case we simply take the limit along a subsequence). The fundamental importance of μ^* has been explained in Section 15.3. In this section we prove, in the setting of Theorem 15.4.4, that the measure μ^* can be computed through a relatively simple object that we call a *determinator* (as it determines all the possible measures μ^*). We then prove the remarkable fact that the extended Ghirlanda-Guerra identities imply a strong invariance property of the determinator. This is the cornerstone of Panchenko's partial solution of Conjecture 15.3.3 that we will present in the next section.

The basic idea of this construction is that as far as the overlaps are concerned, each of the sets $(A_\alpha)_{\alpha\leq r}$ (constructed in Theorem 15.4.4) can (in the limit $N\to\infty$) be replaced by a single point, its barycentre

$$\boldsymbol{\sigma}_\alpha = \frac{1}{G_N(A_\alpha)} \int_{A_\alpha} \boldsymbol{\sigma}\mathrm{d}G_N(\boldsymbol{\sigma}) \ . \tag{15.91}$$

For $\alpha, \gamma \leq r$, let us define

$$q_{\alpha,\gamma} = \frac{\boldsymbol{\sigma}_\alpha \cdot \boldsymbol{\sigma}_\gamma}{N} \ . \tag{15.92}$$

The essential fact (that is not yet proved) is that the overlap of a configuration in A_α and a configuration in A_γ is basically $q_{\alpha,\gamma}$. Therefore we should expect that the data of the weights $w_\alpha = G_N(A_\alpha)$ and the numbers $q_{\alpha,\gamma}$ suffices to nearly compute μ_N^*.

We explain this now. Let us denote by $\mathcal{Q}$ the set of $\mathbb{N}^* \times \mathbb{N}^*$ symmetric positive definite matrices $(q_{\alpha,\gamma})$ with $|q_{\alpha,\gamma}| \le 1$. We provide $\mathcal{Q}$ with the topology induced by $[-1,1]^{\mathbb{N}^* \times \mathbb{N}^*}$, for which it is a compact set. We recall that we denote by $\mathcal{S}$ the set of non-negative, non-increasing sequences with sum at most 1. Consider

$$\rho = ((w_\alpha)_{\alpha \ge 1}, (q_{\alpha,\gamma})) \in \mathcal{S} \times \mathcal{Q} \,. \tag{15.93}$$

We should think of $1, 2, \ldots$ as configurations, the weight of configuration α being w_α. (The sum of the weights need not be one.) Given ρ, if $f = f(\alpha^1, \ldots, \alpha^n)$ is a function of n configurations we define

$$\langle f \rangle_\rho = \sum_{\alpha^1, \ldots, \alpha^n} w_{\alpha^1} \cdots w_{\alpha^n} f(\alpha^1, \ldots, \alpha^n) \,. \tag{15.94}$$

This definition requires caution, since, when $\sum w_\alpha < 1$, we do not get the same expression whether we think of f as a function of n or $(n+1)$ configurations. This is not a real problem because we will be interested only "in the case where $\sum w_\alpha \to 1$" and soon only in the case where $\sum w_\alpha = 1$. So, rather than being pedantic and denoting the left-hand side of (15.94) as $\langle f \rangle_{n,\rho}$ (to indicate the dependence on ρ) let us accept that the value of n actually used will be obvious from the context.

In the expression (15.93) we should think of $q_{\alpha,\gamma}$ as the overlap of configurations α and γ. Given a function f^* on $\mathbb{R}^{n(n-1)/2}$ and given ρ, we define the function f of n configurations by

$$f(\alpha^1, \ldots, \alpha^n) = f^*((q_{\alpha^\ell, \alpha^{\ell'}})_{1 \le \ell < \ell' \le n}) \tag{15.95}$$

so that

$$\langle f \rangle_\rho = \sum_{\alpha^1, \ldots, \alpha^n} w_{\alpha^1} \cdots w_{\alpha^n} f^*((q_{\alpha^\ell, \alpha^{\ell'}})_{1 \le \ell < \ell' \le n}) \,. \tag{15.96}$$

The notation (15.95), (15.96) is used throughout the entire section.

Let us recall that we assume the hypotheses of Theorem 15.4.4, and consider the random element ρ_N of $\mathcal{S} \times \mathcal{Q}$, given by

$$w_\alpha = G_N(A_\alpha) \ \text{ if } \alpha \le r \,; \quad w_\alpha = 0 \ \text{ if } \alpha > r \tag{15.97}$$

$$q_{\alpha,\gamma} = \frac{\boldsymbol{\sigma}_\alpha \cdot \boldsymbol{\sigma}_\gamma}{N} \ \text{ if } \alpha, \gamma \le r \,; \quad q_{\alpha,\gamma} = 0 \ \text{ otherwise.} \tag{15.98}$$

Again, the idea is that from ρ_N one can nearly compute μ_N^*, as the next lemma shows.

Lemma 15.5.1. *Consider a continuous function* f^* *on* $\mathbb{R}^{n(n-1)/2}$*, and define a function* $f^\sim$ *of* $\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n$ *by*

$$f^\sim = f^\sim(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n) = f^*((R_{\ell,\ell'})_{1 \le \ell < \ell' \le n}) . \tag{15.99}$$

Then we have (recalling (15.95) and (15.96))

$$\lim_{N \to \infty} |\mathsf{E}\langle f^\sim \rangle - \mathsf{E}\langle f \rangle_{\rho_N}| = 0 . \tag{15.100}$$

Proof. Let $D_n \subset \Sigma_N^n$ be the set

$$\forall \ell \le n , \quad \boldsymbol{\sigma}^\ell \in \bigcup_{\alpha \le r} A_\alpha .$$

Since

$$G_N^{\otimes n}(D_n) = G_N^n\left(\bigcup_{\alpha \le r} A_\alpha\right) ,$$

it follows from (15.56) that $\mathsf{E} G_N^{\otimes n}(D_n) \to 1$ so to prove (15.100) it suffices to show that

$$\lim_{N \to \infty} \mathsf{E}|\langle f^\sim \mathbf{1}_{D_n} \rangle - \langle f \rangle_{\rho_N}| = 0 . \tag{15.101}$$

Now

$$\langle f^\sim \mathbf{1}_{D_n} \rangle = \sum_{\alpha^1, \ldots, \alpha^n \le r} \int_{A_{\alpha^1} \times \cdots \times A_{\alpha^n}} f^\sim \mathrm{d}G_N(\boldsymbol{\sigma}^1) \cdots \mathrm{d}G_N(\boldsymbol{\sigma}^n)$$

and, by (15.97) and (15.98)

$$\langle f \rangle_{\rho_N} = \sum_{\alpha^1, \ldots, \alpha^n \le r} G_N(A_{\alpha^1}) \cdots G_N(A_{\alpha^n}) f(\alpha^1, \ldots, \alpha^n) .$$

Thus

$$\begin{aligned} &|\langle f^\sim \mathbf{1}_{D_n} \rangle - \langle f \rangle_{\rho_N}| \\ &\le \sum_{\alpha^1, \ldots, \alpha^n \le r} \int_{A_{\alpha^1} \times \cdots \times A_{\alpha^n}} |f^\sim - f(\alpha^1, \ldots, \alpha^n)| \mathrm{d}G_N(\boldsymbol{\sigma}^1) \cdots \mathrm{d}G_N(\boldsymbol{\sigma}^n) . \end{aligned} \tag{15.102}$$

To prove (15.101), by approximation we may assume that for a constant B we have

$$|f^*((x_{\ell,\ell'})_{1 \le \ell < \ell' \le n}) - f^*((y_{\ell,\ell'})_{1 \le \ell < \ell' \le n})| \le B \sum_{1 \le \ell < \ell' \le n} |x_{\ell,\ell'} - y_{\ell,\ell'}| ,$$

so that, since

$$f^\sim = f((R_{\ell,\ell'})_{1 \le \ell < \ell' \le n}) ; \quad f(\alpha^1, \ldots, \alpha^n) = f^*((q_{\alpha^\ell, \alpha^{\ell'}})_{1 \le \ell < \ell' \le n})$$

and since

$$R_{\ell,\ell'} = \frac{\boldsymbol{\sigma}^\ell \cdot \boldsymbol{\sigma}^{\ell'}}{N} \ ; \qquad q_{\alpha,\gamma} = \frac{\boldsymbol{\sigma}_\alpha \cdot \boldsymbol{\sigma}_\gamma}{N} \ ,$$

we deduce from (15.102) that

$$\begin{aligned} &|\langle f \mathbf{1}_{D_n}\rangle - \langle f\rangle_{\rho_N}| \qquad (15.103)\\ &\le B \sum_{1\le \ell<\ell'\le n} \sum_{\alpha^\ell,\alpha^{\ell'}\le r} \int_{A_{\alpha^\ell}\times A_{\alpha^{\ell'}}} \left| \frac{\boldsymbol{\sigma}^\ell \cdot \boldsymbol{\sigma}^{\ell'}}{N} - \frac{\boldsymbol{\sigma}_{\alpha^\ell} \cdot \boldsymbol{\sigma}_{\alpha^{\ell'}}}{N} \right| \mathrm{d}G_N(\boldsymbol{\sigma}^\ell)\mathrm{d}G_N(\boldsymbol{\sigma}^{\ell'})\\ &= B\frac{n(n-1)}{2} \sum_{\alpha,\gamma\le r} \int_{A_\alpha\times A_\gamma} \left| \frac{\boldsymbol{\sigma}^1 \cdot \boldsymbol{\sigma}^2}{N} - \frac{\boldsymbol{\sigma}_\alpha \cdot \boldsymbol{\sigma}_\gamma}{N} \right| \mathrm{d}G_N(\boldsymbol{\sigma}^1)\mathrm{d}G_N(\boldsymbol{\sigma}^2) \ . \end{aligned}$$

It is a consequence of (15.55) and a general theorem about Hilbert space (Proposition 15.9.12 below) that

$$\int_{A_\alpha\times A_\gamma} \left| \frac{\boldsymbol{\sigma}^1 \cdot \boldsymbol{\sigma}^2}{N} - \frac{\boldsymbol{\sigma}_\alpha \cdot \boldsymbol{\sigma}_\gamma}{N} \right| \mathrm{d}G_N(\boldsymbol{\sigma}^1)\mathrm{d}G_N(\boldsymbol{\sigma}^2) \le 8\sqrt{\delta} G_N(A_\alpha) G_N(A_\gamma) \ .$$

Since (15.55) holds for $N \ge N_\delta$, we get from (15.103) that then

$$|\langle f^\sim \mathbf{1}_{D_n}\rangle - \langle f\rangle_{\rho_N}| \le 4\sqrt{\delta} B n^2 \ .$$

Since δ is arbitrary this proves (15.101) and finishes the proof. □

Given a probability measure λ on $\mathcal{S}\times\mathcal{Q}$ and a function $\psi(\rho)$ on $\mathcal{S}\times\mathcal{Q}$, we write

$$\mathsf{E}_\lambda \psi(\rho) = \int \psi(\rho)\mathrm{d}\lambda(\rho) \ , \qquad (15.104)$$

that is, we think of ρ as a random element of law λ. In particular, if f^* is a function on $\mathbb{R}^{n(n-1)/2}$, and recalling (15.96), we have

$$\mathsf{E}_\lambda \langle f\rangle_\rho = \int \langle f\rangle_\rho \mathrm{d}\lambda(\rho) \ . \qquad (15.105)$$

We denote by λ_N the law of the element ρ_N defined by (15.97) and (15.98), so that

$$\mathsf{E}\langle f\rangle_{\rho_N} = \mathsf{E}_{\lambda_N}\langle f\rangle_\rho \ . \qquad (15.106)$$

Since we should think of $q_{\alpha,\gamma}$ as the overlap of the configurations α and γ, it is natural to use the same notation

$$R_{\ell,\ell'} = q_{\alpha^\ell,\alpha^{\ell'}} \ ,$$

as in the case of the overlap of spin configuration, so that we can write (15.95) as $f(\alpha^1,\ldots,\alpha^n) = f^*((R_{\ell,\ell'})_{1\le\ell<\ell'\le n})$ and we may use notation such as $\langle \varphi(R_{1,n+1}) f\rangle_\rho$ to mean

$$\sum_{\alpha^1,\ldots,\alpha^{n+1}} w_{\alpha^1}\cdots w_{\alpha^{n+1}}\varphi(q_{\alpha^1,\alpha^{n+1}})f(\alpha^1,\ldots,\alpha^n)\ .$$

We now come to the main point: the limiting law λ of λ_N allows for the computation of the limit μ^* of μ_N, and inherits the nice properties of the sequence (G_N).

Proposition 15.5.2. *Assume that (along a subsequence if necessary) the sequence (λ_N) converges weakly to a probability measure λ on $\mathcal{S}\times\mathcal{Q}$. Then the following hold, where we recall the notation (15.95).*

1. *For every continuous function f^* on $\mathbb{R}^{n(n-1)/2}$ we have*

$$\lim_{N\to\infty}\mathsf{E}\langle f\rangle_{\rho_N}=\lim_{N\to\infty}\mathsf{E}_{\lambda_N}\langle f\rangle_\rho=\mathsf{E}_\lambda\langle f\rangle_\rho \tag{15.107}$$

   (where the limit is taken along the same subsequence).
2. *For every bounded Borel function f^* on $\mathbb{R}^{n(n-1)/2}$ and every bounded Borel function φ on $[-1,1]$ we have*

$$\begin{aligned}\mathsf{E}_\lambda\langle\varphi(R_{1,n+1})f\rangle_\rho=&\frac{1}{n}\mathsf{E}_\lambda\langle\varphi(R_{1,2})\rangle_\rho\mathsf{E}_\lambda\langle f\rangle_\rho\\&+\frac{1}{n}\sum_{2\le\ell\le n}\mathsf{E}_\lambda\langle\varphi(R_{1,\ell})f\rangle_\rho\ .\end{aligned} \tag{15.108}$$

3. *For each bounded Borel function φ on $[-1,1]$ we have*

$$\mathsf{E}_\lambda\langle\varphi(R_{1,2})\rangle_\rho=\int\varphi(x)\mathrm{d}\mu(x)\ , \tag{15.109}$$

   where we recall that μ denotes the Parisi measure of the sequence (G_N).
4. *Moreover*

$$\text{Under } \lambda \text{ the sequence } (w_\alpha) \text{ has a Poisson-Dirichlet distribution } \Lambda_{1-a}\ . \tag{15.110}$$

5. *Finally, λ a.s. we have*

$$\forall\alpha\ ,\quad q_{\alpha,\alpha}=q^*\ ;\quad \alpha\neq\gamma\quad\Rightarrow\quad q_{\alpha,\gamma}<q^*\ . \tag{15.111}$$

Combining (15.90), (15.100) and (15.107) we observe that if λ is as in Proposition 15.5.2 we can compute $\mu^*=\lim_N\mu_N^*$ by the following formula, where $\mathbf{x}_n=(x_{\ell,\ell'})_{1\le\ell<\ell'\le n}$, and where f^* is a continuous function on $\mathbb{R}^{n(n-1)/2}$:

$$\int f^*(\mathbf{x}_n)\mathrm{d}\mu^*(\mathbf{x})=\mathsf{E}_\lambda\langle f\rangle_\rho\ . \tag{15.112}$$

Let us also observe that this formula determines a unique symmetric probability measure μ^* on $\mathcal{C}$.

In words one could say that "μ^* is the law of $(q_{\alpha^\ell,\alpha^{\ell'}})_{1\leq\ell<\ell'} \in \mathcal{C}$ under λ." (One should observe that the law of $(q_{\alpha^\ell,\alpha^{\ell'}})_{1\leq\ell<\ell'} \in \mathcal{C}$ under λ is automatically a symmetric measure.) Condition (15.108) simply says that μ^* satisfies the extended Ghirlanda-Guerra identities, while condition (15.109) asserts that its 1-dimensional marginals are equal to μ.

In particular the probability measures λ on $\mathcal{S}\times\mathcal{Q}$ determine in this manner all possible measures μ^*. This motivates the following definition, where we recall that $\mathcal{S}_1$ is the subset of $\mathcal{S}$ consisting of sequences that sum to 1.

Definition 15.5.3. *A* determinator *is a probability measure on* $\mathcal{S}_1 \times \mathcal{Q}$.

Let us remind the reader that it is a central problem to decide whether the measure μ^* of (15.92) is ultrametric (in the sense of Definition 15.3.2). When μ^* arises from a sequence (G_N) that satisfies the conditions of Theorem 15.4.4, this question can be decided by studying the determinators that satisfy properties (15.108) to (15.111). We will return to this topic after the proof.

Proof of Proposition 15.5.2. The (small) difficulty in proving (15.107) is that the map $\rho \mapsto \langle f\rangle_\rho$ is not continuous in general on $\mathcal{S}\times\mathcal{Q}$. Given an integer s, let us define

$$\langle f\rangle_{\rho,s} = \sum_{\alpha^1,\dots,\alpha^n\leq s} w_{\alpha^1}\cdots w_{\alpha^n} f(\alpha^1,\dots,\alpha^n)\;.$$

The sum above is finite, so the map $\rho \mapsto \langle f\rangle_{\rho,s}$ is continuous. The basic idea is that this map approximates well the map $\rho \mapsto \langle f\rangle_\rho$ at every point of $\mathcal{S}_1\times\mathcal{Q}$. Assuming without loss of generality that $|f|\leq 1$, we then have

$$|\langle f\rangle_{\rho,s} - \langle f\rangle_\rho| \leq \sum w_{\alpha^1}\cdots w_{\alpha^n}\;, \tag{15.113}$$

where the sum is over all the choices of $\alpha^1,\dots,\alpha^n$ with $\max\{\alpha^\ell\;,\;\ell\leq n\} > s$. This sum is

$$\Big(\sum_{\alpha\geq 1} w_\alpha\Big)^n - \Big(\sum_{\alpha\leq s} w_\alpha\Big)^n$$

so that

$$|\langle f\rangle_{\rho,s} - \langle f\rangle_\rho| \leq 1 - \Big(\sum_{\alpha\leq s} w_\alpha\Big)^n\;, \tag{15.114}$$

and thus

$$|\mathsf{E}\langle f\rangle_{\rho_N,s} - \mathsf{E}\langle f\rangle_{\rho_N}| \leq 1 - \mathsf{E}\Big(\sum_{\alpha\leq s} G_N(A_\alpha)\Big)^n\;. \tag{15.115}$$

We know that the sequence $(G_N(A_\alpha))$ has in the limit a Poisson-Dirichlet distribution Λ_{1-a}. This implies 4), and (15.115) implies

$$\limsup_{N\to\infty} |\mathsf{E}\langle f\rangle_{\rho_N,s} - \mathsf{E}\langle f\rangle_{\rho_N}| \le 1 - b_{s,n} \tag{15.116}$$

where

$$b_{s,n} = \mathsf{E}\Big(\sum_{\alpha\le s} v_\alpha\Big)^n$$

for (v_α) being a Poisson-Dirichlet distribution Λ_{1-a}. Also, from (15.114) and (15.110) we obtain

$$|\mathsf{E}_\lambda\langle f\rangle_{\rho,s} - \mathsf{E}_\lambda\langle f\rangle_\rho| \le 1 - b_{s,n} \ . \tag{15.117}$$

Now the function $\rho \mapsto \langle f\rangle_{\rho,s}$ is continuous on $\mathcal{S}\times\mathcal{Q}$, so

$$\lim_{N\to\infty} \mathsf{E}\langle f\rangle_{\rho_N,s} = \lim_{N\to\infty} \mathsf{E}_{\lambda_N}\langle f\rangle_{\rho,s} = \mathsf{E}_\lambda\langle f\rangle_{\rho,s}$$

because λ_N converges weakly to λ.

Combining with (15.116) and (15.117) we get

$$\limsup_{N\to\infty} |\mathsf{E}\langle f\rangle_{\rho_N} - \mathsf{E}_\lambda\langle f\rangle_\rho| \le 2(1 - b_{n,s})$$

for each s, and since $\lim_{s\to\infty} b_{n,s} = 1$, this proves (15.107).

When f^* and φ are continuous functions, we see from (15.100) and (15.107) that (15.108) and (15.109) follow respectively from the fact that the sequence (G_N) satisfies the extended Ghirlanda-Guerra identities and has μ as Parisi measure. The case where f^* and φ are Borel functions follows by approximation.

It remains to prove (15.111). We recall that given $\delta > 0$, (15.55) holds for $N \ge N_\delta$. This inequality, together with Jensen's inequality and the fact that $\int_{A_\alpha^2} R_{1,2}\mathrm{d}G_N(\boldsymbol{\sigma}^1)\mathrm{d}G_N(\boldsymbol{\sigma}^2) = \|\boldsymbol{\sigma}_\alpha\|^2/N$ implies

$$G_N(A_\alpha) \ne 0 \quad\Rightarrow\quad \left|\frac{\|\boldsymbol{\sigma}_\alpha\|^2}{N} - q^*\right| \le \delta \ .$$

Recalling (15.97) and (15.98), the element ρ_N satisfies

$$\alpha \le r \quad\Rightarrow\quad q_{\alpha,\alpha} = \frac{\|\boldsymbol{\sigma}_\alpha\|^2}{N} \ge q^* - \delta \ .$$

Consider a continuous function $\varphi : [-1,1] \to [0,1]$, and assume that $\varphi = 1$ in a neighborhood of q^*, i.e. $\varphi(q) = 1$ if $|q - q^*| \le \delta$ for some $\delta > 0$. Then, for $N \ge N_\delta$ and any given number s not depending on N, the element ρ_N satisfies

$$\sum_{\alpha\le s} w_\alpha^2 \varphi(q_{\alpha,\alpha}) = \sum_{\alpha\le s} w_\alpha^2 = \sum_{\alpha\le s} G_N(A_\alpha)^2 \ , \tag{15.118}$$

so that, using the notation (15.104)

$$\mathsf{E}_{\lambda_N} \sum_{\alpha \le s} w_\alpha^2 \varphi(q_{\alpha,\alpha}) = \mathsf{E} \sum_{\alpha \le s} G_N(A_\alpha)^2 \ . \tag{15.119}$$

Since the left-hand side of (15.118) is a continuous function of ρ, taking the limit $N \to \infty$ in (15.119) we have

$$\mathsf{E}_\lambda \sum_{\alpha \le s} w_\alpha^2 \varphi(q_{\alpha,\alpha}) = \mathsf{E} \sum_{\alpha \le s} v_\alpha^2 \ , \tag{15.120}$$

where (v_α) is a sequence with distribution Λ_{1-a}, and letting $s \to \infty$

$$\mathsf{E}_\lambda \sum_{\alpha} w_\alpha^2 \varphi(q_{\alpha,\alpha}) = a \ .$$

This is true for any continuous function φ that is equal to 1 in a neighborhood of q^*, so, taking limits, this is true also if $\varphi = \mathbf{1}_{\{x=q^*\}}$. Thus

$$\mathsf{E}_\lambda \sum_{\alpha} w_\alpha^2 \mathbf{1}_{\{q_{\alpha,\alpha}=q^*\}} = a = \mathsf{E}_\lambda \sum_{\alpha} w_\alpha^2 \tag{15.121}$$

using (15.110). Since $w_\alpha > 0$ λ-a.s. this proves $q_{\alpha,\alpha} = q^*$ λ-a.s.

Now (15.109) implies that

$$\mathsf{E}_\lambda \sum_{\alpha,\gamma} w_\alpha w_\gamma \mathbf{1}_{\{q_{\alpha,\gamma} \ge q^*\}} = \mathsf{E}_\lambda \langle \mathbf{1}_{\{R_{1,2} \ge q^*\}} \rangle_\rho = \mu([q^*, \infty[) = a \ .$$

Comparing with (15.121) yields

$$\mathsf{E}_\lambda \sum_{\alpha \ne \gamma} w_\alpha w_\gamma \mathbf{1}_{\{q_{\alpha,\gamma} \ge q^*\}} = 0$$

and this proves that λ a.s. if $\alpha \ne \gamma$ we have $q_{\alpha,\gamma} < q^*$. ☐

Quite naturally, we will say that a determinator λ satisfies the extended Ghirlanda-Guerra identities if (15.108) holds; and that it has μ as Parisi measure if (15.109) holds.

Definition 15.5.4. *We say that a determinator λ is ultrametric if λ a.s. we have*

$$\forall \alpha, \gamma, \delta \qquad q_{\alpha,\delta} \ge \min(q_{\alpha,\gamma}, q_{\gamma,\delta}) \ . \tag{15.122}$$

Lemma 15.5.5. *The determinator λ is ultrametric if and only if the symmetric probability measure μ^* on $\mathcal{C}$ defined by (15.112) is ultrametric in the sense of Definition 15.3.2.*

Proof. Since (15.112) holds for continuous functions, it also holds for

$$f^*(\mathbf{x}) = \mathbf{1}_{\{x_{1,3} < \min(x_{1,2}, x_{2,3})\}}$$

and by Definition 15.3.2, μ^* is ultrametric if and only if $\int f^*(x) \mathrm{d}\mu^*(x) = 0$. Now when

$$0 = \mathsf{E}_\lambda \langle f \rangle_\rho = \mathsf{E}_\lambda \sum_{\alpha,\gamma,\delta} w_\alpha w_\gamma w_\delta \mathbf{1}_{\{q_{\alpha,\gamma} < \min(q_{\alpha,\gamma}, q_{\gamma,\delta})\}},$$

and since $w_\alpha w_\gamma w_\delta > 0$ λ a.s., this implies that (15.122) holds λ a.s. The converse is obvious. □

Conjecture 15.5.6. A determinator that satisfies the conditions of Proposition 15.5.2 is ultrametric.

This of course would lend great support to Conjecture 15.3.3. In the next section we shall prove that the answer is positive if one moreover assumes that the Parisi measure has a finite support. The remainder of the present section is devoted to a truly remarkable technical result of D. Panchenko that will be the key for this.

Consider independent Bernoulli r.v.s $(\eta_\alpha)_{\alpha \geq 1}$ (i.e. $\eta_\alpha = \pm 1$ with probability 1/2). Given $t \in \mathbb{R}$ and

$$\rho = ((w_\alpha)_{\alpha \geq 1}, (q_{\alpha,\gamma})) \in \mathcal{S} \times \mathcal{Q} ,$$

let us define

$$w_{\alpha,t} = \frac{w_\alpha \exp t\eta_\alpha}{\sum_\gamma w_\gamma \exp t\eta_\gamma} . \tag{15.123}$$

For a function $f = f(\alpha^1, \ldots, \alpha^n)$ of n configurations (that might be random) let us define

$$\langle f \rangle_{\rho,t} = \sum_{\alpha^1,\ldots,\alpha^n} w_{\alpha^1,t} \cdots w_{\alpha^n,t} f . \tag{15.124}$$

Theorem 15.5.7. *(Panchenko's invariance theorem [71], [75]) Consider a determinator λ that satisfies the extended Ghirlanda-Guerra identities. Assume moreover that (15.111) holds λ a.e. for a certain number q^*. Then, if f^* is a continuous function on $\mathbb{R}^{n(n-1)/2}$, recalling the notation (15.95) we have*

$$\forall t\,; \; \mathsf{E}_\lambda \langle f \rangle_\rho = \mathsf{E}\mathsf{E}_\lambda \langle f \rangle_{\rho,t} = \mathsf{E} \int \langle f \rangle_{\rho,t} \mathrm{d}\lambda(\rho) , \tag{15.125}$$

where E denotes of course expectation in the r.v.s η_α.

Proof. Consider the function

$$\psi(t) = \mathsf{E}\mathsf{E}_\lambda \langle f \rangle_{\rho,t} \; .$$

We will show that

$$|\psi^{(s)}(t)| \le 2^s n(n+1)\cdots(n+s-1) \sup |f| \tag{15.126}$$

and

$$\forall s \ge 1 \; , \quad \psi^{(s)}(0) = 0 \; . \tag{15.127}$$

We now complete the proof as follows. Assuming that for some t_0 we have

$$\psi(t_0) = \psi(0) \, ; \, \forall \, s \ge 1 \; , \; \psi^{(s)}(t_0) = 0 \; , \tag{15.128}$$

we prove that this holds for each t with $|t - t_0| < 1/2$. Since this holds for $t_0 = 0$ by (15.127) this completes the proof. To prove (15.128) we simply use that by Taylor's formula for each integers $s \ge 0$ and $k \ge 1$ we have

$$|\psi^{(s)}(t) - \psi^{(s)}(t_0)| \le |t - t_0|^k \sup_u \frac{|\psi^{(s+k)}(u)|}{k!} \; ,$$

and as $k \to \infty$ when $|t - t_0| < 1/2$ the right-hand side goes to zero by (15.126).

We now turn to the proof of (15.126) and (15.127). To lighten notation we write $\nu_t(f)$ rather than $\mathsf{E}\mathsf{E}_\lambda \langle f \rangle_{\rho,t}$ and $\nu(f)$ rather than $\nu_0(f)$.

For any $n \ge 0$ and $\ell \ge 1$ we define the number $c(\ell, n)$ as follows

$$c(\ell, n) = 1 \;\text{ if } \ell \le n \; ; \quad c(n+1, n) = -n \; ; \quad c(\ell, n) = 0 \;\text{ if } \ell \ge n+2$$

so that

$$\sum_{\ell \ge 1} c(\ell, n) = 0 \; ; \quad \sum_{\ell \ge 1} |c(\ell, n)| = 2n \; . \tag{15.129}$$

We consider the quantity

$$D_n = \sum_\ell c(\ell, n) \eta_{\alpha^\ell} \; . \tag{15.130}$$

This is a random function of the $n+1$ configurations $\alpha^1, \ldots, \alpha^{n+1}$. If $h = h(\alpha^1, \ldots, \alpha^n)$ is a function of n configurations (that might also depend on the r.v. η_α) we have the identity

$$\frac{\mathrm{d}}{\mathrm{d}t}(\nu_t(h)) = \nu_t(h D_n) \; . \tag{15.131}$$

To see this we write the explicit value $\langle h \rangle_{\rho,t}$:

$$\begin{aligned}
\langle h \rangle_{\rho(t)} &= \sum_{\alpha^1, \ldots, \alpha^n} w_{\alpha^1,t} \cdots w_{\alpha^n,t} h(\alpha^1, \ldots, \alpha^n) \\
&= \frac{\sum_{\alpha^1, \ldots, \alpha^n} w_{\alpha^1} \cdots w_{\alpha^n} \exp(t \sum_{\ell \le n} \eta_{\alpha^\ell}) h(\alpha^1, \ldots, \alpha^n)}{\left(\sum_\alpha w_\alpha \exp t\eta_\alpha\right)^n} \; ,
\end{aligned}$$

and thus (differentiating and repeating the previous step in the reverse order)

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle h\rangle_{\rho(t)} = \langle hD_n\rangle_{\rho(t)} \ ,$$

from which (15.131) follows by taking expectation.

Iteration of the formula (15.131) yields the relation

$$\psi^{(s)}(t) = \nu_t(fD_n \cdots D_{n+s-1}) \ . \tag{15.132}$$

Since $|D_n| \le 2n$ by (15.130) and (15.132) we obtain (15.126), and we turn to the proof of (15.127). First we compute $\psi^{(s)}(0)$, by substitution of the value (15.130) in (15.132). To lighten notation we write

$$\varepsilon_\ell = \eta_{\alpha^\ell} \ , \tag{15.133}$$

which is a random function of the ℓ^{th} configuration. Thus

$$\psi^{(s)}(0) = \nu\biggl(f \sum_{\ell_1,\ldots,\ell_n} c(\ell_1, n)\cdots c(\ell_s, n+s-1)\varepsilon_{\ell_1}\cdots\varepsilon_{\ell_s}\biggr) \ .$$

Now $\nu(h) = \mathsf{E}\mathsf{E}_\lambda\langle h\rangle_\rho$, and since ρ does not depend on the r.v. η_α we have $\nu(h) = \mathsf{E}_\lambda\langle \mathsf{E}h\rangle_\rho = \nu(\mathsf{E}h)$, and this shows that

$$\psi^{(s)}(0) = \nu\biggl(f \sum_{\ell_1,\ldots,\ell_s} c(\ell_1, n)\cdots c(\ell_s, n+s-1)A(\ell_1,\ldots,\ell_s)\biggr) \tag{15.134}$$

where

$$A(\ell_1,\ldots,\ell_s) := \mathsf{E}\varepsilon_{\ell_1}\cdots\varepsilon_{\ell_s} \tag{15.135}$$

is a function of the configurations $\alpha^1,\ldots,\alpha^s$. By parity (i.e. changing each η_α in $-\eta_\alpha$ does not change the law of this family) we see that $A(\ell_1,\ldots,\ell_s) = 0$ when s is odd, so to prove that $\psi^{(s)}(0) = 0$ we only have to consider the case where s is even.

If we think of $\alpha^1,\ldots,\alpha^s$ as fixed, each of the r.v.s $\varepsilon_{\ell_1},\ldots,\varepsilon_{\ell_k}$ is one of the independent Bernoulli r.v.s η_α. Then $A(\ell_1,\ldots,\ell_s) = 1$ if each of the r.v.s η_α that occurs at least one time in the product $\varepsilon_{\ell_1},\ldots,\varepsilon_{\ell_k}$ occurs an even number of times in this product, and $A(\ell_1,\ldots,\ell_s) = 0$ otherwise. It does not seem easy to use this fact directly, but it is what motivates the next step of the proof; this step establishes an explicit formula for $A(\ell_1,\ldots,\ell_s)$ using a kind of expansion. To prepare for this we observe that there exist integers $w(i)$, $i = 1, 3, \ldots$ such that whenever p is an odd integer, we have

$$\sum_{i\le p,\ i\,\mathrm{odd}} w(i)\binom{p}{i} = 1 \ . \tag{15.136}$$

These integers are constructed recursively, setting $w(1) = 1$ and, setting for p odd

$$w(p+2) = 1 - \sum_{i \le p,\, i\,\mathrm{odd}} w(i) \binom{p+2}{i} .$$

For a subset I of $\{1, \dots, s\}$, we define

$$w'(I) = w(\mathrm{card} I - 1)$$

if $\mathrm{card} I$ is even, and $w'(I) = 0$ if $\mathrm{card} I$ is odd. For a subset I of $\{1, \dots, s\}$, we also define

$$a(I, \ell_1, \dots, \ell_s) = 1 \quad \text{if} \quad \forall i, j \in I\ , \quad \alpha^{\ell_i} = \alpha^{\ell_j}\ , \tag{15.137}$$

and $a(I, \ell_1, \dots, \ell_s) = 0$ otherwise. This is a function of the configurations $\alpha^1, \dots, \alpha^s$. Thus if $a(I, \ell_1, \dots, \ell_s) = 1$ then $\varepsilon_{\ell_i} = \varepsilon_{\ell_j}$ for $i, j \in I$ and if $\mathrm{card} I$ is even we have $\prod_{i \in I} \varepsilon_{\ell_i} = 1$. Let us also note that the value of $a(I, \ell_1, \dots, \ell_s)$ does not depend on ℓ_i for $i \notin I$. We claim that

$$A(\ell_1, \dots, \ell_s) = \sum w'(I) a(I, \ell_1, \dots, \ell_s)\, \mathsf{E} \prod_{i \notin I} \varepsilon_{\ell_i}\ , \tag{15.138}$$

where the summation is over all subsets I of $\{1, \dots, s\}$ with $\mathrm{card} I \ge 2$, $\mathrm{card} I$ even and $1 \in I$. To prove this, let us consider

$$I_0 = \{i \le s\ ;\ \alpha^{\ell_i} = \alpha^{\ell_1}\}\ ,$$

so that when $1 \in I$ we have $a(I, \ell_1, \dots, \ell_s) = 1$ if $I \subset I_0$ and $a(I, \ell_1, \dots, \ell_s) = 0$ otherwise. Also, when $I \subset I_0$ and $\mathrm{card} I$ is even, then $\prod_{i \in I} \varepsilon_{\ell_i} = \varepsilon_{\ell_1}^{\mathrm{card} I} = 1$ since $\mathrm{card} I$ is even and $\varepsilon_{\ell_1} = \pm 1$, and thus

$$\mathsf{E} \prod_{i \notin I} \varepsilon_{\ell_i} = \mathsf{E} \prod_{i \notin I} \varepsilon_{\ell_i} \prod_{i \in I} \varepsilon_{\ell_i} = \mathsf{E} \prod_{i \le s} \varepsilon_{\ell_i} = A(\ell_1, \dots, \ell_s)\ .$$

Therefore the right-hand side of (15.138) is

$$A(\ell_1, \dots, \ell_s) \sum_{I \subset I_0} w'(I)$$

where the summation is as in (15.138), i.e. $1 \in I$ and $\mathrm{card} I$ is even. The set I is determined by a subset of $I_0 \setminus \{1\}$ of cardinality $i = \mathrm{card} I - 1$. Thus, if $s' = \mathrm{card} I_0$ we have

$$\sum_I w'(I) = \sum_{i\,\mathrm{odd},\, i \le s'-1} w(i) \binom{s'-1}{i} = 1$$

by (15.136), and this proves (15.138).

Recalling that $A(\ell_1, \dots, \ell_s) = \mathsf{E} \varepsilon_{\ell_1} \cdots \varepsilon_{\ell_s}$, we observe that the quantity $\mathsf{E} \prod_{i \notin I} \varepsilon_{\ell_i}$ is of the same nature, but with a product of fewer quantities ε_{ℓ_s}. Therefore we can iterate the procedure (15.138) to find an expansion

$$A(\ell_1,\ldots,\ell_s) = \sum w(I_1,\ldots,I_r) \prod_{j\le r} a(I_j,\ell_1,\ldots,\ell_s) \ ,$$

where $w(I_1,\ldots,I_r)$ is a number, and where the summation is over all partitions of $\{1,\ldots,s\}$ into sets $I_1,\ldots,I_r$ with cardI_j even. Recalling (15.134), to prove that $\psi^{(s)}(0) = 0$ it suffices to prove that for each partition $I_1,\ldots,I_r$ of $\{1,\ldots,s\}$ as above we have

$$\nu\Bigl(f \sum_{\ell_1,\ldots,\ell_s} c(\ell_1,n)\cdots c(\ell_s,n+s-1) \prod_{j\le r} a(I_j,\ell_1,\ldots,\ell_s)\Bigr) = 0 \ . \tag{15.139}$$

We need to prove (15.139) only when cardI_j is even for each j, but will prove it for any partition $I_1,\ldots,I_r$ of $\{1,\ldots,s\}$ with card$I_j \ge 1$. When card$I = 1$, we define $a(I,\ell_1,\ldots,\ell_s) = 1$. Without loss of generality we can assume that $s \in I_1$.

We first consider the case where card$I_1 = 1$. In that case, $a(I_1,\ell_1,\ldots,\ell_s) = 1$ does not depend on ℓ_s and since $s \notin I_j$ for $j \ge 2$, $a(I_j,\ell_1,\ldots,\ell_s)$ does not depend on ℓ_s either. So in (15.139) the sum $\sum_{\ell_s} c(\ell_s,n+s-1)$ factors out, and this sum is 0. In particular (15.139) holds when $s = 1$.

We consider then the case where card$I_1 \ge 2$, and we consider an element $s' < s$ of I_1. We define $I'_1 = I_1 \setminus \{s\}$ and $I'_j = I_j$ for $j \ge 2$. We observe the identity

$$a(I_1,\ell_1,\ldots,\ell_s) = a(I'_1,\ell_1,\ldots,\ell_{s-1}) V(\ell_{s'},\ell_s) \ , \tag{15.140}$$

where $V(\ell_{s'},\ell_s) = 1$ if $\alpha^{\ell_{s'}} = \alpha^{\ell_s}$ and $V(\ell_{s'},\ell_s) = 0$ otherwise. Also, for $2 \le j \le r$ we have $s \notin I_j$, so

$$a(I_j,\ell_1,\ldots,\ell_s) = a(I'_j,\ell_1,\ldots,\ell_{s-1}) \ . \tag{15.141}$$

The only values of $\ell_1,\ldots,\ell_{s-1}$ that matter in the summations are those $\le n+s-1$. We fix such values and we intend to prove that

$$\begin{aligned} &\nu\Bigl(f \sum_{\ell_s} c(\ell_s,n+s-1) \prod_{j\le r} a(I_j,\ell_1,\ldots,\ell_s)\Bigr) \\ &= b\nu\Bigl(f \prod_{j\le r} a(I'_j,\ell_1,\ldots,\ell_{s-1})\Bigr) \ , \end{aligned} \tag{15.142}$$

where b is a quantity that does not depend on $\ell_1,\ldots,\ell_{s-1}$. This equality shows that the left-hand side of (15.139) is

$$b\nu\Bigl(f \sum_{\ell_1,\ldots,\ell_{s-1}} c(\ell_1,n)\cdots c(\ell_s,n+s-2) \prod_{j\le r} a(I'_j,\ell_1,\ldots,\ell_{s-1})\Bigr) \ .$$

This proves (15.139) by induction over s.

We turn to the proof of (15.142). This proof relies on the extended Ghirlanda-Guerra identities, that we reformulate now for this purpose. Let us define $R_{\ell,\ell} = q^*$. Then we can rewrite (15.108) as

$$\sum_{1\le\ell\le n} \nu(\varphi(R_{1,\ell})f) - n\nu(f\varphi(R_{1,n+1})) = b\nu(f)$$

where $b = \varphi(q^*) - \nu(\varphi(R_{1,2}))$ does not depend on f. Recalling the definition of $c(\ell, n)$ this means that

$$\nu\bigg(f\sum_{\ell} c(\ell,n)\varphi(R_{1,\ell})\bigg) = b\nu(f)\ , \tag{15.143}$$

and more generally by symmetry among sites, if $p \le n$,

$$\nu\bigg(f\sum_{\ell} c(\ell,n)\varphi(R_{p,\ell})\bigg) = b\nu(f)\ . \tag{15.144}$$

Using (15.140) and (15.141) the left-hand side of (15.142) is

$$\nu\bigg(f\sum_{\ell} c(\ell,n+s-1)V(\ell_{s'},\ell)\prod_{j\le p} a(I'_j,\ell_1,\ldots,\ell_{s-1})\bigg)\ . \tag{15.145}$$

We aim to deduce (15.142) from (15.144). We write (15.144) as follows. When $f^{\sim}$ is a continuous function on $\mathbb{R}^{m(m-1)/2}$, and if we define

$$f'(\alpha^1,\ldots,\alpha^m) = f^{\sim}((R_{\ell,\ell'})_{1\le\ell<\ell'\le m}) \tag{15.146}$$

then whenever $p \le m$ we have

$$\nu\bigg(f'\sum_{\ell} c(\ell,m)\varphi(R_{p,\ell})\bigg) = b\nu(f') \tag{15.147}$$

where b depends on φ only. This equality also holds if φ and $f^{\sim}$ are Borel functions. We choose $\varphi(x) = \mathbf{1}_{\{x\ge q^*\}}$, $m = n+s-1$, $p = \ell_{s'}$. For $j \le p$, we define

$$f_j^{\sim}((x_{\ell,\ell'})_{1\le\ell<\ell'\le m}) = \prod_{\ell<\ell',\,\ell,\ell'\in I_j} \mathbf{1}_{\{x_{\ell,\ell'}\ge q^*\}}$$

and we define

$$f^{\sim} = f^*\prod_{j\le p} f_j^{\sim}\ ,$$

and f' by (15.146), so that $f' = f\prod_{j\le p} f_j$ where

$$f_j(\alpha^1,\ldots,\alpha^m) = \prod_{\ell<\ell',\ell,\ell'\in I_j} \mathbf{1}_{\{R_{\ell,\ell'}\ge q^*\}}\ .$$

Using (15.147) for this choice of f' yields

$$\nu\Bigl(f\prod_{j\le p} f_j \sum_{\ell} c(\ell,m)\varphi(R_{p,\ell})\Bigr) = b\nu\Bigl(f\prod_{j\le r} f_j\Bigr) . \tag{15.148}$$

Now, λ a.s. we have

$$\alpha=\gamma \quad\Leftrightarrow\quad q_{\alpha,\gamma}=q^*$$

so that

$$\alpha^p=\alpha^\ell \quad\Leftrightarrow\quad R_{p,\ell}\ge q^* ,$$

and hence

$$\varphi(R_{p,\ell}) = V(p,\ell) \tag{15.149}$$

$$f_j = a(I'_j,\ell_1,\ldots,\ell_{s-1}) , \tag{15.150}$$

so that we can replace $\varphi(R_{p,\ell})$ by $V(p,\ell)$ and f_j by $a(I'_j,\ell_1,\ldots,\ell_{s-1})$ in (15.148) to obtain that the quantity (15.145) is equal to the right-hand side of (15.142). □

It is important to stress the very beautiful idea of Theorem 15.5.7. The natural approach is, instead of defining η_α as Bernoulli r.v.s, to use independent standard Gaussian r.v.s. In that case it is much easier to prove that $\psi^{(s)}(0)=0$; but one does not know how to control the size of $\psi^{(s)}(t)$.

Here is a simple but important fact.

Lemma 15.5.8. *If a determinator λ satisfies the extended Ghirlanda-Guerra identities, under λ the sequence (w_α) has a Poisson-Dirichlet distribution Λ_m where $m=1-\mathsf{E}_\lambda\sum_{\alpha\ge1}w_\alpha^2$.*

Proof. Condition (15.160) means that λ a.s. we have $q_{\alpha,\gamma}\ge q_1 \Leftrightarrow \alpha=\gamma$.

As we have already used in Section 14.3 the extended Ghirlanda-Guerra identities determine the quantities $\mathsf{E}_\lambda\prod_{s\le k}\sum_{\alpha\ge1}w_\alpha^{n_s}$, which in turn determines the law of the quantities (w_α) as is shown in Proposition 15.2.4. □

We now turn to a striking consequence of Theorem 15.5.7. Consider an independent sequence of r.v.s ζ_α with $\mathsf{P}(\zeta_a=1)=1/2$ and $\mathsf{P}(\zeta_a=0)=1/2$, and, given $\rho=((w_\alpha),(q_{\alpha,\gamma}))\in\mathcal{S}\times\mathcal{Q}$, let us define

$$w'_\alpha=\frac{\zeta_\alpha w_\alpha}{\sum_\gamma \zeta_\gamma w_\gamma} .$$

The denominator is a.s. not zero because Lemma 15.5.8 implies that infinitely many of the weights w_γ are not zero. In words, we simply delete at random one half of the terms w_α and we renormalize. For a function f of n configurations, we define

$$\langle f\rangle'_\rho=\sum_{\alpha^1,\ldots,\alpha^n} w'_{\alpha^1}\cdots w'_{\alpha^n} f .$$

Proposition 15.5.9. *(D Panchenko [75]) Consider a determinator λ that satisfies the extended Ghirlanda-Guerra identities. Assume moreover that (15.111) holds λ a.e. for a certain number q^*. Then, if f^* is a continuous function on $\mathbb{R}^{n(n-1)/2}$, recalling the notation (15.95) we have*

$$\mathsf{E}_\lambda\langle f\rangle_\rho = \mathsf{E}\mathsf{E}_\lambda\langle f\rangle'_\rho = \mathsf{E}\int\langle f\rangle'_\rho \mathrm{d}\lambda(\rho) \,. \tag{15.151}$$

Proof. It is a simple matter to deduce this from Theorem 15.5.7 simply by letting $t\to\infty$. The details are left to the reader. □

An important property of Theorem 15.5.7 is that it can be iterated. Consider independent Bernoulli r.v.s $\eta_{\alpha,\ell}$. Given

$$\rho = ((w_\alpha),(q_{\alpha,\gamma}))\in\mathcal{S}\times\mathcal{Q}\,,$$

let us define

$$w^s_\alpha = \frac{w_\alpha \exp\sum_{\ell\le s}\eta_{\alpha,\ell}}{\sum_\gamma w_\gamma \exp\sum_{\ell\le s}\eta_{\gamma,\ell}}\,. \tag{15.152}$$

Theorem 15.5.10. *Consider a determinator λ that satisfies the extended Ghirlanda-Guerra identities. Assume moreover that (15.111) holds λ a.e. for a certain number q^*. Then for each Borel function f^* on $\mathbb{R}^{n(n-1)/2}$, recalling the notation (15.95), for each integer s we have*

$$\mathsf{E}_\lambda\langle f\rangle_\rho = \mathsf{E}\mathsf{E}_\lambda\sum_{\alpha^1,\dots,\alpha^n} w^s_{\alpha^1}\cdots w^s_{\alpha^n} f\,. \tag{15.153}$$

Proof. It suffices to prove (15.153) when f^* is continuous. The proof is by induction on s. For $s=1$ this is a consequence of Theorem 15.5.7 used for $t=1$. We now perform the induction step from s to $s+1$. Given $\rho\in\mathcal{S}\times\mathcal{Q}$ and the r.v.s $\eta_{\alpha,\ell}$ for $\alpha\ge1$ and $\ell\le s$, let us consider a permutation τ of $\mathbb{N}^*$ such that the sequence $w^s_{\tau(\alpha)}$ is not decreasing, and

$$\rho^* := ((w^s_{\tau(\alpha)}),(q_{\tau(\alpha),\tau(\gamma)}))\in\mathcal{S}\times\mathcal{Q}\,.$$

When f^* is a continuous function on $\mathbb{R}^{n(n-1)/2}$, recalling the notation (15.95), we have by definition

$$\begin{aligned}\langle f\rangle_{\rho^*} &= \sum_{\alpha^1,\dots,\alpha^n} w^s_{\tau(\alpha^1)}\cdots w^s_{\tau(\alpha^n)} f^*((q_{\tau(\alpha^\ell),\tau(\alpha^{\ell'})})_{1\le\ell<\ell'\le n})\\ &= \sum_{\alpha^1,\dots,\alpha^n} w^s_{\alpha^1}\cdots w^s_{\alpha^n} f((q_{\alpha^\ell,\alpha^{\ell'}})_{1\le\ell<\ell'\le n})\,,\end{aligned} \tag{15.154}$$

and combining with the induction hypothesis (15.153) we get

$$\mathsf{E}\mathsf{E}_\lambda\langle f\rangle_{\rho^*} = \mathsf{E}_\lambda\langle f\rangle_\rho\,. \tag{15.155}$$

Let us now think of ρ^* as a random element of $\mathcal{S} \times \mathcal{Q}$, that depends on ρ (to which we think of as random element of $\mathcal{S}\times\mathcal{Q}$) and on the r.v.s $(\eta_{\alpha,\ell})_{\alpha\geq 1,\ell\leq s}$. Let us denote by λ^* the law of ρ^*. Thus, using the definition of λ^* in the first equality and (15.155) in the second one, we get

$$\mathsf{E}_{\lambda^*}\langle f\rangle_\rho = \mathsf{E}\mathsf{E}_\lambda\langle f\rangle_{\rho^*} = \mathsf{E}_\lambda\langle f\rangle_\rho \ . \tag{15.156}$$

This implies that λ^* satisfies the extended Ghirlanda-Guerra identities, because λ does. (In fact one can prove that $\lambda^* = \lambda$, see [71].) It should be clear by construction that (15.111) holds λ^* a.e. Therefore we can use Theorem 15.5.7 for λ^* and $t = 1$. Thus, when f^* is a continuous function on $\mathbb{R}^{n(n-1)/2}$ and $(\eta_\alpha)_{\alpha\geq 1}$ are independent Bernoulli r.v.s (independent of everything else) we have

$$\begin{aligned}
&\mathsf{E}_{\lambda^*}\sum_{\alpha^1,\ldots,\alpha^n} w_{\alpha^1}\cdots w_{\alpha^n} f^*((q_{\alpha^\ell,\alpha^{\ell'}})_{1\leq\ell<\ell'\leq n})\\
&= \mathsf{E}_{\lambda^*}\sum_{\alpha^1,\ldots,\alpha^n} \frac{w_{\alpha^1}\cdots w_{\alpha^n}\exp\left(\sum_{\ell\leq n}\eta_{\alpha^\ell}\right)}{\left(\sum_\gamma w_\gamma\exp(\eta_\gamma)\right)^n} f^*((q_{\alpha^\ell,\alpha^{\ell'}})_{1\leq\ell<\ell'\leq n}) \ .
\end{aligned} \tag{15.157}$$

By definition of λ^* (as being the law of ρ^*) this means that

$$\begin{aligned}
&\mathsf{E}\mathsf{E}_\lambda\sum_{\alpha^1,\ldots,\alpha^n} w^s_{\tau(\alpha^1)}\cdots w^s_{\tau(\alpha^n)} f^*((q_{\tau(\alpha^\ell),\tau(\alpha^{\ell'})})_{1\leq\ell<\ell'\leq n}) = \\
&\mathsf{E}\mathsf{E}_\lambda\sum_{\alpha^1,\ldots,\alpha^n} \frac{w^s_{\tau(\alpha^1)}\cdots w^s_{\tau(\alpha^n)}\exp\left(\sum_{\ell\leq n}\eta_{\alpha^\ell}\right)}{\left(\sum_\gamma w^s_\gamma\exp(\eta_\gamma)\right)^n} f^*((q_{\tau(\alpha^\ell),\tau(\alpha^{\ell'})})_{1\leq\ell<\ell'\leq n}) \ ,
\end{aligned} \tag{15.158}$$

and consequently

$$\begin{aligned}
&\mathsf{E}\mathsf{E}_\lambda\sum_{\alpha^1,\ldots,\alpha^n} w^s_{\alpha^1}\cdots w^s_{\alpha^n} f^*((q_{\alpha^\ell,\alpha^{\ell'}})_{1\leq\ell<\ell'\leq n}) = \\
&\mathsf{E}\mathsf{E}_\lambda\sum_{\alpha^1,\ldots,\alpha^n} \frac{w^s_{\alpha^1}\cdots w^s_{\alpha^n}\exp\left(\sum_{\ell\leq n}\eta_{\tau^{-1}(\alpha^\ell)}\right)}{\left(\sum_\gamma w^s_\gamma\exp(\eta_\gamma)\right)^n} f^*((q_{\alpha^\ell,\alpha^{\ell'}})_{1\leq\ell<\ell'\leq n})\\
&= \mathsf{E}\mathsf{E}_\lambda\sum_{\alpha^1,\ldots,\alpha^n} w^*_{\alpha^1}\cdots w^*_{\alpha^n} f^*((q_{\alpha^\ell,\alpha^{\ell'}})_{1\leq\ell<\ell'\leq n}) \ ,
\end{aligned} \tag{15.159}$$

where

$$w^*_\alpha = \frac{1}{Z} w_\alpha \exp\left(\sum_{\ell\leq s}\eta_{\alpha,\ell} + \eta_{\tau^{-1}(\alpha)}\right) \ ,$$

where Z is the normalizing factor $\sum_\gamma w_\gamma \exp(\sum_{\ell\leq s}\eta_{\gamma,\ell} + \eta_{\tau^{-1}(\gamma)})$. Given ρ the permutation τ is determined by the r.v.s $\eta_{\alpha,\ell}$, and since the sequence (η_α) is independent of these we have, denoting by E' expectation in the r.v.s η_α only,

$$\begin{aligned}&\mathsf{E}'\sum_{\alpha^1,\ldots,\alpha^n} w^*_{\alpha^1}\cdots w^*_{\alpha^n} f^*((q_{\alpha^\ell,\alpha^{\ell'}})_{1\le\ell<\ell'\le n})\\ &=\mathsf{E}'\sum_{\alpha^1,\ldots,\alpha^n} w^{s+1}_{\alpha^1}\cdots w^{s+1}_{\alpha^n} f^*((q_{\alpha^\ell,\alpha^{\ell'}})_{1\le\ell<\ell'\le n})\ ,\end{aligned}$$

so that (15.159) implies

$$\mathsf{E}\mathsf{E}_\lambda\sum_{\alpha^1,\ldots,\alpha^n} w^s_{\alpha^1}\cdots w^s_{\alpha^n} f=\mathsf{E}\mathsf{E}_\lambda\sum_{\alpha^1,\ldots,\alpha^n} w^{s+1}_{\alpha^1}\cdots w^{s+1}_{\alpha^n} f\ ,$$

and this concludes the induction. □

15.6 Panchenko's Ultrametricity Theorem

In this section we prove the following, where we recall Definitions 15.5.3 and 15.5.4.

Theorem 15.6.1. *Consider a determinator λ that satisfies the extended Ghirlanda-Guerra identities. Assume that there is a finite set $F=\{q_r,\ldots,q_1\}$ with $q_r<\ldots<q_1$ such that λ a.s. $q_{\alpha,\gamma}\in F$ for each α,γ. Assume moreover that λ a.s. we have*

$$q_{\alpha,\alpha}=q_1\ ;\qquad \alpha\ne\gamma\quad\Rightarrow\quad q_{\alpha,\gamma}<q_1\ . \tag{15.160}$$

Then λ is ultrametric.

We recall the set $\mathcal{C}$ of Section 15.3 and Definitions 15.4.1 and 15.3.2.

Corollary 15.6.2. *Consider a sequence (G_N) of random measures on Σ_N. Assume that it satisfies the extended Ghirlanda-Guerra identities, and that it has a Parisi measure μ with a finite support. Then for each continuous function f^* on $\mathbb{R}^{n(n-1)/2}$ the limit*

$$\lim_{N\to\infty}\mathsf{E}\langle f^*((R_{\ell,\ell'})_{1\le\ell<\ell'\le n})\rangle \tag{15.161}$$

exists. Moreover, setting $\mathbf{x}_n=(x_{\ell,\ell'})_{1\le\ell<\ell'\le n}$, the probability measure μ^* on $\mathcal{C}$ defined by

$$\int f^*(\mathbf{x}_n)\mathrm{d}\mu^*(\mathbf{x})=\lim_{N\to\infty}\mathsf{E}\langle f^*((R_{\ell,\ell'})_{1\le\ell<\ell'\le n})\rangle \tag{15.162}$$

for every such function f^* is ultrametric.

Proof. By hypothesis μ is supported by a finite set $\{q_r, \ldots, q_1\}$, with $q_r < \ldots < q_1$. Without loss of generality we may assume that $\mu(\{q_1\}) > 0$. We consider a subsequence along which all the limits exist. We then use Theorem 15.4.4 and Proposition 15.5.2 to construct a determinator λ that satisfies the extended Ghirlanda-Guerra identities, that has μ as a Parisi measure, and that satisfies (15.160). Since μ has a finite support, λ is ultrametric by Theorem 15.6.1. By Lemma 15.5.5 the measure μ^* generated by λ through (15.112) is ultrametric. It follows from (15.107) that it coincides with the measure μ^* given by (15.162), which is therefore ultrametric. Since μ^* also satisfies the extended Ghirlanda-Guerra identities, Theorem 14.4.4 shows that the measure μ^* is completely determined by μ so that it does not depend on the subsequence, and the limit exists. □

Can we construct interesting sequences (G_N) satisfying the conditions of Corollary 15.6.2 through Proposition 15.4.3? Equivalently, can we choose $\boldsymbol{\beta}$ such that the Hamiltonian (14.406) has a Parisi measure with finite support (not reduced to one or two points)? Since it is very difficult to compute the Parisi measure, the answer to this question is not known, but partial arguments given in [96] incline the author to believe that the answer is yes.

Research Problem 15.6.3. (Level 3) Find all the probability measures on $[0,1]$ that arise as a Parisi measure of a Hamiltonian (14.406).

Our proof of Theorem 15.6.1 will use some elementary probabilistic estimates.

Proposition 15.6.4. *Consider independent r.v.s* $X, X_1, X_2, X_3 \geq 0$*, and assume that there is a number* M *and a number* C *with the following properties*

$$\mathsf{P}(X + X_1 + X_2 + X_3 \leq M) \geq \frac{1}{2} \tag{15.163}$$

$$\forall t > 0\,, \quad \mathsf{P}(X \leq tM) \leq Ct^{10}\,. \tag{15.164}$$

Then we have

$$\mathsf{E} \prod_{\ell \leq 3} \frac{X_\ell}{X + X_1 + X_2 + X_3} \leq K(C) \prod_{\ell \leq 3} \mathsf{E} \frac{X_\ell}{X + X_1 + X_2 + X_3} \tag{15.165}$$

where $K(C)$ *depends on* C *only.*

Proof. First we note that

$$\prod_{\ell \leq 3} \frac{X_\ell}{X + X_1 + X_2 + X_3} \leq \prod_{\ell \leq 3} \frac{X_\ell}{X + X_\ell}\,,$$

so that denoting by E_0 expectation given X we obtain

$$\mathsf{E}_0 \prod_{\ell \le 3} \frac{X_\ell}{X + X_1 + X_2 + X_3} \le \prod_{\ell \le 3} \mathsf{E}_0 \frac{X_\ell}{X + X_\ell} \ . \tag{15.166}$$

Let us define the functions

$$\varphi_\ell(x) = \mathsf{E}\frac{X_\ell}{x + X_\ell} \ ,$$

so that taking expectation in (15.166) and using Hölder's inequality yields

$$A := \mathsf{E} \prod_{\ell \le 3} \frac{X_\ell}{X + X_1 + X_2 + X_3} \le \mathsf{E} \prod_{\ell \le 3} \varphi_\ell(X) \le \Bigl(\prod_{\ell \le 3} \mathsf{E}\varphi_\ell(X)^3 \Bigr)^{1/3} \ . \tag{15.167}$$

Now we observe that the function φ_ℓ is decreasing, and that for $t < 1$ we have

$$\varphi_\ell(tx) = \mathsf{E}\left(\frac{X_\ell}{tx + X_\ell} \right) \le \frac{1}{t} \mathsf{E}\left(\frac{X_\ell}{x + X_\ell} \right) = \frac{1}{t}\varphi_\ell(x) \ .$$

Therefore using this for $x = M$ we get

$$\varphi_\ell(X) = \varphi\ell\Bigl(\frac{X}{M}M\Bigr) \le \varphi_\ell(M) \max\left(1, \frac{M}{X}\right) \ .$$

Thus using (15.164) for $t = \varphi_\ell(M)/u$ in the last inequality, for $u > \varphi_\ell(M)$ we have

$$\mathsf{P}(\varphi_\ell(X) > u) \le \mathsf{P}\Bigl(\varphi_\ell(M)\frac{M}{X} \ge u\Bigr) = \mathsf{P}\Bigl(X \le M\frac{\varphi_\ell(M)}{u}\Bigr) \le C\Bigl(\frac{\varphi_\ell(M)}{u}\Bigr)^{10} .$$

Therefore (A.28) implies

$$\begin{aligned} \mathsf{E}\varphi_\ell(X)^3 &= \int_0^\infty 3t^2 \mathsf{P}(\varphi_\ell(X) \ge t)\mathrm{d}t \\ &\le \int_0^{\varphi_\ell(M)} 3t^2 \mathrm{d}t + 3C\int_{\varphi_\ell(M)}^\infty t^{-8}\varphi_\ell(M)^{10}\mathrm{d}t \\ &\le K(C)\varphi_\ell(M)^3 \ , \end{aligned} \tag{15.168}$$

and thus

$$A \le K(C) \prod_{\ell \le 3} \varphi_\ell(M) \ . \tag{15.169}$$

Next, consider

$$\Omega = \{X + X_2 + X_3 \le M\} \ ,$$

so $\mathsf{P}(\Omega) \ge 1/2$ by (15.163). Using independence,

$$\mathsf{E}\frac{X_1}{X + X_1 + X_2 + X_3} \ge \mathsf{E}\left(\mathbf{1}_\Omega \frac{X_1}{M + X_1} \right) = \mathsf{P}(\Omega)\mathsf{E}\frac{X_1}{M + X_1} \ge \frac{1}{2}\varphi_1(M) \ ,$$

and similarly for $\ell = 2, 3$. Combining with (15.169) and recalling the value of A proves the result. $\square$

Proposition 15.6.5. *Consider* $0 < m < 1$ *and a r.v.* $U \geq 0$ *with* $\mathsf{E}U^m < \infty$. *Consider i.i.d. copies* $(U_k)_{k\geq 1}$ *of* U *and define*

$$S = \sum_{k\geq 1} k^{-1/m} U_k \ . \tag{15.170}$$

Then for $m' < m$ *we have*

$$(\mathsf{E}S^{m'})^{1/m'} \leq K(m,m')(\mathsf{E}U^m)^{1/m} \tag{15.171}$$

where $K(m,m')$ *depends only on* m *and* m'.

Proof. By homogeneity we may assume that $\mathsf{E}U^m = 1$. Let $Z_k = k^{-1/m}U_k$, so

$$\begin{aligned}\sum_{k\geq 1} \mathsf{P}(Z_k > t) &= \sum_{k\geq 1} \mathsf{P}(k^{-1/m}U > t) = \sum_{k\geq 1} \mathsf{P}\Big(\Big(\frac{U}{t}\Big)^m \geq k\Big) \\ &\leq \int_0^\infty \mathsf{P}\Big(\Big(\frac{U}{t}\Big)^m \geq x\Big)\mathrm{d}x = \mathsf{E}\frac{U^m}{t^m} = \frac{1}{t^m} \ . \end{aligned} \tag{15.172}$$

For $r \in \mathbb{Z}$, define

$$S_r = \sum \{Z_k \ ; \ 2^r \leq Z_k \leq 2^{r+1}\} \leq 2^{r+1}\mathrm{card}\{k \ ; \ Z_k \geq 2^r\} \ ,$$

so that, since $n^{m'} \leq n$ for any integer $n \geq 0$,

$$S_r^{m'} \leq (2^{r+1})^{m'}\mathrm{card}\{k \ ; \ Z_k \geq 2^r\} \ ,$$

and using (15.172),

$$\begin{aligned}\mathsf{E}S_r^{m'} &\leq (2^{r+1})^{m'}\mathsf{E}\mathrm{card}\{k \ ; \ Z_k \geq 2^r\} \\ &= (2^{r+1})^{m'} \sum_{k\geq 1} \mathsf{P}(Z_k \geq 2^r) \leq \frac{(2^{r+1})^{m'}}{2^{rm}} \end{aligned} \tag{15.173}$$

and, similarly

$$\mathsf{E}S_r \leq \frac{2^{r+1}}{2^{rm}} = 2^{r(1-m)+1} \ .$$

Thus, if $T = \sum_{r\leq 0} S_r$, we have $\mathsf{E}T \leq K(m,m')$, and therefore since $T^{m'} \leq 1 + T$ we have $\mathsf{E}T^{m'} \leq K(m,m')$. We then simply observe that $S = T + \sum_{r\geq 1} S_r$, so that, since $m' < 1$, we have

$$S^{m'} \leq T^{m'} + \sum_{r\geq 1} S_r^{m'} \ ,$$

we take expectation and we use (15.173). □

Proposition 15.6.6. *Consider* $m < 1$ *and a r.v.* $U \geq 0$ *with* $\mathsf{E}U^m < \infty$*. Then for* $v > 0$ *the r.v.* S *of (15.170) satisfies*

$$\mathsf{P}(S < v(\mathsf{E}U^m)^{1/m}) \leq 3\exp(-v^{-m}) \,. \tag{15.174}$$

Proof. We write $b = (\mathsf{E}U^m)^{1/m}$, so that

$$\begin{aligned}\mathsf{P}(S < vb) &\leq \mathsf{P}(\forall k\,, k^{-1/m}U_k < vb) \\ &= \prod_{k\geq 1}(1 - \mathsf{P}(k^{-1/m}U_k \geq vb)) \\ &\leq \exp\Bigg(-\sum_{k\geq 1}\mathsf{P}\bigg(\Big(\frac{U}{vb}\Big)^m \geq k\bigg)\Bigg)\end{aligned}$$

and

$$\begin{aligned}\sum_{k\geq 1}\mathsf{P}\bigg(\Big(\frac{U}{vb}\Big)^m \geq k\bigg) &\geq \int_0^\infty \mathsf{P}\bigg(\Big(\frac{U}{vb}\Big)^m \geq x\bigg)\mathrm{d}x - 1 \\ &= \mathsf{E}\Big(\frac{U}{vb}\Big)^m - 1 = \frac{1}{v^m} - 1\,,\end{aligned}$$

and the proof is finished. □

Proposition 15.6.7. *Consider* $0 < m < 1$ *and a non-increasing enumeration* $(u_k)_{k\geq 1}$ *of the realization of a Poisson point process of intensity measure* $x^{-m-1}\mathrm{d}x$*. Then a.s. we have*

$$0 < \inf_k k^{1/m}u_k \leq \sup_k k^{1/m}u_k < \infty \,. \tag{15.175}$$

Proof. This is an easy consequence of the fact that a Poisson r.v. of large expectation is sharply concentrated around its expectation. The number N_r of points of the Poisson point process that belong to the interval $[2^{-r}, 2^{-r+1}[$ satisfies

$$\mathsf{E}N_r = \int_{2^{-r}}^{2^{-r+1}} x^{-m-1}\mathrm{d}x = \frac{1}{m}((2^{-r+1})^{-m} - (2^{-r})^{-m}) = c_m 2^{rm}$$

for a certain number c_m, and using (A.60) yields

$$\sum_r \mathsf{P}(|N_r - \mathsf{E}N_r| \geq \mathsf{E}N_r/2) < \infty$$

and (15.175) holds true when $|N_r - \mathsf{E}N_r| \leq \mathsf{E}N_r/2$ for all r large enough. □

The final ingredient of the proof of Theorem 15.6.1 is the following geometrical fact.

Proposition 15.6.8. *Given number $\delta > 0$, there exists a number $\varepsilon > 0$ with the following property. Consider a probability π on the unit ball of a Hilbert space. Assume that for a certain number $q \geq 0$ we have*

$$\pi^{\otimes 2}(\{(x,y)\ ;\ x \cdot y \geq q + \varepsilon\}) \leq \varepsilon\ . \tag{15.176}$$

Then we have

$$\pi^{\otimes 3}(\{(x,y,z)\ ;\ x \cdot y \geq q - \varepsilon\ ,\ |x \cdot z - y \cdot z| \geq \delta\}) \leq \delta\ . \tag{15.177}$$

Proof. This is an immediate consequence of Corollary 15.9.17 of Section 15.9. □

We now start the proof of Theorem 15.6.1. Without loss of generality we may assume that

$$\forall p\,,\ 1 \leq p \leq r\ ,\ \lambda(\{\exists \alpha\,, \gamma\,;\ \ q_{\alpha,\gamma} = q_p\}) > 0\ , \tag{15.178}$$

for otherwise we simply remove the "irrelevant values of q_p".

As a consequence of Proposition 15.6.7, given a number $\kappa > 0$, there exists a number $K(\kappa)$ such that with probability $\geq 1 - \kappa$ for all α we have

$$\frac{1}{K(\kappa)} \alpha^{-1/m} \leq w_\alpha \leq K(\kappa) \alpha^{-1/m}\ . \tag{15.179}$$

The central argument of the proof is as follows.

Lemma 15.6.9. *Consider the set*

$$W = \{(\alpha_1, \alpha_2, \alpha_3)\ ;\ q_{\alpha_1,\alpha_2} = q_2\ ,\ q_{\alpha_1,\alpha_3} < q_2\ ,\ q_{\alpha_2,\alpha_3} \neq q_{\alpha_1,\alpha_3}\}\ . \tag{15.180}$$

Then λ a.s. W is empty.

Proof. Consider a r.v. $U \geq 0$ with $\mathsf{E}U^m < \infty$ and i.i.d. copies $(U_\alpha)_{\alpha \geq 1}$ of U. Through the proof we denote by E_0 expectation in the randomness of the sequence U_α and P_0 the corresponding probability. Until the very end of the proof (where we will take expectation in the randomness of the sequence (w_α)), all the estimates are done given a realization of the sequence (w_α); we assume that this given realization satisfies (15.179). Let

$$Z = \sum_{\alpha \geq 1} w_\alpha U_\alpha\ ,$$

so that $Z \leq K(\kappa) \sum_{\alpha \geq 1} \alpha^{-1/m} U_\alpha$ by (15.179). Using (15.171) for $m' = m/2$ gives

$$(\mathsf{E}_0 Z^{m'})^{1/m'} \leq K(m,\kappa)(\mathsf{E}U^m)^{1/m}\ . \tag{15.181}$$

Here and in the sequel $K(m,\kappa)$ denotes a number depending only on m and κ, that need not to be the same at each occurrence. Now (15.181) implies

$$\mathsf{P}_0(Z \le M) \ge \frac{1}{2} \tag{15.182}$$

where

$$M = K(m,\kappa)(\mathsf{E}U^m)^{1/m} . \tag{15.183}$$

Consider different indices $\alpha_1, \alpha_2, \alpha_3$ and let

$$X = \sum_{\alpha \ne \alpha_1, \alpha_2, \alpha_3} w_\alpha U_\alpha .$$

If we remove the terms $w_{\alpha_1}, w_{\alpha_2}, w_{\alpha_3}$ from the sequence (w_α) and relabel the remaining terms as $(w^*_\alpha)_{\alpha \ge 1}$, we see from (15.179) that

$$\frac{1}{K'(m,\kappa)}\alpha^{-1/m} \le w^*_\alpha \le K'(m,\kappa)\alpha^{-1/m} . \tag{15.184}$$

Since the sequence (U_α) is i.i.d. in distribution we have

$$X \stackrel{\mathcal{D}}{=} \sum_\alpha w^*_\alpha U_\alpha ,$$

where the equality is in distribution. Therefore by (15.184) we have $\mathsf{P}_0(X \le u) \le \mathsf{P}_0(X' \le u)$, where

$$X' = \frac{1}{K'(m,\kappa)} \sum_{\alpha \ge 1} \alpha^{-1/m} U_\alpha .$$

Consequently if

$$M' = \frac{1}{K'(m,\kappa)}(\mathsf{E}U^m)^{1/m}$$

by (15.174) we have

$$\mathsf{P}_0(X \le vM') \le 3\exp(-v^{-m}) \le K(m)v^{10} ,$$

and thus if M is as in (15.183), for $t > 0$ it holds

$$\mathsf{P}_0(X \le tM) \le K(m,\kappa)t^{10} . \tag{15.185}$$

For $\ell \le 3$ let us set $X_\ell = w_{\alpha_\ell} U_{\alpha_\ell}$, so that

$$Z = X + X_1 + X_2 + X_3 .$$

Let us write

$$w'_\alpha = \frac{w_\alpha U_\alpha}{Z} ; \qquad v_\alpha = \mathsf{E}_0 w'_\alpha ,$$

where E_0 denotes expectation in the randomness of the sequence (U_α) only. We observe that

$$w'_{\alpha_\ell} = \frac{X_\ell}{X + X_1 + X_2 + X_3} \, .$$

Conditions (15.182) and (15.185) imply (15.163) and (15.164). Then (15.165) yields

$$\mathsf{E}_0 w'_{\alpha_1} w'_{\alpha_2} w'_{\alpha_3} \le K_0(\kappa, m) v_{\alpha_1} v_{\alpha_2} v_{\alpha_3} \, . \tag{15.186}$$

The important fact here is that the constant $K_0(\kappa, m)$ does not depend on the r.v. U. The idea is now to consider $U_\alpha = \exp h_\alpha$ where $h_\alpha = \sum_{\ell \le s} \eta_{\alpha,\ell}$ where $(\eta_{\alpha,\ell})_{\alpha,\ell \ge 1}$ are independent Bernoulli r.v.s and s a large enough integer to be determined later. The important point is that $\max_\alpha v_\alpha$ becomes small as s becomes large. To see this we observe that by the choice of U

$$v_\alpha = \mathsf{E}_0 w_\alpha \frac{\exp h_\alpha}{Z}$$

where $Z = \sum_\gamma w_\gamma \exp h_\gamma$. Now

$$(\mathsf{E} \exp m h_\alpha)^{1/m} = A^s$$

where $A = (\mathrm{ch} m)^{1/m} > 1$ is independent of s. For large s we have $A^{s/2} = v A^s$ where $v = A^{-s/2}$ is very small. By (15.179) and (15.174) it becomes very rare for large s that $Z \le A^{s/2}$. By (A.19) (used for s rather than N) we have $\mathsf{E} \exp(h_\alpha^2 / s) \le \sqrt{2}$ so that for large s it is also very rare that $h_\alpha \ge (s/4) \log A$. Thus the quantity $w_\alpha \exp h_\alpha / Z$ is very rarely $\ge A^{-s/4}$, and since it is always ≤ 1, its expectation becomes very small for large s. Recalling the constant $K_0(\kappa, m)$ of (15.186), let us consider $0 < \delta < \kappa / K_0(\kappa, m)$, that is also small enough that

$$\delta < \min_{1 \le p < r} |q_p - q_{p+1}| \, . \tag{15.187}$$

Let us then consider $\varepsilon > 0$ (with $\varepsilon < \delta$) that is provided by Proposition 15.6.8 for this value of δ. Let us finally fix s large enough (depending only on m and κ) that $\max_\alpha v_\alpha \le \varepsilon$. Since λ a.s. we have $q_{\alpha,\gamma} \le q_2$ for $\alpha \ne \gamma$, we see that

$$\sum \{ v_\alpha v_\gamma \, ; \, q_{\alpha,\gamma} > q_2 \} = \sum_\alpha v_\alpha^2 \le (\max_\alpha v_\alpha) \sum_{\alpha \ge 1} v_\alpha \le \varepsilon \, . \tag{15.188}$$

Since the matrix $(q_{\alpha,\gamma})$ is positive definite, there exists a sequence (x_α) of points in a Hilbert space such that $q_{\alpha,\gamma} = x_\alpha \cdot x_\gamma$. Consider the probability measure $\pi = \sum_{\alpha \ge 1} v_\alpha \delta_{x_\alpha}$ on H. By (15.188) it satisfies (15.176) for $q = q_2$ and hence (15.177). Now, by (15.187) and since $q_{\alpha,\gamma} \in \{q_r, \dots, q_1\}$, if $q_{\alpha_2,\alpha_3} \ne q_{\alpha_1,\alpha_3}$ we have $|q_{\alpha_1,\alpha_3} - q_{\alpha_2,\alpha_3}| \ge \delta$, and (15.177) implies that

$$\sum \{ v_{\alpha_1} v_{\alpha_2} v_{\alpha_3} \, ; \, (\alpha_1, \alpha_2, \alpha_3) \in W \} \le \delta \le \kappa / K_0(\kappa, m)$$

and thus by (15.186),

$$\mathsf{E}_0 \sum \{w'_{\alpha_1} w'_{\alpha_2} w'_{\alpha_3} \; ; \; (\alpha_1, \alpha_2, \alpha_3) \in W\} \le \kappa \; . \tag{15.189}$$

This is true provided (15.179) holds for each α, which occurs with λ-probability $\ge 1 - \kappa$. Thus, taking expectation E_λ in (15.189), we obtain

$$\mathsf{E}\mathsf{E}_\lambda \sum \{w'_{\alpha_1} w'_{\alpha_2} w'_{\alpha_3} \; ; \; (\alpha_1, \alpha_2, \alpha_3) \in W\} \le 2\kappa \; .$$

Now we observe that by the choice of U, w'_α is the quantity w^s_α of (15.152), and thus

$$\mathsf{E}\mathsf{E}_\lambda \sum \{w^s_{\alpha_1} w^s_{\alpha_2} w^s_{\alpha_3} \; ; \; (\alpha_1, \alpha_2, \alpha_3) \in W\} \le 2\kappa \; . \tag{15.190}$$

Consider now the function f^* on $\mathbb{R}^3$ valued in $\{0,1\}$ such that

$$f^*(x_{1,2}, x_{1,3}, x_{2,3}) = 1 \quad \Leftrightarrow \quad x_{1,2} = q_2 \; , \quad x_{1,3} < q_2 \; , \quad x_{2,3} \ne x_{1,3} \; .$$

Then use of (15.153) and (15.190) for this function show that

$$\mathsf{E}_\lambda \sum \{w_{\alpha_1} w_{\alpha_2} w_{\alpha_3} \; ; \; (\alpha_1, \alpha_2, \alpha_3) \in W\} \le 2\kappa$$

and since κ is arbitrary we have

$$\mathsf{E}_\lambda \sum \{w_{\alpha_1} w_{\alpha_2} w_{\alpha_3} \; ; \; (\alpha_1, \alpha_2, \alpha_3) \in W\} = 0$$

so that W is empty λ a.s. □

Since W is empty λ a.s., λ a.s. we have

$$q_{\alpha_1,\alpha_2} = q_2 \; , \quad q_{\alpha_1,\alpha_3} < q_2 \quad \Rightarrow \quad q_{\alpha_2,\alpha_3} = q_{\alpha_1,\alpha_3} \; . \tag{15.191}$$

Lemma 15.6.10. *When (15.191) occurs, the relation*

$$\alpha \sim \gamma \quad \Leftrightarrow \quad q_{\alpha,\gamma} \ge q_2 \tag{15.192}$$

is an equivalence relation. Moreover, if $\alpha \sim \alpha'$, $\gamma \sim \gamma'$ but $\alpha \not\sim \gamma$, then $q_{\alpha,\gamma} = q_{\alpha',\gamma'}$.

Proof. First we have $\alpha \sim \alpha$ because $q_{\alpha,\alpha} = q_1 \ge q_2$. Next, consider $\alpha_1, \alpha_2, \alpha_3$ with $\alpha_1 \sim \alpha_2$ and $\alpha_2 \sim \alpha_3$. If $q_{\alpha_1,\alpha_2} = q_1$, then by (15.160) we have $\alpha_1 = \alpha_2$ and then $\alpha_1 \sim \alpha_3$. Otherwise we have $\alpha_1 \ne \alpha_2$, so $q_{\alpha_1,\alpha_2} = q_2$. Then we must have $q_{\alpha_1,\alpha_3} \ge q_2$, for otherwise (15.191) shows that $q_{\alpha_2,\alpha_3} = q_{\alpha_1,\alpha_3} < q_2$, and we cannot have $\alpha_2 \sim \alpha_3$. Thus $\sim$ is an equivalence relation.

If $\alpha \sim \alpha'$ and $\alpha \not\sim \gamma$, then $q_{\alpha,\alpha'} \ge q_2$ and $q_{\alpha,\gamma} < q_2$. If $\alpha = \alpha'$, then $q_{\alpha',\gamma} = q_{\alpha,\gamma}$. Otherwise, $q_{\alpha,\alpha'} = q_2$ and (15.191) shows that $q_{\alpha',\gamma} = q_{\alpha,\gamma}$. So always $q_{\alpha',\gamma} = q_{\alpha,\gamma}$. If now $\gamma \sim \gamma'$, since $\alpha' \not\sim \gamma$ the previous argument shows that $q_{\alpha',\gamma} = q_{\alpha',\gamma'}$, and thus $q_{\alpha,\gamma} = q_{\alpha',\gamma'}$. □

Let us enumerate the equivalence classes of the relation $\sim$ as $C_1, \ldots, C_k, \ldots$ (The notation is a little bit abusive because we have not proved yet that there is an infinite number of such equivalence classes.) We set

$$v_k = \sum_{\alpha \in C_k} w_\alpha \tag{15.193}$$

and without loss of generality we may assume that the sequence (v_k) is non-increasing.

Lemma 15.6.11. *The equivalence classes C_k are infinite λ a.s.*

Proof. Let us define $q'_{\alpha,\gamma} = \max(q_3, q_{\alpha,\gamma})$, and let us define the measure $\mu^\sim$ on the set $\mathcal{C}$ as

$$\text{“ the law of the element } (q'_{\alpha^\ell,\alpha^{\ell'}})_{1\le\ell<\ell'} \in \mathcal{C} \text{ ”.} \tag{15.194}$$

That is, for a function f^* on $\mathbb{R}^{n(n-1)/2}$ we have, recalling the notation $\mathbf{x}_n = (x_{\ell,\ell'})_{1\le\ell<\ell'\le n}$,

$$\int f^*(\mathbf{x}_n) \mathrm{d}\mu^\sim(\mathbf{x}) = \mathsf{E}_\lambda \sum_{\alpha^1,\ldots,\alpha^n} w_{\alpha^1} \cdots w_{\alpha^n} f^*((q'_{\alpha^\ell,\alpha^{\ell'}})_{1\le\ell<\ell'\le n}) \, .$$

Let us consider the subset $\mathcal{C}'$ of $\mathcal{C}$ that consists of the sequences $\mathbf{x} = (x_{\ell,\ell'})$ with the following properties

$$\forall \ell < \ell' \ , \ x_{\ell,\ell'} \in \{q_3, q_2, q_1\}$$

$$\text{The relation } \ell \mathcal{R} \ell' \Leftrightarrow x_{\ell,\ell'} \ge q_1 \text{ is an equivalence relation on } \mathbb{N}^*$$

$$\text{The relation } \ell \mathcal{R}' \ell' \Leftrightarrow x_{\ell,\ell'} \ge q_2 \text{ is an equivalence relation on } \mathbb{N}^*$$

$$\text{At least one of the equivalences classes for } \mathcal{R}' \text{ contains only finitely many classes for } \mathcal{R} \, .$$

When one of the classes C_k is finite, and when at least one of the points α^ℓ for $\ell \ge 1$ belongs to C_k the sequence $(q'_{\alpha^\ell,\alpha^{\ell'}})_{1\le\ell<\ell'}$ belongs to $\mathcal{C}'$. This is because

$$\ell \mathcal{R} \ell' \Leftrightarrow \alpha^\ell = \alpha^{\ell'}$$

so that the set $\{\ell \ ; \ \alpha^\ell \in C_k\}$ is an equivalence class for $\mathcal{R}'$, that contains only finitely many classes for $\mathcal{R}$. So, it suffices to prove that $\mu^\sim(\mathcal{C}') = 0$.

It is obvious that $\mu^\sim$ is symmetric (in the sense of Definition 15.3.1). It is obvious that $\mu^\sim$ satisfies the extended Ghirlanda-Guerra identities (as in Definition 15.3.4), as this simply amounts to use (15.108) in the case where the function depends on each $x_{\ell,\ell'}$ only through $\max(x_{\ell,\ell'}, q_3)$. Moreover, the numbers $q'_{\alpha,\gamma} = \max(q_3, q_{\alpha,\gamma})$ satisfy

$$q'_{\alpha,\delta} \geq \min(q'_{\alpha,\gamma}, q'_{\gamma,\delta}) \ ,$$

and, as in the proof of Theorem 15.3.6 this shows that $\mu^\sim$ is ultrametric. The one dimensional marginal μ of $\mu^\sim$ is supported by the set $\{q_3, q_2, q_1\}$, and, by (15.178), it gives positive mass to each of these points.

Now comes the punch line. In the "existence part" of Theorem 15.3.6, we have used Poisson-Dirichlet cascades to construct a symmetric measure μ^*on $\mathcal{C}$, that satisfies the extended Ghirlanda-Guerra identities, and that has μ as one dimensional marginal. The "uniqueness part" of Theorem 15.3.6 proves that $\mu^\sim = \mu^*$, so it suffices to prove that $\mu^*(\mathcal{C}') = 0$. This, however, is obvious by (15.43), since the weights v_α, $\alpha \in A = \mathbb{N}^{*3}$ are a.s. all > 0 by construction, and since on A all the equivalence classes of the relation

$$\alpha \mathcal{R}'' \gamma \Leftrightarrow q_{\alpha,\gamma} \geq q_2 \Leftrightarrow (\alpha,\gamma) \geq 2$$

are infinite, because they are exactly the sets $\{(j_1, j_2, j) \ ; \ j \in \mathbb{N}^*\}$ where j_1 and j_2 are given integers. □

Let us define $\overline{q}_{k,k} = q_2$ and if $k \neq k'$ let us define $\overline{q}_{k,k'} = q_{\alpha,\gamma}$ where $\alpha \in C_k$ and $\gamma \in C_{k'}$ (a quantity that does not depend on the choice of α and γ by the second part of Lemma 15.6.10.)

Lemma 15.6.12. *The matrix* $(\overline{q}_{k,k'})$ *is positive definite.*

Proof. Since the matrix $(q_{\alpha,\gamma})$ is positive definite, there exists vectors $(x_\alpha)_{\alpha \geq 1}$ in a Hilbert space such that $x_\alpha \cdot x_\gamma = q_{\alpha,\gamma}$. For each k consider $(\alpha(k,i))_{i \geq 1}$ distinct points of C_k. Given n consider $y_k = n^{-1} \sum_{i \leq n} x_{\alpha(k,i)}$ so $y_k \cdot y_{k'} = \overline{q}_{k,k'}$ if $k \neq k'$ (since then $x_\alpha \cdot x_\gamma = \overline{q}_{k,k'}$ for $\alpha \in C_k$ and $\gamma \in C_{k'}$), while for $k = k'$

$$y_k \cdot y_k = q_2 + \frac{1}{n}(q_1 - q_2)$$

since $x_\alpha \cdot x_\alpha = q_1$ and $x_\alpha \cdot x_\gamma = q_2$ for $\alpha \neq \gamma$, $\alpha, \gamma \in C_k$. Thus for each n the matrix $(\overline{q}_{k,k'} + (q_1 - q_2)/n \mathbf{1}_{\{k=k'\}})$ is positive definite. □

Proof of Theorem 15.6.1. The proof is by induction over r. The case $r = 2$ is obvious. The map

$$((w_\alpha)_{\alpha \geq 1}, (q_{\alpha,\gamma})) \mapsto ((v_k)_{k \geq 1}, (\overline{q}_{k,k'})) \in \mathcal{S} \times \mathcal{Q}$$

is well defined λ a.s. Let us call λ' the image measure of λ, and observe that $\overline{q}_{k,k'} \in \{q_r, \ldots, q_2\}$. For a function f^* on $\mathbb{R}^{n(n-1)/2}$ consider the function $\widehat{f}$ on $\mathbb{R}^{n(n-1)/2}$ given by

$$\widehat{f}((x_{\ell,\ell'})_{1 \leq \ell < \ell' \leq n}) = f^*((\min(q_2, x_{\ell,\ell'}))_{1 \leq \ell < \ell' \leq n}) \ .$$

Then since for $\alpha \in C_k$ and $\gamma \in C_{k'}$ we have $\overline{q}_{k,k'} = \min(q_2, q_{\alpha,\gamma})$, we see that

$$\sum_{k^1,\dots,k^n} v_{k^1} \cdots v_{k^n} f^*((\overline{q}_{k^\ell,k^{\ell'}})_{1\le\ell<\ell'\le n}) = \sum_{\alpha^1,\dots,\alpha^n} w_{\alpha^1} \cdots w_{\alpha^n} \widehat{f}((q_{\alpha^\ell,\alpha^{\ell'}})_{1\le\ell<\ell'\le n})$$

and by definition of λ',

$$\mathsf{E}_{\lambda'} \sum_{k^1,\dots,k^n} v_{k^1} \cdots v_{k^n} f^*((\overline{q}_{k^\ell,k^{\ell'}})_{1\le\ell<\ell'\le n}) = \mathsf{E}_{\lambda} \sum_{\alpha^1,\dots,\alpha^n} w_{\alpha^1} \cdots w_{\alpha^n} \widehat{f}((q_{\alpha^\ell,\alpha^{\ell'}})_{1\le\ell<\ell' n}) \, .$$

This identity should make it obvious that λ' satisfies the extended Ghirlanda-Guerra identities since this is the case for λ.

Finally, by construction we have $\overline{q}_{k,k} = q_2$ and $\overline{q}_{k,k'} < q_2$ if $k \neq k'$ (as is shown by the definition of the equivalence relation $\sim$).

This concludes the proof of the inductive step from $r-1$ to r and of Theorem 15.6.1. □

Remark 15.6.13. It is not difficult to see that Theorem 15.6.1 remains true without assuming condition (15.160). To see this one simply shows that the relation

$$\alpha \sim \gamma \quad \Leftrightarrow \quad q_{\alpha,\gamma} = q_1$$

is an equivalence relation (which is obvious of we think that there exist vectors x_α in Hilbert space with $q_{\alpha,\gamma} = x_\alpha \cdot x_\gamma$, since then $\alpha \sim \gamma$ if and only if $x_\alpha = x_\gamma$). One then "merges all the elements in an equivalence class" to obtain a new determinator λ' that satisfies (15.160), the extended Ghirlanda-Guerra identities and is ultrametric if and only if λ is ultrametric.

15.7 Problems: Strong Ultrametricity and Chaos

It is quite natural to consider the following stronger version of the ultrametricity conjecture (Conjecture 15.3.3).

Conjecture 15.7.1. (Level 3) If $\varepsilon > 0$, we have

$$\nu(\mathbf{1}_{\{R_{1,3}+\varepsilon\le\min(R_{1,2},R_{2,3})\}}) \le K \exp\left(-\frac{N}{K}\right) \tag{15.195}$$

where K is independent of N.

Given three numbers $u_{1,2}, u_{2,3}, u_{1,3}$, a natural approach to this conjecture is to try to find a bound for

$$p_N(u_{1,2}, u_{2,3}, u_{1,3}) := \frac{1}{N}\mathsf{E}\log\sum\exp\Bigg(-H_N(\boldsymbol{\sigma}^1) - H_N(\boldsymbol{\sigma}^2) - H_N(\boldsymbol{\sigma}^3) + \sum_{i\le N}\sum_{\ell\le 3} h_i\sigma_i^\ell\Bigg)\,, \tag{15.196}$$

where the summation is over the set of triplets $(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2, \boldsymbol{\sigma}^3)$ of configurations for which $R_{\ell,\ell'} = u_{\ell,\ell'}$ for $\ell, \ell' = 1,2,3$, $\ell \ne \ell'$.
The most natural route is to copy the scheme of Lemma 14.6.1, and, since it requires no extra work, we will write this scheme in the greatest generality we can. Consider an integer n, and jointly Gaussian Hamiltonians $H_1, \ldots, H_n$ on Σ_N. We assume that for $\ell, \ell' \le n$ there is a convex function $\xi_{\ell,\ell'}$ such that

$$\frac{1}{N}\mathsf{E}H_\ell(\boldsymbol{\sigma}^1)H_{\ell'}(\boldsymbol{\sigma}^2) = \xi_{\ell,\ell'}(R_{1,2})\,. \tag{15.197}$$

Consider numbers $(u_{\ell,\ell'})$ for $1 \le \ell, \ell' \le n$ and assume

$$u_{\ell,\ell} = 1\,; \qquad u_{\ell,\ell'} = u_{\ell',\ell}\,.$$

Consider the set B of n-tuples of configurations $(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n)$ such that

$$\forall \ell, \ell' \le n\,,\quad R_{\ell,\ell'} = u_{\ell,\ell'}\,.$$

We try to find an upper bound for

$$p_{N,B} = \frac{1}{N}\log\sum_B\exp\Bigg(-\sum_{\ell\le n} H_\ell(\boldsymbol{\sigma}^\ell) - H_0\Bigg)\,, \tag{15.198}$$

where $H_0 = H_0(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n)$ is a random Hamiltonian probabilistically independent of $H_1, \ldots, H_n$. Consider a set A, and a Gaussian Hamiltonian $H(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n, \alpha)$ where $\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n \in \Sigma_N$ and $\alpha \in A$. Let us assume that there exists numbers $q_{\alpha,\gamma}^{\ell,\ell'}$ ($\ell, \ell' \le n$, $\alpha, \gamma \in A$) such that, for any $\alpha, \gamma \in A$ and any $\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n, \boldsymbol{\tau}^1, \ldots, \boldsymbol{\tau}^n \in \Sigma_N$ we have

$$\frac{1}{N}\mathsf{E}H(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n, \alpha)H(\boldsymbol{\tau}^1, \ldots, \boldsymbol{\tau}^n, \gamma) = \sum_{\ell,\ell'\le n} R^{\ell,\ell'}\xi'_{\ell,\ell'}(q_{\alpha,\gamma}^{\ell,\ell'}) \tag{15.199}$$

where $R^{\ell,\ell'} = N^{-1}\sum_{i\le N}\sigma_i^\ell\tau_i^{\ell'}$. Assume moreover that

$$\forall \ell, \ell' \le n\,,\ \forall \alpha \in A\,,\ q_{\alpha,\alpha}^{\ell,\ell'} = u_{\ell,\ell'}\,. \tag{15.200}$$

For $0 \le s \le 1$ consider the interpolating Hamiltonian

$$H_s(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2, \ldots, \boldsymbol{\sigma}^n, \alpha) = \sqrt{s} \sum_{\ell \le n} H_\ell(\boldsymbol{\sigma}^\ell) + \sqrt{1-s} H(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n, \alpha) + H_0(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n) .$$

Consider random weights w_α on A; we assume of course that the randomnesses of H_ℓ, H, H_0, w_α are all independent of each other. Given a function $f(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n, \alpha)$ we define

$$\langle f(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n, \alpha) \rangle_s = \frac{1}{Z_s} \sum_{\alpha, B} w_\alpha f(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n, \alpha) \exp(-H_s(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n, \alpha)) ,$$

where Z_s is the normalizing factor, and where the sum is over $\alpha \in A$ and all $(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n) \in B$.

Lemma 15.7.2. *The function*

$$\varphi(s) = \frac{1}{N} \mathsf{E} \log \sum_{\alpha, B} w_\alpha \exp(-H_s(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n, \alpha))$$

satisfies

$$\varphi'(s) \le -\frac{1}{2} \sum_{\ell, \ell' \le n} \theta_{\ell,\ell'}(u_{\ell,\ell'}) + \frac{1}{2} \sum_{\ell, \ell' \le n} \mathsf{E} \langle \theta_{\ell,\ell'}(q_{\alpha,\gamma}^{\ell,\ell'}) \rangle_s \tag{15.201}$$

where $\theta_{\ell,\ell'}(x) = x\xi'_{\ell,\ell'}(x) - \xi_{\ell,\ell'}(x)$.

Proof. This is the same proof as in Lemma 14.6.1, although the greater generality makes the proof easier to write. Let us define

$$U(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n, \boldsymbol{\tau}^1, \ldots, \boldsymbol{\tau}^n, \alpha, \gamma) = \frac{1}{2N} \left(\mathsf{E} \Big(\sum_{\ell \le n} H_\ell(\boldsymbol{\sigma}^\ell) \Big) \Big(\sum_{\ell \le n} H_\ell(\boldsymbol{\tau}^\ell) \Big) - \mathsf{E} H(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n, \alpha) H(\boldsymbol{\tau}^1, \ldots, \boldsymbol{\tau}^n, \gamma) \right).$$

Use of (15.197) and (15.199) show that

$$\begin{aligned} U(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n, \boldsymbol{\tau}^1, \ldots, \boldsymbol{\tau}^n, \alpha, \gamma) &= \frac{1}{2} \left(\sum_{\ell, \ell'} \xi_{\ell,\ell'}(R^{\ell,\ell'}) - R^{\ell,\ell'} \xi'_{\ell,\ell'}(q_{\alpha,\gamma}^{\ell,\ell'}) \right) \\ &= -\frac{1}{2} \sum_{\ell,\ell'} \theta_{\ell,\ell'}(q_{\alpha,\gamma}^{\ell,\ell'}) + \frac{1}{2} \sum_{\ell,\ell'} S_{\alpha,\gamma}^{\ell,\ell'} , \end{aligned} \tag{15.202}$$

where

$$S_{\alpha,\gamma}^{\ell,\ell'} = \xi_{\ell,\ell'}(R^{\ell,\ell'}) - R^{\ell,\ell'} \xi'_{\ell,\ell'}(q_{\alpha,\gamma}^{\ell,\ell'}) + \theta_{\ell,\ell'}(q_{\alpha,\gamma}^{\ell,\ell'}) .$$

The quantity $S_{\alpha,\gamma}^{\ell,\ell'}$ has two remarkable properties:

$$S_{\alpha,\gamma}^{\ell,\ell'} \geq 0 \tag{15.203}$$

$$(\boldsymbol{\tau}^1, \ldots, \boldsymbol{\tau}^n) = (\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n) \in B \;,\; \gamma = \alpha \quad \Rightarrow \quad S_{\alpha,\gamma}^{\ell,\ell'} = 0 \,. \tag{15.204}$$

This is because then $R^{\ell,\ell'} = u_{\ell,\ell'} = q_{\alpha,\gamma}^{\ell,\ell'}$ by (15.200) and the definition of B. Now, differentiation in s and integration by parts show that

$$\begin{aligned}\varphi'(s) &= \mathsf{E}\langle U(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n, \boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n, \alpha, \alpha)\rangle_s \\ &\quad - \mathsf{E}\langle U(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n, \boldsymbol{\tau}^1, \ldots, \boldsymbol{\tau}^n, \alpha, \gamma)\rangle_s\end{aligned}$$

and (15.202), (15.203) and (15.204) imply the result. □

Let us next consider an integer κ, and, for $0 \leq p \leq \kappa$ jointly Gaussian n-tuples $(y_p^\ell)_{\ell \leq n}$, such that, for certain numbers $(\rho_p^{\ell,\ell'})$ $(1 \leq p \leq \kappa + 1; \ell, \ell' \leq n)$ we have

$$\mathsf{E} y_p^\ell y_p^{\ell'} = \xi'_{\ell,\ell'}(\rho_{p+1}^{\ell,\ell'}) - \xi'_{\ell,\ell'}(\rho_p^{\ell,\ell'}) \,. \tag{15.205}$$

Let us further assume that

$$\rho_0^{\ell,\ell'} = 0 \,; \qquad \rho_{\kappa+1}^{\ell,\ell'} = u_{\ell,\ell'} \,. \tag{15.206}$$

For $\alpha \in \mathbb{N}^{*\kappa}$ we define the n-tuples $(y_{i,p,\alpha}^\ell)_{\ell \leq n}$ as in (14.134) and we use the Hamiltonian

$$H(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n, \alpha) = \sum_{\ell \leq n} \sum_{i \leq N} \sigma_i^\ell \left(\sum_{0 \leq p \leq \kappa} y_{i,p,\alpha}^\ell \right) ,$$

so that (15.199) holds for $q_{\alpha,\gamma}^{\ell,\ell'} = \rho_{(\alpha,\gamma)}^{\ell,\ell'}$. Assume that the weights (w_α) form a Poisson-Dirichlet cascade associated to the sequence $0 < n_1 < \ldots < n_{\kappa-1} < n_\kappa = 1$ (and we will now denote these weights by (v_α)). By (14.38) we have $\langle \mathsf{E} \mathbf{1}_{\{(\alpha,\gamma)=p\}} \rangle_s = n_p - n_{p-1}$ and (15.201) yields, using (15.200)

$$\begin{aligned}\varphi'(s) &\leq -\frac{1}{2} \sum_{\ell,\ell' \leq n} \theta(u_{\ell,\ell'}) + \frac{1}{2} \sum_{\ell,\ell' \leq n} \sum_{1 \leq p \leq \kappa} \theta_{\ell,\ell'}(\rho_p^{\ell,\ell'})(n_{p+1} - n_p) \\ &= -\frac{1}{2} \sum_{\ell,\ell' \leq n} \sum_{1 \leq p \leq \kappa} n_p (\theta_{\ell,\ell'}(\rho_{p+1}^{\ell,\ell'}) - \theta_{\ell,\ell'}(\rho_p^{\ell,\ell'})) \,.\end{aligned}$$

To bound $\varphi(0)$, let us assume that

$$H_0(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n) = \sum_{i \leq N} \sum_{\ell \leq n} h_i^\ell \sigma_i^\ell \tag{15.207}$$

where $(h_i^\ell)_{\ell \leq n,\, i \leq N}$ are i.i.d. copies of a random n-tuple $(h^1, \ldots, h^n)$. Then, given numbers $\lambda_{\ell,\ell'}$, and since $N R_{\ell,\ell'} = \sum_{i \leq n} \sigma_i^\ell \sigma_i^{\ell'}$, we have

$$\varphi(0) = \frac{1}{N} \mathsf{E} \log \sum_{\alpha} v_\alpha \sum_{B} \exp \sum_{i \le N} \left(\sum_{\ell \le n} \sigma_i^\ell \left(h_i^\ell + \sum_{0 \le p \le \kappa} y_{i,p,\alpha}^\ell \right) \right.$$
$$\left. + \sum_{1 \le \ell, \ell' \le n} \lambda_{\ell,\ell'} \sigma_i^\ell \sigma_i^{\ell'} \right) - \sum_{1 \le \ell < \ell' \le n} \lambda_{\ell,\ell'} u_{\ell,\ell'} \; .$$

We bound this expression by replacing the sum over $(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n) \in B$ by the sum over all configurations; we transform it using Theorem 14.2.1 and we then decouples over the sites as in (14.141) to obtain the following, where we recall the notation $p_{N,B}$ of (15.198).

Theorem 15.7.3. *Under the previous conditions we have*

$$p_{N,B} \le Y_0 - \sum_{1 \le \ell < \ell' \le n} \lambda_{\ell,\ell'} u_{\ell,\ell'}$$
$$- \frac{1}{2} \sum_{\ell,\ell' \le n} \sum_{0 \le p \le \kappa} n_p (\theta_{\ell,\ell'}(\rho_{p+1}^{\ell,\ell'}) - \theta_{\ell,\ell'}(\rho_p^{\ell,\ell'})) \; , \qquad (15.208)$$

where the quantity Y_0 is defined as follows. Starting with

$$Y_{\kappa+1} = \log \sum_{\varepsilon_1, \ldots \varepsilon_n = \pm 1} \exp \left(\sum_{\ell \le n} \varepsilon_\ell \left(h^\ell + \sum_{0 \le p \le \kappa} y_p^\ell \right) + \sum_{1 \le \ell < \ell' \le n} \lambda_{\ell,\ell'} \varepsilon_\ell \varepsilon_{\ell'} \right) \; ,$$

we define recursively

$$Y_p = \frac{1}{n_p} \log \mathsf{E}_p \exp n_p Y_{p+1} \; , \qquad (15.209)$$

where E_p denotes expectation in the r.v.s y_n^ℓ for $n \ge p$ and we set $Y_0 = \mathsf{E} Y_1$.

Research Problem 15.7.4. (Level 3) Is the bound (15.208) sharp? That is, is it true that, as $N \to \infty$, the limit of the left-hand side is the infimum of the right-hand side of (15.208) over the choices of parameters?

If we assume (15.207), and that each function $\xi_{\ell,\ell}$ satisfies the conditions (14.101), then replacing in (15.198) the summation $(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n) \in B$ by the summation over all values of $(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n)$, and using Theorem 14.5.1 we see that with obvious notation

$$p_{N,B} \le \sum_{\ell \le n} \mathcal{P}(\xi_{\ell,\ell}, h^\ell) \; .$$

Research Problem 15.7.5. (Level 3) Given $\varepsilon > 0$, does there exist a choice of parameters such that the right-hand side of (15.208) is $\le \sum_{\ell\le n}\mathcal{P}(\xi_{\ell,\ell},h^\ell)+\varepsilon$?

When $n=2$, $H_1=H_2$ and $h^1=h^2$, the answer is yes, a fact that played a fundamental part in the proof of Theorem 14.5.1. Following the same idea, the reader is strongly encouraged to solve the following very beautiful exercise.

Exercise 15.7.6. Assume that $n=3$, $H_1=H_2=H_3$, $h^1=h^2=h^3$, $u_{1,2}=u_{1,3}<u_{2,3}$. Prove that in this case, Problem 15.7.4 has a positive solution.

This unfortunately does not help towards Conjecture 15.7.1 since the relation $u_{1,2}=u_{1,3}<u_{2,3}$ is precisely the one allowed by ultrametricity.

If we have found the correct bound for $p_{N,B}$, we certainly expect that Problem 15.7.5 has a positive solution. On the other hand, it is probably too naive to expect that this problem will yield to direct attack (i.e. a simple method to choose the parameters). The reason is that a choice of parameters giving a good bound in (15.208) should be related to the structure of the Gibbs measure on B with Hamiltonian

$$\sum_{\ell\le n}\left(-H_\ell(\boldsymbol{\sigma}^\ell)+\sum_{i\le N}h_i^\ell\sigma_i^\ell\right) . \tag{15.210}$$

When B is an "infinitesimally small" subset of Σ_N^n one really see no reasons why the structure of the Gibbs measure on B should be simply related to the structure of the Gibbs measure on the whole of Σ_N^n. Understanding what happens is possibly the main remaining open question for this class of models.

It must also be said that even when the ultrametricity condition is violated, say $u_{1,2}<u_{2,3}<u_{1,3}$ it is not absolutely certain (recalling (15.196)) that

$$\limsup_{N\to\infty} p_N(u_{1,2},u_{2,3},u_{3,1}) < 3\mathcal{P}(\xi,h) . \tag{15.211}$$

An example (albeit a rather special one and for a different model) is provided in [81]. Still, one might hope that (15.210) holds "most of the time". Here is a much simpler question that would deserve to be completely clarified.

Research Problem 15.7.7. (Level 1^+) Under the conditions of Theorem 14.5.1, prove the following is "typically true". If μ is a Parisi measure (as defined in the previous section) and if u does not belong to the support of μ (i.e. $\mu(]u-\varepsilon,u+\varepsilon[)=0$ when ε is small enough) then for ε small enough we have

$$\limsup_{N\to\infty}\frac{1}{N}\mathsf{E}\log\sum_{|R_{1,2}-u|\le\varepsilon}\exp\left(-\sum_{\ell=1,2}\left(-H(\boldsymbol{\sigma}^\ell)+\sum_{i\le N}h_i\sigma_i^\ell\right)\right) < 2\mathcal{P}(\xi,h) . \tag{15.212}$$

By "typically true" we mean that the set of parameters $\boldsymbol{\beta}$ for which this fails for the Hamiltonian (14.406) should be very small in some sense. On the other hand, we cannot expect that (15.212) will be true for all values of $\boldsymbol{\beta}$. This can be seen by considering the case of the Hamiltonian (14.406) when all the coefficients β_r are zero, except for $r = p$ (a given large enough integer), and when there is no external field. In that case, it should become clear to the reader after studying the next chapter that, as β_p grows from zero, the Parisi measure μ is first concentrated at zero, then, when β_p crosses a certain value β^* (and is not too large) the Parisi measure is concentrated at 0 and a point $q(\beta_p)$. Then (although the details have not been checked) it seems very likely that for $\beta_p = \beta^*$, the point $u = \lim_{\beta_p \to \beta^*_+} q(\beta_p)$ fails (15.212) (while the Parisi measure is concentrated at zero).

We now turn to the statement of another major open question, the so-called Chaos Problem, that we state in a very general form. Consider two Hamiltonians $H_1 = H_{1,N}$ and $H_2 = H_{2,N}$. Consider on Σ_N^2 the Hamiltonian

$$-H_1(\boldsymbol{\sigma}^1) - H_2(\boldsymbol{\sigma}^2) \tag{15.213}$$

and $\langle \cdot \rangle$ an average for the corresponding Gibbs measure.

Definition 15.7.8. *There is* chaos between the Hamiltonians H_1 *and* H_2 *when there exists a number* u *(depending only on* H_1 *and* H_2*) such that*

$$\forall \varepsilon > 0 \, , \qquad \lim_{N \to \infty} \mathsf{E} \langle \mathbf{1}_{\{|R_{1,2} - u| \geq \varepsilon\}} \rangle = 0 \, . \tag{15.214}$$

In words "the overlap $R_{1,2}$ *takes a unique value".*

To discuss this idea, let us assume that each of the Hamiltonians H_ℓ is of the type (14.406). Each of these Hamiltonians then creates an intricate structure (that we have been studying in the present chapter). However the structures created by H_1 and H_2 are not well related. The "pure states" created by these two Hamiltonians are not "aligned" but rather in completely different directions.

A natural approach to the chaos problem would be to find bounds for

$$p_{N,u} = \frac{1}{N} \mathsf{E} \log \sum_{R_{1,2} = u} \exp\left(\sum_{\ell = 1,2} -H_\ell(\boldsymbol{\sigma}^\ell) + \sum_{i \leq N} h_i^\ell \sigma_i^\ell \right) ,$$

and in particular to show that there exists a value u_0 such that for $u \neq u_0$ we have, for some $\varepsilon > 0$ that

$$p_{N,u} < \mathcal{P}(\xi_{1,1}, h^1) + \mathcal{P}(\xi_{2,2}, h^2) - \varepsilon \, .$$

Of course, on has little chance to make this approach work unless on can first solve the corresponding case of Problem 15.7.5.

One might get the feeling that there should be chaos whenever the Parisi measures associated to H_1 and H_2 are "fundamentally different". Let us denote by μ_1 and μ_2 these Parisi measures.

Conjecture 15.7.9. (Level 3) There is chaos between H_1 and H_2 **unless** for a certain $t > 0$ the restrictions of the Parisi measures μ_1 and μ_2 to a certain interval $[0, t]$ coincide, while $\mu_1([0, t]) = \mu_2([0, t]) > 0$.

The reason why we state this conjecture in terms of Parisi measures is simply that so little is known about the way the Parisi measure depends on the parameters of the Hamiltonian. One of course would like to approach this conjecture through Theorem 15.7.3, but, as explained, this does not seem to be easy.

Research Problem 15.7.10. (Level 3) For the Hamiltonians of the type (14.406), is there a one to one correspondence between the Parisi measure and the parameters of the Hamiltonian? And can the condition of Conjecture 15.7.9 be satisfied unless the parameters of the two Hamiltonians are the same?

The following does not have any special consequence, but seems rather a curiosity.

Research Problem 15.7.11. Consider two Hamiltonians as in (15.213), and the function

$$\psi(\lambda) = \lim_{N \to \infty} \frac{1}{N} \mathsf{E} \log \sum_{\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2} \exp(-H_1(\boldsymbol{\sigma}^1) - H_2(\boldsymbol{\sigma}^2) - \lambda R_{1,2}) .$$

If ψ is differentiable at $\lambda = 0$ then (it is an exercise to show that) there is chaos between the Hamiltonians H_1 and H_2. But can it ever happen that ψ fails to be differentiable at some point $\lambda_0 \neq 0$?

Since the author has already stuck his neck out so much by stating conjectures that are supported by little else than wishful thinking, there is really no reason to be shy any more. Consider two sequences $\boldsymbol{\beta}_\ell = (\beta_{\ell,p})_{p \geq 1}$, $\ell = 1, 2$; consider a pair (h^1, h^2) of r.v.s and i.i.d. copies $(h_i^1, h_i^2)_{i \leq N}$ of this pair. Consider the Hamiltonian

$$-H_\ell(\boldsymbol{\sigma}) = \sum_{i \leq N} h_i^\ell \sigma_i + \sum_{p \geq 1} \frac{\beta_{\ell,p}}{N^{(p-1)/2}} \sum_{i_1, \ldots, i_p} g_{i_1, \ldots, i_p} \sigma_{i_1} \cdots \sigma_{i_p} . \tag{15.215}$$

This is exactly the Hamiltonian (14.406), except that we "do not restrict the summation to even values of p".

Conjecture 15.7.12. (Level 3) (The Generalized Chaos conjecture for Ising spins) Assume it either of the following:

$$\boldsymbol{\beta}_1 \neq \boldsymbol{\beta}_2 \tag{15.216}$$

or

$$\text{it is not true that } h^1 = h^2 \text{ a.e.} \tag{15.217}$$

Then there is chaos between the Hamiltonians H_1 and H_2.

The word 'chaos' reflects the idea that if we start with the Hamiltonian H_1 and make a change (however small) in one value of β_p or in the external field, the structure of the Gibbs measure changes totally, in a chaotic rather than in a smooth way.

Let us point out that in the case where $\boldsymbol{\beta}_1 = \boldsymbol{\beta}_2$ and (15.217) holds, while h^1 and h^2 have the same law then H_1 and H_2 have the same Parisi measure. The chaos is not created by a change of $\boldsymbol{\beta}$ but by a "change in the disorder of the external field".

Research Problem 15.7.13. (Level 1) When there are only terms for even p in (15.215) for H_1 and H_2, prove that Conjecture 15.7.12 holds in the case where $\boldsymbol{\beta}_1 = \boldsymbol{\beta}_2$ and h^1 and h^2 have the same distribution.

The reason that this is rated only level 1 is that in this case it is really easy to solve Problem 15.7.5. More hints about a possible scheme of proof will be given shortly.

In a direction similar to the case of Problem 15.7.13, S. Chatterjee considers in a recent paper [31] a more general situation of "Chaos in disorder". Let us consider two jointly Gaussian Hamiltonians H_1 and H_2, that have the same distribution, but that are correlated in such a manner that, for some $t \neq 1$ and a certain function ξ we have

$$\mathsf{E}H_1(\boldsymbol{\sigma}^1)H_2(\boldsymbol{\sigma}^2) = t\xi(R_{1,2}) \,, \tag{15.218}$$

while

$$\xi(R_{1,2}) = \mathsf{E}H_1(\boldsymbol{\sigma}^1)H_1(\boldsymbol{\sigma}^2) = \mathsf{E}H_2(\boldsymbol{\sigma}^1)H_2(\boldsymbol{\sigma}^2) \,.$$

We also assume that the function ξ satisfies the hypothesis of Theorem 14.5.1 (for otherwise our methods are currently powerless).

Research Problem 15.7.14. (Level 1+) Prove that there is chaos between the Hamiltonians $-H_1(\boldsymbol{\sigma}) + h\sum_{i\le N}\sigma_i$ and $-H_2(\boldsymbol{\sigma}) + h\sum_{i\le N}\sigma_i$.

One can also consider a situation with random external field, but the random external fields must have the same distribution to maintain the level 1 rating. This rating is intended modulo the ideas toward a solution that we explain now. First we explain why in the situation of Problem 15.7.14 it is not difficult to solve Problem 15.7.5. In that case, $n = 2$, and we choose $\lambda_{1,2} = 0$.

Given numbers $0 = q_0 \le q_1 \le \ldots \le q_\kappa \le q_{\kappa+1} = 1$ and $0 = m_0 < \ldots < m_\kappa \le 1$, and given an integer $\tau < \kappa$ let us then chose $\rho_p^{\ell,\ell'} = q_p$ if either $p < \tau$ or $\ell = \ell'$, and $\rho_p^{1,2} = \rho_p^{2,1} = q_\tau$ if $p \ge \tau$. Then if $p < \tau$ we have

$$\sum_{\ell,\ell'\le 2} (\theta_{\ell,\ell'}(\rho_{p+1}^{\ell,\ell'}) - \theta_{\ell,\ell'}(\rho_p^{\ell,\ell'})) = 2(1+t)(\theta(q_{p+1}) - \theta(q_p)) \,,$$

while if $p \ge \tau$ we have

$$\sum_{\ell,\ell'\le 2}(\theta_{\ell,\ell'}(\rho_{p+1}^{\ell,\ell'})-\theta_{\ell,\ell'}(\rho_p^{\ell,\ell'}))=2(\theta(q_{p+1})-\theta(q_p))\ .$$

Let us also choose

$$n_p=\frac{m_p}{t+1}$$

if $p<\tau$ while $n_p=m_p$ otherwise. Then the last term in (15.208) is

$$-\sum_{0\le p\le\kappa}m_p(\theta(q_{p+1})-\theta(q_p))\ .$$

The variables $y_{p,\alpha}^{\ell}$ of (15.208) satisfy $\mathsf{E}(y_{p,\alpha}^1)^2=\mathsf{E}(y_{p,\alpha}^2)^2=\xi'(q_{p+1})-\xi'(q_p)$ and

$$p<\tau\Rightarrow \mathsf{E}y_{p,\alpha}^1y_{p,\alpha}^2=t(\xi'(q_{p+1})-\xi'(q_p))\ ;\ p\ge\tau\Rightarrow \mathsf{E}y_{p,\alpha}^1y_{p,\alpha}^2=0\ .$$

This suggests the following result to control the first term of the right-hand side of (15.208), when we take $\lambda_{1,2}=0$.

Exercise 15.7.15. Consider two jointly Gaussian r.v.s y_1 and y_2 such that $\mathsf{E}y_1^2=\mathsf{E}y_2^2$ and $\mathsf{E}y_1y_2=t\mathsf{E}y_1^2$. Consider a function F such that all the derivatives of F are uniformly bounded. Then for any values of x_1, x_2 and $m>0$ we have

$$\frac{1+t}{m}\log\mathsf{E}\exp\frac{m}{1+t}(F(x_1{+}y_1){+}F(x_2{+}y_2))\le\sum_{j=1,2}\frac{1}{m}\log\mathsf{E}\exp mF(x_j{+}y_j)\ . \tag{15.219}$$

Hint: Consider the functions

$$\varphi(u)=\frac{1+t}{m}\log\mathsf{E}\exp\frac{m}{1+t}(F(x_1+\sqrt{u}y_1)+F(x_2+\sqrt{u}y_2))$$

$$\varphi_j(u)=\frac{1}{m}\log\mathsf{E}\exp mF(x_j+\sqrt{u}y_j)\ .$$

Prove that $\varphi'(0)\le\varphi_1'(0){+}\varphi_2'(0)$, and hence that $\varphi(u)\le\varphi_1(u){+}\varphi_2(u){+}Cu^2$, where C depends only on F. Then observe that both sides of (15.219) can be obtained "by iterating the case where $\mathsf{E}y_1^2$ is very small".

Using independence for $p\ge\tau$ and Exercise 15.7.15 when $p<\tau$ we show recursively that the quantities Y_p of (15.209) satisfy $Y_p\le 2X_p$ where

$$X_{\kappa+1}=\log 2\mathrm{ch}\Bigl(h^1+\sum_{0\le p\le\kappa}y_p^1\Bigr)\ ,$$

and where $X_p=m_p^{-1}\log\mathsf{E}_p\exp m_pX_{p+1}$. In this manner we show that for every $u\ge 0$ one has

$$p_{N,u}\le 2\mathcal{P}(\xi,h)\ . \tag{15.220}$$

Moreover in (15.219) there can be equality only when $x_1 = x_2$. Let us denote by c the smallest point of the support of the Parisi measure μ of H_1 or H_2. We expect that techniques similar to the ones presented at the end of Section 14.12 will allow to show that when $u > c$, there is inequality in (15.220) (even in the limit $N \to \infty$). (Of course the expression "we expect" means that the author has not checked in detail whether this can be actually done). We also expect that these techniques can show that this is also true for $u < -c$. After these steps, it remains only to understand the case where $|u| \le c$. In that case one has to prove there is a unique value of u for which one cannot get inequality in (15.220) by making a small variation of λ. We feel that we also have given arguments allowing to prove this.

15.8 The Aizenman-Sims-Starr Scheme

In this section we go back to the fundamental Problem 15.7.4. What could be the structure of the proof? We have to find a mechanism that would somehow produce the right parameters. As far as producing random weights, a Gibbs measure is certainly an intricate tool. In this line of thought, we must discuss a rather canonical scheme brought forward in [11].

Consider a Gaussian Hamiltonian H on Σ_N such that

$$\frac{1}{N}\mathsf{E}H(\boldsymbol{\sigma}^1)H(\boldsymbol{\sigma}^2) = \xi(R_{1,2}) \tag{15.221}$$

for a convex function ξ. (In fact using the method of Theorem 14.4.4 it would suffice for some of the results of this section to assume that ξ is convex on $\mathbb{R}^+$.) Consider a set A and random weights $(w_\alpha)_{\alpha\in A}$. We consider two Gaussian Hamiltonians $H^a(\boldsymbol{\sigma},\alpha)$, $H^b(\alpha)$ which are independent of these random weights and we assume that for certain numbers $q_{\alpha,\gamma}$, for any values of $\alpha, \gamma, \boldsymbol{\sigma}^1$ and $\boldsymbol{\sigma}^2$ we have (recalling the notation $\theta(x) = x\xi'(x) - \xi(x)$)

$$\frac{1}{N}\mathsf{E}H^a(\boldsymbol{\sigma}^1,\alpha)H^a(\boldsymbol{\sigma}^2,\gamma) = R_{1,2}\xi'(q_{\alpha,\gamma}) \tag{15.222}$$

$$\frac{1}{N}\mathsf{E}H^b(\alpha)H^b(\gamma) = \theta(q_{\alpha,\gamma}) \tag{15.223}$$

$$q_{\alpha,\alpha} = 1\ . \tag{15.224}$$

We consider yet another random Hamiltonian $H_0(\boldsymbol{\sigma})$. We assume that H, H^a, H^b, H_0 are independent of each other.

Proposition 15.8.1. *We have*

$$\mathsf{E}\log\sum_{\boldsymbol{\sigma}}\exp(-H(\boldsymbol{\sigma}) - H_0(\boldsymbol{\sigma})) \tag{15.225}$$

$$\le \mathsf{E}\log\sum_{\alpha,\boldsymbol{\sigma}} w_\alpha \exp(-H^a(\boldsymbol{\sigma},\alpha) - H_0(\boldsymbol{\sigma})) - \mathsf{E}\log\sum_{\alpha} w_\alpha\exp(-H^b(\alpha))\ .$$

Proof. Consider the Hamiltonian

$$H_s(\boldsymbol{\sigma},\alpha)=\sqrt{s}(H(\boldsymbol{\sigma})+H^b(\alpha))-\sqrt{1-s}H^a(\boldsymbol{\sigma},\alpha)+H_0(\boldsymbol{\sigma})\ ,$$

and let us denote by $\langle\cdot\rangle_s$ an average for the corresponding Gibbs measure. Let us denote

$$\varphi(s)=\mathsf{E}\log\sum_{\alpha,\boldsymbol{\sigma}}w_\alpha\exp(-H_s(\boldsymbol{\sigma},\alpha))\ .$$

Then differentiation and integration by parts show that

$$\begin{aligned}\varphi'(s)&=\frac{1}{2}(\xi(1)+\theta(1)-\xi'(1))-\frac{1}{2}\langle\xi(R_{1,2})-R_{1,2}\xi'(q_{\alpha,\gamma})+\theta(q_{\alpha,\gamma})\rangle_s\\&\le 0\end{aligned}$$

because $\theta(1)=\xi'(1)-\xi(1)$ and $\xi(x)-x\xi'(q)+\theta(q)\ge 0$. Writing that $\varphi(1)\le\varphi(0)$ means exactly (15.225). □

The natural way to achieve (15.222) is to write

$$H^a(\boldsymbol{\sigma},\alpha)=\sum_{i\le N}\sigma_i z_{i,\alpha} \tag{15.226}$$

where $(z_{i,\alpha})_{i\le N}$ are i.i.d. copies of a Gaussian family (z_α) that satisfies

$$\mathsf{E}z_\alpha z_\gamma=\xi'(q_{\alpha,\gamma})\ . \tag{15.227}$$

It turns out that it is a good choice to assume that $H^b(\alpha)=\sqrt{N}\eta_\alpha$, where the jointly Gaussian r.v.s η_α satisfy

$$\mathsf{E}\eta_\alpha\eta_\gamma=\theta(q_{\alpha,\gamma})\ . \tag{15.228}$$

So, in that case we get the bound

$$\begin{aligned}p_N:=&\frac{1}{N}\mathsf{E}\log\sum_{\boldsymbol{\sigma}}\exp\left(-H(\boldsymbol{\sigma})+\sum_{i\le N}h_i\sigma_i\right)\\\le&\frac{1}{N}\mathsf{E}\log\sum_\alpha w_\alpha\sum_{\sigma_1,\dots,\sigma_N=\pm1}\exp\sum_{i\le N}\sigma_i(z_{i,\alpha}+h_i)\\-&\frac{1}{N}\mathsf{E}\log\sum_\alpha w_\alpha\exp\sqrt{N}\eta_\alpha\ .\end{aligned} \tag{15.229}$$

Let us also observe that it would not really matter if we had only approximate equality in (15.227), e.g. $|\mathsf{E}z_\alpha z_\gamma-\xi'(q_{\alpha,\gamma})|\le\varepsilon$ and similarly in (15.228) as this would introduce only an error term $2\varepsilon N$ in (15.225).

From now on we assume that the Hamiltonian H is given by

$$-H(\boldsymbol{\sigma})=H_N(\boldsymbol{\sigma})=\sum_{p\ge1}\frac{\beta_p}{N^{p-\frac{1}{2}}}\sum_{i_1,\dots,i_{2p}\le N}g_{i_1\dots i_{2p}}\sigma_{i_1}\cdots\sigma_{i_{2p}}\ . \tag{15.230}$$

Copying the argument of Theorem 1.3.9, we see that $p^* = \lim_N p_N$ exists, so that

$$p^* \le \liminf B_N \ , \tag{15.231}$$

where B_N denotes the infimum of the right-hand side of (15.229) over all choices of parameters (that is, the set A, the random weights (w_α) and the numbers $(q_{\alpha,\gamma})$).

The most interesting property of the bound (15.231) is that in this case it can be reversed. (Is is absolutely inessential in (15.230) to restrict the sum to even values of p. The only reason for doing this is that we know a bit more about the existence certain limits.)

Proposition 15.8.2. *When the Hamiltonian H is given by (15.230) then*

$$\limsup_N B_N \le p^* \ . \tag{15.232}$$

The proof is based on comparing the $(M+N)$-spin system with the M-spin system when M is large and $N \ll M$. We perform the following decomposition, where we write $\boldsymbol{\sigma} = (\sigma_i)_{i \le M+N}$, $\boldsymbol{\rho} = (\sigma_i)_{i \le M}$

$$-H_{N+M}(\boldsymbol{\sigma}) = -H^*(\boldsymbol{\rho}) + \sum_{M < i \le M+N} \sigma_i g_i(\boldsymbol{\rho}) + g(\boldsymbol{\sigma}) \ . \tag{15.233}$$

Here $H^*(\boldsymbol{\rho})$ is the sum of the terms in the right-hand side of (15.230) that do not contain σ_i for $i > M$, i.e.

$$-H^*(\boldsymbol{\rho}) = \sum_{p \ge 1} \frac{\beta_p}{(M+N)^{p-\frac{1}{2}}} \sum_{i_1,\ldots,i_{2p} \le M} g_{i_1 \ldots i_{2p}} \sigma_{i_1} \cdots \sigma_{i_{2p}} \ . \tag{15.234}$$

It is this Hamiltonian (after adding the external field) that will produce the weights w_α. The term $\sigma_i g_i(\boldsymbol{\rho})$ in (15.233) collects all the terms in the right-hand side of (15.230) that contain σ_i but not other σ_j for $M < j \le M+N$, and the term $g(\boldsymbol{\sigma})$ gathers the other terms, those that contain at least two factors σ_i for $i > M$.

A first idea is that, when $N \ll M$, the term $g(\boldsymbol{\sigma})$ is "small compared to $\sqrt{N}$". Indeed, among the $(M+N)^{2p}$ choices of $i_1, \ldots, i_{2p} \le M+N$ there are at most $(2p)^2 N^2 (M+N)^{2p-2}$ choices where two of the indices are $> M$. Thus

$$\mathsf{E} g^2(\boldsymbol{\sigma}) \le \sum_{p \ge 1} \beta_p^2 (2p)^2 \frac{N^2 (M+N)^{2p-2}}{(M+N)^{2p-1}} \le \frac{N^2}{M+N} \sum_{p \ge 1} 4 \beta_p^2 p^2 \ . \tag{15.235}$$

The r.v. $g_i(\boldsymbol{\rho})$ are independent of each other and of H^*. If we keep in mind that there are $2p$ ways one of the indexes $i_1, \ldots, i_{2p}$ can be equal to i, we see that

$$\mathsf{E} g_i(\boldsymbol{\rho}^1) g_i(\boldsymbol{\rho}^2) = \sum_{p\geq 1} 2p\beta_p^2 \left(\frac{M}{M+N}\right)^{2p-1} R(\boldsymbol{\rho}^1,\boldsymbol{\rho}^2)^{2p-1} \tag{15.236}$$

where $R(\boldsymbol{\rho}^1,\boldsymbol{\rho}^2) = M^{-1}\sum_{i\leq M}\sigma_i^2\sigma_i^2$. If we define

$$\xi(x) = \sum_{p\geq 1}\beta_p^2 x^{2p} \ , \tag{15.237}$$

then (15.233) almost means that $\mathsf{E} g_i(\boldsymbol{\rho}^1) g_i(\boldsymbol{\rho}^2) = \xi'(R(\boldsymbol{\rho}^1,\boldsymbol{\rho}^2))$.

A basic observation is that we have the (obvious) identity

$$\begin{aligned} &\log\sum_{\boldsymbol{\sigma}}\exp\Bigl(-H_{N+M}(\boldsymbol{\sigma}) + \sum_{i\leq M+N} h_i\sigma_i\Bigr) \\ &-\log\sum_{\boldsymbol{\rho}}\exp\Bigl(-H^*(\boldsymbol{\rho}) + \sum_{i\leq M} h_i\sigma_i\Bigr) \\ &= \log\Bigl\langle \sum_{\sigma_{M+1},\ldots,\sigma_{M+N}=\pm 1}\exp\Bigl(\sum_{M<i\leq M+N}\sigma_i(g_i(\boldsymbol{\rho})+h_i) + g(\boldsymbol{\sigma})\Bigr)\Bigr\rangle \ , \end{aligned} \tag{15.238}$$

where $\langle\cdot\rangle$ denotes an average for the Gibbs measure with Hamiltonian $-H(\boldsymbol{\rho}) = -H^*(\boldsymbol{\rho}) + \sum_{i\leq M} h_i\sigma_i$.

Let us write $A = \{-1,1\}^M$, and, for $\boldsymbol{\rho}^1,\boldsymbol{\rho}^2\in A$, let us define

$$q_{\boldsymbol{\rho}^1,\boldsymbol{\rho}^2} = R(\boldsymbol{\rho}^1,\boldsymbol{\rho}^2) \ ,$$

so that (15.224) holds. Denoting by w_α the weight of $\alpha\in A$ for the Gibbs measure with Hamiltonian $-H(\boldsymbol{\rho})$, we write the right-hand side of (15.238) as

$$\log\sum_A w_\alpha \sum_{\sigma_{M+1},\ldots,\sigma_{M+N}=\pm 1}\exp\Bigl(\sum_{M<i\leq M+N}\sigma_i(z_{i,\alpha}+h_i) + g(\boldsymbol{\sigma})\Bigr) \ , \tag{15.239}$$

where $z_{i,\alpha} = g_i(\boldsymbol{\rho})$ when $\alpha = \boldsymbol{\rho}$.

Let us now compare $-H^*(\boldsymbol{\rho})$ and

$$-H_M(\boldsymbol{\rho}) = \sum_{p\geq 1}\frac{\beta_p}{M^{p-\frac{1}{2}}}\sum_{i_1,\ldots,i_{2p}\leq M} g_{i_1\ldots i_{2p}}\sigma_{i_1}\cdots\sigma_{i_{2p}} \ .$$

If $g'_{i_1\ldots i_{2p}}$ denote now independent standard Gaussian r.v.s then in distribution we have

$$-H_M(\boldsymbol{\rho}) = -H^*(\boldsymbol{\rho}) - H^{\sim}(\boldsymbol{\rho}) \tag{15.240}$$

where

$$H^{\sim}(\boldsymbol{\rho}) = \sum_{p\geq 1} \beta_p \left(\frac{1}{M^{2p-1}} - \frac{1}{(N+M)^{2p-1}}\right)^{1/2} \sum_{i_1,\ldots,i_{2p}\leq M} g'_{i_1\ldots i_{2p}} \sigma_{i_1}\cdots\sigma_{i_{2p}} \ . \tag{15.241}$$

Since when N/M is small we have

$$\frac{1}{M^{2p-1}} - \frac{1}{(N+M)^{2p-1}} \sim \frac{(2p-1)N}{M^{2p}} \ ,$$

it is almost true that

$$\begin{aligned} \mathsf{E}H^{\sim}(\boldsymbol{\rho}^1)H^{\sim}(\boldsymbol{\rho}^2) &= N\sum_{p\geq 1}(2p-1)\beta_{2p}^2 R(\boldsymbol{\rho}^1,\boldsymbol{\rho}^2)^{2p} \\ &= N\theta(R(\boldsymbol{\rho}^1,\boldsymbol{\rho}^2)) \ . \end{aligned}$$

Moreover, by (15.238) we have the identity

$$\begin{aligned} &\mathsf{E}\log\sum_{\boldsymbol{\rho}}\exp\Big(-H_M(\boldsymbol{\rho})+\sum_{i\leq M}h_i\sigma_i\Big) - \mathsf{E}\log\sum_{\boldsymbol{\rho}}\exp\Big(-H^*(\boldsymbol{\rho})+\sum_{i\leq M}h_i\sigma_i\Big) \\ &= \mathsf{E}\log\langle\exp(-H^{\sim}(\boldsymbol{\rho})\rangle \\ &= \mathsf{E}\log\sum_{\alpha\in A} w_\alpha \exp\sqrt{N}\eta_\alpha \ , \end{aligned} \tag{15.242}$$

where if $\alpha = \boldsymbol{\rho}$ we write $\eta_\alpha = N^{-1/2}H^{\sim}(\boldsymbol{\rho})$. Combining (15.238), (15.239) and (15.242), we have obtained the identity

$$\begin{aligned} &(N+M)p_{N+M} - Mp_M \\ &= \mathsf{E}\log\sum_{\alpha} w_\alpha \sum_{\sigma_{M+1},\ldots,\sigma_{M+N}=\pm 1}\exp\Bigg(\sum_{M<i\leq M+N}\sigma_{M+i}(z_{i,\alpha}+h_i)+g(\boldsymbol{\sigma})\Bigg) \\ &- \mathsf{E}\log\sum_{\alpha} w_\alpha \exp\sqrt{N}\eta_\alpha \ . \end{aligned} \tag{15.243}$$

We now take M large, and N a small proportion of M, $N = \varepsilon M$, so the left-hand side is about Np^*. To finish the proof we shall show that when ε is small the right-hand side is nearly NB_N. Besides the small error created by the term $g(\boldsymbol{\sigma})$, it is not exactly true that $\mathsf{E}z_{i,\alpha}z_{i,\gamma} = \xi'(q_{\alpha,\gamma})$ and $\mathsf{E}\eta_\alpha\eta_\gamma = \theta(q_{\alpha,\gamma})$. Let us examine what happens to the term $\mathsf{E}\log\sum_\alpha w_\alpha \exp\sqrt{N}\eta_\alpha$ when replacing (keeping the notation $\alpha = \boldsymbol{\rho}$)

$$\eta_\alpha = N^{-1/2}H^{\sim}(\boldsymbol{\rho})$$

by

$$\eta'_\alpha = \sum_{p\geq 1}\beta_p\frac{\sqrt{2p-1}}{M^p}\sum_{i_1,\ldots,i_{2p}\leq M} g_{i_1\ldots i_{2p}}\sigma_{i_1}\cdots\sigma_{i_{2p}} \ ,$$

(so that it is exactly true that $\mathsf{E}\eta'_\alpha\eta'_\gamma = \theta(q_{\alpha\gamma})$). For this we interpolate as usual. If (η^*_α) denotes an independent copy of (η'_α), we consider the function

$$\varphi(s) = \frac{1}{N}\mathsf{E}\log\sum_\alpha w_\alpha \exp\sqrt{N}(\sqrt{s}\eta_\alpha + \sqrt{1-s}\eta^*_\alpha) \ ,$$

and we show that

$$|\varphi'(s)| \le \frac{1}{2}\sup_{\alpha,\gamma}|\mathsf{E}\eta_\alpha\eta_\gamma - \mathsf{E}\eta^*_\alpha\eta^*_\gamma| = \frac{1}{2}\sup_{\alpha,\gamma}|\mathsf{E}\eta_\alpha\eta_\gamma - \theta(q_{\alpha,\gamma})| \ ,$$

and using (15.241) it is straightforward to see that this quantity goes to zero with ε. In this manner we prove that

$$\mathsf{E}\log\sum_\alpha w_\alpha \exp\sqrt{N}\eta_\alpha \simeq \mathsf{E}\log\sum_\alpha w_\alpha \exp\sqrt{N}\eta'_\alpha \ ,$$

where $\simeq$ means that the error is a proportion of N that vanishes with ε. In a similar manner one proves that

$$\begin{aligned}
&\mathsf{E}\log\sum_\alpha w_\alpha \sum_{\sigma_{M+1},\ldots,\sigma_{M+N}=\pm1}\exp\Bigg(\sum_{M<i\le M+N}\sigma_{M+i}(z_{i,\alpha}+h_i) + g(\boldsymbol{\sigma})\Bigg) \\
&\simeq \mathsf{E}\log\sum_\alpha w_\alpha \sum_{\sigma_{M+1},\ldots,\sigma_{M+N}=\pm1}\exp\sum_{M<i\le M+N}\sigma_{M+i}(z'_{i,\alpha}+h_i) \\
&= \mathsf{E}\log\sum_\alpha w_\alpha \sum_{\sigma_1,\ldots,\sigma_N=\pm1}\exp\sum_{1\le i\le N}\sigma_i(z'_{i,\alpha}+h_i) \ ,
\end{aligned}$$

where the jointly Gaussian r.v.s $z'_{i,\alpha}$ satisfy $\mathsf{E}z'_{i,\alpha}z'_{i,\gamma} = \xi'(q_{\alpha,\gamma})$ and are independent as i varies. Thus given any $\delta > 0$ we can find $\varepsilon > 0$ small enough that for large N the right-hand side of (15.243) is $\ge N(B_N - \delta)$, and this completes the proof of Proposition 15.8.2.

While the A.S.S. scheme is completely natural and canonical, it seems very difficult to understand anything at all about the right-hand side of (15.229). For example, since we take expectation, its infimum over all parameters is the same whether we take random weights w_α or deterministic weights. It does not give a clue either as to why the infimum should be attained (at least under the condition (14.101)) for the very special structure brought forward by Theorem 14.5.1.

Let us now try to use the A.S.S. scheme in the setting of Conjecture 15.7.12. This will reveal why we study this scheme despite the shortcomings we just explained. We consider a pair H_1, H_2 of jointly Gaussian Hamiltonians such that for $\ell, \ell' = 1, 2$ we have

$$\frac{1}{N}\mathsf{E}H_\ell(\boldsymbol{\sigma}^1)H_{\ell'}(\boldsymbol{\sigma}^2) = \xi_{\ell,\ell'}(R_{1,2}) \ ,$$

where the function $\xi_{\ell,\ell'}$ is convex.

Proposition 15.8.3. *Suppose we have a set A, and pairs of Gaussian r.v.s $(\zeta_\alpha^1, \zeta_\alpha^2)_{\alpha\in A}$, $(\eta_\alpha^1, \eta_\alpha^2)_{\alpha\in A}$ with the following properties*

$$\mathsf{E}\zeta_\alpha^\ell \zeta_\gamma^{\ell'} = \xi'_{\ell,\ell'}(q_{\alpha,\gamma}^{\ell,\ell'}) \; ; \qquad \mathsf{E}\eta_\alpha^\ell \eta_\gamma^{\ell'} = \theta_{\ell,\ell'}(q_{\alpha,\gamma}^{\ell,\ell'}) \; , \tag{15.244}$$

where the numbers $q_{\alpha,\gamma}^{\ell,\ell'}$ satisfy

$$q_{\alpha,\alpha}^{1,1} = q_{\alpha,\alpha}^{2,2} = 1 \; ; \qquad |q_{\alpha,\alpha}^{1,2} - u| \le \varepsilon \; ; \qquad |q_{\alpha,\alpha}^{2,1} - u| \le \varepsilon \; . \tag{15.245}$$

Consider independent copies $(\zeta_{i,\alpha}^1, \zeta_{i,\alpha}^2)_\alpha$ of the families $(\zeta_\alpha^1, \zeta_\alpha^2)$. Consider random weights w_α on A, that are independent of these r.v.s. Consider i.i.d. copies $(h_i^1, h_i^2)_{i\ge 1}$ of a pair (h^1, h^2) of r.v.s, and assume that these are independent of all the other forms of randomness. Then, for each ε

$$\begin{aligned} p_N(u,\varepsilon) &:= \frac{1}{N}\mathsf{E}\log \sum_{|R_{1,2}-u|\le\varepsilon} \exp\Bigg(\sum_{\ell=1,2}\Big(-H_\ell(\boldsymbol{\sigma}^\ell) + \sum_{i\le N} h_i^\ell \sigma_i^\ell\Big)\Bigg) \\ &\le C\varepsilon + \frac{1}{N}\mathsf{E}\log\sum_\alpha w_\alpha \sum_{|R_{1,2}-u|\le\varepsilon} \exp\Bigg(\sum_{i\le N}\sum_{\ell=1,2} \sigma_i^\ell(\zeta_{i,\alpha}^\ell + h_i^\ell)\Bigg) \\ &\quad - \frac{1}{N}\mathsf{E}\log\sum_\alpha w_\alpha \exp(\sqrt{N}(\eta_\alpha^1 + \eta_\alpha^2)) \; , \end{aligned} \tag{15.246}$$

where C depends only on the functions $\xi_{\ell,\ell'}$.

The ideas necessary to adapt the proof of (15.229) have been explained so many times that it seems appropriate to leave the proof to the reader. The term $C\varepsilon$ occurs from the fact that

$$|\xi_{1,2}(R_{1,2}) - R_{1,2}\xi'_{1,2}(q_{\alpha,\alpha}^{1,2}) - \theta_{1,2}(q_{\alpha,\alpha}^{1,2})| \le C\varepsilon$$

when $|R_{1,2} - u| \le \varepsilon$ and $|q_{\alpha,\alpha}^{1,2} - u| \le \varepsilon$.

One would like of course to use the bound (15.246) to attack Conjecture 15.7.12. Unfortunately, the problem is (again) that one has no clue about how to choose the parameters in this bound. However we demonstrate in Theorem 15.8.4 below that the bound (15.246) **is** "a sharp bound". This implies that useful choices of parameters *do exist*, even though we do not know how to find them. (In contrast, we were not really sure that this was the case for the bound (15.208).)

Let us turn to the investigation of the converse of (15.246). We assume that both H_1 and H_2 are of the type (15.230), for possibly different values of the parameters. Recalling the definition of $p_N(u,\varepsilon)$ in (15.246) let

$$p^*(u) = \lim_{\varepsilon\to 0}\liminf_{N\to\infty} p_N(u,\varepsilon) = \inf_{\varepsilon>0}\liminf_{N\to\infty} p_N(u,\varepsilon) \; . \tag{15.247}$$

The existence of the limit as $N \to \infty$ can be shown using the methods of [52], but this is really only a side story.

Theorem 15.8.4. *(D. Panchenko [64]) Given $\varepsilon' > 0$, we can find $\varepsilon > 0$, N arbitrarily large, a set A, pairs of r.v.s $(\zeta^1_\alpha, \zeta^2_\alpha)_{\alpha\in A}$, $(\eta^1_\alpha, \eta^2_\alpha)_{\alpha\in A}$ that satisfy (15.244) and (15.245), and weights w_α such that*

$$\frac{1}{N}\mathsf{E}\log\sum_\alpha w_\alpha \sum_{|R_{1,2}-u|\le\varepsilon} \exp\left(\sum_{i\le N}\sum_{\ell=1,2}\sigma^\ell_i(\zeta^\ell_{i,\alpha}+h^\ell_i)\right)$$
$$-\frac{1}{N}\mathsf{E}\log\sum_\alpha w_\alpha \exp(\sqrt{N}(\eta^1_\alpha+\eta^2_\alpha)) \le p^*(u)+\varepsilon' \,. \tag{15.248}$$

It follows in particular that given $\varepsilon > 0$ there exists a choice of the parameters for which the left-hand side is not larger than $\mathcal{P}(\xi_{1,1}, h^1)+\mathcal{P}(\xi_{2,2}, h^2)+\varepsilon$.

Research Problem 15.8.5. Find an explicit construction of parameters such that the left-hand side of (15.245) is not larger than $\mathcal{P}(\xi_{1,1}, h^1) + \mathcal{P}(\xi_{2,2}, h^2) + \varepsilon$.

The proof of Theorem 15.8.4 will make clear what is the difficulty: the weights are created by the restriction of the Gibbs' measure to a small set, and as already pointed out there seems to be no reason why the structure of this restriction should be simply related to the structure of the Gibbs measure itself.

To lighten notation let us now write

$$p^*_\ell = \mathcal{P}(\xi_{\ell,\ell}, h^\ell) \,.$$

To study the chaos problem we are particularly interested in the values of u for which

$$p^*(u) = p^*_1 + p^*_2 \,. \tag{15.249}$$

Indeed, if we can show that this inequality can occur for a unique value of u, using concentration of measure as in Proposition 13.4.3 this shows that there is chaos between the Hamiltonians $-H_\ell + \sum_{i\le N}\sigma_i h^\ell_i$ for $\ell = 1, 2$.

In the case where u is a candidate to satisfy (15.249) we can do better than (15.248).

Theorem 15.8.6. *Assume that*

$$\forall\varepsilon > 0\,, \qquad \limsup_{N\to\infty} p_N(u,\varepsilon) = p^*_1 + p^*_2 \,.$$

Then, given any $\varepsilon' > 0$, we can replace (15.248) by

$$\frac{1}{N}\mathsf{E}\log\sum_\alpha w_\alpha \sum_{\sigma^\ell_i=\pm1} \exp\left(\sum_{i\le N}\sum_{\ell=1,2}\sigma^\ell_i(\zeta^\ell_{i,\alpha}+h^\ell_i)\right)$$
$$-\frac{1}{N}\mathsf{E}\log\sum_\alpha w_\alpha \exp(\sqrt{N}(\eta^1_\alpha+\eta^2_\alpha)) \le p^*_1+p^*_2+\varepsilon' \,. \tag{15.250}$$

The difference with (15.248) is that the summation is now over all values of $\sigma_i^\ell = \pm 1$. The whole point of (15.250) is that $|\mathsf{E}\zeta_\alpha^1\zeta_\alpha^2 - \xi'_{1,2}(u)| \le \varepsilon$ and $|\mathsf{E}\eta_\alpha^1\eta_\alpha^2 - \theta_{1,2}(u)| \le \varepsilon$. We did succeed in creating a quantity that we can control while exhibiting this particular correlation structure. (The control of the correlation structure of these variables is precisely the main obstacle in using bounds such as that of Theorem 15.7.3.)

Proof of Theorem 15.8.4. Consider $\varepsilon_1 > 0$, and let $\varepsilon > 0$ be small enough that

$$\liminf_{N\to\infty} p_N(u,\varepsilon) < p^*(u) + \varepsilon_1^2 \ . \tag{15.251}$$

By definition of $p^*(u)$, we have $\liminf_{N\to\infty} p_N(u,\varepsilon) \ge p^*(u)$, so that for large M we have $p_M(u,\varepsilon) \ge p^*(u) - \varepsilon_1^2$.

We can find N_1 arbitrarily large with $p_{N_1}(u,\varepsilon) < p^*(u) + \varepsilon_1^2$. Let us decompose $N_1 = N + M$ where $N \simeq \varepsilon_1 M$. Then, since $M \le N_1 \le 2N/\varepsilon_1$ we have

$$\begin{aligned} N_1 p_{N_1}(u,\varepsilon) - M p_M(u,\varepsilon) &\le N_1(p^*(u) + \varepsilon_1^2) - M(p^*(u) - \varepsilon_1^2) \\ &= Np^*(u) + (N_1 + M)\varepsilon_1^2 \\ &\le Np^*(u) + 4N\varepsilon_1 \ . \end{aligned} \tag{15.252}$$

Let us define the Hamiltonians H_ℓ^* for $\ell = 1,2$ as in (15.234) and let

$$A = \{(\boldsymbol{\rho}^1, \boldsymbol{\rho}^2) \in \Sigma_M^2 \ ; \ |R(\boldsymbol{\rho}^1,\boldsymbol{\rho}^2) - u| < \varepsilon\} \ .$$

The weights w_α are produced by the Gibbs measure on A with Hamiltonian

$$- H(\boldsymbol{\rho}^1,\boldsymbol{\rho}^2) = -H_1^*(\boldsymbol{\rho}^1) - H_2^*(\boldsymbol{\rho}^2) + \sum_{i\le M}(h_i^1\sigma_i^1 + h_i^2\sigma_i^2) \ . \tag{15.253}$$

We write (15.252) as

$$\mathrm{I} + \mathrm{II} \le Np^*(u) + 4N\varepsilon_1 \ ,$$

where

$$\mathrm{I} = (M+N)p_{N+M}(u,\varepsilon) - \mathsf{E}\log \sum_{(\boldsymbol{\rho}^1,\boldsymbol{\rho}^2)\in A} \exp(-H(\boldsymbol{\rho}^1,\boldsymbol{\rho}^2)) \tag{15.254}$$

$$\mathrm{II} = -Mp_M(u,\varepsilon) + \mathsf{E}\log \sum_{(\boldsymbol{\rho}^1,\boldsymbol{\rho}^2)\in A} \exp(-H(\boldsymbol{\rho}^1,\boldsymbol{\rho}^2)) \ . \tag{15.255}$$

Let us write

$$B = \left\{(\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2) \in \Sigma_{M+N}^2 \ ; \ (\boldsymbol{\rho}^1,\boldsymbol{\rho}^2) \in A \ ; \ \left|\frac{1}{N}\sum_{M<i\le M+N} \sigma_i^1\sigma_i^2 - u\right| \le \varepsilon\right\} ,$$

so that for $(\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2) \in B$ we have

$$\left|\frac{1}{M+N}\sum_{i\le M+N}\sigma_i^1\sigma_i^2-u\right|\le\varepsilon$$

and thus

$$(N+M)p_{N+M}(u,\varepsilon)\ge \mathsf{E}\log\sum_{(\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2)\in B}\exp\Big(-H_{1,N+M}(\boldsymbol{\sigma}^1)-H_{2,N+M}(\boldsymbol{\sigma}^2)$$
$$+\sum_{i\le N+M}(h_i^1\sigma_i^1+h_i^2\sigma_i^2)\Big)\ .$$

Using a decomposition as in (15.233) for the Hamiltonians $H_{\ell,N+M}$, we get

$$\mathrm{I}\ge\mathsf{E}\log\left\langle\sum\exp\left(\sum_{\ell=1,2}\left(\sum_{M<i\le M+N}\sigma_i^\ell(g_i^\ell(\boldsymbol{\rho}^\ell)+h_i^\ell)+g^\ell(\boldsymbol{\sigma}^\ell)\right)\right)\right\rangle\ ,\quad(15.256)$$

where the first summation is over the set $\left|N^{-1}\sum_{M<i\le M+N}\sigma_i^1\sigma_i^2-u\right|\le\varepsilon$, and where the bracket refers to the Gibbs measure with Hamiltonian (15.253).

For $\alpha=(\boldsymbol{\rho}^1,\boldsymbol{\rho}^2)\in A$ and $\gamma=(\boldsymbol{\tau}^1,\boldsymbol{\tau}^2)\in A$ we define

$$q_{\alpha,\gamma}^{\ell,\ell'}=R(\boldsymbol{\rho}^\ell,\boldsymbol{\tau}^\ell)\ ,$$

so that $|q_{\alpha,\alpha}^{1,2}-u|\le\varepsilon$ and $q_{\alpha,\alpha}^{1,1}=q_{\alpha,\alpha}^{2,2}=1$. From this point on the proof follows the argument of (15.232), so we do not give all the details. We have $\mathsf{E}g_i^\ell(\boldsymbol{\rho}^\ell)g_i^{\ell'}(\boldsymbol{\rho}^{\ell'})\simeq\xi'_{\ell,\ell'}(R(\boldsymbol{\rho}^\ell,\boldsymbol{\rho}^{\ell'}))$, so the right-hand side of (15.256) is nearly

$$\mathsf{E}\log\sum w_\alpha\sum_{\left|N^{-1}\sum_{i\le N}\sigma_i^1\sigma_i^2-u\right|\le\varepsilon}\exp\left(\sum_{\ell=1,2}\sum_{i\le N}\sigma_i^\ell(\zeta_{i,\alpha}^\ell+h_i^\ell)\right)\ ,$$

where the jointly Gaussian r.v.s $\zeta_{i,\alpha}^\ell$ satisfy $\mathsf{E}\zeta_{i,\alpha}^\ell\zeta_{i,\gamma}^{\ell'}=\xi'_{\ell,\ell'}(q_{\alpha,\gamma}^{\ell,\ell'})$. The term II is bounded just as in the case of (15.232). □

Proof of Theorem 15.8.6. Consider $\varepsilon_1>0$, so that for M large enough we have $p_{1,M}+p_{2,M}\le p_1^*+p_2^*+\varepsilon_1^2$. By hypothesis we can find M arbitrarily large with $p_M(u,\varepsilon)\ge p_1^*+p_2^*-\varepsilon_1^2$. Let $N\simeq\varepsilon_1 M$, so that when M is large enough we have $p_{1,N+M}+p_{2,N+M}\le p_1^*+p_2^*+\varepsilon_1^2$, and therefore

$$\begin{aligned}(N+M)(p_{1,N+M}+p_{2,N+M})-Mp_M(u,\varepsilon)&\le(N+M)(p_1^*+p_2^*+\varepsilon_1^2)\\&\quad-M(p_1^*+p_2^*-\varepsilon_1^2)\\&\le N(p_1^*+p_2^*)+4\varepsilon_1 N\ .\end{aligned}$$

Instead of (15.254) we define

$$\mathrm{I}=(N+M)(p_{1,N+M}+p_{2,N+M})-\mathsf{E}\log\sum_{(\boldsymbol{\rho}_1,\boldsymbol{\rho}^2)\in A}\exp(-H(\boldsymbol{\rho}^1,\boldsymbol{\rho}^2))\ ,$$

and then we have

$$\mathrm{I} \geq \mathsf{E} \log \left\langle \sum \exp \left(\sum_{\ell=1,2} \left(\sum_{M \leq i \leq M+N} \sigma_i^\ell (g_i^\ell(\boldsymbol{\rho}^\ell) + h_i^\ell) + g^\ell(\boldsymbol{\sigma}^\ell) \right) \right) \right\rangle ,$$

where the first sum is over all choices of σ_i^ℓ, $M \leq i \leq M+N$. We then proceed as in the proof of Theorem 15.8.4 to produce the set A and the weights w_α. □

15.9 Probability Measures on Hilbert Space

In this section we prove Theorem 15.4.5 and Proposition 15.6.8. We write q rather than q^*.

We fix $\delta > 0$ and consider numbers $a, \gamma, \tau > 0$ depending on δ only that will be specified later. For the time being we think of them as small parameters. Let us assume that

$$\pi^{\otimes 2}(\{(x,y) \; ; \; (x \cdot y) \geq q + \gamma\}) \leq \gamma\tau^2 \,. \tag{15.257}$$

We recall that B denotes the unit ball of the Hilbert space.

Definition 15.9.1. *A set $S \subset B$ will be called* adequate *if $\pi(S) \geq a$ and*

$$\int_{S\times S} |x \cdot y - q| \mathrm{d}\pi(x) \mathrm{d}\pi(y) \leq 8\gamma\pi(S)^2 \,. \tag{15.258}$$

The overall strategy is to prove in a first stage that there exists a large collection of disjoint adequate sets. In a second stage we will construct the sets (A_α) from these. First of all, we must understand how one can obtain condition (15.258). We recall the notation $x^+ = \max(x, 0)$.

Lemma 15.9.2. *If $\pi(S) \geq \tau$ we have*

$$\int_{S\times S} (x \cdot y - q)^+ \mathrm{d}\pi(x) \mathrm{d}\pi(y) \leq 2\gamma\pi(S)^2 \,. \tag{15.259}$$

Proof. Since $x, y \in B$ and $q \geq 0$ we have $(x \cdot y - q)^+ \leq 1$. Using (15.257), we have

$$\begin{aligned}\int_{S\times S} (x \cdot y - q)^+ \mathrm{d}\pi(x) \mathrm{d}\pi(y) &\leq \gamma\pi(S)^2 + \pi^{\otimes 2}(\{(x,y) \; ; \; x \cdot y \geq q + \gamma\}) \\ &\leq \gamma\pi(S)^2 + \gamma\tau^2 \leq 2\gamma\pi(S)^2 \,,\end{aligned}$$

and this concludes the proof. □

Definition 15.9.3. *For a set S, we define the barycentre of π over S by*

$$b_S = \frac{1}{\pi(S)} \int_S x \mathrm{d}\pi(x) . \tag{15.260}$$

Lemma 15.9.4. *If $\pi(S) \geq \tau$ we have*

$$\|b_S\|^2 \leq q + 2\gamma . \tag{15.261}$$

Proof. We note that

$$\pi(S)^2 \|b_S\|^2 = \left\| \int_S x \mathrm{d}\pi(x) \right\|^2 = \int_{S\times S} x \cdot y \mathrm{d}\pi(x)\mathrm{d}\pi(y) . \tag{15.262}$$

We have, using (15.262) and (15.259) in the second line

$$\begin{aligned} \pi(S)^2 \|b_S\|^2 &\leq \pi(S)^2 q + \int_{S\times S} (x \cdot y - q)^+ \mathrm{d}\pi(x)\mathrm{d}\pi(y) \\ &\leq (q + 2\gamma)\pi(S)^2 , \end{aligned}$$

and this finishes the proof. □

Lemma 15.9.5. *If $\pi(S) \geq \tau$ we have*

$$\int_{S\times S} |x \cdot y - q| \mathrm{d}\pi(x)\mathrm{d}\pi(y) \leq \pi(S)^2 (q + 2\gamma - \|b_S\|^2) . \tag{15.263}$$

Proof. For any number t we have $|t| = 2t^+ - t$ so that

$$|x \cdot y - q| \leq 2(x \cdot y - q)^+ + q - x \cdot y .$$

Integrating over x and y for μ, and using (15.259) and (15.262) we get

$$\int |x \cdot y - q| \mathrm{d}\pi(x)\mathrm{d}\pi(y) \leq 2\gamma\pi(S)^2 + q\pi(S)^2 - \pi(S)^2 \|b_S\|^2 . \qquad \square$$

We now assume

$$\tau \leq a \leq 1/4 . \tag{15.264}$$

Lemma 15.9.6. *Consider $S, S' \subset B$, and assume that*

$$\pi(S) \geq a ; \qquad \pi(S') \geq \tau , \qquad b_S \cdot b_{S'} \geq q - 2\gamma . \tag{15.265}$$

Then S is adequate.

Proof. We shall prove that

$$\|b_S\|^2 \geq q - 6\gamma \; . \tag{15.266}$$

Combining with (15.263) this implies (15.258) and completes the proof. To establish (15.266), since the left-hand side is ≥ 0 we may assume that $q \geq 2\gamma$. Using (15.261) for S' then yields

$$(q-2\gamma)^2 \leq (b_S \cdot b_{S'})^2 \leq \|b_S\|^2 \|b_{S'}\|^2 \leq \|b_S\|^2 (q+2\gamma)$$

and therefore

$$\|b_S\|^2 \geq \frac{(q-2\gamma)^2}{q+2\gamma} \geq q - 6\gamma \; . \qquad \square$$

We now state the main tool to construct adequate sets.

Proposition 15.9.7. *There exists a number L with the following property. Assume that*

$$a^{1+L/\gamma^2} = \tau \; . \tag{15.267}$$

Consider a subset D of B and assume that

$$\pi^{\otimes 2}((B \times D) \cap \{(x,y) \; ; \; x \cdot y \geq q - \gamma\}) \geq 2a \; . \tag{15.268}$$

Then D contains an adequate set.

This will use the following, which is in a sense the central point of the argument.

Lemma 15.9.8. *There exists a number L such that one can cover the set*

$$W = \{b_S \; ; \; \pi(S) \geq a\} \tag{15.269}$$

by at most a^{-L/γ^2} balls of radius γ.

Proof of Proposition 15.9.7. Given x in B, let

$$C(x) = \{y \in D \; ; \; x \cdot y \geq q - \gamma\} \; , \tag{15.270}$$

so that, using (15.268) in the last line

$$\begin{aligned}\int \pi(C(x)) \mathrm{d}\pi(x) &= \int_B \pi(\{y \in D \; ; \; x \cdot y \geq q - \gamma\}) \mathrm{d}\pi(x) \\ &= \pi^{\otimes 2}((B \times D) \cap \{(x,y) \; ; \; x \cdot y \geq q - \gamma\}) \geq 2a \; .\end{aligned}$$

Let $T = \{x \; ; \; \pi(C(x)) \geq a\}$ so that

$$2a \le \int \pi(C(x))\mathrm{d}\pi(x) \le \int_T \pi(C(x))\mathrm{d}\pi(x) + \int_{T^c} \pi(C(x))\mathrm{d}\pi(x)$$
$$\le \pi(T) + a$$

and therefore

$$\pi(T) \ge a \; . \tag{15.271}$$

By Lemma 15.9.8, W can be covered by a^{-L/γ^2} balls of radius $\gamma/2$, which implies that T can be covered by a^{-L/γ^2} sets of the type

$$S' = \{x \; ; \; \|b_{C(x)} - b_{C(x')}\| \le \gamma\}$$

where $x' \in T$. Therefore, we can find $x' \in T$ such that the corresponding set S' satisfies

$$a \le \pi(T) \le a^{-L/\gamma^2}\pi(S') \; . \tag{15.272}$$

We set $S = C(x')$ so $\pi(S) \ge a$ since $x' \in T$. For $x \in S'$, we have

$$x \cdot b_{C(x)} = \frac{1}{\pi(C(x))}\int_{C(x)} x \cdot y \mathrm{d}\pi(y) \ge q - \gamma \tag{15.273}$$

since $x \cdot y \ge q - \gamma$ for y in $C(x)$. Since $x \in S'$ we have

$$\|b_{C(x)} - b_S\| = \|b_{C(x)} - b_{C(x')}\| \le \gamma \; ,$$

and since $\|x\| \le 1$ this yields $|x \cdot b_{C(x)} - x \cdot b_S| \le \gamma$ and (15.273) implies

$$x \cdot b_S \ge q - 2\gamma \; .$$

Averaging for x over S' we get $b_{S'} \cdot b_S \ge q - 2\gamma$. Since $\pi(S') \ge \tau$ by (15.272) and (15.267), it follows from (15.265) that S is adequate. □

Proof of Lemma 15.9.8. The proof relies on Gaussian tools. The starting point is that there exists a Gaussian process $(X(x))_{x\in H}$ with the property that $\mathsf{E}X(x)X(y) = x \cdot y$. One simple way to see this is to assume (without loss of generality) that H is a Gaussian Hilbert space, i.e. consists of jointly Gaussian r.v.s, in which case $X(x) = x$. Another way is to set $X(x) = \sum g_i x_i$ where (g_i) are i.i.d. standard Gaussian r.v.s and $x = \sum x_i e_i$ is the decomposition of x in a given orthonormal basis.

Hölder's inequality implies that if p and q are conjugate exponents,

$$\left|\int_S X(x)\mathrm{d}\pi(x)\right| \le \pi(S)^{1/q}\left(\int_B |X(x)|^p \mathrm{d}\pi(x)\right)^{1/p} \; .$$

By linearity

$$|X(b_S)| = \left|\frac{1}{\pi(S)}\int_S X(x)\mathrm{d}\mu(x)\right| = \frac{1}{\pi(S)}\left|\int_S X(x)\mathrm{d}\mu(x)\right|$$
$$\le \pi(S)^{-1/p}\left(\int_B |X(x)|^p \mathrm{d}\pi(x)\right)^{1/p}$$

so that

$$\sup_{b\in W} |X(b)| \le a^{-1/p} \left(\int_B |X(x)|^p \mathrm{d}\pi(x) \right)^{1/p}$$

and

$$\begin{aligned} \mathsf{E} \sup_{b\in W} |X(b)| &\le a^{-1/p} \mathsf{E} \left(\int_B |X(x)|^p \mathrm{d}\pi(x) \right)^{1/p} \\ &\le a^{-1/p} \left(\int_B \mathsf{E}|X(x)|^p \mathrm{d}\pi(x) \right)^{1/p} . \end{aligned} \tag{15.274}$$

For a Gaussian r.v. X with $\mathsf{E}X^2 \le 1$, we have $\mathsf{E}|X|^p \le (Lp)^{p/2}$. So, for x in B we have $\mathsf{E}|X(x)|^p \le (Lp)^{p/2}$ and (15.274) implies

$$\mathsf{E} \sup_{b\in W} |X(b)| \le La^{-1/p}\sqrt{p} .$$

This bound holds for every $p > 1$. Since $a \le 1/4$ by (15.264), we have $\log(1/a) \ge 1$. Taking $p = \log(1/a)$ we get

$$\mathsf{E} \sup_{b\in W} |X(b)| \le L\sqrt{\log(1/a)} .$$

To conclude we appeal to a key result about Gaussian processes, Sudakov minoration (see [58] Theorem 3.18): The set W can be covered by

$$\exp\left(\frac{L\sqrt{\log(1/a)}}{\gamma} \right)^2 \le a^{-L/\gamma^2}$$

balls of radius $\le \gamma$. □

If B contains no adequate set we set $s = 0$. Otherwise we select such a set C_1. If $B \setminus C_1$ contains no adequate set we define $s = 1$. Otherwise we select an adequate set C_2 contained in $B \setminus C_1$. Continuing in this manner we construct recursively disjoint adequate sets $C_1, C_2, \ldots, C_s$ for as long as we can. We note that since $\pi(C_k) \ge a$ we have

$$s \le \frac{1}{a} . \tag{15.275}$$

We define

$$C = C_1 \cup \ldots \cup C_s .$$

Lemma 15.9.9. *We have*

$$\pi^{\otimes 2}(\{(x,y)\ ;\ x\cdot y \ge q-\gamma\} \setminus C \times C) \le 4a . \tag{15.276}$$

Proof. Let $D = B \setminus C$. By construction D does not contain an adequate set, so that by Proposition 15.9.7 we have

$$\pi^{\otimes 2}((B \times D) \cap \{(x,y)\ ;\ x \cdot y \geq q - \gamma\}) \leq 2a$$

and the result follows since $B^2 \setminus (C \times C) \subset (B \times D) \cup (D \times B)$ and "symmetry around the diagonal". □

The sets A_α will be constructed by suitably grouping the sets C_k.

Definition 15.9.10. *We say that a pair* (k, k') *is* close *if*

$$\pi^{\otimes 2}(\{(x,y)\ ;\ x \cdot y \geq q - \gamma\} \cap (C_k \times C_{k'})) \geq a\pi(C_k)\pi(C_{k'})\ .$$

On the set $\{1, \ldots, s\}$ we define the equivalence relation $k\mathcal{R}k'$ by

$$\exists\, \ell_1, \ldots, \ell_p,\ \ell_1 = k,\ \ell_p = k',\ \text{all pairs } (\ell_n, \ell_{n+1}) \text{ are close for } 1 \leq n < p.$$

The set $\{1, \ldots, s\}$ is then divided into equivalence classes $I_1, \ldots, I_r$. For $\alpha \leq r$ we set $A_\alpha = \bigcup_{k \in I_\alpha} C_k$.

Lemma 15.9.11. *We have*

$$\pi^{\otimes 2}\left(\{(x,y)\ ;\ x \cdot y \geq q - \gamma\} \setminus \bigcup_{\alpha \leq r} A_\alpha^2\right) \leq 5a\ . \tag{15.277}$$

Proof. Let

$$U = \{(x,y)\ ;\ x \cdot y \geq q - \gamma\} \setminus \bigcup_{\alpha \leq r} A_\alpha^2$$

and note that

$$\pi^{\otimes 2}(U \cap (C \times C)) = \sum_{k,\ell \leq s} \pi^{\otimes 2}(U \cap (C_k \times C_\ell))\ .$$

When $k\mathcal{R}\ell$, then $C_k \times C_\ell \subset A_\alpha^2$ for a certain α, so that $U \cap (C_k \times C_\ell) = \emptyset$. Otherwise k and ℓ are not close, and thus

$$\pi^{\otimes 2}(U \cap (C_k \times C_\ell)) \leq a\pi(C_k)\pi(C_\ell)$$

and by summation over k and ℓ,

$$\pi^{\otimes 2}(U \cap (C \times C)) \leq a\ .$$

Now (15.276) implies

$$\pi^{\otimes 2}(U \setminus (C \times C)) \leq 4a$$

and this completes the proof. □

In the last stage of the proof, we will show that $x \cdot y \simeq q$ on each set A_α. We start with the inequality

$$\int_{A_\alpha^2} |x \cdot y - q| \mathrm{d}\pi(x)\mathrm{d}\pi(y) = \sum_{k,\ell \in I_\alpha} \int_{C_k \times C_\ell} |x \cdot y - q| \mathrm{d}\pi(x)\mathrm{d}\pi(y) \,. \quad (15.278)$$

Denoting by b_ℓ the barycentre of π over C_ℓ, we write

$$\int_{C_k \times C_\ell} |x \cdot y - q| \mathrm{d}\pi(x)\mathrm{d}\pi(y) \le \int_{C_k \times C_\ell} |x \cdot y - b_k \cdot b_\ell| \mathrm{d}\pi(x)\mathrm{d}\pi(y) + |b_k \cdot b_\ell - q| \pi(C_k)\pi(C_\ell) \,. \quad (15.279)$$

The general idea is that $b_k \cdot b_\ell \sim q$ for $k, \ell \in I_\alpha$. To prove this, we will appeal to a general principle, that will also help to control the first term on the right-hand side of (15.279).

Proposition 15.9.12. *Consider two probability measures ν and ν' on B, and assume that for certain numbers q and q' they satisfy*

$$\int |x \cdot y - q| \mathrm{d}\nu(x)\mathrm{d}\nu(y) \le \gamma \,; \qquad \int |x \cdot y - q'| \mathrm{d}\nu'(x)\mathrm{d}\nu'(y) \le \gamma \,. \quad (15.280)$$

Then, if b and b' denote the barycenters of ν and ν' respectively, we have

$$\int |x \cdot y - b \cdot b'| \mathrm{d}\nu(x)\mathrm{d}\nu'(y) \le 8\sqrt{\gamma} \,. \quad (15.281)$$

Of course, the barycentre b of ν is given by the formula $b = \int x \mathrm{d}\nu(x)$.

Lemma 15.9.13. *Consider a probability measure ν on B and assume that*

$$\int |x \cdot y - q| \mathrm{d}\nu(x)\mathrm{d}\nu(y) \le \gamma \,. \quad (15.282)$$

Consider a probability measure ν' on B, and assume that it has a density w with respect to ν. Then

$$\left\| \int x \mathrm{d}\nu(x) - \int x \mathrm{d}\nu'(x) \right\| \le 2\sqrt{\gamma} \|w\|_\infty \,. \quad (15.283)$$

Proof. We write $\left\| \int x \mathrm{d}\nu(x) - \int x \mathrm{d}\nu'(x) \right\|^2$ as

$$\int (x \cdot y - q) \mathrm{d}\nu(x)\mathrm{d}\nu(y) - 2\int (x \cdot y - q) \mathrm{d}\nu(x)\mathrm{d}\nu'(y) + \int (x \cdot y - q) \mathrm{d}\nu'(x)\mathrm{d}\nu'(y) \,,$$

and we bound each term using (15.282). For example,

$$\int |x \cdot y - q| \mathrm{d}\nu(x) \mathrm{d}\nu'(y) = \int |x \cdot y - q| w(y) \mathrm{d}\nu(x) \mathrm{d}\nu(y) \le \gamma \|w\|_\infty .$$

The other two terms are bounded respectively by γ and $\gamma\|w\|_\infty^2$, and, since $\|w\|_\infty \ge 1$,

$$\gamma + 2\gamma\|w\|_\infty + \gamma\|w\|_\infty^2 = \gamma(1 + \|w\|_\infty)^2 \le 4\gamma\|w\|_\infty^2 . \qquad \square$$

Proof of Proposition 15.9.12. Consider the set $A = \{(x,y)\, ; x \cdot y \ge b \cdot b'\}$ and for $y \in B$ let $A(y) = \{x\, ;\ (x,y) \in A\}$. Then

$$\begin{aligned} \int (x \cdot y - b \cdot b')^+ \mathrm{d}\nu(x) \mathrm{d}\nu'(y) &= \int_A (x \cdot y - b \cdot b') \mathrm{d}\nu(x) \mathrm{d}\nu'(y) \\ &= \int_A x \cdot y \mathrm{d}\nu(x) \mathrm{d}\nu'(y) - b \cdot b' \nu \otimes \nu'(A) \\ &= \int \Big(\int_{A(y)} x \mathrm{d}\nu(x) \Big) \cdot y \mathrm{d}\nu'(y) \\ &\quad - b \cdot b' \nu \otimes \nu'(A) . \end{aligned} \tag{15.284}$$

By (15.283),

$$\left\| \int x \mathrm{d}\nu(x) - \frac{1}{\nu(A(y))} \int_{A(y)} x \mathrm{d}\nu(x) \right\| \le \frac{2\sqrt{\gamma}}{\nu(A(y))}$$

so that

$$\left\| \int_{A(y)} x \mathrm{d}\nu(x) - \nu(A(y)) \int x \mathrm{d}\nu(x) \right\| \le 2\sqrt{\gamma}$$

and therefore

$$\begin{aligned} \int \Big(\int_{A(y)} x \mathrm{d}\nu(x) \Big) \cdot y \mathrm{d}\nu'(y) &\le 2\sqrt{\gamma} + \int \Big(\int \nu(A(y)) \int x \mathrm{d}\nu(x) \Big) \cdot y \mathrm{d}\nu'(y) \\ &\le 2\sqrt{\gamma} + \int x \mathrm{d}\nu(x) \cdot \int \nu(A(y)) y \mathrm{d}\nu'(y) . \end{aligned} \tag{15.285}$$

We observe now that $\int \nu(A(y)) \mathrm{d}\nu'(y) = \nu \otimes \nu'(A)$, so using (15.283) for $w(y) = \nu(A(y))/\nu \otimes \nu'(A)$ and ν' instead of ν we get

$$\left\| \int y \mathrm{d}\nu'(y) - \frac{1}{\nu \otimes \nu'(A)} \int \nu(A(y)) y \mathrm{d}\nu'(y) \right\| \le \frac{2\sqrt{\gamma}}{\nu \otimes \nu'(A)}$$

and

$$\left\| \int \nu(A(y)) y \mathrm{d}\nu'(y) - \nu \otimes \nu'(A) \int y \mathrm{d}\nu'(y) \right\| \le 2\sqrt{\gamma} .$$

Therefore

$$\int x\mathrm{d}\nu(x)\cdot\int \nu(A(y))y\mathrm{d}\nu'(y) \le \nu\otimes\nu'(A)\int x\mathrm{d}\nu(x)\cdot\int y\mathrm{d}\nu'(y)+2\sqrt{\gamma}$$
$$= \nu\otimes\nu'(A)b\cdot b'+2\sqrt{\gamma}\;.$$

Combining with (15.284) and (15.285) we have shown that

$$\int (x\cdot y-b\cdot b')^+\mathrm{d}\nu(x)\mathrm{d}\nu'(y)\le 4\sqrt{\gamma}$$

and we proceed in the same way for the "other half". □

Lemma 15.9.14. *If a pair* (k,ℓ) *is close then*

$$b_k\cdot b_\ell \ge q-\gamma-12\sqrt{\gamma}a^{-1}\;. \tag{15.286}$$

Proof. We may assume that $q-\gamma > b_k\cdot b_\ell$, for otherwise there is nothing to prove. Then

$$\mathbf{1}_{\{x\cdot y\ge q-\gamma\}}\le \frac{1}{q-\gamma-b_k\cdot b_\ell}(x\cdot y-b_k\cdot b_\ell)^+\;. \tag{15.287}$$

Using that the pair (k,ℓ) is close in the first line and (15.287) in the second line, we have

$$\begin{aligned} a\pi(C_k)\pi(C_\ell) &\le \pi^{\otimes 2}((C_k\times C_\ell)\cap\{(x,y)\;;\;x\cdot y\ge q-\gamma\}) \\ &= \int_{C_k\times C_\ell}\mathbf{1}_{\{x\cdot y\ge q-\gamma\}}\mathrm{d}\pi(x)\mathrm{d}\pi(y) \\ &\le \frac{1}{q-\gamma-b_k\cdot b_\ell}\int_{C_k\times C_\ell}|x\cdot y-b_k\cdot b_\ell|\mathrm{d}\pi(x)\mathrm{d}\pi(y)\;. \end{aligned} \tag{15.288}$$

Since C_k and C_ℓ are adequate the conditional measures ν and ν' of π on C_k and C_ℓ satisfy (15.280) for 8γ rather than γ. Then (15.281) implies

$$\frac{1}{\pi(C_k)\pi(C_\ell)}\int_{C_k\times C_\ell}|x\cdot y-b_k\cdot b_\ell|\mathrm{d}\pi(x)\mathrm{d}\pi(y)\le 8\sqrt{8\gamma}\le 12\sqrt{\gamma}$$

and combining with (15.288) we get

$$a\le\frac{12\sqrt{\gamma}}{q-\gamma-b_k\cdot b_\ell}\;.$$ □

Lemma 15.9.15. *If* $k\mathcal{R}\ell$ *then*

$$|b_k\cdot b_\ell-q|\le 16\sqrt{\gamma}a^{-3}\;. \tag{15.289}$$

Proof. Since the set C_k is adequate, it follows from (15.258) and Jensen's inequality (i.e. integration inside the absolute value rather than outside) that

$$|\|b_k\|^2 - q| \le 8\gamma \,. \tag{15.290}$$

When the pair (k,ℓ) is close, using (15.286) and (15.290) we get, since $a^{-1} \ge 4$,

$$\|b_k - b_\ell\|^2 = \|b_k\|^2 + \|b_\ell\|^2 - 2b_k \cdot b_\ell \le 16\gamma + 2\gamma + 24\sqrt{\gamma}a^{-1} \le 29\sqrt{\gamma}a^{-1} \,,$$

so that $\|b_k - b_\ell\| \le (29\sqrt{\gamma}a^{-1})^{1/2}$. When $k\mathcal{R}\ell$, k and ℓ are connected by a chain of pairs (ℓ_n, ℓ_{n+1}) that are close, and since $s \le a^{-1}$ we get

$$\|b_k - b_\ell\| \le a^{-1}(29\sqrt{\gamma}a^{-1})^{1/2} \,,$$

so that $\|b_k - b_\ell\|^2 \le 29\sqrt{\gamma}a^{-3}$. Therefore, using (15.290) again, (and since $a^{-3} \ge 64$)

$$\begin{aligned} b_k \cdot b_\ell &= \frac{1}{2}(\|b_k\|^2 + \|b_\ell\|^2 - \|b_k - b_\ell\|^2) \\ &\ge q - 8\gamma - 15\sqrt{\gamma}a^{-3} \ge q - 16\sqrt{\gamma}a^{-3} \end{aligned}$$

and since $b_k \cdot b_\ell \le \|b_k\|\|b_\ell\| \le q + 8\gamma$ we have proved (15.289). □

Lemma 15.9.16. *For each α we have*

$$\int_{A_\alpha^2} |x \cdot y - q| \mathrm{d}\pi(x)\mathrm{d}\pi(y) \le 17\sqrt{\gamma}a^{-3}\pi(A_\alpha)^2 \,. \tag{15.291}$$

Proof. Using (15.281) to control the first term in the right-hand side of (15.279) and (15.289) to control the second term we obtain

$$\begin{aligned} \int_{C_k \times C_\ell} |x \cdot y - q| \mathrm{d}\pi(x)\mathrm{d}\pi(y) &\le 8\sqrt{8\gamma} + 16\sqrt{\gamma}a^{-3}\pi(C_k)\pi(C_\ell) \\ &\le 17\sqrt{\gamma}a^{-3}\pi(C_k)\pi(C_\ell) \,, \end{aligned}$$

and since $\pi(A_\alpha) = \sum_{k \in I_\alpha} \pi(C_k)$ substitution in (15.278) finishes the proof. □

Proof of Theorem 15.4.5. We define a by $5a = \delta$, γ by $17\sqrt{\gamma}a^{-3} = \delta$ and τ by (15.267). Then if $\varepsilon = \gamma\tau$, (15.61) implies (15.257). Then (15.62) follows from (15.277) and (15.63) from (15.291). □

Corollary 15.9.17. *Under the hypothesis of Theorem 15.4.5, given u with $\|u\| \le 1$ we have*

$$\int_{\{x \cdot y \ge q - \varepsilon\}} |x \cdot u - y \cdot u| \mathrm{d}\pi(x)\mathrm{d}\pi(y) \le 18\sqrt{\delta} \,. \tag{15.292}$$

Proof. Since by (15.63)

$$\pi^{\otimes 2}\Big(\{(x,y)\ ;\ x\cdot y \geq q-\varepsilon\}\setminus \bigcup A_\alpha^2\Big) \leq \delta\ ,$$

and since $|x\cdot u - y\cdot u| \leq 2$, we may bound the left-hand side of (15.292) by

$$2\delta + \sum_\alpha \int_{A_k^2} |x\cdot u - y\cdot u| \mathrm{d}\pi(x)\mathrm{d}\pi(y)\ . \tag{15.293}$$

Using Proposition 15.9.12 with ν' the point mass at u, and denoting by b the barycentre of π over A_k, we obtain

$$\int_{A_k} |x\cdot u - b\cdot u| \mathrm{d}\pi(x) \leq 8\sqrt{\delta}\pi(A_k)$$

so that, since $|x\cdot u - y\cdot u| \leq |x\cdot u - b\cdot u| + |y\cdot u - b\cdot u|$,

$$\int_{A_k^2} |x\cdot u - y\cdot u| \mathrm{d}\pi(x)\mathrm{d}\pi(y) \leq 16\sqrt{\delta}\pi(A_k)^2$$

and the quantity (15.293) is indeed bounded by $18\sqrt{\delta}$. □

15.10 Notes and Comments

Although the methods are different, Panchenko's work [71] (and also the author's [113]) has been influenced by the earlier works of M. Aizenman and P. Contucci [8] and M. Aizenman and P. Arguin [5]. The idea of stochastic stability (of which Theorem 15.5.7 is a particular instance) is put forward in [8], and an ultrametricity theorem is proved (under different hypotheses) in [5]. Determinators are introduced in that paper under the name of "Random Overlap Structures." The author wishes he had been able to explain here the ideas of [5], but, despite significant efforts, he was defeated by the style of writing.

16. The p-Spin Interaction Model

16.1 Overview

In this chapter, given an integer p, we study the system with Hamiltonian

$$-H_p(\boldsymbol{\sigma}) = \beta\left(\frac{p!}{N^{p-1}}\right)^{1/2} \sum_{i_1<\ldots<i_p} g_{i_1\ldots i_p}\sigma_{i_1}\cdots\sigma_{i_p} \,. \tag{16.1}$$

This Hamiltonian is closely related to the Hamiltonian (14.57). The use of the formula (16.1) is motivated by the fact that is slightly easier then than in the case (14.57) to use the cavity method, which will be an essential tool in this chapter. It could be shown that results similar to those of this chapter hold for the Hamiltonian (14.57).

Physicists usually write the Hamiltonian (16.1) with a normalization factor $(p!/2N^{p-1})^{1/2}$. The only purpose of the factor 2 is to make the definition coincide with the definition of the SK model when $p = 2$. Since we will be interested in the case where p is large, there is no point to use the physicists' normalization.

We shall concentrate on the case where p is odd. This case is not covered by the results of Chapter 14 because the function $\xi(x) = \beta^2 x^p$ is convex on $\mathbb{R}^+$ but not on $\mathbb{R}$. It is conceivable that it will require a single good idea to be able to use the methods of Chapter 14 to study the system with Hamiltonian (16.1) when p is odd, but, for the time being, the only result from Chapter 14 that can be applied to the Hamiltonian (16.1) is the upper bound of Theorem 14.4.4. Some of the methods we will use might look unsophisticated compared to the work of Chapter 14, but we do not know how to do better. A major obstacle is that we do not know how to prove results similar to those of Section 14.6 in the present case.

The real motivation for looking at the Hamiltonian governed by the Hamiltonian (16.1) is not to be able to say something about this case at any cost. It is that the proofs will shed a completely different light on the model than the methods of Chapter 14. In the range of β where we can control the model, that is, up to values of β that are exponentially large in p, (and for β not too small) it has a structure "with one level of replica-symmetry breaking". This means here that the configuration space is the union of a sequence of "pure states". The overlap of two configurations in two different

M. Talagrand, *Mean Field Models for Spin Glasses*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys in Mathematics 55, DOI 10.1007/978-3-642-22253-5_9,

pure states is essentially 0 and the overlap of two configurations in the same pure state is essentially a certain number q depending on the system. The proof closely follows the intuition provided by this picture, trying in several steps to prove that it holds true.

In the first part of the proof (Section 16.3) we use elementary estimates to prove that the overlap can essentially take only values that are either in a small interval around 0 or in a small interval around 1, where these small intervals do not depend on N. If one thinks a minute about this situation, one sees (in Section 16.4) that it strongly constrains the structure of the Gibbs measure. This is proved by a construction in the spirit of that of Section 14.12, but quite simpler. The configuration space Σ_N can be divided in a sequence of pieces such that the overlap of two configurations in the same piece is in a small interval around 1, while the overlap of two configurations in different pieces is in a small interval around 0. We call these pieces the *lumps*. We expect of course that they will be the pure states that we are looking for, but we have not proved it at this stage.

The main step in gaining control over the model is to prove that (for β not too large) the overlap takes only two values (one of which turns out to be 0). The proof relies on the cavity method, with the remarkable feature that the construction of the interpolating Hamiltonian depends critically on the decomposition of the $(N-2)$-spin system into lumps. It is done in Section 16.5.

In order to be able to fully use the cavity method, we however need to control the distribution of the sequence of the Gibbs weights of the lumps. The only way we know how to do this is to add the perturbation term (12.32) to the Hamiltonian, so that we can benefit from the extended Ghirlanda-Guerra identities. These identities imply that the sequence of the Gibbs weights of the lumps has in the limit a Poisson-Dirichlet distribution, a crucial ingredient of the cavity method.

Once we have the precise information that the overlap take only two values, we can use again the cavity method to prove that the upper bound of Theorem 14.4.4 is also asymptotically a lower bound. The perturbation term (12.32) is irrelevant in the limit for this computation. We however do not know how to gain a detailed control of the model.

16.2 Poisson-Dirichlet Distribution and Ghirlanda-Guerra Identities

In Section 15.2 we proved that a sequence of random weights that satisfies a certain version of the Ghirlanda-Guerra identities has a Poisson-Dirichlet distribution. On the other hand, a sequences of random weights with a Poisson-Dirichlet distribution has the remarkable properties spelled out by Theorem 13.1.6. Therefore we expect that a sequence of weights that nearly satisfies the Ghirlanda-Guerra identities nearly satisfies Theorem 13.1.6. This is what we show in this section. We recall that we define the numbers

$$S^{(m)}(n_1,\ldots,n_k) := \mathsf{E}\prod_{s\le k}\sum_{\alpha} v_\alpha^{n_s}\ , \tag{16.2}$$

where the random weights (v_α) have a Poisson-Dirichlet distribution Λ_m.

Theorem 16.2.1. *Consider two numbers A, $\varepsilon' > 0$. Then we can find a number n and a number $\varepsilon > 0$ with the following property. Consider a random sequence $(w_\alpha)_{\alpha\ge1}$ with $w_1 \ge w_2 \ge \ldots$ and $\sum_{\alpha\ge1} w_\alpha = 1$. Consider a number $0 < m < 1$, and assume that, given any integer k, any integers $n_1,\ldots,n_k$ with $n_s \ge 2$, $\sum_{s\le k} n_s \le n$, we have*

$$\left|\mathsf{E}\prod_{s\le k}\sum_{\alpha\ge1} w_\alpha^{n_s} - S^{(m)}(n_1,\ldots,n_k)\right| \le \varepsilon\ . \tag{16.3}$$

Consider a pair (U,V) of r.v.s with $V\ge1$, $\mathsf{E}U^2 \le A$ and $\mathsf{E}V^4 \le A^2$. Consider an independent sequence of pairs $((U_\alpha,V_\alpha))_{\alpha\ge1}$ that are distributed like the pair (U,V) and that are independent of the sequence (w_α). Then

$$\left|\mathsf{E}\,\frac{\sum_{\alpha\ge1} w_\alpha U_\alpha}{\sum_{\alpha\ge1} w_\alpha V_\alpha} - \frac{\mathsf{E}\,UV^{m-1}}{\mathsf{E}\,V^m}\right| \le \varepsilon' \tag{16.4}$$

$$\left|\mathsf{E}\,\frac{\sum_{\alpha\ne\gamma} w_\alpha\, w_\gamma\, U_\alpha\, U_\gamma}{(\sum_{\alpha\ge1} w_\alpha\, V_\alpha)^2} - m\left(\frac{\mathsf{E}\,UV^{m-1}}{\mathsf{E}\,V^m}\right)^2\right| \le \varepsilon' \tag{16.5}$$

$$\left|\mathsf{E}\,\frac{\sum_{\alpha\ge1} w_\alpha^2 U_\alpha^2}{(\sum_{\alpha\ge1} w_\alpha V_\alpha)^2} - (1-m)\frac{\mathsf{E}\,U^2V^{m-2}}{\mathsf{E}\,V^m}\right| \le \varepsilon'\ . \tag{16.6}$$

Proof. We will consider only the case of (16.4). The others are similar. It is possible to prove (16.4) through a rather straightforward compactness argument; but we find it more fun to give a somewhat constructive proof as follows. Given ε and A, we will find a quantity $C(A,\varepsilon)$ and a number $n = n(A,\varepsilon)$, depending on ε and A only, with the following property. For any pair (U,V) of r.v.s as above, we can find numbers $c(n_1,\ldots,n_k,U,V)$ for $n_1,\ldots,n_k \ge 2$, $n_1+\cdots+n_k \le n(A,\varepsilon) = n$ with

$$|c(n_1,\ldots,n_k,U,V)| \le C(A,\varepsilon) \tag{16.7}$$

and

$$\left|\mathsf{E}\frac{\sum_{\alpha\ge1} w_\alpha U_\alpha}{\sum_{\alpha\ge1} w_\alpha V_\alpha} - \sum_{n_1+\cdots+n_k\le n} c(n_1,\ldots,n_k,U,V)\mathsf{E}\prod_{s\le k}\sum_{\alpha\ge1} w_\alpha^{n_s}\right| \le \varepsilon\ . \tag{16.8}$$

This inequality holds in particular in the case $w_\alpha = v_\alpha$, a sequence with distribution Λ_m, in which case (13.13) implies

$$\left|\frac{\mathsf{E}UV^{m-1}}{\mathsf{E}V^m} - \sum_{n_1+\cdots+n_k\le n} c(n_1,\ldots,n_k,U,V)S^{(m)}(n_1,\ldots,n_k)\right| \le \varepsilon$$

and combining with (16.7) and (16.8) we get

$$\begin{aligned}&\left|\mathsf{E}\frac{\sum_{\alpha\ge1} w_\alpha U_\alpha}{\sum_{\alpha\ge1} w_\alpha V_\alpha} - \frac{\mathsf{E}UV^{m-1}}{\mathsf{E}V^m}\right| \\ &\le 2\varepsilon + C(A,\varepsilon)\sum_{n_1+\cdots+n_k\le n}\left|S^{(m)}(n_1,\ldots,n_k) - \mathsf{E}\prod_{s\le k}\sum_{\alpha\ge1} w_\alpha^{n_s}\right|, \quad (16.9)\end{aligned}$$

from which the theorem follows.

We now turn to the proof of (16.8). The argument will actually show that this inequality holds uniformly over non-random sequences (w_α). We first reduce to the case where U and V are bounded by a truncation argument. Consider a truncation level $M \ge 1$ and define

$$V'_\alpha = \min(V_\alpha, M) \ \ ; \ \ U'_\alpha = U_\alpha \mathbf{1}_{\{|U_\alpha|\le M\}} .$$

Since $\sum_{\alpha\ge1} w_\alpha = 1$ and $V_\alpha, V'_\alpha \ge 1$, we have

$$\begin{aligned}\left|\frac{\sum_{\alpha\ge1} w_\alpha U_\alpha}{\sum_{\alpha\ge1} w_\alpha V_\alpha} - \frac{\sum_{\alpha\ge1} w_\alpha U'_\alpha}{\sum_{\alpha\ge1} w_\alpha V'_\alpha}\right| &\le \sum_{\alpha\ge1} w_\alpha |U_\alpha - U'_\alpha| \quad (16.10)\\ &+ \Big(\sum_{\alpha\ge1} w_\alpha U_\alpha\Big)\Big(\sum_{\alpha\ge1} w_\alpha |V_\alpha - V'_\alpha|\Big) .\end{aligned}$$

Using the Cauchy-Schwarz inequality and that $(\sum_\alpha w_\alpha x_\alpha)^2 \le \sum_\alpha w_\alpha x_\alpha^2$ by convexity of the function $x \mapsto x^2$, we obtain (using the obvious definition of U and U')

$$\begin{aligned}&\mathsf{E}\Big(\sum_{\alpha\ge1} w_\alpha U_\alpha\Big)\Big(\sum_{\alpha\ge1} w_\alpha |V_\alpha - V'_\alpha|\Big) \\ &\le \Big(\mathsf{E}\sum_{\alpha\ge1} w_\alpha U_\alpha^2\Big)^{1/2}\Big(\mathsf{E}\sum_{\alpha\ge1} w_\alpha (V_\alpha - V'_\alpha)^2\Big)^{1/2} \\ &= (\mathsf{E}U^2)^{1/2}(\mathsf{E}(V-V')^2)^{1/2} .\end{aligned}$$

Therefore the expectation of the left-hand side of (16.10) is at most

$$\mathsf{E}|U-U'| + (\mathsf{E}U^2)^{1/2}(\mathsf{E}(V-V')^2)^{1/2} .$$

Now,

$$\mathsf{E}(V-V')^2 \le \mathsf{E}V^2\mathbf{1}_{\{V\ge M\}} \le \frac{1}{M^2}\mathsf{E}V^4 \le \frac{A^2}{M^2} ,$$

and similarly we have $\mathsf{E}|U - U'| \le \mathsf{E}|U|\mathbf{1}_{\{|U|\ge M\}} \le \mathsf{E}U^2/M \le A/M$. Therefore the expectation of the left-hand side of (16.10) is $\le (A + A^{3/2})/M$. This should make it clear that it suffices to prove (16.8) when U and V are bounded. Assuming that $|U|, |V| \le M$ we prove that we can find a quantity $C(M,\varepsilon)$ and a number $n = n(M,\varepsilon)$ depending only on M and ε, such that (16.8) holds for numbers $c(n_1,\dots,n_k,U,V)$ with

$$|c(n_1,\dots,n_k,U,V)| \le C(M,\varepsilon)\ .$$

On the compact set

$$\{x,y\ ;\ |x| \le M\ ,\ 1 \le y \le M\}$$

the function x/y is approximated within ε by a polynomial in x and y, so that to prove (16.8) it suffices to prove it when we replace the ratio $\sum w_\alpha U_\alpha / \sum w_\alpha V_\alpha$ by a monomial $(\sum w_\alpha U_\alpha)^a (\sum w_\alpha V_\alpha)^b$. In particular it suffices to show that, given any two integers a and b we have

$$\mathsf{E}\left(\left(\sum w_\alpha U_\alpha\right)^a \left(\sum w_\alpha V_\alpha\right)^b\right) = \sum c(n_1,\dots,n_k,U,V)\mathsf{E}\prod_{s\le k}\prod_{\alpha\ge 1} w_\alpha^{n_s}\ , \tag{16.11}$$

where the summation is over $n_1 + \cdots + n_k = a + b$, and where the numbers $c(n_1,\dots,n_k,U,V)$ satisfy

$$|c(n_1,\dots,n_k,U,V)| \le C(a,b)M^{a+b}\ .$$

First we observe that, using independence in the last equality,

$$\begin{aligned}
&\mathsf{E}\left(\left(\sum w_\alpha U_\alpha\right)^a \left(\sum w_\alpha V_\alpha\right)^b\right) \qquad (16.12)\\
&= \mathsf{E}\sum w_{\alpha_1}\cdots w_{\alpha_a} w_{\alpha_{a+1}}\cdots w_{\alpha_{a+b}} U_{\alpha_1}\cdots U_{\alpha_a} V_{\alpha_{a+1}}\cdots V_{\alpha_{a+b}}\\
&= \sum \mathsf{E}(w_{\alpha_1}\cdots w_{\alpha_a} w_{\alpha_{a+1}}\cdots w_{\alpha_{a+b}})\mathsf{E}(U_{\alpha_1}\cdots U_{\alpha_a} V_{\alpha_{a+1}}\cdots V_{\alpha_{a+b}})\ ,
\end{aligned}$$

where the summations are over all values of $\alpha_1,\dots,\alpha_{a+b} \ge 1$. Next, consider a partition $\mathcal{J}$ of the set $\{1,\dots,a+b\}$ into subsets $J_1,\dots,J_k$, and for $s \le k$ set

$$m_s = \mathrm{card}(J_s \cap \{1,\dots,a\})\ ;\ r_s = \mathrm{card}(J_s \cap \{a+1,\dots,a+b\})\ .$$

Let us consider the following property of a sequence $(\alpha_1,\dots,\alpha_{a+b})$:

$$\forall \ell\ ,\ \ell' \le a+b\ , \alpha_\ell = \alpha'_\ell\ \Leftrightarrow\ \exists p \le n\ \ ;\ \ell\ ,\ \ell' \in J_p\ . \tag{16.13}$$

When the sequence $(\alpha_1,\dots,\alpha_{a+b})$ satisfies (16.13), then

$$\mathsf{E}U_{\alpha_1}\cdots U_{\alpha_a} V_{\alpha_{a+1}}\cdots V_{\alpha_{a+b}} = \mathsf{E}\prod_{s\le k} U_s^{m_s} V_s^{r_s}$$

is a number $C(\mathcal{J})$ that depends only on $\mathcal{J}$ (and U and V). Moreover $|C(\mathcal{J})| \le M^{a+b}$.

Let us denote by $\sum_{\mathcal{J}}$ a summation over all the choices of $\alpha_1, \ldots, \alpha_{a+b}$ that satisfy (16.13). Then

$$\sum_{\mathcal{J}} \mathsf{E}(w_{\alpha_1} \cdots w_{\alpha_{a+b}})\mathsf{E}(U_{\alpha_1} \cdots U_{\alpha_a} V_{\alpha_{a+1}} \cdots V_{\alpha_{a+b}}) \tag{16.14}$$

$$= C(\mathcal{J})\mathsf{E} \sum_{\mathcal{J}} w_{\alpha_1} \cdots w_{\alpha_{a+b}} \ . \tag{16.15}$$

It then follows from (16.13) that

$$\mathsf{E}\left(\left(\sum w_\alpha U_\alpha\right)^a \left(\sum w_\alpha V_\alpha\right)^b\right) = \sum_{\text{all choices of } \mathcal{J}} C(\mathcal{J})\mathsf{E} \sum_{\mathcal{J}} w_{\alpha_1} \cdots w_{\alpha_{a+b}} \ .$$

Moreover we have

$$\sum_{\mathcal{J}} w_{\alpha_1} \cdots w_{\alpha_{a+b}} = \sum_{\alpha_1, \ldots, \alpha_k \text{ all different}} w_{\alpha_1}^{m_1+r_1} \cdots w_{\alpha_k}^{m_k+r_k} \ .$$

Thus to prove (16.11), it suffice to show that the numbers

$$\mathsf{E} \sum_{\alpha_1, \ldots, \alpha_k \text{ all different}} w_{\alpha_1}^{m_1+r_1} \cdots w_{\alpha_k}^{m_k+r_k}$$

can be recovered as linear combinations of the numbers $\mathsf{E}\prod_{s \le k'} \prod_{\alpha \ge 1} w_\alpha^{n_s}$. This follows from the "inclusion-exclusion formula" as in $\sum_{\alpha_1 \ne \alpha_2} w_{\alpha_1}^3 w_{\alpha_2}^2 = (\sum_\alpha w_\alpha^3)(\sum_\alpha w_\alpha^2) - \sum_\alpha w_\alpha^5$. □

16.3 A Priori Estimates

First, we observe that from (16.1) we have

$$\begin{aligned} \mathsf{E} H_N(\boldsymbol{\sigma}^1) H_N(\boldsymbol{\sigma}^2) &= \frac{\beta^2 p!}{N^{p-1}} \sum_{i_1 < \ldots < i_p} \sigma_{i_1}^1 \sigma_{i_1}^2 \cdots \sigma_{i_p}^1 \sigma_{i_p}^2 \\ &= \frac{\beta^2}{N^{p-1}} \sum_d \sigma_{i_1}^1 \sigma_{i_1}^2 \cdots \sigma_{i_p}^1 \sigma_{i_p}^2 \end{aligned}$$

where $\sum_d$ means that the summation is over all possible choices of indices that are all distinct. Denoting throughout the chapter by K a number depending on p and β only, there are at most KN^{p-1} choices of indices $i_1, \ldots, i_p$ that are not all distinct. Therefore, since $\sum_{i_1, \ldots, i_p} \sigma_{i_1}^1 \sigma_{i_1}^2 \cdots \sigma_{i_p}^1 \sigma_{i_p}^2 = N^p R_{1,2}^p$ we have

$$|\mathsf{E} H_N(\boldsymbol{\sigma}^1) H_N(\boldsymbol{\sigma}^2) - N\beta^2 R_{1,2}^p| \le K \ . \tag{16.16}$$

We recall the function $\mathcal{I}(t)$ of (A.22),

$$\mathcal{I}(t) = \frac{1}{2}\big((1+t)\log(1+t) + (1-t)\log(1-t)\big) .$$

We define the number γ_p by

$$\gamma_p^2 = \inf_{0<t<1}(1+t^{-p})\mathcal{I}(t) .$$

We observe that $\gamma_p^2 \le 2\log 2$ as is seen by taking $t \to 1$. We recall the notation

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

Theorem 16.3.1. *If $\beta \le \gamma_p$ then*

$$\lim_{N\to\infty} p_N(\beta) = \log 2 + \frac{\beta^2}{2} . \tag{16.17}$$

Proof. Jensen's inequality implies

$$p_N(\beta) \le \frac{1}{N}\log\sum_{\boldsymbol{\sigma}}\mathsf{E}\exp(-H_N(\boldsymbol{\sigma})) .$$

Using (16.16) for $\boldsymbol{\sigma}^1 = \boldsymbol{\sigma}^2$ and (A.1) yields $\mathsf{E}\exp(-H_N(\boldsymbol{\sigma})) \le \exp(N\beta^2/2 + K)$, so for every β we have

$$p_N(\beta) \le \log 2 + \frac{\beta^2}{2} + \frac{K}{N} . \tag{16.18}$$

To prove a lower bound for $p_N(\beta)$, consider $\beta \le \gamma_p$ and the r.v.

$$X = \mathrm{card}\{\boldsymbol{\sigma}\ ;\ -H_N(\boldsymbol{\sigma}) \ge N\beta^2\} .$$

The key to the proof is the estimate

$$\mathsf{E}X^2 \le KN(\mathsf{E}X)^2\ ; \quad \mathsf{E}X \ge \frac{2^N}{K\sqrt{N}}\exp\left(-\frac{N\beta^2}{2}\right) . \tag{16.19}$$

Combining with the Paley-Zygmund inequality (A.61) we get

$$\mathsf{P}\left(X \ge \frac{2^N}{2K\sqrt{N}}\exp\left(-\frac{N\beta^2}{2}\right)\right) \ge \mathsf{P}\left(X \ge \frac{\mathsf{E}X}{2}\right) \ge \frac{1}{KN} .$$

Since

$$Z_N(\beta) = \sum_{\boldsymbol{\sigma}}\exp(-H_N(\boldsymbol{\sigma})) \ge X\exp N\beta^2 ,$$

this implies

$$\mathsf{P}\left(Z_N(\beta) \geq \frac{2^N}{K\sqrt{N}} \exp \frac{N\beta^2}{2}\right) \geq \frac{1}{KN}$$

and

$$\mathsf{P}\left(\frac{1}{N}\log Z_N(\beta) \geq \log 2 + \frac{\beta^2}{2} - K\frac{\log N}{N}\right) \geq \frac{1}{KN} \,. \tag{16.20}$$

As in (1.53) we deduce from Proposition 1.3.5 that for $t > 0$ we have

$$\mathsf{P}\left(\left|\frac{1}{N}\log Z_N(\beta) - \frac{1}{N}\mathsf{E}\log Z_N(\beta)\right| \geq t\right) \leq 2\exp\Bigl(-\frac{t^2 N}{2\beta^2}\Bigr) \,,$$

and in particular

$$\mathsf{P}\left(\frac{1}{N}\log Z_N(\beta) \geq \frac{1}{N}\mathsf{E}\log Z_N(\beta) + t\right) \leq 2\exp\Bigl(-\frac{t^2 N}{2\beta^2}\Bigr) \,.$$

We choose $t = K_1\sqrt{\log N}/N$ such that the right-hand side above is less than the right-hand side of (16.20) to infer that

$$p_N(\beta) = \frac{1}{N}\mathsf{E}\log Z_N(\beta) \geq \log 2 + \frac{\beta^2}{2} - K\sqrt{\frac{\log N}{N}} \,,$$

completing the proof of (16.17).

To prove (16.19) we first observe that

$$X = \sum_{\boldsymbol{\sigma}} \mathbf{1}_{\{-H_N(\boldsymbol{\sigma}) \geq N\beta^2\}} \,. \tag{16.21}$$

Using (16.16) for $\boldsymbol{\sigma}^1 = \boldsymbol{\sigma}^2$ and (A.5) yields

$$\mathsf{E}X = \sum_{\boldsymbol{\sigma}} \mathsf{P}(-H_N(\boldsymbol{\sigma}) \geq N\beta^2) \geq \frac{2^N}{K\sqrt{N}} \exp\left(-\frac{N^2\beta^4}{2(N\beta^2 - K)}\right) \,.$$

Moreover,

$$\frac{N^2\beta^4}{2(N\beta^2 - K)} = \frac{N\beta^2}{2}\left(\frac{1}{1 - K/N}\right) \leq \frac{N\beta^2}{2} + K \,, \tag{16.22}$$

and this proves the second part of (16.19). To compute $\mathsf{E}X^2$, we deduce from (16.21) that

$$X^2 = \sum_{\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2} \mathbf{1}_{\{-H_N(\boldsymbol{\sigma}^1) \geq N\beta^2\}} \mathbf{1}_{\{-H_N(\boldsymbol{\sigma}^2) \geq N\beta^2\}}$$

and therefore

$$\mathsf{E}X^2 = \sum_{\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2} A(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) \,, \tag{16.23}$$

where

$$A(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) = \mathsf{P}(-H_N(\boldsymbol{\sigma}^1) \geq N\beta^2, -H_N(\boldsymbol{\sigma}^2) \geq N\beta^2)$$
$$\leq \mathsf{P}(-H_N(\boldsymbol{\sigma}^1) - H_N(\boldsymbol{\sigma}^2) \geq 2N\beta^2) \; .$$

For a Gaussian r.v. g and $t > 0$ we have $\mathsf{P}(g \geq t) \leq \exp(-t^2/2\mathsf{E}g^2)$, and consequently

$$A(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) \leq \exp\left(-\frac{2N^2\beta^4}{\mathsf{E}(H_N(\boldsymbol{\sigma}^1) + H_N(\boldsymbol{\sigma}^2))^2}\right) \; .$$

Combining with (16.16) yields

$$\mathsf{E}(H_N(\boldsymbol{\sigma}^1) + H_N(\boldsymbol{\sigma}^2))^2 \leq 2N\beta^2(1 + R_{1,2}^p) + K \leq 2N\beta^2(1 + |R_{1,2}|^p) + K \; .$$

Proceeding as in (16.22) we obtain

$$A(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) \leq K \exp\left(-\frac{N\beta^2}{(1 + |R_{1,2}|^p)}\right)$$

and therefore

$$\sum A(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) \leq K \sum_{0 \leq k \leq N} \exp\left(-\frac{N\beta^2}{1 + (k/N)^p}\right) \operatorname{card}\{(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) \; ; \; |R_{1,2}| = k/N\} \; .$$

Now, (A.24) implies

$$\operatorname{card}\{(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) \; ; \; |R_{1,2}| = k/N\} \leq 2^{2N+1} \exp\left(-N\mathcal{I}\left(\frac{k}{N}\right)\right)$$

and therefore

$$\sum A(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) \leq K 2^{2N+1} \sum_{0 \leq k \leq N} \exp\left(-N\left(\frac{\beta^2}{1 + (k/N)^p} + \mathcal{I}\left(\frac{k}{N}\right)\right)\right) \; . \tag{16.24}$$

Since $\beta \leq \gamma_p$ we deduce for each t that

$$\beta^2 \leq \left(1 + \frac{1}{t^p}\right) \mathcal{I}(t)$$

and consequently

$$\frac{\beta^2}{1 + t^p} + \mathcal{I}(t) \geq \beta^2 \; ,$$

so that (16.24) and (16.23) imply that $\mathsf{E}X^2 \leq KN2^{2N} \exp(-N\beta^2)$. □

We now perform a bit of calculus to gather information about γ_p. The results we present do not aim at sharpness, we try simply to give clean statements.

Proposition 16.3.2. *If p is large enough then*

$$\sqrt{2\log 2} \geq \gamma_p \geq \sqrt{2\log 2}(1 - 2^{-p}) \,. \tag{16.25}$$

Lemma 16.3.3. *If p is large enough, the following holds. If $v \geq 2^{-p/2}$, then*

$$\inf_{0 \leq t \leq 1-v} (1 + t^{-p})\mathcal{I}(t) \geq 2\log 2 + \frac{pv}{8} \,. \tag{16.26}$$

Moreover

$$\gamma_p^2 \geq 2\log 2 \left(1 - \frac{2^{-p}}{\log 2}\right) + \mathcal{R}_p \tag{16.27}$$

where $|\mathcal{R}_p| \leq Lp^2 2^{-2p}$.

One should first observe that (16.27) is stronger than the second inequality of (16.25). Indeed, since $\sqrt{1-x} \geq 1 - x/2$, (16.27) implies

$$\gamma_p \geq \sqrt{2\log 2}\left(1 - \frac{2^{-p}}{2\log 2}\right) + \mathcal{R}_p$$

where $\mathcal{R}_p$ is as above, and since $2\log 2 > 1$, for large p this proves (16.25).

Proof. We set $u = 1 - t$. We shall prove that

$$u \geq 2^{-p/2} \Rightarrow (1 + t^{-p})\mathcal{I}(t) \geq 2\log 2 + \frac{pu}{8} \,. \tag{16.28}$$

Therefore if $v \geq 2^{-p/2}$,

$$1 - t = u \geq v \geq 2^{-p/2} \Rightarrow (1 + t^{-p})\mathcal{I}(t) \geq 2\log 2 + \frac{pv}{8} \,, \tag{16.29}$$

which is (16.26). Since $\mathcal{I}(t) \geq t^2/2$, we have

$$2(1 + t^{-p})\mathcal{I}(t) \geq 2t^{-p}\mathcal{I}(t) \geq t^{2-p} \,,$$

and by convexity

$$\begin{aligned} t^{2-p} = (1-u)^{2-p} &\geq 1 + (p-2)u \\ &= 1 + \frac{p}{4}u + \left(\frac{3p}{4} - 2\right)u \,. \end{aligned}$$

If $u \geq 3/p$, this is at least

$$1 + \frac{p}{4}u + \left(\frac{9}{4} - \frac{6}{p}\right) \geq 4\log 2 + \frac{p}{4}u$$

for large p, since $13/4 > 4\log 2$. Consequently it holds that

$$2(1+t^{-p})\mathcal{I}(t) \geq 4\log 2 + \frac{p}{4}u \ ,$$

and to prove (16.28) we may assume $u \leq 3/p$. We observe that, by convexity,

$$(1+t)\log(1+t) = (2-u)\log(2-u) \geq 2\log 2 - (1+\log 2)u$$

so that

$$2\mathcal{I}(t) \geq 2\log 2 - (1+\log 2)u + u\log u \ .$$

By convexity again,

$$1 + t^{-p} = 1 + (1-u)^{-p} \geq 2 + pu$$

and therefore

$$2(1+t^{-p})\mathcal{I}(t) \geq \psi(u) := (2+pu)(2\log 2 - (1+\log 2)u + u\log u) \ . \quad (16.30)$$

Moreover

$$\psi(u) = 4\log 2 + u\Big(2(p-1)\log 2 - 2 + 2\log u - pu(1+\log 2) + pu\log u\Big) \quad (16.31)$$

by straightforward algebra. Since we assume $u \leq 3/p$, we have

$$u\log u \geq \frac{3}{p}\log\frac{3}{p} = -\frac{3}{p}\log\frac{p}{3} \ ,$$

and therefore $pu\log u \geq -3\log(p/3)$. Since $pu \leq 3$ we get

$$\psi(u) \geq 4\log 2 + u\left(2(p-1)\log 2 - 2 + 2\log u - 3(1+\log 2) - 3\log\frac{p}{3}\right) \ .$$

When we assume $u \geq 2^{-p/2}$, we have $2\log u \geq -p\log 2$ so that for a certain number L

$$2(p-1)\log 2 - 2 + 2\log u - 3(1+\log 2) - 3\log\frac{p}{3} \geq p\log 2 - L - 3\log\frac{p}{3} \ .$$

Since $\log 2 > 1/2$, when p is large enough we have $p\log 2 - L - 3\log\frac{p}{3} \geq p/2$ and therefore $\psi(u) \geq 4\log 2 + pu/2$ whenever $u \geq 2^{-p/2}$. This proves (16.28) and hence (16.26).

As a consequence of (16.26), to prove (16.27) it suffices to show that

$$\inf_{0\leq u<2^{-p/2}} 2(1+(1-u)^{-p})\mathcal{I}(1-u) \geq 4\log 2\left(1 - \frac{2^{-p}}{\log 2}\right) + \mathcal{R}_p \ ,$$

and therefore by (16.30) that for $u < 2^{-p/2}$ we have

$$\psi(u) \geq 4\log 2 - 2^{-p+2} + \mathcal{R}_p \ . \quad (16.32)$$

This occupies the rest of the proof. For $u < 2^{-p/2}$ we have $pu \leq p2^{-p/2}$ and $pu\log u \geq p2^{-p/2}\log 2^{-p/2} = -p2^{-p/2}\log 2^{p/2}$, so (16.31) implies

$$\psi(u) \ge 4\log 2 + u\Big(2(p-1)\log 2 - 2 + 2\log u - p2^{-p/2}(1+\log 2) - p2^{-p/2}\log 2^{p/2}\Big) . \tag{16.33}$$

If $u \ge 2^{-p+3}$ we have $2\log u \ge -2(p-3)\log 2$, and since $4\log 2 > 2$, (16.33) shows that

$$\psi(u) \ge 4\log 2 + u(4\log 2 - 2 - p2^{-p/2}(1+\log 2) - p2^{-p/2}\log 2^{p/2}) \ge 4\log 2$$

(for large p) and (16.32) holds. Now, if $u \le 2^{-p-3}$, we have

$$|pu^2 \log u| \ , \ |pu^2| \le Lp^2 2^{-2p}$$

and (16.31) yields

$$\psi(u) \ge \varphi(u) + \mathcal{R}_p$$

with $|\mathcal{R}_p| \le Lp^2 2^{-2p}$ and

$$\varphi(u) = 4\log 2 + 2u((p-1)\log 2 - 1) + 2u\log u \ .$$

Now $\varphi(u)$ has its minimum at $u = 2^{-p+1}$ and $\varphi(2^{-p+1}) = 4\log 2 - 2^{-p+2}$. This completes the proof of (16.32). □

Proposition 16.3.4. *If p is large enough then for $\beta \ge \gamma_p$ we have*

$$\liminf_{N\to\infty} p_N(\beta) \ge \gamma_p\beta \ge \beta\sqrt{2\log 2}(1-2^{-p}) \ . \tag{16.34}$$

Proof. The function $\beta \mapsto p_N(\beta)$ is convex, and Theorem 16.3.1 implies that for $\beta < \gamma_p$ we have $\lim_{N\to\infty} p_N(\beta) = \log 2 + \beta^2/2$ so that by Griffiths' lemma

$$\liminf_{N\to\infty} p'_N(\gamma_p) \ge \gamma_p \ .$$

Using convexity again, for $\beta \ge \gamma_p$,

$$p_N(\beta) \ge p_N(\gamma_p) + (\beta - \gamma_p)p'_N(\gamma_p) \ ,$$

and thus

$$\begin{aligned} \liminf_{N\to\infty} p_N(\beta) &\ge \log 2 + \frac{\gamma_p^2}{2} + \gamma_p(\beta-\gamma_p) \\ &= \log 2 - \frac{\gamma_p^2}{2} + \gamma_p\beta \ge \gamma_p\beta \ . \end{aligned}$$

This proves the first inequality in (16.34). The second one follows from (16.25). □

Definition 16.3.5. *We say that a subset I of $\mathbb{R}$ is negligible if $\mathsf{E}G_N^{\otimes 2}(\{R_{1,2} \in I\}) \le K\exp(-N/K)$, where as usual K denotes a number independent of N.*

The main result of this section is the following.

Theorem 16.3.6. *There exists an integer p_0 such that if $p \geq p_0$, for each β the intervals $[-1, -2^{-p/4}]$ and $[2^{-p/4}, 1-2^{-p/2}]$ are negligible.*

In other words, the overlap is essentially always in the small interval $[-2^{-p/4}, 2^{-p/4}]$ around 0 or in the small interval $[1-2^{-p/2}, 1]$ around 1. There is nothing magic about the numbers $2^{-p/4}$, etc. These are simply convenient choices. The proof relies on the following.

Proposition 16.3.7. *We have*

$$\varphi_N(u) := \frac{1}{N} \mathsf{E} \log \sum_{R_{1,2}=u} \exp(-H_N(\boldsymbol{\sigma}^1) - H_N(\boldsymbol{\sigma}^2)) \leq \xi(u, \beta) + \frac{1}{N} \tag{16.35}$$

where

$$\xi(u, \beta) = \beta^2(1+u^p) + 2\log 2 - \mathcal{I}(|u|) \tag{16.36}$$

if

$$\beta^2(1+u^p) \leq 2\log 2 - \mathcal{I}(|u|) \tag{16.37}$$

and otherwise

$$\xi(u, \beta) = 2\beta\sqrt{(1+u^p)(2\log 2 - \mathcal{I}(|u|))}\ . \tag{16.38}$$

We will then prove that $\xi(u, \beta) < 2 \liminf p_N(\beta)$ for u in the intervals $[-1, -2^{-p/4}]$ and $[2^{-p/4}, 1-2^{-p/2}]$ of Theorem 16.3.6 to conclude as usual the proof of Theorem 16.3.6 using concentration of measure. The bound (16.35) is very primitive compared with the bounds of Section 14.6 but we do not know how to do better.

Proof. There are at most $2^{2N} \exp(-N\mathcal{I}(|u|))$ pairs in the summation (16.35); for each of them

$$\mathsf{E}(H_N(\boldsymbol{\sigma}^1) + H_N(\boldsymbol{\sigma}^2))^2 \leq 2\beta^2 N(1+u^p) + K$$

by (16.16). Therefore Proposition 16.3.7 follows from Lemma A.2.1. □

Let us also point out that one can easily deduce from Lemma A.2.1 that $p_N(\beta) \leq \beta\sqrt{2\log 2} + K/N$, to be compared with (16.34).

Before proving Theorem 16.3.6, we need one more dash of calculus.

Lemma 16.3.8. *For p large enough, if $u \in [-1, -2^{-p/4}]$ or $u \in [2^{-p/4}, 1-2^{-p/2}]$ we have*

$$2u^p \log 2 - (1+u^p)\mathcal{I}(u) \leq -2^{-p/2-2}\ . \tag{16.39}$$

Proof. Assume first that $u < 0$, and observe that $1 + u^p \geq 0$, $\mathcal{I}(u) \geq 0$, so that (16.39) is obvious if $u^p \leq -1/2$. Observe also that $u^p < 0$ because $u < 0$ and p is odd. If $u^p \geq -1/2$, since $\mathcal{I}(u) \geq u^2/2$ we have

$$2u^p \log 2 - (1 + u^p)\mathcal{I}(u) \leq -\frac{1}{2}\mathcal{I}(u) \leq -\frac{u^2}{4} \leq -2^{-p/2-2}$$

when $u \leq -2^{-p/4}$. This proves (16.39) when $u < 0$. When $u > 0$, let us first consider the case where $u^{p-2} \leq 1/8$. Then $2u^p \leq u^2/4$ and, since $\log 2 \leq 1$ and $\mathcal{I}(u) \geq u^2/2$,

$$2u^p \log 2 - (1 + u^p)\mathcal{I}(u) \leq \frac{u^2}{4}\log 2 - \mathcal{I}(u) \leq \frac{u^2}{4} - \frac{u^2}{2} = -\frac{u^2}{4} \leq -2^{-p/2-2}$$

provided $u \geq 2^{-p/4}$. So it suffices to consider the case where $u^{p-2} \geq 1/8$ and $u \leq 1 - 2^{-p/2}$. Using (16.26) with $v = 2^{-p/2}$ and since $u \leq 1 - 2^{-p/2}$, we deduce that

$$(1 + u^p)\mathcal{I}(u) \geq \left(2\log 2 + \frac{p2^{-p/2}}{8}\right) u^p$$

and thus

$$2u^p \log 2 - (1 + u^p)\mathcal{I}(u) \leq u^p \left(-\frac{p2^{-p/2}}{8}\right) .$$

Since we assume that $u \geq 8^{-1/(p-2)}$, for p large enough the right-hand side is $\leq -2^{-p/2-2}$. □

Lemma 16.3.9. *If p is large enough then for $u \in [-1, -2^{-p/4}]$ or $u \in [2^{-p/4}, 1 - 2^{-p/2}]$ the function $\xi(u, \beta)$ of Proposition 16.3.7 satisfies*

$$\beta \geq \gamma_p \Rightarrow \xi(u, \beta) \leq \beta(2\sqrt{2\log 2} - 2^{-p/2-3}) \tag{16.40}$$

$$\beta \leq \gamma_p \Rightarrow \xi(u, \beta) \leq \beta^2 + 2\log 2 - 2^{-p/2}/L . \tag{16.41}$$

Proof. Assume first that $\beta^2(1 + u^p) > 2\log 2 - \mathcal{I}(u)$ so $\xi(u, \beta)$ is given by (16.38). We write

$$\begin{aligned}(1 + u^p)(2\log 2 - \mathcal{I}(u)) &= 2\log 2 + 2u^p \log 2 - (1 + u^p)\mathcal{I}(u) \\ &\leq 2\log 2 - 2^{-p/2-2}\end{aligned}$$

by (16.39). Combining the inequality $\sqrt{1 + x} \leq 1 + x/2$ with (16.38) implies

$$\begin{aligned}\xi(u, \beta) &= 2\beta\sqrt{(1 + u^p)(2\log 2 - \mathcal{I}(u))} \\ &\leq 2\beta\sqrt{2\log 2 - 2^{-p/2-2}} \\ &\leq 2\beta\sqrt{2\log 2}\left(1 - \frac{2^{-p/2-4}}{\log 2}\right) \\ &\leq \beta(2\sqrt{2\log 2} - 2^{-p/2-3}) .\end{aligned}$$

This proves (16.40) when $\beta^2(1+u^p) > 2\log 2 - \mathcal{I}(u)$. In that case, and since $\mathcal{I}(u) \le \log 2$, we have $\beta^2 \ge (\log 2)/2$ and since $2\beta\sqrt{2\log 2} \le \beta^2 + 2\log 2$ we have also proved (16.41).

Assume next that

$$\beta^2(1+u^p) \le 2\log 2 - \mathcal{I}(u)\ , \tag{16.42}$$

so that $\xi(u,\beta)$ is given by (16.36), and then

$$\begin{aligned}\xi(u,\beta) &= \beta^2(1+u^p) + 2\log 2 - \mathcal{I}(u)\\ &\le 4\log 2 - 2\mathcal{I}(u) \le 4\log 2 - u^2 \le 4\log 2 - 2^{-p/2}\ ,\end{aligned}$$

because $|u| \ge 2^{-p/4}$. When $\beta \ge \gamma_p$, for p large, we have, using (16.25) in the last 2 inequalities,

$$\begin{aligned}4\log 2 - 2^{-p/2} &\le \sqrt{2\log 2}(1-2^{-p})(2\sqrt{2\log 2} - 2^{-p/2-2})\\ &\le \gamma_p(2\sqrt{2\log 2} - 2^{-p/2-2})\\ &\le \beta(2\sqrt{2\log 2} - 2^{-p/2-2})\ ,\end{aligned}$$

and this proves (16.40). When $\beta \le \gamma_p$ we write, using (16.42) in the third line and (14.87) in the last line,

$$\begin{aligned}\xi(u,\beta) &= \beta^2(1+u^p) + 2\log 2 - \mathcal{I}(u)\\ &= \beta^2 + 2\log 2 + \beta^2 u^p - \mathcal{I}(u)\\ &\le \beta^2 + 2\log 2 + \frac{u^p}{u^p+1}(2\log 2 - \mathcal{I}(u)) - \mathcal{I}(u)\\ &\le \beta^2 + 2\log 2 + \frac{1}{u^p+1}(2u^p\log 2 - \mathcal{I}(u))\\ &\le \beta^2 + 2\log 2 - 2^{-p/2-3}\ ,\end{aligned}$$

and this completes the proof of (16.41). □

Proof of Theorem 16.4.3. For N the relations (16.17), (16.41), (16.34) and (16.40) imply

$$\xi(u,\beta) \le 2p_N(\beta) - \frac{2^{-p/2}}{L} \tag{16.43}$$

whenever u belongs to one of the intervals $[-1, -2^{-p/4}]$ or $[2^{-p/4}, 1-2^{-p/2}]$, and then using concentration of measure as in Proposition 13.4.3 we deduce that $\mathsf{E}G_N^{\otimes 2}(\{R_{1,2} = u\}) \le K\exp(-K/N)$. □

Inspecting the proof of Theorem 16.4.3 we observe that we have actually proved a stronger statement, namely that if p is large enough, for each β_0 there exists a constant $K(\beta_0)$ depending only on β_0 and p, such that whenever $\beta \le \beta_0$ then

$$\mathsf{E}G_N^{\otimes 2}(\{R_{1,2} \in [-1, -2^{-p/4}]\}) \le K(\beta_0)\exp(-K(\beta_0)/N) \tag{16.44}$$

$$\mathsf{E}G_N^{\otimes 2}(\{R_{1,2} \in [2^{-p/4}, 1-2^{-p/2}]\}) \le K(\beta_0)\exp(-K(\beta_0)/N)\ . \tag{16.45}$$

This will be used in Section 16.7.

Exercise 16.3.10. Consider

$$c_p = \max_{-1 \le u \le 1} (1 + u^p)(2 \log 2 - \mathcal{I}(u)) \ ,$$

so that c_p is slightly larger than $2 \log 2$. Define

$$\gamma'_p = \sqrt{c_p} - \sqrt{c_p - 2 \log 2} \ .$$

Prove that for $\beta \le \gamma'_p$ and every u we have

$$\xi(u, \beta) \le \beta^2 + 2 \log 2 \ ,$$

where $\xi(u, \beta)$ is the function of Proposition 16.3.7. Use the method of proof of Theorem 13.4.1 to show that for $\beta \le \gamma'_p$ one has $\lim_{N \to \infty} p_N(\beta) = \beta^2/2 + \log 2$. How does γ'_p compare to γ_p?

16.4 The Lumps and Their Weights

In this section we prove that the configuration space Σ_N decomposes into a sequence of well separated pieces called the lumps. We then show that the extended Ghirlanda-Guerra inequalities determine the properties of the sequence of the Gibbs weights of these lumps. The following key result is completely deterministic.

Theorem 16.4.1. *Consider a probability measure μ on Σ_N. Consider two numbers a, a', $0 \le a' \le 1/2$, $0 < a < 1/28$ and set*

$$\varepsilon = \mu^{\otimes 2}(\{(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) \ ; \ |R_{1,2}| \ge a' \ , \ R_{1,2} \le 1 - 2a\}) \ . \tag{16.46}$$

Then we can find a partition $(C_\alpha)_{\alpha \ge 1}$ of Σ_N with the following properties:

$$\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2 \in C_\alpha \Rightarrow R_{1,2} \ge 1 - 12a \tag{16.47}$$

$$\mu^{\otimes 2}\left(\{(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) \ ; \ |R_{1,2}| \ge a'\} \setminus \bigcup_{\alpha \ge 1} C_\alpha^2 \right) \le 3 \, \varepsilon^{1/3} \ . \tag{16.48}$$

This will be used for ε exponentially small in N. Of course at most 2^N of the sets $C_\alpha \subset \Sigma_N$ are non-empty.

Theorem 16.4.1 means essentially the following: basically the only way that $|R_{1,2}| > a'$ (or, which is essentially equivalent, that $R_{1,2} \ge 1 - 2a$) is that $\boldsymbol{\sigma}^1$ and $\boldsymbol{\sigma}^2$ lie in the same set C_α. Conversely, if $\boldsymbol{\sigma}^1$ and $\boldsymbol{\sigma}^2$ belong to C_α, then $R_{1,2}$ is close to 1. Consequently,

$$\mu^{\otimes 2}(\{|R_{1,2}| \ge 1 - 2a\}) \simeq \mu^{\otimes 2}(\{|R_{1,2}| \ge a'\}) \simeq \sum_{\alpha \ge 1} \mu(C_\alpha)^2 \ . \tag{16.49}$$

It is *not* a consequence of the hypothesis of Theorem 16.4.1 that the quantity (16.49) is not very small. For example, the uniform probability on Σ_N satisfies

$$\mu^{\otimes 2}(\{|R_{1,2}| \geq t\}) \leq 2\exp\left(-\frac{Nt^2}{2}\right)$$

by (A.16). But, when the quantity (16.49) is not too small, some of the small sets (C_α) must have a significant mass. For this reason we will call the sets C_α the *lumps* (even though it might happen under the hypothesis of Theorem 16.4.1 that they all have a very small mass).

Theorem 16.4.1 is close in spirit to Theorem 15.4.4, but it is a bit different (and much easier). It has a version for probability measures on the unit ball of the Hilbert space (with the same proof). The reason for which we stated it for Σ_N is simply that this is the setting we need, and that the Hilbert space structure is not used in the proof (in contrast with the situation of Theorem 15.4.4).

Proof. We recall that

$$d(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) = \frac{1}{N}\mathrm{card}(\{i \leq N\,;\ \sigma_i^1 \neq \sigma_i^2\}) = \frac{1}{2}\,(1 - R_{1,2}) \tag{16.50}$$

is a distance on Σ_N. We denote by $B(\boldsymbol{\sigma}, r)$ the ball of radius r and center $\boldsymbol{\sigma}$ for this distance. We set

$$B_0(\boldsymbol{\sigma}) = B(\boldsymbol{\sigma}, a)\,;\ B(\boldsymbol{\sigma}) = B(\boldsymbol{\sigma}, 3a)\,;\ D(\boldsymbol{\sigma}) = B(\boldsymbol{\sigma}, 6a)\backslash B(\boldsymbol{\sigma}, 2a)\ . \tag{16.51}$$

We observe that

$$\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2 \in B(\boldsymbol{\sigma}) \Rightarrow d\,(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) \leq 6a \Rightarrow R_{1,2} \geq 1 - 12a\ . \tag{16.52}$$

We will say that a set $B(\boldsymbol{\sigma})$ is a *lump* if $\mu(B_0(\boldsymbol{\sigma})) \geq \varepsilon^{1/3}$. We consider a maximal disjoint family $(C_\alpha)_{\alpha \leq M}$ of lumps (this family might well be empty).

We will prove that (16.48) holds when $\bigcup_{\alpha \geq 1} C_\alpha$ is replaced by $\bigcup_{\alpha \leq M} C_\alpha$. By (16.52), the sets C_α satisfy (16.47), but they do not form a partition of Σ_N. To obtain a partition of Σ_N it suffices to decompose $\Sigma_N \backslash \bigcup_{\alpha \leq M} C_\alpha$ into sets consisting of a single point.

To start, we prove that

$$B(\boldsymbol{\sigma}) \text{ is a lump } \Rightarrow \mu(D(\boldsymbol{\sigma})) \leq \varepsilon^{2/3}\ . \tag{16.53}$$

This is because if $\boldsymbol{\sigma}^1 \in B_0(\boldsymbol{\sigma})$ and $\boldsymbol{\sigma}^2 \in D(\boldsymbol{\sigma})$, then $a \leq d\,(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) \leq 7a$, so that $1 - 14a \leq R_{1,2} \leq 1 - 2a$, and since $a < 1/28$ and $a' \leq 1/2$ we have

$$B_0(\boldsymbol{\sigma}) \times D(\boldsymbol{\sigma}) \subset \{a' < R_{1,2} \leq 1 - 2a\}\ ,$$

so that by (16.46)

$$\varepsilon^{1/3}\mu(D(\boldsymbol{\sigma})) \leq \mu(B_0(\boldsymbol{\sigma}))\,\mu(D(\boldsymbol{\sigma})) = \mu^{\otimes 2}(B_0(\boldsymbol{\sigma}) \times D(\boldsymbol{\sigma})) \leq \varepsilon\ .$$

This proves (16.53).

Since the sets C_α are disjoint and since $\mu(C_\alpha) \geq \varepsilon^{1/3}$ by definition of a lump, we have $M \leq \varepsilon^{-1/3}$. For $\alpha \leq M$, we consider $\boldsymbol{\sigma}_\alpha$ such that $C_\alpha = B(\boldsymbol{\sigma}_\alpha)$, and we set $D = \bigcup_{\alpha \leq M} D(\boldsymbol{\sigma}_\alpha)$. By (16.53) for each α we have $\mu(D(\boldsymbol{\sigma}_\alpha)) \leq \varepsilon^{2/3}$, so that $\mu(D) \leq \varepsilon^{1/3}$. Consider the set

$$A = \{(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2)\,;\ \boldsymbol{\sigma}^2 \in B_0(\boldsymbol{\sigma}^1), \mu(B_0(\boldsymbol{\sigma}^1)) \leq \varepsilon^{1/3}\}\ .$$

Then, by Fubini theorem, we have $\mu^{\otimes 2}(A) \leq \varepsilon^{1/3}$. To prove (16.48), keeping (16.46) in mind, it suffices to prove that

$$R_{1,2} \geq 1 - 2a \Rightarrow (\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) \in A \cup (D \times \Sigma_N) \cup \bigcup_{\alpha \leq M} C_\alpha^2\ .$$

If $R_{1,2} \geq 1 - 2a$, then $d(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) \leq a$, so $\boldsymbol{\sigma}^2 \in B_0(\boldsymbol{\sigma}^1)$. If $\mu(B_0(\boldsymbol{\sigma}^1)) \leq \varepsilon^{1/3}$, then $(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) \in A$ and we are done. Thus we may assume that $\mu(B_0(\boldsymbol{\sigma}^1)) \geq \varepsilon^{1/3}$ and then $B(\boldsymbol{\sigma}^1)$ is a lump. Since the family $(C_\alpha)_{\alpha \leq M}$ is a maximum disjoint family of lumps, the lump $B(\boldsymbol{\sigma}^1)$ is not disjoint from this family. That is, there exists $\alpha \leq M$ with $B(\boldsymbol{\sigma}^1) \cap B(\boldsymbol{\sigma}_\alpha) \neq \emptyset$, so that $d(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}_\alpha) \leq 6a$. If $\boldsymbol{\sigma}^1 \in D(\boldsymbol{\sigma}_{\boldsymbol{\alpha}})$ we are done because then $(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) \in D \times \Sigma_N$. Thus we may assume that $\boldsymbol{\sigma}^1 \notin D(\boldsymbol{\sigma}_{\boldsymbol{\alpha}})$. Now $D(\boldsymbol{\sigma}_{\boldsymbol{\alpha}}) = \{\boldsymbol{\sigma}\,;\ 2a \leq d(\boldsymbol{\sigma}_\alpha, \boldsymbol{\sigma}) \leq 6a\}$, so that since $d(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}_\alpha) \leq 6a$ we have $d(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}_\alpha) \leq 2a$ and, since $d(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) \leq a$, we have $d(\boldsymbol{\sigma}^2, \boldsymbol{\sigma}_\alpha) \leq 3a$. This proves that $(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) \in C_\alpha^2$. □

Theorem 16.4.2. *There exists a number p_0 with the following property. If $p \geq p_0$ is odd, then we may find a partition $(C_\alpha)_{\alpha \geq 1}$ of Σ_N (depending on the disorder) such that*

$$\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2 \in C_\alpha \Rightarrow R_{1,2} \geq 1 - 2^{-p/2+4} \tag{16.54}$$

$$\mathsf{E}\, G_N^{\otimes 2}\left(\{|R_{1,2}| \geq 2^{-p/4}\} \setminus \bigcup_{\alpha \geq 1} C_\alpha^2\right) \leq K \exp\left(-\frac{N}{K}\right)\ . \tag{16.55}$$

Proof. We use Theorem 16.4.1 at given disorder with $a' = 2^{-p/4}$ and $a = 2^{-p/2}$ and Theorem 16.3.6. □

Let us stress again that this contains *no* information when $\mathsf{E}\, G_N^{\otimes 2}(\{|R_{1,2}| \geq 2^{-p/4}\}) \leq K \exp(-N/K)$.

The sets $(C_\alpha)_{\alpha \geq 1}$ will be called the lumps. They are well defined for every realization of the randomness, and depend on this randomness.

It can be shown that the lumps are all of very small Gibbs' measure when $\beta \leq \gamma_p$. On the other hand there is a simple observation showing that this is not the case for large β. Indeed, by Lemma A.2.1 we have

$$\frac{1}{N} \mathsf{E} \max_{\boldsymbol{\sigma}}(-H_N(\boldsymbol{\sigma})) \leq \beta\sqrt{2\log 2}$$

because $\mathsf{E}H_N^2(\boldsymbol{\sigma}) \le \beta^2 N$ and since there are 2^N values of $\boldsymbol{\sigma}$. Now

$$p_N'(\beta) = \frac{1}{N}\nu\biggl(\Bigl(\frac{p!}{N^{p-1}}\Bigr)^{1/2} \sum_{i_1<\ldots<i_p} g_{i_1\ldots i_p}\sigma_{i_1}\cdots\sigma_{i_p}\biggr)$$
$$= \frac{1}{N}\nu\Bigl(\frac{-H_N(\boldsymbol{\sigma})}{\beta}\Bigr) \le \frac{1}{\beta N}\mathsf{E}\max_{\boldsymbol{\sigma}}(-H_N(\boldsymbol{\sigma})) \le \sqrt{2\log 2}\ .$$

On the other hand integration by parts yields

$$p_N'(\beta) \ge \beta(1-\nu(R_{1,2}^p)) - \frac{K}{N} \tag{16.56}$$

so that

$$\liminf_{N\to\infty} \nu(R_{1,2}^p) \ge 1 - \frac{\sqrt{2\log 2}}{\beta}$$

and for large β some part of $\nu(R_{1,2}^p)$ must come from values of $R_{1,2}$ that are $\ge 2^{-p/4}$ (and hence from values of $R_{1,2}$ that are close to 1).

In the remainder of the present chapter, we add a perturbation term $H_{N,\boldsymbol{\beta}}^{per}$ of the type (12.32) to the Hamiltonian. Of course this term is probabilistically independent of H_N. It is of smaller order, as is shown by Lemma 12.2.1. Consequently Theorem 16.3.6 remains valid for the Hamiltonian $H_{N,\boldsymbol{\beta}} = H_N + H_{N,\boldsymbol{\beta}}^{per}$ because the perturbation term changes both sides of (16.43) by quantity that vanishes as $N\to\infty$.

The perturbation term $H_{N,\boldsymbol{\beta}}^{per}$ depends on the parameter $\boldsymbol{\beta} = (\beta_s)_{s\ge1}$. Throughout the rest of the chapter, we denote by δ a quantity (depending on N and $\boldsymbol{\beta}$) such that

$$\lim_{N\to\infty}\int |\delta| \mathrm{d}\boldsymbol{\beta} = 0\ , \tag{16.57}$$

where $\int \mathrm{d}\boldsymbol{\beta}$ means that each β_s is integrated from 0 to 1.

Let us denote by

$(w_\alpha)_{\alpha\ge1}$ the sequence of Gibbs' weights of the lumps of the system,

as constructed by Theorem 16.4.2. We will always assume that the lumps are numbered in non-increasing order of their Gibbs' measure. Let us define

$$m = m_N(\beta,\boldsymbol{\beta}) = \nu(\mathbf{1}_{\{R_{1,2}\le 1/2\}})\ . \tag{16.58}$$

Let us recall the numbers $S^{(m)}(n_1,\ldots,n_k)$ of (16.2), so that $S^{(m)}(2) = 1-m$.

Theorem 16.4.3. *Given β and any integers $k, n_1,\ldots,n_k$, we have*

$$\mathsf{E}\prod_{s\le k}\sum_{\alpha\ge1} w_\alpha^{n_s} = S^{(m)}(n_1,\ldots,n_k) + \delta\ . \tag{16.59}$$

Since, as explained in Section 16.1, the quantities in the left-hand side of (16.59) determine the distribution of the sequence (w_α), this result can be interpreted as saying that "in the limit, the distribution of the sequence (w_α) is the Poisson-Dirichlet distribution Λ_m where m is given by (16.58)". One must be cautious however because $m = m(N, \boldsymbol{\beta})$ depends on N and $\boldsymbol{\beta}$ and we have not proved that m has a limit as $N \to \infty$; even the possibility that $m = m_N(\beta, \boldsymbol{\beta})$ varies widely with $\boldsymbol{\beta}$ remains open.

Proof. Let us use the notation

$$S(n_1, \ldots, n_k) = \mathsf{E} \prod_{s \le k} \sum_{\alpha \ge 1} w_\alpha^{n_s} .$$

It follows from Theorem 16.4.2 that

$$\nu(\mathbf{1}_{\{R_{1,2} > 1/2\}}) = \mathsf{E} \sum_{\alpha \ge 1} w_\alpha^2 + \delta$$

where δ satisfies (16.57) (and is even exponentially small in N). Thus since $S^{(m)}(2) = 1 - m$, (16.58) implies

$$S(2) = \mathsf{E} \sum_{\alpha \ge 1} w_\alpha^2 = 1 - m + \delta = S^{(m)}(2) + \delta .$$

To prove (12.54) we shall argue that, for each integers $k, n_1, \ldots, n_k$, we have

$$\begin{aligned} S(n_1 + 1, \ldots, n_k) = {} & \frac{n_1 - m}{n} S(n_1, \ldots, n_k) \\ & + \sum_{2 \le s \le k} \frac{n_s}{n} S(n_2, \ldots, n_{s-1}, n_s + n_1, n_{s+1}, \ldots, n_k) + \delta . \end{aligned} \tag{16.60}$$

Since these relations are the same as the relations (15.19), and since, as explained after (15.19), these relations determine the numbers $S(n_1, \ldots, n_k)$, this will prove (16.59).

The proof of (16.60) relies on (12.28). The argument is a very much simplified version of the proof of (15.54). Consider a continuous function ψ with $\psi(x) = 0$ for $x \le 1/2$ and $\psi(x) = 1$ for $x \ge 3/4$. Then if $|f| \le 1$ is a function on Σ_N, and since it essentially never happens that $R_{1,2} \in [1/2, 3/4]$ we have

$$\nu(\psi(R_{\ell,\ell'}) f) = \nu(\mathbf{1}_{\{R_{\ell,\ell'} \ge 3/4\}} f) + \delta ,$$

and also $\nu(\mathbf{1}_{\{R_{1,2} \ge 3/4\}}) = 1 - m + \delta$. Using these in (12.28) yields

$$\nu(\mathbf{1}_{\{R_{1,n+1} \ge 3/4\}} f) = \frac{1 - m}{n} \nu(f) + \frac{1}{n} \sum_{2 \le \ell \le n} \nu(\mathbf{1}_{\{R_{1,\ell} \ge 3/4\}} f) + \delta . \tag{16.61}$$

Consider now disjoint sets $I_1, \ldots, I_k$ of $\{1, \ldots, n\}$, where $\mathrm{card} I_s = n_s$ and $n = \sum_{s \le k} n_s$; assume that $1 \in I_1$, and consider the function f on Σ_N^n given by $f(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n) = 1$ if for each $s \leq k$, and each $\ell, \ell' \in I_s$ we have $R_{\ell,\ell'} \geq 3/4$, and by $f(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n) = 0$ otherwise. It is essentially true that

$$R_{\ell,\ell'} \geq \frac{3}{4} \quad \Leftrightarrow \quad \exists \alpha\,, \quad \boldsymbol{\sigma}^\ell, \boldsymbol{\sigma}^{\ell'} \in C_\alpha$$

and this should make it obvious that

$$\nu(f) \simeq S(n_1, \ldots, n_k)$$

within exponentially small error. Proceeding similarly for the other terms of (16.61) we obtain (16.60). □

16.5 One Step of Replica-Symmetry Breaking

In this section we prove that the overlaps take only two possible values. We recall the notation δ of (16.57).

Theorem 16.5.1. *There exists a number p_0 such that if $p \geq p_0$ and $\gamma_p \leq \beta \leq p^{-2}2^{p/12}$, there exists a number $q = q_N(\beta, \boldsymbol{\beta})$ such that for any $\varepsilon > 0$ we have*

$$\nu(\mathbf{1}_{\{|R_{1,2}| \geq \varepsilon,\, |R_{1,2} - q| \geq \varepsilon\}}) \leq \delta\,. \tag{16.62}$$

This means that for the typical choice of $\boldsymbol{\sigma}^1$ and $\boldsymbol{\sigma}^2$, we have either $R_{1,2} \simeq 0$ or else $R_{1,2} \simeq q$. Of course one would like to prove that (16.62) holds for a number q independent of N and $\boldsymbol{\beta}$, but unfortunately this is not part of the statement of the theorem. The condition $\beta \leq p^{-2}2^{p/12}$ is simply a convenient choice, showing that we may take β exponentially large in p. We made no effort to get a sharp condition, since in any case our methods do not lend themselves to this. For example, our proof yields in fact that the condition $\beta \leq p^{-1}2^{p/12}/L$ suffices, but we see no reason why the exponent $1/12$ should be anywhere close to optimal. It is conjectured however that, given p, if β is large enough then (16.62) fails.

We will for the moment maintain the suspense on how to find q, revealing only that $q \geq 3/4$. We set

$$q_{1,2} = \begin{cases} q \text{ if } R_{1,2} \geq \dfrac{1}{2}\,, \\[2ex] 0 \text{ if } R_{1,2} < \dfrac{1}{2}\,. \end{cases} \tag{16.63}$$

To prove (16.62) it suffices to show that

$$\nu\big((R_{1,2} - q_{1,2})^2\big) = \delta \tag{16.64}$$

because since $q \geq 3/4$ when $\varepsilon < 1/4$ we have $(R_{1,2} - q_{1,2})^2 \geq \varepsilon^2$ when $|R_{1,2}| \geq \varepsilon$ and $|R_{1,2} - q| \geq \varepsilon$. Expanding the square

$$(R_{1,2} - q_{1,2})^2 = \left(\frac{1}{N}\sum_{i\leq N}(\sigma_i^1\sigma_i^2 - q_{1,2})\right)^2$$

and using symmetry between sites, we get

$$\begin{aligned}\nu\big((R_{1,2} - q_{1,2})^2\big) = &\left(1 - \frac{1}{N}\right)\nu((\sigma_N^1\sigma_N^2 - q_{1,2})(\sigma_{N-1}^1\sigma_{N-1}^2 - q_{1,2})) \\ &+ \frac{1}{N}\nu\big((\sigma_N^1\sigma_N^2 - q_{1,2})^2\big)\ .\end{aligned}$$

Using the notation

$$\varepsilon_\ell = \sigma_N^\ell\ , \qquad \eta_\ell = \sigma_{N-1}^\ell\ ,$$

we therefore have

$$\nu\big((R_{1,2} - q_{1,2})^2\big) \leq \nu((\varepsilon_1\varepsilon_2 - q_{1,2})(\eta_1\eta_2 - q_{1,2})) + \frac{L}{N}\ . \tag{16.65}$$

We will use a kind of cavity method to study this quantity. As already hinted by (16.65), we will distinguish the last two spins, rather than only the last spin as we did most of the time. To do this, we have to make explicit the contribution of these spins to the Hamiltonian. This is the purpose of the following identity, that spells out the dependence of H_N on σ_N and σ_{N-1}:

$$-H_N(\boldsymbol{\sigma}) = -H_{N-2}(\boldsymbol{\rho}) + \sigma_N g_0(\boldsymbol{\rho}) + \sigma_{N-1} g_0^*(\boldsymbol{\rho}) + \sigma_N\sigma_{N-1} g_0^{\sim}(\boldsymbol{\rho}) \tag{16.66}$$

where $\boldsymbol{\rho} = (\sigma_1, \dots, \sigma_{N-2})$,

$$H_{N-2}(\boldsymbol{\rho}) = \beta\left(\frac{p!}{N^{p-1}}\right)^{1/2} \sum_{i_1<\dots<i_p\leq N-2} g_{i_1\dots i_p}\sigma_{i_1}\cdots\sigma_{i_p}$$

$$g_0(\boldsymbol{\rho}) = \beta\left(\frac{p!}{N^{p-1}}\right)^{1/2} \sum_{i_1<\dots<i_{p-1}\leq N-2} g_{i_1\dots i_{p-1}N}\sigma_{i_1}\cdots\sigma_{i_{p-1}}$$

$$g_0^*(\boldsymbol{\rho}) = \beta\left(\frac{p!}{N^{p-1}}\right)^{1/2} \sum_{i_1<\dots<i_{p-1}\leq N-2} g_{i_1\dots i_{p-1}N-1}\sigma_{i_1}\cdots\sigma_{i_{p-1}}$$

$$g_0^{\sim}(\boldsymbol{\rho}) = \beta\left(\frac{p!}{N^{p-1}}\right)^{1/2} \sum_{i_1<\dots<i_{p-2}\leq N-2} g_{i_1\dots i_{p-2}N-1N}\sigma_{i_1}\cdots\sigma_{i_{p-2}}\ .$$

Thus $H_{N-2}(\boldsymbol{\rho})$ is the Hamiltonian of an $(N-2)$-spin system, except that we have replaced β by β' such that

$$\frac{\beta'}{(N-2)^{(p-1)/2}} = \frac{\beta}{N^{(p-1)/2}}\ . \tag{16.67}$$

Let us define

$$R^-_{1,2} = R^-(\boldsymbol{\rho}^1, \boldsymbol{\rho}^2) = \frac{1}{N} \sum_{i \le N-2} \sigma_i^2 \sigma_i^2 \; .$$

Exactly as in (16.16) we show that

$$|\mathsf{E} g_0(\boldsymbol{\rho}^1) g_0(\boldsymbol{\rho}^2) - p\beta^2 (R^-_{1,2})^{p-1}| \le \frac{K}{N} \tag{16.68}$$

$$|\mathsf{E} g_0^*(\boldsymbol{\rho}^1) g_0^*(\boldsymbol{\rho}^2) - p\beta^2 (R^-_{1,2})^{p-1}| \le \frac{K}{N} \tag{16.69}$$

$$|\mathsf{E} g_0^\sim(\boldsymbol{\rho})^2| \le \frac{K}{N} \; . \tag{16.70}$$

In particular (16.70) implies as expected that the term $g_0^\sim(\boldsymbol{\rho})$ is a lower order term.

We need a formula similar to (16.66) for the Hamiltonian $H_{N,\boldsymbol{\beta}} = H_N + H^{per}_{N,\boldsymbol{\beta}}$. The perturbation term is "a combination of s-spin interaction models". For each of them we proceed as in (16.66), and combining the results we get an expression

$$-H_{N,\boldsymbol{\beta}}(\boldsymbol{\sigma}) = -H_{N-2,\boldsymbol{\beta}'}(\boldsymbol{\rho}) + \sigma_N g(\boldsymbol{\rho}) + \sigma_{N-1} g^*(\boldsymbol{\rho}) + \sigma_N \sigma_{N-1} g^\sim(\boldsymbol{\rho}) \; . \tag{16.71}$$

In this formula, $\boldsymbol{\beta}' = (\beta'_s)_{s \ge 1}$ where

$$\frac{\beta'_s}{(N-2)^{(s-1)/2}} = \frac{\beta_s}{N^{(s-1)/2}} \; ,$$

and the jointly Gaussian families $g(\boldsymbol{\rho})$, $g^*(\boldsymbol{\rho})$ and $g^\sim(\boldsymbol{\rho})$ of r.v.s are independent of each other and of the randomness of $-H_{N-2,\boldsymbol{\beta}'}(\boldsymbol{\rho})$. The difference between this formula and (16.66) is that $g(\boldsymbol{\rho})$ and $g^*(\boldsymbol{\rho})$ (as opposed to $g_0(\boldsymbol{\rho})$ and $g_0^*(\boldsymbol{\rho})$) now also account for terms created by the perturbation terms. Since these all have a coefficient $c_N = N^{-1/8}$, we have

$$|\mathsf{E} g(\boldsymbol{\rho}^1) g(\boldsymbol{\rho}^2) - \mathsf{E} g_0(\boldsymbol{\rho}^1) g_0(\boldsymbol{\rho}^2)| \le K c_N^2$$

$$|\mathsf{E} g^*(\boldsymbol{\rho}^1) g^*(\boldsymbol{\rho}^2) - \mathsf{E} g_0^*(\boldsymbol{\rho}^1) g_0^*(\boldsymbol{\rho}^2)| \le K c_N^2$$

$$|\mathsf{E} g^\sim(\boldsymbol{\rho})^2 - \mathsf{E} g_0^\sim(\boldsymbol{\rho})^2| \le \frac{K}{N} c_N^2 \; .$$

Combining with (16.68) and (16.70) entails

$$|\mathsf{E} g(\boldsymbol{\rho}^1) g(\boldsymbol{\rho}^2) - p\beta^2 (R^-_{1,2})^{p-1}| \le K c_N^2 \tag{16.72}$$

$$|\mathsf{E} g^*(\boldsymbol{\rho}^1) g^*(\boldsymbol{\rho}^2) - p\beta^2 (R^-_{1,2})^{p-1}| \le K c_N^2 \tag{16.73}$$

$$|\mathsf{E} g^\sim(\boldsymbol{\rho})^2| \le \frac{K}{N} \; . \tag{16.74}$$

We will now define a suitable interpolating Hamiltonian. The remarkable feature of the Hamiltonian is that it depends in a crucial way on the Gibbs measure relative to the Hamiltonian $H_{N-2,\boldsymbol{\beta}'}$ on Σ_{N-2} that occurs in (16.71). We consider the partition $(C_\alpha)_{\alpha\geq 1}$ of Σ_{N-2} that is obtained by application of Theorem 16.4.2 to the $(N-2)$-spin system with Hamiltonian $H_{N-2,\boldsymbol{\beta}'}$. This partition of course depends on the randomness of this Hamiltonian. Consider two independent sequences (z_α), (z^*_α) of standard Gaussian r.v.s, that are independent of all other sources of randomness. Given $\boldsymbol{\rho} \in \Sigma_{N-2}$, consider the unique α such that $\boldsymbol{\rho} \in C_\alpha$ and define

$$Y(\boldsymbol{\rho}) = Y_\alpha := \beta z_\alpha \sqrt{pq^{p-1}}$$

(where $q = q_N(\beta, \boldsymbol{\beta})$ will be chosen later), and define similarly $Y^*(\boldsymbol{\rho}) = Y^*_\alpha := \beta z^*_\alpha \sqrt{pq^{p-1}}$. We further define the Hamiltonian

$$\begin{aligned} H_t(\boldsymbol{\sigma}) &= H_{N-2,\boldsymbol{\beta}'}(\boldsymbol{\rho}) + \sigma_N(\sqrt{t}g(\boldsymbol{\rho}) + \sqrt{1-t}Y(\boldsymbol{\rho})) \\ &\quad + \sigma_{N-1}(\sqrt{t}g^*(\boldsymbol{\rho}) + \sqrt{1-t}Y^*(\boldsymbol{\rho})) + \sqrt{t}\sigma_N\sigma_{N-1}g^{\sim}(\boldsymbol{\rho}) \ . \end{aligned} \tag{16.75}$$

We denote by $\langle\cdot\rangle_t$ an average for the Gibbs measure with Hamiltonian H_t, and we write as usual $\nu_t(f) = \mathsf{E}\langle f\rangle_t$. We can expect from this construction that ν_0 is a rather non-trivial object.

We consider a number q with $|q| \leq 1$, that might depend on N, β and $\boldsymbol{\beta}$. We define $q_{1,2}$ by (16.63). Next, we consider the number m^- that is to the $(N-2)$-spin system what the number m of (16.58) is to the N-spin system. That is, if

$$\langle\cdot\rangle_- \tag{16.76}$$

denotes an average for the Hamiltonian $H_{N-2,\boldsymbol{\beta}'}$ on Σ_{N-2}, then

$$m^- = \mathsf{E}\langle \mathbf{1}_{\{\widehat{R}_{1,2}\leq 1/2\}}\rangle_- \ ,$$

where

$$\widehat{R}_{1,2} = \frac{1}{N-2}\sum_{i\leq N-2} \sigma^1_i\sigma^2_i \ . \tag{16.77}$$

It is very much unimportant whether we put a factor N or $N-2$ in the denominator in (16.77). A more touchy question is that we do not know well how to relate m and m^-.

Proposition 16.5.2. *We have*

$$\nu_0((\varepsilon_1\varepsilon_2 - q_{1,2})(\eta_1\eta_2 - q_{1,2})) = (1-m^-)\left(q - \frac{\mathsf{E}\mathrm{th}^2 Y \mathrm{ch}^{m^-} Y}{\mathsf{E}\mathrm{ch}^{m^-} Y}\right)^2 + \delta \tag{16.78}$$

where $Y = \beta z\sqrt{pq^{p-1}}$.

This formula reveals that it would be a good idea that

$$q = \frac{\mathsf{E}\mathrm{th}^2 Y \mathrm{ch}^{m^-} Y}{\mathsf{E}\mathrm{ch}^{m^-} Y} \ . \tag{16.79}$$

This equation has always a root, since the right-hand side, seen as function of q, is a continuous map from $[0,1]$ to itself. For a technical reason however, we must find a solution q close to 1.

Lemma 16.5.3. *If p is large enough, and if $\gamma_p \le \beta \le p^{-2}2^{p/2}$, then for N large enough the equation (16.79) has a solution q with*

$$1 - 2^{-p/2} \le q \le 1 \ . \tag{16.80}$$

Any such solution will satisfy (16.62). This solution might depend on N, β and $\boldsymbol{\beta}$ through m^-.

The proof of Lemma 16.5.3 is tedious, so we delay it until the end of the present section.

For a function f on Σ_N^2, we have the algebraic identity

$$\langle f \rangle_0 = \frac{1}{Z_0^2} \left\langle \mathrm{Av} f \exp\left(\sum_{\ell \le 2} \big(\varepsilon_\ell Y(\boldsymbol{\rho}^\ell) + \eta_\ell Y^*(\boldsymbol{\rho}^\ell) \big) \right) \right\rangle_- \tag{16.81}$$

holds, where

$$Z_0 = \langle \mathrm{Av} \exp(\varepsilon Y(\boldsymbol{\rho}) + \eta Y^*(\boldsymbol{\rho})) \rangle_- \ ,$$

and where Av denotes average over $\varepsilon, \eta, \varepsilon_\ell, \eta_\ell = \pm 1$. To compute Z_0, we denote by w_α the mass of C_α, i.e. $w_\alpha = \langle \mathbf{1}_{C_\alpha} \rangle_-$, and since $Y(\boldsymbol{\rho}) = Y_\alpha$ for $\boldsymbol{\rho} \in C_\alpha$ (and similarly for $Y^*(\boldsymbol{\rho})$), we get

$$Z_0 = \langle \mathrm{ch} Y(\boldsymbol{\rho}) \mathrm{ch} Y^*(\boldsymbol{\rho}) \rangle_- = \sum_{\alpha \ge 1} w_\alpha \mathrm{ch} Y_\alpha \mathrm{ch} Y_\alpha^* \ . \tag{16.82}$$

Definition 16.5.4. *We say that two functions f_1 and f_2 on Σ_N^2 are* essentially equal *and we write $f_1 \simeq f_2$ if the set A where these two functions are not equal satisfies*

$$\nu_0(\mathbf{1}_A) \le K \exp\left(-\frac{N}{K} \right) \ . \tag{16.83}$$

We consider the set

$$B = \bigcup_{\alpha \ge 1} C_\alpha^2 \subset \Sigma_{N-2}^2 \ . \tag{16.84}$$

The function $\mathbf{1}_B$ will be considered either as a function on Σ_N^2 or on Σ_{N-2}^2.

Lemma 16.5.5. *We have*

$$q_{1,2} \simeq q \mathbf{1}_B \ . \tag{16.85}$$

Proof. Recalling (16.77) let us set

$$A = \{\widehat{R}_{1,2} \geq 2^{-p/4}\} \setminus B \ ,$$

so that (16.55), used for the $(N-2)$-spin system means that

$$\mathsf{E}\langle \mathbf{1}_A \rangle_- \leq K \exp\left(-\frac{N}{K}\right) . \tag{16.86}$$

Let us now use (16.81) for the function $f = \mathbf{1}_A$, seen as a function on Σ_N^2. Since $Z_0 \geq 1$ we have

$$\nu_0(\mathbf{1}_A) = \mathsf{E}\langle \mathbf{1}_A \rangle_0 \leq \mathsf{E}\langle \mathbf{1}_A \mathrm{ch} Y(\boldsymbol{\rho}^1) \mathrm{ch} Y(\boldsymbol{\rho}^2) \rangle_- \leq K \mathsf{E}\langle \mathbf{1}_A \rangle_- \leq K \exp\left(-\frac{N}{K}\right) , \tag{16.87}$$

as is seen by taking first expectation in the r.v.s z_α and z_α^* and then using (16.86).

Let us define

$$B' = \{R_{1,2} \geq 1/2\} \ ,$$

so that $q_{1,2} = q\mathbf{1}_{B'}$. Using (16.54) for the $(N-2)$-spin system,

$$B \subset \{\widehat{R}_{1,2} \geq 1 - 2^{-p/2+4}\} \ .$$

For $N \geq 10$ and p large we have

$$R_{1,2} \geq 1 - 2^{-p/2+4} \Rightarrow R_{1,2} \geq \frac{1}{2} \Rightarrow \widehat{R}_{1,2} \geq 2^{-p/4} \ ,$$

and thus

$$B \subset \{\widehat{R}_{1,2} \geq 1 - 2^{-p/2+4}\} \subset B' \subset \{\widehat{R}_{1,2} \geq 2^{-p/4}\} \ .$$

Consequently,

$$B \bigtriangleup B' = B' \setminus B \subset A \ ,$$

and therefore

$$|q_{1,2} - q\mathbf{1}_B| = q|\mathbf{1}_{B'} - \mathbf{1}_B| \leq q\mathbf{1}_A \ ,$$

and the conclusion follows from (16.87). □

Proof of Proposition 16.5.2. First we use Lemma 16.5.5 to see that

$$\nu_0((\varepsilon_1\varepsilon_2 - q_{1,2})(\eta_1\eta_2 - q_{1,2})) = \nu_0((\varepsilon_1\varepsilon_2 - q\mathbf{1}_B)(\eta_1\eta_2 - q\mathbf{1}_B)) + \delta \ .$$

Therefore

$$\nu_0((\varepsilon_1\varepsilon_2 - q_{1,2})(\eta_1\eta_2 - q_{1,2})) = \mathrm{I} + \mathrm{II} + \mathrm{III} + \mathrm{IV} + \delta \ , \tag{16.88}$$

where

$$\begin{aligned}
\mathrm{I} &= \mathsf{E}\langle\varepsilon_1\varepsilon_2\eta_1\eta_2\rangle_0 \\
\mathrm{II} &= -q\mathsf{E}\langle\varepsilon_1\varepsilon_2\mathbf{1}_B\rangle_0 \\
\mathrm{III} &= -q\mathsf{E}\langle\eta_1\eta_2\mathbf{1}_B\rangle_0 \\
\mathrm{IV} &= q^2\mathsf{E}\langle\mathbf{1}_B\rangle_0 \ .
\end{aligned}$$

Using (16.81) we obtain

$$\langle\varepsilon_1\varepsilon_2\eta_1\eta_2\rangle_0 = \frac{1}{Z_0^2}\langle \mathrm{sh}Y(\boldsymbol{\rho}^1)\mathrm{sh}Y(\boldsymbol{\rho}^2)\mathrm{sh}Y^*(\boldsymbol{\rho}^1)\mathrm{sh}Y^*(\boldsymbol{\rho}^2)\rangle_- \ ,$$

and proceeding as in (16.82) yields

$$\langle \mathrm{sh}Y(\boldsymbol{\rho}^1)\mathrm{sh}Y(\boldsymbol{\rho}^2)\mathrm{sh}Y^*(\boldsymbol{\rho}^1)\mathrm{sh}Y^*(\boldsymbol{\rho}^2)\rangle_- = \sum_{\alpha,\gamma} w_\alpha w_\gamma \mathrm{sh}Y_\alpha \mathrm{sh}Y_\gamma \mathrm{sh}Y_\alpha^* \mathrm{sh}Y_\gamma^* \ .$$

Let us try to compute

$$\mathrm{I} = \mathsf{E}\,\frac{\sum_{\alpha,\gamma} w_\alpha w_\gamma \mathrm{sh}Y_\alpha \mathrm{sh}Y_\gamma \mathrm{sh}Y_\alpha^* \mathrm{sh}Y_\gamma^*}{\left(\sum w_\alpha \mathrm{ch}Y_\alpha \mathrm{ch}Y_\alpha^*\right)^2} \ .$$

First we observe that the terms on the numerator for which $\alpha \neq \gamma$ give a zero contribution. This is seen by changing Y_α into $-Y_\alpha$ and leaving all the other Y_γ for $\gamma \neq \alpha$ unchanged. Thus

$$\mathrm{I} = \mathsf{E}\,\frac{\sum_\alpha w_\alpha^2 \mathrm{sh}^2 Y_\alpha \mathrm{sh}^2 Y_\alpha^*}{\left(\sum w_\alpha \mathrm{ch}Y_\alpha \mathrm{ch}Y_\alpha^*\right)^2} \ .$$

To compute this the idea is to combine Theorem 16.4.3 and (16.6) with $U = \mathrm{sh}Y\mathrm{sh}Y^*$ and $V = \mathrm{ch}Y\mathrm{sh}Y^*$ (where of course $Y = \beta z\sqrt{pq^{p-1}}$, $Y^* = \beta z^*\sqrt{pq^{p-1}}$ for z and z^* independent standard Gaussian r.v.s). There is a familiar little obstacle: Theorem 16.4.3 states "Given β...", while here we work with the $(N-2)$-spin system with β' given by (16.67) and dependent on N; but the dependence of β' on N is irrelevant because the conclusion of Theorem 16.4.3 holds uniformly over β as β remains bounded. We then proceed as follows. First we observe that there is a number A depending only on β such that $\mathsf{E}U^2 \le A^2$ and $\mathsf{E}V^4 \le A^2$. Given $\varepsilon' > 0$ let us then consider ε and n as provided by Theorem 16.4.3. we may then write (16.6) as

$$\left|\mathsf{E}\,\frac{\sum_{\alpha\ge 1} w_\alpha^2 U_\alpha^2}{(\sum_{\alpha\ge 1} w_\alpha V_\alpha)^2} - (1-m)\frac{\mathsf{E}\,U^2 V^{m-2}}{\mathsf{E}\,V^m}\right| \le \varepsilon' + A\sum \delta_{n_1,\ldots,n_k} \ , \qquad (16.89)$$

where the sum is over all (finitely many choices of) k and integers $n_1,\ldots,n_k \ge 2$ with $\sum_{s\le k} n_k \le n$, and where $\delta_{n_1,\ldots,n_k} = 0$ when (16.3) holds and $= 1$ otherwise. Now, for a r.v. $X \ge 0$ we have $\mathsf{E}\mathbf{1}_{\{X\ge\varepsilon\}} \le \varepsilon^{-1}\mathsf{E}X$. It then follows from Theorem 16.4.3 that each of the quantiles $\delta_{n_1,\ldots,n_k}$ satisfies (16.57). Since ε' is arbitrary, the right-hand side of (16.89) satisfies (16.57) and therefore

$$\mathrm{I} = (1-m^-)\frac{\mathsf{E}\mathrm{th}^2 Y \mathrm{th}^2 Y^* \mathrm{ch}^{m^-} Y \mathrm{ch}^{m^-} Y^*}{\mathsf{E}\mathrm{ch}^{m^-} Y \mathrm{ch}^{m^-} Y^*} + \delta$$
$$= (1-m^-)\left(\frac{\mathsf{E}\mathrm{th}^2 Y \mathrm{ch}^{m^-} Y}{\mathsf{E}\mathrm{ch}^{m^-} Y}\right)^2 + \delta$$

by independence of Y and Y^*.

To estimate the term II, we note that (16.81) implies

$$\langle \varepsilon_1\varepsilon_2 \mathbf{1}_B\rangle_0 = \frac{1}{Z_0^2}\langle \mathbf{1}_B \mathrm{sh}Y(\boldsymbol{\rho}^1)\mathrm{sh}Y(\boldsymbol{\rho}^2)\mathrm{ch}Y^*(\boldsymbol{\rho}^1)\mathrm{ch}Y^*(\boldsymbol{\rho}^2)\rangle_- \ ,$$

and proceeding as in (16.82) we get

$$\langle \mathbf{1}_B \mathrm{sh}Y(\boldsymbol{\rho}^1)\mathrm{sh}Y(\boldsymbol{\rho}^2)\mathrm{ch}Y^*(\boldsymbol{\rho}^1)\mathrm{ch}Y^*(\boldsymbol{\rho}^2)\rangle_- = \sum w_\alpha^2 \mathrm{sh}^2 Y_\alpha \mathrm{ch}^2 Y_\alpha^* \ .$$

Therefore

$$\mathrm{II} = -q\mathsf{E}\frac{1}{Z_0^2}\sum w_\alpha^2 \mathrm{sh}^2 Y_\alpha \mathrm{ch}^2 Y_\alpha^* \ ,$$

and we appeal now to (16.6) with $U = \mathrm{sh}Y\mathrm{ch}Y^*$ and $V = \mathrm{ch}Y\mathrm{ch}Y^*$ to obtain

$$\mathrm{II} = -q(1-m^-)\frac{\mathsf{E}\mathrm{th}^2 Y \mathrm{ch}^{m^-} Y \mathrm{ch}^{m^-} Y^*}{\mathsf{E}\mathrm{ch}^{m^-} Y \mathrm{ch}^{m^-} Y^*} + \delta$$
$$= -q(1-m^-)\frac{\mathsf{E}\mathrm{th}^2 Y \mathrm{ch}^{m^-} Y}{\mathsf{E}\mathrm{ch}^{m^-} Y} + \delta \ ,$$

using independence of Y and Y^*. We find the same value for III and, similarly, $\mathrm{IV} = q^2(1-m^-)+\delta$. □

We now study $\nu_t'(f) = \mathrm{d}\nu_t(f)/\mathrm{d}t$. We denote by E' expectation given the Hamiltonian $H_{N-2,\boldsymbol{\beta}'}$ and we set

$$\begin{aligned}\Delta_{\ell,\ell'} = \Delta(\boldsymbol{\rho}^\ell,\boldsymbol{\rho}^{\ell'}) &= \mathsf{E}'g(\boldsymbol{\rho}^\ell)g(\boldsymbol{\rho}^{\ell'}) - \mathsf{E}'Y(\boldsymbol{\rho}^\ell)Y(\boldsymbol{\rho}^{\ell'}) \\ &= \mathsf{E}'g^*(\boldsymbol{\rho}^\ell)g^*(\boldsymbol{\rho}^{\ell'}) - \mathsf{E}'Y^*(\boldsymbol{\rho}^\ell)Y^*(\boldsymbol{\rho}^{\ell'}) \ . \end{aligned} \tag{16.90}$$

Proposition 16.5.6. *If f is a function on Σ_N^n, then*

$$\begin{aligned}\nu_t'(f) = &\sum_{1\le\ell<\ell'\le n}\nu_t((\varepsilon_\ell\varepsilon_{\ell'}+\eta_\ell\eta_{\ell'})\Delta_{\ell,\ell'}f) \\ &- n\sum_{\ell\le n}\nu_t((\varepsilon_\ell\varepsilon_{n+1}+\eta_\ell\eta_{n+1})\Delta_{\ell,n+1}f) \\ &+ \frac{n(n+1)}{2}\nu_t((\varepsilon_{n+1}\varepsilon_{n+2}+\eta_{n+1}\eta_{n+2})\Delta_{n+1,n+2}f) \\ &+ \mathcal{R}(t) \ , \end{aligned} \tag{16.91}$$

where $|\mathcal{R}(t)|\,,\ |\mathcal{R}'(t)|\,,\ |\mathcal{R}''(t)| \le K(n)\sup|f|/N$.

Here $\mathcal{R}'(t) = \mathrm{d}\mathcal{R}(t)/\mathrm{d}t$, etc.

Proof. We compute $\nu_t'(f)$ using (16.75) and straightforward differentiation. We then integrate by parts, keeping in mind that the quantities $g(\boldsymbol{\rho}), g^*(\boldsymbol{\rho}), Y(\boldsymbol{\rho}), Y^*(\boldsymbol{\rho}), g^\sim(\boldsymbol{\rho})$ are all independent of each other. The quantity $\mathcal{R}(t)$ collects the terms created after integration by parts by the (lower order) quantities $\sqrt{t}\sigma_N\sigma_{N-1}g^\sim(\boldsymbol{\rho})$ of (16.75). It is of the form $\mathcal{R}(t) = N^{-1}\nu_t(f')$ where f' is a function on Σ_N^{n+2} with $|f'| \le K|f|$. This makes it obvious that $|\mathcal{R}(t)| \le K(n)\sup|f|/N$. Let us observe that (16.91) shows in particular that

$$|\nu_t'(f)| \le K(n)\nu_t(|f|) \,. \tag{16.92}$$

Using this for f' rather than f this implies

$$|\mathcal{R}'(t)| \le K(n)N^{-1}\nu_t(|f'|) \le K(n)N^{-1}\sup|f| \,.$$

Iteration of these facts proves that $|\mathcal{R}''(t)| \le K(n)N^{-1}\sup|f|$. □

Let us observe that integration of (16.92) yields for $f \ge 0$ that

$$\nu_t(f) \le K(n)\nu_0(f) \,. \tag{16.93}$$

Our next goal is to examine closely the function $\Delta_{\ell,\ell'}$. We recall Definition 16.5.4.

Lemma 16.5.7. *We have*

$$\mathsf{E}'Y(\boldsymbol{\rho}^1)Y(\boldsymbol{\rho}^2) \simeq \beta^2 p q_{1,2}^{p-1} \,. \tag{16.94}$$

Proof. Recalling the set B of (16.84), by definition, we have

$$\mathsf{E}'Y(\boldsymbol{\rho}^1)Y(\boldsymbol{\rho}^2) = \beta^2 p q^{p-1}\mathbf{1}_B \,.$$

Moreover, by Lemma 16.5.5 we have $q_{1,2} \simeq q\mathbf{1}_B$, so that $q_{1,2}^{p-1} \simeq q^{p-1}\mathbf{1}_B$. □

For two functions f_1 and f_2 on Σ_N^2, let us write $f_1 \lesssim f_2$ if the set A of configurations where $f_1 > f_2$ is as in (16.83). The following summarizes the properties of $\Delta_{\ell,\ell'}$ that we need.

Lemma 16.5.8. *We have*

$$|\Delta_{1,2} - \beta^2 p(R_{1,2}^{p-1} - q_{1,2}^{p-1})| \lesssim K c_N^2 \,, \tag{16.95}$$

and thus

$$|\Delta_{1,2}| \lesssim \beta^2 p|R_{1,2}^{p-1} - q_{1,2}^{p-1}| + K c_N^2 \,. \tag{16.96}$$

We have

$$|R_{1,2}^{p-1} - q_{1,2}^{p-1}| \le p|R_{1,2} - q_{1,2}| \tag{16.97}$$

and

$$|R_{1,2}^{p-1} - q_{1,2}^{p-1}| \lesssim 2p2^{-p/2} \ . \tag{16.98}$$

Consequently,

$$|\Delta_{1,2}| \lesssim \beta^2 p^2 |R_{1,2} - q_{1,2}| + Kc_N^2 \tag{16.99}$$

$$|\Delta_{1,2}| \lesssim 2\beta^2 p^2 2^{-p/2} + Kc_N^2 \ . \tag{16.100}$$

Moreover

$$(1 - R_{1,2})|R_{1,2}^{p-1} - q_{1,2}^{p-1}| \lesssim p2^{-p/2}|R_{1,2} - q_{1,2}| \ . \tag{16.101}$$

The bad news is that when we try to compare $\Delta_{1,2}$ with $|R_{1,2} - q_{1,2}|$, we get a very large coefficient $\beta^2 p^2$ in (16.99). The good news is that $|\Delta_{1,2}|$ is always very small by (16.100), at least as long as $\beta \le 2^{p/8}$.

Proof. Combining (16.90) with (16.94) and (16.68), and since $|R_{1,2} - R_{1,2}^-| \le L/N$ we get

$$|\Delta_{1,2} - \beta^2 p(R_{1,2}^{p-1} - q_{1,2}^{p-1})| \lesssim Kc_N^2 \ ,$$

which is (16.95). Since $|x^p - y^p| \le p|x - y|$ for $|x|, |y| \le 1$, (16.97) holds. Also, Theorem 16.3.6 yields

$$|R_{1,2}^{p-1} - q_{1,2}^{p-1}| \lesssim (\mathbf{1}_{\{|R_{1,2}| \le 2^{-p/4}\}} + \mathbf{1}_{\{R_{1,2} \ge 1-2^{-p/2}\}})|R_{1,2}^{p-1} - q_{1,2}^{p-1}| \ . \tag{16.102}$$

If $|R_{1,2}| \le 2^{-p/4}$ we have $q_{1,2} = 0$ so that $|R_{1,2}^{p-1} - q_{1,2}^{p-1}| = |R_{1,2}|^{p-1} \le 2^{-p}$ and if $R_{1,2} \ge 1 - 2^{-p/2}$ then $q_{1,2} = q$ so that

$$|R_{1,2}^{p-1} - q_{1,2}^{p-1}| = |R_{1,2}^{p-1} - q^{p-1}| \le |R_{1,2}^{p-1} - 1| + |q^{p-1} - 1| \le 2p2^{-p/2}$$

because $|x^{p-1} - 1| \le (p-1)|x - 1|$ for $|x| \le 1$, and since $q \ge 1 - 2^{-p/2}$ by (16.80). This proves (16.98). To prove (16.101) we use (16.102) again. If $|R_{1,2}| \le 2^{-p/4}$ then (for $p \ge 4$),

$$\begin{aligned}(1 - R_{1,2})|R_{1,2}^{p-1} - q_{1,2}^{p-1}| &\le 2|R_{1,2}^{p-1}| \le 2 \cdot 2^{-(p-2)p/4}|R_{1,2}| \\ &\le p2^{-p/2}|R_{1,2}| = p2^{-p/2}|R_{1,2} - q_{1,2}| \ ,\end{aligned}$$

while, if $R_{1,2} \ge 1 - 2^{-p/2}$,

$$\begin{aligned}(1 - R_{1,2})|R_{1,2}^{p-1} - q_{1,2}^{p-1}| &\le (1 - R_{1,2})p|R_{1,2} - q_{1,2}| \\ &\le p2^{-p/2}|R_{1,2} - q_{1,2}| \ ,\end{aligned}$$

which finishes the proof. □

Corollary 16.5.9. *Consider a function $f \ge 0$ on Σ_N^2 with $f \le 1$. Then if $\beta^2 p^2 2^{-p/2} \le 1$ we have*

$$\nu_t(f) \le L\nu(f) + Kc_N^2 \ , \tag{16.103}$$

where L is a universal constant.

Proof. First we observe that if $f_1 \lesssim f_2$ and $|f_1|, |f_2| \le 1$ then $\nu_t(f_1) \le \nu_t(f_2) + K\exp(-N/K)$ where K does not depend on t or N. This is simply because (16.83) and (16.93) imply that $\nu_t(\mathbf{1}_A) \le K\exp(-N/K)$. Combining (16.91) and (16.100) we deduce

$$|\nu_t'(f)| \le L\nu_t(f) + Kc_N^2$$

and we conclude using Lemma A.11.1 . □

Lemma 16.5.10. *Consider a function f on Σ_N^2 such that $|f| \le 2$. Then, if $\beta^2 p^2 2^{-p/2} \le 1$ we have*

$$|\nu_t^{(3)}(f)| \le L\beta^6 p^6 2^{-p/2}\nu\big((R_{1,2} - q_{1,2})^2\big) + Kc_N^2 \ . \tag{16.104}$$

Proof. We iterate the formula (16.91) to compute $\nu^{(3)}(f)$. The terms $\mathcal{R}(t)$ (and their derivatives) create only a contribution $\le K/N$. In the other terms "each derivation creates a factor $\Delta_{\ell,\ell'}$", so that use of Hölder's inequality yields

$$|\nu_t^{(3)}(f)| \le L\nu_t(|\Delta_{1,2}|^3) + \frac{K}{N} \ .$$

Using (16.99) for two of the factors $\Delta_{1,2}$ and (16.100) for the third yields

$$\nu_t(|\Delta_{1,2}|^3) \le L\beta^6 p^6 2^{-p/2}\nu_t\big((R_{1,2} - q_{1,2})^2\big) + Kc_N^2$$

and we conclude with (16.103). □

Despite this success, it seems very difficult to usefully control $\nu_t'(f)$ for a general function f. But we remember (recalling (16.65)) that we are really interested in the case where

$$f = (\varepsilon_1\varepsilon_2 - q_{1,2})(\eta_1\eta_2 - q_{1,2}) \ . \tag{16.105}$$

It turns out that in that case we may use site-symmetry to obtain a good control of $\nu_1'(f)$ and $\nu_1''(f)$. We will then use the formula

$$\nu(f) = \nu_0(f) + \nu_1'(f) - \frac{1}{2}\nu_1''(f) + \int_0^1 \frac{t^2}{2}\nu_t^{(3)}(f)\mathrm{d}t \ , \tag{16.106}$$

that follows from simple integration by parts. Needless to say, it took a very long time to find that approach.

Lemma 16.5.11. *If f is as in (16.105) we have*

$$|\nu_1'(f)| \le L\beta^2 p^2 2^{-p/2}\nu((R_{1,2} - q_{1,2})^2) + Kc_N^2 \ .$$

Proof. Using (16.91), we have to bound terms such as

$$\nu((\varepsilon_1\varepsilon_2 - q_{1,2})(\eta_1\eta_2 - q_{1,2})\varepsilon_\ell\varepsilon_{\ell'}\Delta_{\ell,\ell'}) = \mathrm{I} + \mathrm{II} + \mathrm{III}$$

where (please admire the trick)

$$\begin{aligned} \mathrm{I} &= \nu\big((\varepsilon_1\varepsilon_2 - q_{1,2})(\eta_1\eta_2 - q_{1,2})\varepsilon_\ell\varepsilon_{\ell'}(\Delta_{\ell,\ell'} - \beta^2 p(R^{p-1}_{\ell,\ell'} - q^{p-1}_{\ell,\ell'})\big) \\ \mathrm{II} &= \beta^2 p\nu\big((\varepsilon_1\varepsilon_2 - q_{1,2})(\eta_1\eta_2 - q_{1,2})(R^{p-1}_{\ell,\ell'} - q^{p-1}_{\ell,\ell'})\big) \\ \mathrm{III} &= -\beta^2 p\nu\big((1 - \varepsilon_\ell\varepsilon_{\ell'})(\varepsilon_1\varepsilon_2 - q_{1,2})(\eta_1\eta_2 - q_{1,2})(R^{p-1}_{\ell,\ell'} - q^{p-1}_{\ell,\ell'})\big) \,. \end{aligned}$$

First, (16.95) implies

$$|\mathrm{I}| \le Kc_N^2 \,.$$

Using site-symmetry as in the proof of (16.65) we get

$$\begin{aligned} \mathrm{II} &\le \beta^2 p\nu\big((R_{1,2} - q_{1,2})^2(R^{p-1}_{\ell,\ell'} - q^{p-1}_{\ell,\ell'})\big) + \frac{K}{N} \\ &\le 2\beta^2 p^2 2^{-p/2}\nu((R_{1,2} - q_{1,2})^2) + Kc_N^2 \end{aligned}$$

using (16.98) in the second line.

Using site-symmetry a second time (i.e. that we can replace $\eta_1\eta_2 = \sigma^1_{N-1}\sigma^2_{N-1}$ by any $\sigma^1_i\sigma^2_i$) we obtain

$$\big|\mathrm{III} + \beta^2 p\nu\big((1 - \varepsilon_\ell\varepsilon_{\ell'})(\varepsilon_1\varepsilon_2 - q_{1,2})(R_{1,2} - q_{1,2})(R^{p-1}_{\ell,\ell'} - q^{p-1}_{\ell,\ell'})\big)\big| \le \frac{K}{N}$$

and since $1 - \varepsilon_\ell\varepsilon_{\ell'} \ge 0$ and $|\varepsilon_1\varepsilon_2 - q_{1,2}| \le 2$,

$$\begin{aligned} \mathrm{III} &\le 2\beta^2 p\nu\big((1 - \varepsilon_\ell\varepsilon_{\ell'})|R_{1,2} - q_{1,2}||R^{p-1}_{\ell,\ell'} - q^{p-1}_{\ell,\ell'}|\big) + \frac{K}{N} \\ &\le 2\beta^2 p\nu\big((1 - R_{\ell,\ell'})|R_{1,2} - q_{1,2}||R^{p-1}_{\ell,\ell'} - q^{p-1}_{\ell,\ell'}|\big) + \frac{K}{N} \,, \end{aligned}$$

using again site-symmetry in the second line. Now (16.101) yields

$$\mathrm{III} \le 2\beta^2 p^2 2^{-p/2}\nu(|R_{1,2} - q_{1,2}||R_{\ell,\ell'} - q_{\ell,\ell'}|) + \frac{K}{N} \,,$$

and the conclusion by the Cauchy-Schwarz inequality. □

The proof of the following is entirely similar.

Lemma 16.5.12. *If f is as in (16.105) then*

$$|\nu_1''(f)| \le L\beta^2 p^4 2^{-p/2}\nu\big((R_{1,2} - q_{1,2})^2\big) + Kc_N^2 \,.$$

Proof of Theorem 16.5.1. By (16.65) we have

$$\nu\big((R_{1,2}-q_{1,2})^2\big) \le \nu(f) + \frac{L}{N}\,. \tag{16.107}$$

Using (16.106) and Lemmas 16.5.10 to 16.5.12 we get

$$\nu(f) \le \nu_0(f) + L\beta^6 p^6 2^{-p/2}\nu\big((R_{1,2}-q_{1,2})^2\big) + Kc_N^2\,,$$

so that if $\beta \le 2^{-p/12}/Lp$, it follows from (16.107) that

$$\nu\big((R_{1,2}-q_{1,2})^2\big) \le \nu_0(f) + \frac{1}{2}\nu\big((R_{1,2}-q_{1,2})^2\big) + Kc_N^2\,.$$

Now $\nu(f) = \delta$ by (16.78) and the choice of q. □

Proof of Lemma 16.5.3. It suffices to prove that if we define

$$\Phi(q) = \frac{\mathsf{E}\mathrm{th}^2 Y \mathrm{ch}^{m^-} Y}{\mathsf{E}\mathrm{ch}^{m^-} Y}$$

(where $Y = \beta z\sqrt{pq^{p-1}}$), then (for large p) Φ maps the interval $[1-2^{-p/2}, 1]$ into itself. We have, since $\mathrm{ch}x \ge e^x/2$ and $m^- \le 1$,

$$1-\Phi(q) = \frac{\mathsf{E}\mathrm{ch}^{m^- -2} Y}{\mathsf{E}\mathrm{ch}^{m^-} Y} \le \frac{1}{\mathsf{E}\mathrm{ch}^{m^-} Y} \le \frac{2}{\mathsf{E}\exp m^- Y} = 2\exp\left(-\frac{\beta^2 pq^{p-1}m^{-2}}{2}\right).$$

To conclude the proof it suffices to show that $\beta m^- \ge 1$, because for large p we have $2\exp(-p(1-2^{-p/2})^{p-1}/2) \le 2^{-p/2}$. To prove that $\beta m^- \ge 1$ we will for convenience of notation work first with the N-spin system rather than with the $(N-1)$-spin system. The idea is to reverse the argument given just after the proof of Theorem 16.4.2, by which we showed that $m < 1$ for large β. Instead of (16.56) we write

$$p_N'(\beta) \le \beta(1-\nu(R_{1,2}^p)) + \frac{K}{N}$$

so that

$$\nu(R_{1,2}^p) \le 1 - \frac{p_N'(\beta)}{\beta} + \frac{K}{N} \le 1 - \frac{p_N'(\gamma_p)}{\beta} + \frac{K}{N}\,, \tag{16.108}$$

because $p_N'(\beta)$ is increasing by convexity of p_N. Letting $m = \nu(\mathbf{1}_{\{R_{1,2}<1/2\}})$, since it is essentially true that either $|R_{1,2}| \le 2^{-p/4}$ or $R_{1,2} \ge 1-2^{-p/2}$ we get

$$\begin{aligned}\nu(R_{1,2}^p) &= \nu(R_{1,2}^p \mathbf{1}_{\{R_{1,2}\le 1/2\}}) + \nu(R_{1,2}^p \mathbf{1}_{\{R_{1,2}>1/2\}}) \\ &\ge -m2^{-p^2/4} + (1-m)(1-2^{-p/2})^p - \frac{K}{N} := -cm + (1-m)d - \frac{K}{N}\end{aligned}$$

and combining with (16.108) yields

$$d - m(c+d) \le 1 - \frac{p'_N(\gamma_p)}{\beta} + \frac{K}{N}$$

i.e.

$$d - 1 + \frac{p'_N(\gamma_p)}{\beta} - \frac{K}{N} \le m(c+d) .$$

Now $d = (1-2^{-p/2})^p \ge 1 - p2^{-p/2}$ and, for large p, we have $d \le 1 - p2^{-p/2}/2$ so that $c + d \le 1$ and

$$m \ge \frac{p'_N(\gamma_p)}{\beta} - p2^{-p/2} - \frac{K}{N} .$$

Thus for $\beta \le p^{-2}2^{p/2}$ we have

$$m\beta \ge p'_N(\gamma_p) - \frac{1}{p} - \frac{K}{N} .$$

Since $p'_N(\gamma_p) \to \gamma_p \ge \sqrt{2\log 2}(1-2^{-p})$, since $2\log 2 > 1$, if p and N are large, and $\beta \le p^{-2}2^{-p/2}$, this does imply that for a certain number a independent of N we have $m\beta \ge a > 1$.

Thus (changing N into $N-2$) we see that $m^-\beta^- \ge a$ where $\beta^- = \beta((N-2)/N)^{p-1/2}$, so $m^-\beta \ge 1$ for large N. □

16.6 Computing $p_N(\beta)$

The function $\xi(x) = \beta^2 x^p$ is convex for $x \ge 0$; therefore by Theorem 14.4.4 we have

$$\limsup_{N\to\infty} p_N(\beta) \le \mathcal{P}_k(\mathbf{m}, \mathbf{q}) , \tag{16.109}$$

where $\mathcal{P}_k(\mathbf{m}, \mathbf{q})$ is given by (14.77). Here

$$\theta(x) = x\xi'(x) - \xi(x) = (p-1)\beta^2 x^p .$$

When $k = 2$, $q_0 = q_1 = 0$, $q_2 = q$, $q_3 = 1$, $m_0 = 0$, $m_1 = m$, $m_2 = 1$, then

$$\mathcal{P}_2(\mathbf{m}, \mathbf{q}) = \log 2 - \frac{1}{2}m(p-1)\beta^2 q^p - \frac{1}{2}(p-1)\beta^2(1-q^p) + X_0 , \tag{16.110}$$

where X_0 is defined as follows: consider independent Gaussian r.v.s z_1 and z_2 with $\mathsf{E}z_1^2 = \beta^2 p q^{p-1}$ and $\mathsf{E}z_1^2 = \beta^2 p(1-q^{p-1})$. Let

$$X_3 = \log \mathrm{ch}(z_1 + z_2)$$
$$X_2 = \log \mathsf{E}_2 \mathrm{ch}(z_1 + z_2)$$
$$X_0 = X_1 = \frac{1}{m} \log \mathsf{E} \exp m X_2 .$$

Equivalently,

$$X_0 = \frac{\beta^2 p}{2}(1 - q^{p-1}) + \frac{1}{m} \log \mathsf{E}\mathrm{ch}^m(\beta z \sqrt{pq^{p-1}}) \, .$$

Combining with (16.110) we have shown that for this choice of $\mathbf{m}$ and $\mathbf{q}$, and if we set $Y = \beta z \sqrt{pq^{p-1}}$, then

$$\begin{aligned} \mathcal{P}_2(\mathbf{m}, \mathbf{q}) = F(\beta, q, m) := \log 2 + \frac{\beta^2}{2}(1 - pq^{p-1}) + \frac{\beta^2}{2}(p-1)q^p(1-m) \\ + \frac{1}{m} \log \mathsf{E}\mathrm{ch}^m Y \, . \end{aligned} \tag{16.111}$$

Let

$$p_\infty(\beta) = \inf_{q,m} F(\beta, q, m) \, ,$$

where the infimum is over $0 \le q \le 1$, $0 \le m \le 1$. It follows from (16.109) that

$$\limsup_N p_N(\beta) \le p_\infty(\beta) \, . \tag{16.112}$$

Theorem 16.6.1. *There exists p_0 and L with the following property. If $\gamma_p \le \beta \le p^{-3} 2^{p/2}$ and $p \ge p_0$ we have*

$$\lim_{N \to \infty} p_N(\beta) = p_\infty(\beta) \, . \tag{16.113}$$

Exercise 16.6.2. It follows from (16.113) and (16.109) that for each $k, \mathbf{m}, \mathbf{q}$ we have $\mathcal{P}_k(\mathbf{m}, \mathbf{q}) \ge p_\infty(\beta) = \inf_{q,m} F(\beta, q, m)$. In other words, the Parisi solution is the solution with one level of replica-symmetry breaking, that is the infimum of $\mathcal{P}_k(\mathbf{m}, \mathbf{q})$ over all values of $\mathbf{m}, \mathbf{q}$ and k is obtained for the value $k = 2$. Find a direct proof of this fact, using only the definition of $\mathcal{P}_k(\mathbf{m}, \mathbf{q})$. (The author does not know how to solve this exercise.)

To prove Theorem 16.6.1, we use the Hamiltonian $H_{N,\boldsymbol{\beta}} = H_{N,\beta,\boldsymbol{\beta}}$ including the perturbation term, and we write

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

We know from Lemma 12.2.1 that

$$p_N(\beta) \le p_N(\beta, \boldsymbol{\beta}) \le p_N(\beta) + c_N^2 \, . \tag{16.114}$$

The idea is to prove the inequality

$$(N+1)p_{N+1}(\beta, \boldsymbol{\beta}) - N p_N(\beta, \boldsymbol{\beta}) \ge p_\infty(\beta) + \delta \tag{16.115}$$

where $\delta = \delta_N(\beta)$ is as usual a quantity such that $\lim_{N\to\infty}\int \delta \mathrm{d}\boldsymbol{\beta} = 0$. Summation of this relations entails

$$\liminf_{N\to\infty} \int p_{N+1}(\beta, \boldsymbol{\beta})\mathrm{d}\boldsymbol{\beta} \geq p_\infty(\beta) ,$$

and together with (16.114) this implies that $\liminf_N p_N(\beta) \geq p_\infty(\beta)$. Combining with (16.112) concludes the proof.

The overall approach is as in Proposition 1.6.8. Let us define

$$\beta' = \beta\left(\frac{N+1}{N}\right)^{(p-1)/2} \quad ; \quad \boldsymbol{\beta}' = \left(\beta_s\left(\frac{N+1}{N}\right)^{(s-1)/2}\right)_{s\geq 1} .$$

Proposition 16.6.3. *We have*

$$(N+1)p_{N+1}(\beta', \boldsymbol{\beta}') - Np_N(\beta, \boldsymbol{\beta}) = \log 2 + \frac{\beta^2 p}{2}(1 - q^{p-1}) \\ + \frac{1}{m}\log \mathsf{E}\mathrm{ch}^m Y + \delta , \qquad (16.116)$$

where $Y = \beta z\sqrt{pq^{p-1}}$, $q = q_N(\beta, \boldsymbol{\beta})$ is as in Theorem 16.5.1 and

$$m = m_N(\beta, \boldsymbol{\beta}) = \nu(\mathbf{1}_{\{R_{1,2}\leq 1/2\}}) .$$

Proof. The choice of β' is of course such that

$$\frac{\beta'}{(N+1)^{(p-1)/2}} = \frac{\beta}{N^{(p-1)/2}} .$$

If we remove in $H_{N+1,\beta',\boldsymbol{\beta}'}$ the terms containing σ_{N+1} we get exactly $H_{N,\beta,\boldsymbol{\beta}}$. Thus, as in (1.172) we deduce the identity

$$(N+1)p_{N+1}(\beta', \boldsymbol{\beta}') - Np_N(\beta, \boldsymbol{\beta}) = \log 2 + \mathsf{E}\log\langle \mathrm{Av}\exp \sigma_{N+1} g(\boldsymbol{\sigma})\rangle , \qquad (16.117)$$

where Av denotes average over $\sigma_{N+1} = \pm 1$, where $\langle\cdot\rangle$ denotes an average for the Gibbs measure with Hamiltonian $H_{N,\boldsymbol{\beta}}$, and where $(g(\boldsymbol{\sigma}))_{\boldsymbol{\sigma}}$ is a Gaussian family, independent of the randomness of $\langle\cdot\rangle$, that satisfies

$$|\mathsf{E}g(\boldsymbol{\sigma}^1)g(\boldsymbol{\sigma}^2) - \beta^2 p R_{1,2}^{p-1}| \leq Kc_N^2 . \qquad (16.118)$$

As in Section 16.5 we use the decomposition $(C_\alpha)_{\alpha\geq 1}$ of Σ_N provided by Theorem 16.4.2. We consider an independent sequence (z_α) of standard Gaussian r.v.s and we define

$$Y(\boldsymbol{\sigma}) = Y_\alpha := \beta z_\alpha\sqrt{pq^{p-2}} ,$$

where α is the unique integer for which $\boldsymbol{\sigma} \in C_\alpha$. Thus $\mathsf{E}'Y(\boldsymbol{\sigma}^1)Y(\boldsymbol{\sigma}^2) = \beta^2 qp^{p-1}$ if $\boldsymbol{\sigma}^1$ and $\boldsymbol{\sigma}^2$ belong to the same set C_α, and $= 0$ otherwise. Let us define

$$g_t(\boldsymbol{\sigma}) = \sqrt{t}g(\boldsymbol{\sigma}) + \sqrt{1-t}Y(\boldsymbol{\sigma})$$

and

$$\varphi(t) = \mathsf{E}\log\langle \mathrm{Av}\exp \sigma_{N+1} g_t(\boldsymbol{\sigma})\rangle = \mathsf{E}\log\langle \mathrm{ch} g_t(\boldsymbol{\sigma})\rangle \ .$$

Then, writing $g'_t(\boldsymbol{\sigma}) = \mathrm{d}g_t(\boldsymbol{\sigma})/\mathrm{d}t$, we have

$$\varphi'(t) = \mathsf{E}\frac{\langle g'_t(\boldsymbol{\sigma})\mathrm{sh} g_t(\boldsymbol{\sigma})\rangle}{\langle \mathrm{ch} g_t(\boldsymbol{\sigma})\rangle} \ ,$$

and integration by parts yields

$$\begin{aligned}\varphi'(t) &= \mathsf{E}\frac{\langle \mathsf{E}'(g'_t(\boldsymbol{\sigma})g_t(\boldsymbol{\sigma}))\mathrm{ch} g_t(\boldsymbol{\sigma})\rangle}{\langle \mathrm{ch} g_t(\boldsymbol{\sigma})\rangle} \\ &\quad - \mathsf{E}\frac{1}{\langle \mathrm{ch} g_t(\boldsymbol{\sigma})\rangle^2}\langle \mathsf{E}'(g'_t(\boldsymbol{\sigma}^1)g_t(\boldsymbol{\sigma}^2))\mathrm{sh} g_t(\boldsymbol{\sigma}^1)\mathrm{sh} g_t(\boldsymbol{\sigma}^2)\rangle \ . \qquad (16.119)\end{aligned}$$

Now (16.118) and Lemma 16.5.7 imply

$$\left|\mathsf{E}'(g'_t(\boldsymbol{\sigma})g_t(\boldsymbol{\sigma})) - \frac{\beta^2 p}{2}(1-q^{p-1})\right| \le Kc_N^2$$

and

$$\left|\mathsf{E}'(g'_t(\boldsymbol{\sigma}^1)g_t(\boldsymbol{\sigma}^2)) - \frac{\beta^2 p}{2}(R_{1,2}^{p-1} - q_{1,2}^{p-1})\right| \lesssim Kc_N^2$$

where $q_{1,2} = q\mathbf{1}_{\{R_{1,2}\ge 1/2\}}$, and where $\lesssim$ means that the set A of configurations where the inequality might fail satisfies $\nu(\mathbf{1}_A) \le K\exp(-N/K)$. Therefore, using that $\mathrm{ch} g_t(\boldsymbol{\sigma}) \ge 1$ in the second line, (16.119) implies

$$\begin{aligned}&\left|\varphi'(t) - \frac{\beta^2 p}{2}(1-q^{p-1})\right| \\ &\le \mathsf{E}\frac{1}{\langle \mathrm{ch} g_t(\boldsymbol{\sigma})\rangle_t^2}\left\langle \frac{\beta^2 p}{2}|R_{1,2}^{p-1} - q_{1,2}^{p-1}||\mathrm{sh} g_t(\boldsymbol{\sigma}^1)\mathrm{sh} g_t(\boldsymbol{\sigma}^2)|\right\rangle + Kc_N^2 \\ &\le \mathsf{E}\left\langle \frac{\beta^2 p}{2}|R_{1,2}^{p-1} - q_{1,2}^{p-1}||\mathrm{sh} g_t(\boldsymbol{\sigma}^1)\mathrm{sh} g_t(\boldsymbol{\sigma}^2)|\right\rangle + Kc_N^2 \\ &\le K\mathsf{E}\left\langle |R_{1,2} - q_{1,2}||\mathrm{sh} g_t(\boldsymbol{\sigma}^1)\mathrm{sh} g_t(\boldsymbol{\sigma}^2)|\right\rangle + Kc_N^2 \\ &\le K\mathsf{E}\langle |R_{1,2} - q_{1,2}|\rangle + Kc_N^2 \ ,\end{aligned}$$

using that $|x^{p-1} - y^{p-1}| \le (p-1)|x-y|$ in the third inequality, and integrating in $g_t(\boldsymbol{\sigma})$ inside the bracket rather than outside in the last inequality. Therefore, using the Cauchy-Schwarz inequality,

$$\begin{aligned}\left|\varphi(1) - \varphi(0) - \frac{\beta^2 p}{2}(1-q^{p-1})\right| &\le K\mathsf{E}\langle |R_{1,2} - q_{1,2}|\rangle + Kc_N^2 \\ &\le K\nu((R_{1,2} - q_{1,2})^2)^{1/2} + Kc_N^2 \ , \qquad (16.120)\end{aligned}$$

and (16.64) implies

$$\varphi(1) = \mathsf{E}\log\langle \mathrm{ch} g(\boldsymbol{\sigma})\rangle = \frac{\beta^2 p}{2}(1 - q^{p-1}) + \varphi(0) + \delta\ .$$

It remains to estimate $\varphi(0)$. If w_α is the Gibbs weight of C_α, then

$$\varphi(0) = \mathsf{E}\log\sum_{\alpha\geq 1} w_\alpha \mathrm{ch} Y_\alpha\ .$$

We shall show that $\varphi(0) = m^{-1}\mathsf{E}\log\mathrm{ch}^m Y + \delta$. This follows, in a simplified manner, the arguments of Theorem 16.2.1. First, consider a truncation level M; and let $X_\alpha = \min(M, \mathrm{ch}Y_\alpha)$. Then, using Jensen's inequality in the third line,

$$\begin{aligned}
\mathsf{E}\log\sum_{\alpha\geq 1} w_\alpha \mathrm{ch}Y_\alpha - \mathsf{E}\log\sum_{\alpha\geq 1} w_\alpha X_\alpha &= \mathsf{E}\log\left(1 + \frac{\sum_{\alpha\geq 1} w_\alpha(\mathrm{ch}Y_\alpha - X_\alpha)}{\sum_{\alpha\geq 1} w_\alpha X_\alpha}\right)\\
&\leq \mathsf{E}\log\left(1 + \sum_{\alpha\geq 1} w_\alpha(\mathrm{ch}Y_\alpha - X_\alpha)\right)\\
&\leq \log\left(1 + \mathsf{E}\sum_{\alpha\geq 1} w_\alpha(\mathrm{ch}Y_\alpha - X_\alpha)\right)\\
&= \log(1 + \mathsf{E}(\mathrm{ch}Y - \min(M, \mathrm{ch}Y))\\
&\leq \mathsf{E}(\mathrm{ch}Y - \min(M, \mathrm{ch}Y))
\end{aligned}$$

goes to 0 as $M \to \infty$; so it suffices to prove that if X is a r.v. with $0 \leq X \leq M$, and $(X_\alpha)_{\alpha\geq 1}$ are i.i.d copies of X, then

$$\mathsf{E}\log\sum_{\alpha\geq 1} w_\alpha X_\alpha = \frac{1}{m}\log(\mathsf{E}X^m) + \delta\ .$$

Consider a sequence $(v_\alpha)_{\alpha\geq 1}$ with Poisson-Dirichlet distribution Λ_m. It suffices to show that

$$\mathsf{E}\log\sum_{\alpha\geq 1} w_\alpha X_\alpha = \mathsf{E}\log\sum_{\alpha\geq 1} v_\alpha X_\alpha + \delta\ .$$

On the interval $[1, M]$ the function log can be uniformly approximated by polynomials, so that it suffices to prove that for any k we have

$$\mathsf{E}\left(\left(\sum_{\alpha\geq 1} w_\alpha X_\alpha\right)^k\right) = \mathsf{E}\left(\left(\sum_{\alpha\geq 1} v_\alpha X_\alpha\right)^k\right) + \delta\ .$$

This is obtained by the argument of Theorem 16.2.1 and by Theorem 16.4.3. □

Proposition 16.6.4. *We have*

$$(N+1)p_{N+1}(\beta', \boldsymbol{\beta}') - (N+1)p_{N+1}(\beta, \boldsymbol{\beta}) = \beta^2 \frac{p-1}{2}(1-(1-m_{N+1})q_{N+1}^p) + \delta \; . \tag{16.121}$$

Here the indices $N+1$ on m_{N+1} and q_{N+1} stress that these quantities are the same as in Proposition 16.6.3, but for the $(N+1)$-spin system, i.e. $m_{N+1} = m_{N+1}(\beta, \boldsymbol{\beta})$ and $q_{N+1} = q_{N+1}(\beta, \boldsymbol{\beta})$. We of course expect that $m_{N+1} \simeq m_N$ and $q_{N+1} \simeq q_N$ but we do not know how to prove it.

Proof of Theorem 16.6.1. Let us define

$$A_N = \log 2 + \frac{\beta^2 p}{2}(1 - q^{p-1}) + \frac{1}{m} \log \mathsf{E}\mathrm{ch}^m Y \tag{16.122}$$

$$B_N = \beta^2 \frac{p-1}{2}(1 - (1-m)q^p) \; , \tag{16.123}$$

where m and q are as in Proposition 16.6.3. Thus (16.116) and (16.121) mean respectively that

$$(N+1)p_{N+1}(\beta', \boldsymbol{\beta}') - Np_N(\beta, \boldsymbol{\beta}) = A_N + \delta$$
$$(N+1)p_{N+1}(\beta', \boldsymbol{\beta}') - (N+1)p_{N+1}(\beta, \boldsymbol{\beta}) = B_{N+1} + \delta$$

so that

$$(N+1)p_{N+1}(\beta, \boldsymbol{\beta}) - Np_N(\beta, \boldsymbol{\beta}) = A_N - B_{N+1} + \delta \; .$$

Summation of these relations yields

$$\begin{aligned} p_{N+1}(\beta, \boldsymbol{\beta}) &= \delta + \frac{1}{N+1}\left(\sum_{1 \le M \le N} A_M - \sum_{2 \le M \le N+1} B_M\right) \\ &= \delta + \frac{1}{N}\sum_{2 \le M \le N}(A_M - B_M) \; . \end{aligned} \tag{16.124}$$

Now, (16.122) and (16.124) show that, recalling (16.112)

$$A_N - B_N = F(\beta, q, m) \ge p_\infty(\beta)$$

and therefore by (16.124)

$$p_{N+1}(\beta, \boldsymbol{\beta}) \ge p_\infty(\beta) + \delta \; . \qquad \square$$

Proof of Proposition 16.6.4. For simplicity of notation we prove (16.121) for N rather than $N+1$, i.e. we prove that

$$Np_N(\beta^\sim, \boldsymbol{\beta}^\sim) - Np_N(\beta, \boldsymbol{\beta}) = (p-1)\frac{\beta^2}{2}(1 - (1-m)q^p) + \delta$$

where

$$\beta^\sim = \beta\left(\frac{N}{N-1}\right)^{(p-1)/2} \quad ; \quad \boldsymbol{\beta}^\sim = \left(\beta_s\left(\frac{N}{N-1}\right)^{(p-1)/2}\right) .$$

The idea is the same as in (1.175). In distribution,

$$\begin{aligned}
&\beta^\sim\left(\frac{p!}{N^{p-1}}\right)^{1/2} \sum_{i_1<\ldots<i_p} g_{i_1\ldots i_p}\sigma_{i_1}\cdots\sigma_{i_p} \\
&\stackrel{\mathcal{D}}{=} \beta\left(\frac{p!}{N^{p-1}}\right)^{1/2} \sum_{i_1<\ldots<i_p} g_{i_1\ldots i_p}\sigma_{i_1}\cdots\sigma_{i_p} \\
&+ a \sum_{i_1<\ldots<i_p} g'_{i_1\ldots i_p}\sigma_{i_1}\cdots\sigma_{i_p}
\end{aligned}$$

where $g'_{i_1\ldots i_p}$ are standard Gaussian r.v.s, independent of the $g_{i_1\ldots i_p}$ and where

$$a^2 = \frac{p!}{N^{p-1}}(\beta^{\sim 2} - \beta^2) \simeq \beta^2 \frac{p-1}{2}\frac{p!}{N^p} .$$

Using the same trick for the perturbation term, we get the identity

$$Np_N(\beta^\sim, \boldsymbol{\beta}^\sim) - Np_N(\beta, \boldsymbol{\beta}) = \mathsf{E}\log\langle \exp g(\boldsymbol{\sigma})\rangle ,$$

where

$$|\mathsf{E}g(\boldsymbol{\sigma}^1)g(\boldsymbol{\sigma}^2) - \beta^2\frac{p-1}{2}R_{1,2}^p| \le Kc_N^2$$

and we interpolate as in the proof of Proposition 16.6.3, using now

$$Y(\boldsymbol{\sigma}) = Y_\alpha = \beta\sqrt{\frac{p-1}{2}}z_\alpha$$

for $\alpha \in C_\alpha$. Let $g_t(\boldsymbol{\sigma}) = \sqrt{t}g(\boldsymbol{\sigma}) + \sqrt{1-t}Y(\boldsymbol{\sigma})$, and

$$\varphi(t) = \mathsf{E}\log\langle \exp g_t(\boldsymbol{\sigma})\rangle ,$$

so that by integration by parts

$$\begin{aligned}
\varphi'(t) &= \mathsf{E}\frac{\langle g'_t(\boldsymbol{\sigma})\exp g_t(\boldsymbol{\sigma})\rangle}{\langle \exp g_t(\boldsymbol{\sigma})\rangle} \\
&= \mathsf{E}\frac{\langle \mathsf{E}'(g'_t(\boldsymbol{\sigma})g_t(\boldsymbol{\sigma}))\exp g_t(\boldsymbol{\sigma})\rangle}{\langle \exp g_t(\boldsymbol{\sigma})\rangle} \\
&- \mathsf{E}\frac{\langle \mathsf{E}'(g'_t(\boldsymbol{\sigma}^1)g_t(\boldsymbol{\sigma}^2))\exp g_t(\boldsymbol{\sigma}^1)\exp g_t(\boldsymbol{\sigma}^2)\rangle}{\langle \exp g_t(\boldsymbol{\sigma})\rangle^2} .
\end{aligned} \tag{16.125}$$

A slight difference with the case of Proposition 16.6.3 is that we cannot use that $\exp g_t(\boldsymbol{\sigma})$ is bounded below. Instead we use that $\langle \exp g_t(\boldsymbol{\sigma})\rangle \ge$

$\exp\langle g_t(\boldsymbol{\sigma})\rangle$, and taking first expectation in the $g_t(\boldsymbol{\sigma})$ in the last term of (16.125) we obtain as in Proposition 16.6.3 that

$$\varphi(1) = \varphi(0) + \beta^2 \frac{p-1}{2}(1-q^p) + \delta$$

and, as in this proposition, we obtain, for $Y = \beta z_\alpha \sqrt{(p-1)/2}$

$$\varphi(0) = \frac{1}{m} \log \exp mY + \delta = m\beta^2 \frac{p-1}{2} + \delta \, . \qquad \square$$

Of course, this study leaves open a zillion of natural questions, such as the following.

Research Problem 16.6.5. (Level 2) Find a clean proof that for β in the range considered here, there is a unique pair (m, q) (with q close to 1) for which $F(\beta, m, q) = p_\infty(\beta)$. (An ugly proof of this fact can be found in [103].) Prove that Theorem 16.5.1 holds for this value of q.

Inspection of our proof of Theorem 16.6.1 shows that this value of q must work "for most N", but it is another matter to reach every value of N. A related question is to prove that $m_N(\beta, \boldsymbol{\beta}) \to m$ as $N \to \infty$.

Another sore point is as follow. Since (recalling the definition (16.111) of F)

$$p_\infty(\beta) = \inf_{m,q} F(\beta, m, q)$$

the equations

$$\frac{\partial F}{\partial q}(\beta, m, q) = 0 \tag{16.126}$$

$$\frac{\partial F}{\partial m}(\beta, m, q) = 0 \tag{16.127}$$

hold true. A straightforward computation (done many times already) shows that equation (16.126) is equivalent to the relation

$$q = \frac{\mathsf{E}\mathrm{th}^2 Y \mathrm{ch}^m Y}{\mathsf{E}\mathrm{ch}^m Y} \, , \tag{16.128}$$

for which the cavity method (as in Section16.5) provides a good explanation.

Research Problem 16.6.6. (Level 2) Find a "physical" explanation for the equation (16.127).

This is a very good question, because, apparently, the physicists failed to solve it. We do not know of a real "physical" explanation, but the following analytical argument provides elements of an answer. (We used it in [103], where we managed to prove Theorem 16.6.1 without knowing the upper bound (16.109).) The argument goes roughly as follows. Assuming that $m(\beta) = \lim_{N\to\infty} m_N(\beta)$ exists and $q(\beta) = \lim_{N\to\infty} q_N(\beta, \boldsymbol{\beta})$ exists independently of $\boldsymbol{\beta}$, the argument by which we proved Theorem 16.6.1 yields the relation

$$p_\infty(\beta) = \lim_{N\to\infty} p_N(\beta) = F(\beta, m(\beta), q(\beta)) \; .$$

Since $q(\beta)$ satisfies (16.128) and hence (16.127), differentiation in β yields

$$p'_\infty(\beta) = \frac{\partial F}{\partial \beta}(\beta, m(\beta), q(\beta)) + m'(\beta)\frac{\partial F}{\partial m}(\beta, m(\beta), q(\beta)) \; . \tag{16.129}$$

Now

$$\frac{\partial}{\partial \beta} p'_N(\beta, \boldsymbol{\beta}) = \beta(1 - \nu(R^p_{1,2}))$$

and since $R_{1,2}$ takes essentially only the values 0 and q, with $\nu(\mathbf{1}_{\{R_{1,2}=q\}}) = 1 - m$, we expect that

$$p'_\infty(\beta) = \beta(1 - (1-m)q^p) \; . \tag{16.130}$$

On the other hand, since $Y = \beta z \sqrt{qp^{q-1}}$

$$\frac{\partial F}{\partial \beta}(\beta, m, q) = \beta\big(1 - pq^{p-1} + (p-1)q^p(1-m)\big) + \frac{\mathsf{E} z\sqrt{pq^{p-1}}\mathrm{th}Y\mathrm{ch}^m Y}{\mathsf{E}\mathrm{ch}^m Y} \; , \tag{16.131}$$

and integration by parts yields

$$\begin{aligned}\mathsf{E} z\sqrt{pq^{p-1}}\mathrm{th}Y\mathrm{ch}^m Y &= \beta pq^{p-1}\mathsf{E}(\mathrm{ch}^{m-2}Y + m\,\mathrm{th}^2 Y\mathrm{ch}^m Y)\\ &= \beta pq^{p-1}(\mathsf{E}\mathrm{ch}^m Y + (m-1)\mathsf{E}\mathrm{th}^2 Y\mathrm{ch}^m Y) \; .\end{aligned}$$

Using (16.128) and (16.131) we get

$$\frac{\partial F}{\partial \beta}(\beta, m, q) = \beta(1 - (1-m)q^p) \; .$$

Comparing with (16.129) and (16.130) forces the relation

$$m'(\beta)\frac{\partial F}{\partial m}(\beta, m(\beta), q(\beta)) = 0 \; .$$

It is most unlikely that $m'(\beta) = 0$, because in the relation (16.130) $p'_\infty(\beta)$ is close to $\sqrt{2\log 2}$ and q is close to 1, so that $m(\beta) \simeq \sqrt{2\log 2}/\beta$. Therefore it must be true that (16.127) holds.

16.7 A Research Problem: The Dynamical Transition

The problems raised in this section concern not only the case p even, but in a slightly different form, the case where p is odd, which is much better understood. Recalling the notation $p_\infty(\beta)$ of Theorem 16.6.1, consider

$$\gamma_p^* = \sup\left\{\beta\ ;\ p_\infty(\beta) = \frac{\beta^2}{2} + \log 2\right\}\ .$$

We expect (and this should be easy to prove) that

$$\gamma_p < \gamma_p^* < \sqrt{2\log 2}\ .$$

We also expect that $\lim_{N\to\infty} \mathsf{E}\langle R_{1,2}^2\rangle = 0$ for $\beta < \gamma_p^*$. Our last exercise is a bit challenging.

Exercise 16.7.1. Prove that given $\beta < \gamma_p^*$ then for N large enough we have

$$\mathsf{E}\log G_N^{\otimes 2}\{(R_{1,2} \geq 1 - 2^{-p/2}\}) \geq -8N\gamma_p^*(\gamma_p^* - \beta)\ . \tag{16.132}$$

(Hint: think of the left-hand side as a function of β and compute its derivative. What happens for $\beta > \gamma_p$?)

On the other hand, it follows from (16.44) and (16.45) that there exists a number K_0, depending on p only, such that for $\beta \leq \gamma_p^*$, setting $a' = 2^{-p/4}$ and $a = 2^{-p/2-1}$ we have

$$\mathsf{E}G_N^{\otimes 2}(\{|R_{1,2}| \geq a'\ ;\ R_{1,2} \leq 1 - 2a\}) \leq K\exp(-N/K)\ . \tag{16.133}$$

Comparing (16.132) and (16.133) we expect that when β is close to γ_p^* then for the typical disorder the quantities

$$U = \log G_N^{\otimes 2}\{(R_{1,2} \geq 1 - 2a\})$$

and

$$\varepsilon = G_N^{\otimes 2}(\{|R_{1,2}| \geq a'\ ;\ R_{1,2} \leq 1 - 2a\})$$

satisfy

$$U \geq \exp(NK_0/2)\varepsilon^{1/3}\ .$$

When this is the case, the construction of Theorem 16.4.1 gives a non trivial result. Even though all the lumps it constructs might have a Gibbs measure that is exponentially small in N, the "gaps" between the lumps have much smaller Gibbs measure than the lumps themselves.

Research Problem 16.7.2. Describe, for the typical disorder, the sequence of weights of the lumps thus constructed.

The physicists predict that there is a "dynamical transition temperature" $\gamma_{p,d} < \gamma_p^*$ such that for $\gamma_{p,d} < \beta < \gamma_p^*$ the Gibbs measure "decomposes in a union of exponentially many small lumps roughly of the same size". There is little hard evidence that this is indeed the case, since as usual the physicists (quite sensibly) assume from the beginning that "matters should be as simple as consistent with the known facts" (where "known facts" seem largely determined by consensus of the experts). The following could be simpler than Problem 16.7.2, and would provide support for the physicists beliefs.

Research Problem 16.7.3. Prove that if p is large, there exists $\gamma_p^\sim < \gamma_p^*$ and $\delta > 0$ such that for $\gamma_p^\sim < \beta$, for the typical disorder we have

$$G_N\big(\big\{\boldsymbol{\sigma}^1 \ ; \ G_N(\{\boldsymbol{\sigma}^2 \ ; \ R_{1,2} \ge 1 - 2^{-p/2}\}) \ge \exp(-N\delta)\}\big)\big\} \big) \simeq 1 \ .$$

The following is closely related to Problem 1.12.11.

Research Problem 16.7.4. Consider the function

$$\varphi(\boldsymbol{\sigma}^1) = \frac{1}{N} \log G_N(\{\boldsymbol{\sigma}^2 \ ; \ R_{1,2} \ge 1 - 2^{-p/2}\}) \ .$$

Is it true the for any β we have

$$\lim_{N \to \infty} \mathsf{E}\big\langle \big(\varphi(\boldsymbol{\sigma}) - \mathsf{E}\langle \varphi(\boldsymbol{\sigma})\rangle\big)^2 \big\rangle = 0 \ ?$$

The physicists calculations towards Problem 16.7.2 assume beforehand a positive answer to Problem 16.7.4 (among many other things).

Research Problem 16.7.5. Find a rigorous definition of the transition value $\gamma_{p,d}$.

16.8 Notes and Comments

It is known how to obtain results comparable to those of Section 16.5 without doing adding a perturbation term to the Hamiltonian, but the proof is then very much more difficult [104]. In the same paper it is also shown (when adding the perturbation term to the Hamiltonian) how to handle the case of small external field.

A. Appendix: Elements of Probability Theory

A.1 How to Use This Appendix

This appendix reproduces some parts of the appendix of Volume I which should be useful here.

A.2 Gaussian Random Variables

A (centered) Gaussian r.v. g has a density of the type

$$\frac{1}{\sqrt{2\pi}\tau}\exp\left(-\frac{t^2}{2\tau^2}\right)$$

so that $\mathsf{E}\, g^2 = \tau^2$. When $\tau = 1$, g is called standard Gaussian. We hardly ever use non-centered Gaussian r.v., so that the expression "consider a Gaussian r.v. z" means "consider a centered Gaussian r.v. z". A fundamental fact is that

$$\mathsf{E}\exp ag = \exp\frac{a^2\tau^2}{2}\ . \tag{A.1}$$

Indeed,

$$\begin{aligned}
\mathsf{E}\exp ag &= \frac{1}{\sqrt{2\pi}\tau}\int_{-\infty}^{\infty}\exp\left(at - \frac{t^2}{2\tau^2}\right)\mathrm{d}t \\
&= \left(\exp\frac{a^2\tau^2}{2}\right)\frac{1}{\sqrt{2\pi}\tau}\int_{-\infty}^{\infty}\exp\left(-\frac{(t-a\tau^2)^2}{2\tau^2}\right)\mathrm{d}t \\
&= \exp\frac{a^2\tau^2}{2}\ .
\end{aligned}$$

For a r.v. $Y \geq 0$ and $s > 0$ we have Markov's inequality

$$\mathsf{P}(Y \geq s) \leq \frac{1}{s}\mathsf{E}Y\ . \tag{A.2}$$

Using this for $Y = \exp(\lambda X)$, where X is any r.v., we obtain for any $\lambda \geq 0$ the following fundamental inequality:

M. Talagrand, *Mean Field Models for Spin Glasses*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys in Mathematics 55, DOI 10.1007/978-3-642-22253-5,

$$\mathsf{P}(X \geq t) = \mathsf{P}(\exp(\lambda X) \geq e^{\lambda t}) \leq e^{-\lambda t}\mathsf{E}\exp(\lambda X) \; . \tag{A.3}$$

Changing X into $-X$ and t into $-t$, we get the following equally useful fact:

$$\mathsf{P}(X \leq t) \leq e^{\lambda t}\mathsf{E}\exp(-\lambda X) \; .$$

Combining (A.1) with (A.3) we get for any $t \geq 0$ that

$$\mathsf{P}(g \geq t) \leq \exp\Big(-\lambda t + \frac{\lambda^2\tau^2}{2}\Big) \; ,$$

and taking $\lambda = t/\tau^2$

$$\mathsf{P}(g \geq t) \leq \exp\left(-\frac{t^2}{2\tau^2}\right) \; . \tag{A.4}$$

Elementary estimates show that for $t > 0$ we have, for some number L,

$$\mathsf{P}(g \geq t) \geq \frac{1}{L(1+t/\tau)}\exp\left(-\frac{t^2}{2\tau^2}\right) \; . \tag{A.5}$$

There is of course a more precise understanding of the tails of g than (A.4) and (A.5); but (A.4) and (A.5) will mostly suffice here. Another fundamental formula is that when $\mathsf{E}g^2 = \tau^2$ then for $2a\tau^2 < 1$ and any b we have

$$\mathsf{E}\exp(ag^2 + bg) = \frac{1}{\sqrt{1-2a\tau^2}}\exp\frac{\tau^2 b^2}{2(1-2a\tau^2)} \; . \tag{A.6}$$

Indeed,

$$\mathsf{E}\exp(ag^2 + bg) = \frac{1}{\sqrt{2\pi}\tau}\int_{-\infty}^{\infty}\exp\left(at^2 - \frac{t^2}{2\tau^2} + bt\right)\mathrm{d}t \; .$$

We then complete the squares by writing

$$at^2 - \frac{t^2}{2\tau^2} + bt = -\frac{1-2a\tau^2}{2\tau^2}\left(t - \frac{b\tau^2}{1-2a\tau^2}\right)^2 - \frac{b\tau^2}{2(1-2a\tau^2)}$$

and conclude by making the change of variable

$$t = \frac{b\tau^2}{1-2a\tau^2} + u\frac{\tau}{\sqrt{1-2a\tau^2}} \; .$$

The following is also important.

Lemma A.2.1. *Consider M Gaussian r.v.s $(g_i)_{i\leq M}$ with $\mathsf{E}g_i^2 \leq \tau$ for each $i \leq N$. We do NOT assume that they are independent. Then we have*

$$\mathsf{E}\max_{i\leq M} g_i \leq \tau\sqrt{2\log M} \; . \tag{A.7}$$

Proof. Consider $\beta > 0$. Using Jensen's inequality (1.23) as in (1.24) and (A.1) we have

$$\begin{aligned}\mathsf{E}\log\Big(\sum_{i\le M}\exp\beta g_i\Big) &\le \log\Big(\mathsf{E}\sum_{i\le M}\exp\beta g_i\Big)\\ &\le \log\left(M\exp\left(\frac{1}{2}\beta^2\tau^2\right)\right)\\ &= \frac{\beta^2\tau^2}{2}+\log M\ . \end{aligned} \tag{A.8}$$

Now

$$\beta\max_{i\le M} g_i \le \log\Big(\sum_{i\le M}\exp\beta g_i\Big)\ ,$$

so that, using (A.8),

$$\beta\mathsf{E}\max_{i\le M} g_i \le \mathsf{E}\log\Big(\sum_{i\le M}\exp\beta g_i\Big) \le \frac{\beta^2\tau^2}{2}+\log M\ .$$

Taking $\beta=\sqrt{2\log M}/\tau$ yields (A.7). □

Given independent standard Gaussian r.v.s $g_1,\ldots,g_M$, their joint law has density $(2\pi)^{-M/2}\exp(-\|\mathbf{x}\|^2/2)$, where $\|\mathbf{x}\|^2=\sum_{i\le M}x_i^2$. This density is invariant by rotation, and, as a consequence, the law of every linear combination $z=\sum_{i\le M}a_ig_i$ is Gaussian. The set $\mathcal{G}$ of these linear combinations is a vector space, each element of which is a Gaussian r.v. Such a space is often called a Gaussian space. It has a natural dot product, given with obvious notation by $\mathsf{E}\,zz'=\sum_{k\le M}a_ka'_k$. Given two linear subspaces F_1,F_2 of F, if these spaces are orthogonal, i.e. $\mathsf{E}\,z_1z_2=0$ whenever $z_1\in F_1$, $z_2\in F_2$, they are probabilistically independent. This is obvious from rotational invariance, since after a suitable rotation these spaces are spanned by two disjoint subsets of $g_1,\ldots,g_M$.

We say that a family $z_1,\ldots,z_N$ of r.v.s is jointly Gaussian if the law of every linear combination $\sum_{k\le N}a_kz_k$ is Gaussian. If $z_1,\ldots,z_N$ belong to a Gaussian space $\mathcal{G}$ as above, then obviously the family $z_1,\ldots,z_N$ is jointly Gaussian. All the jointly Gaussian families considered in this book will obviously be of this type, since they are defined by explicit formulas such as $z_k=\sum_{i\le M}a_{k,i}g_i$ where $g_1,\ldots,g_M$ are independent standard Gaussian r.v.s, a formula that we abbreviate by $z_k=\mathbf{g}\cdot\mathbf{a}_k$ where $\mathbf{g}=(g_1,\ldots,g_M)$, $\mathbf{a}_k=(a_{k,1},\ldots,a_{k,M})$ and $\cdot$ denotes the dot product in $\mathbb{R}^M$. For the beauty of it, let us mention that, in distribution, any jointly Gaussian family $z_1,\ldots,z_N$ can be represented as above as $z_k=\mathbf{a}_k\cdot\mathbf{g}$ (with $M=N$). This is simply because the joint law of a jointly Gaussian family $z_1,\ldots,z_k$ is determined by the numbers $\mathsf{E}z_kz_\ell$, so that it suffices to find the vectors $\mathbf{a}_k$ in such a manner that $\mathsf{E}z_kz_\ell=\mathbf{a}_k\cdot\mathbf{a}_\ell$. If we think of the linear span of the r.v.s $z_1,\ldots,z_N$ provided with the dot product $z \cdot z' = \mathsf{E} zz'$ as an Euclidean space, and of $z_1, \ldots, z_N$ as points in this space, they provide exactly such a family of vectors.

Another interesting fact is the following. If $(q_{u,v})_{u,v \le n}$ is a symmetric positive definite matrix, there exists jointly Gaussian r.v.s $(Y_u)_{u \le n}$ such that $\mathsf{E}\, Y_u Y_v = q_{u,v}$. This is obvious when the matrix $(q_{u,v})$ is diagonal; the general case follows from the fact that a symmetric matrix diagonalizes in an orthogonal basis.

A.3 Gaussian Integration by Parts

Given a continuously differentiable function F on $\mathbb{R}$ (that satisfies the growth condition at infinity stated below in (A.10)) and a centered Gaussian r.v. g we have the integration by parts formula

$$\mathsf{E}\, gF(g) = \mathsf{E}\, g^2 \mathsf{E}\, F'(g) \; . \tag{A.9}$$

To see this, if $\mathsf{E}\, g^2 = \tau^2$, we have

$$\begin{aligned} \mathsf{E}\, gF(g) &= \frac{1}{\sqrt{2\pi}\tau} \int_{\mathbb{R}} t \exp\left(-\frac{t^2}{2\tau^2}\right) F(t) \mathrm{d}t \\ &= \frac{\tau^2}{\sqrt{2\pi}\tau} \int_{\mathbb{R}} \exp\left(-\frac{t^2}{2\tau^2}\right) F'(t) \mathrm{d}t \\ &= \mathsf{E}\, g^2 \mathsf{E}\, F'(g) \end{aligned}$$

provided

$$\lim_{|t| \to \infty} F(t) \exp(-t^2/2\tau^2) = 0 \; . \tag{A.10}$$

This formula is used over and over in this work. As a first application, if $\mathsf{E} g^2 = \tau^2$ and $2a\tau^2 < 1$ we have

$$\mathsf{E} g^2 \exp ag^2 = \mathsf{E} g(g \exp ag^2) = \tau^2 (\mathsf{E} \exp ag^2 + \mathsf{E}\; 2ag \exp ag^2) \; , \tag{A.11}$$

so that

$$(1 - 2a\tau^2) \mathsf{E} g \exp ag^2 = \tau^2 \mathsf{E} \exp ag^2 = \tau \frac{1}{\sqrt{1 - 2a\tau^2}}$$

by (A.6) and $\mathsf{E} g \exp ag^2 = \tau^2 (1 - 2a\tau^2)^{-3/2}$. As another application, if $k \ge 2$

$$\mathsf{E} g^k = \mathsf{E} g g^{k-1} = \tau^2 (k-1) \mathsf{E} g^{k-2} \; ,$$

so that in particular $\mathsf{E} g^4 = 3\tau^2$, and one can recursively compute all the moments of g. All kinds of Gaussian integrals can be computed effortlessly in this manner.

Condition (A.10) holds in particular if F is of moderate growth in the sense that $\lim_{|t| \to \infty} F(t) \exp(-at^2) = 0$ for each $a > 0$. A function F (with a regular behavior as will be the case of all the functions we consider) fails to be of moderate growth if "it grows as fast as $\exp(at^2)$ for some $a > 0$". The functions to which we will apply the integration by parts formula typically do not "grow faster than $\exp(At)$" for a certain number A (except in the case of certain very explicit functions such as in (A.11)).

Formula (A.9) generalizes as follows. Given g, $z_1, \ldots, z_n$ in a Gaussian space $\mathcal{G}$, and a function F of n variables (with a moderate behavior at infinity to be stated in (A.13) below), we have

$$\mathsf{E} gF(z_1, \ldots, z_n) = \sum_{\ell \le n} \mathsf{E}(gz_\ell) \mathsf{E} \frac{\partial F}{\partial x_\ell}(z_1, \ldots, z_n) . \tag{A.12}$$

This is probably the single most important formula in this work. For a proof, consider the r.v.s

$$z'_\ell = z_\ell - g \frac{\mathsf{E}\, z_\ell g}{\mathsf{E}\, g^2} .$$

They satisfy $\mathsf{E}\, z'_\ell g = 0$; thus g is independent of the family $(z'_1, \ldots, z'_n)$. We then apply (A.9) at $(z'_\ell)_{\ell \le n}$ given. Since $z_\ell = z'_\ell + g \mathsf{E}\, g z_\ell / \mathsf{E}\, g^2$, (A.12) follows whenever the following is satisfied to make the use of (A.9) legitimate (and to allow the interchange of the expectation in z and in the family $(z'_1, \ldots, z'_n)$): for each number $a > 0$, we have

$$\lim_{\|\mathbf{x}\| \to \infty} |F(\mathbf{x})| \exp(-a\|\mathbf{x}\|^2) = 0 . \tag{A.13}$$

A.4 Tail Estimates

We recall that given any r.v. X and $\lambda > 0$, by (A.3) we have

$$\mathsf{P}(X \ge t) \le e^{-\lambda t} \mathsf{E} \exp \lambda X .$$

If $X = \sum_{i \le N} X_i$ where $(X_i)_{i \le N}$ are independent, then

$$\mathsf{E} \exp \lambda X = \prod_{i \le N} \mathsf{E} \exp \lambda X_i ,$$

so that

$$\mathsf{P}(X \ge t) \le e^{-\lambda t} \prod_{i \le N} \mathsf{E} \exp \lambda X_i = \exp\Bigl(-\lambda t + \sum_{i \le N} \log \mathsf{E} \exp \lambda X_i\Bigr) . \tag{A.14}$$

If $(\eta_i)_{i \le N}$ are independent Bernoulli r.v.s, i.e. $\mathsf{P}(\eta_i = \pm 1) = 1/2$, then $\mathsf{E} \exp \lambda\, a_i \eta_i = \operatorname{ch} \lambda a_i$, and thus

$$\mathsf{P}\Bigl(\sum_{i \le N} a_i \eta_i \ge t\Bigr) \le \exp\Bigl(-\lambda t + \sum_{i \le N} \log \operatorname{ch} \lambda a_i\Bigr) . \tag{A.15}$$

It is obvious on power series expansions that $\operatorname{ch} t \le \exp(t^2/2)$, so that

$$\mathsf{P}\Big(\sum_{i\le N} a_i\eta_i \ge t\Big) \le \exp\Big(-\lambda t + \frac{\lambda^2}{2}\sum_{i\le N} a_i^2\Big)\,,$$

and by optimization over λ, for all $t \ge 0$,

$$\mathsf{P}\Big(\sum_{i\le N} a_i\eta_i \ge t\Big) \le \exp\Big(-\frac{t^2}{2\sum_{i\le N} a_i^2}\Big)\,. \tag{A.16}$$

This inequality is often called the subgaussian inequality. By symmetry, $\mathsf{P}\big(\sum_{i\le N} a_i\eta_i \le -t\big)$ is bounded by the same expression, so that

$$\mathsf{P}\Big(\Big|\sum_{i\le N} a_i\eta_i\Big| \ge t\Big) \le 2\exp\Big(-\frac{t^2}{2\sum_{i\le N} a_i^2}\Big)\,. \tag{A.17}$$

As a consequence of (A.16) we have the following

$$\operatorname{card}\{(\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2)\in \Sigma_N^2\ ;\ R_{1,2}\ge t\} \le 2^{2N}\exp\left(-\frac{Nt^2}{2}\right)\,. \tag{A.18}$$

This is seen by taking $a_i = 1/N$, by observing that for the uniform measure on Σ_N^2 the sequence $\eta_i = \sigma_i^1\sigma_i^2$ is an independent Bernoulli sequence and that $R_{1,2} = \sum_{i\le N} a_i\eta_i$. Related to (A.16) is the fact that

$$\mathsf{E}\exp\frac{1}{2}\Big(\sum_{i\le N} a_i\eta_i\Big)^2 \le \frac{1}{\sqrt{1-\sum_{i\le N} a_i^2}}\,. \tag{A.19}$$

Equivalently,

$$\sum \exp\frac{1}{2}\Big(\sum_{i\le N} a_i\sigma_i\Big)^2 \le \frac{2^N}{\sqrt{1-\sum_{i\le N} a_i^2}}\,,$$

where the summation is over all sequences $(\sigma_i)_{i\le N}$ with $\sigma_i = \pm 1$. To prove (A.19) we consider a standard Gaussian r.v. g independent of the r.v.s η_i and, using (A.1), we have, denoting by E_g expectation in g only, and using again that $\log\operatorname{ch} t \le t^2/2$,

$$\begin{aligned}\mathsf{E}\exp\frac{1}{2}\Big(\sum_{i\le N} a_i\eta_i\Big)^2 &= \mathsf{E}\,\mathsf{E}_g\exp\sum_{i\le N} g a_i\eta_i\\ &= \mathsf{E}_g\exp\sum_{i\le N}\log\operatorname{ch} g a_i\\ &\le \mathsf{E}_g\exp\frac{g^2}{2}\sum_{i\le N} a_i^2\\ &= \frac{1}{\sqrt{1-\sum_{i\le N} a_i^2}}\,.\end{aligned}$$

It follows from (A.19) that if $S = \sum_{i\le N} a_i^2$, then, if $b_i = a_i/\sqrt{2S}$, we have $\sum_{i\le N} b_i^2 = 1/2$ and

$$\mathsf{E}\exp \frac{1}{4S}\Big(\sum_{i\le N} a_i\eta_i\Big)^2 = \mathsf{E}\exp\frac{1}{2}\Big(\sum_{i\le N} b_i\eta_i\Big)^2 \le \frac{1}{\sqrt{1/2}} \le 2 \,.$$

Since $\exp x \ge x^n/n! \ge x^n/n^n$ for each n and $x \ge 0$ we see that

$$\mathsf{E}\Big(\sum_{i\le N} a_i\eta_i\Big)^{2n} \le 2(4n)^n S^n = 2(4n)^n\Big(\sum_{i\le N} a_i^2\Big)^n \,, \tag{A.20}$$

a relation known as Khinchin's inequality.

Going back to (A.15), if $a_i = 1$ for each $i \le N$, changing t into Nt, we get

$$\mathsf{P}\Big(\sum_{i\le N}\eta_i \ge Nt\Big) \le \exp N(-\lambda t + \log\operatorname{ch}\lambda) \,.$$

If $0 \le t < 1$, the exponent is minimized for $\operatorname{th}\lambda = t$, i.e.

$$\frac{e^{\lambda}-e^{-\lambda}}{e^{\lambda}+e^{-\lambda}} = \frac{e^{2\lambda}-1}{e^{2\lambda}+1} = t \,,$$

so that $e^{2\lambda} = (1+t)/(1-t)$ and

$$\lambda = \frac{1}{2}(\log(1+t) - \log(1-t)) \,.$$

Also, $\operatorname{ch}^{-2}\lambda = 1 - \operatorname{th}^2\lambda = 1 - t^2$, so that

$$\log\operatorname{ch}\lambda = -\frac{1}{2}\log(1-t^2) \,,$$

and

$$\begin{aligned}\min_{\lambda}\,(-\lambda t + \log\operatorname{ch}\lambda) &= -\frac{1}{2}(t\log(1+t) - t\log(1-t)) \\ &\quad -\frac{1}{2}\log(1-t) - \frac{1}{2}\log(1+t) \\ &= -\mathcal{I}(t) \end{aligned} \tag{A.21}$$

where

$$\mathcal{I}(t) = \frac{1}{2}((1+t)\log(1+t) + (1-t)\log(1-t)) \,. \tag{A.22}$$

The function $\mathcal{I}(t)$ is probably better understood by noting that

$$\mathcal{I}(0) = \mathcal{I}'(0) = 0\,,\ \mathcal{I}''(t) = \frac{1}{1-t^2} \,. \tag{A.23}$$

It follows from (A.21) that

$$\mathsf{P}\Big(\sum_{i\le N}\eta_i \ge Nt\Big) \le \exp(-N\mathcal{I}(t)) ,$$

or, equivalently, that

$$\mathrm{card}\Big\{\boldsymbol{\sigma}\in\Sigma_N\,;\ \sum_{i\le N}\sigma_i \ge tN\Big\} \le 2^N\exp(-N\mathcal{I}(t)) . \tag{A.24}$$

If k is an integer, then $\sum_{i\le N}\sigma_i = k$ exactly when the sequence $(\sigma_i)_{i\le N}$ contains $(N+k)/2$ times 1 and $(N-k)/2$ times -1. This is impossible when $N+k$ is odd. When $N+k$ is even, using Stirling's formula $n! \sim n^n e^{-n}\sqrt{2\pi n}$, we obtain

$$\begin{aligned}\mathrm{card}\Big\{\boldsymbol{\sigma}\in\Sigma_N\,;\ \sum_{i\le N}\sigma_i = k\Big\} &= \binom{N}{\frac{N+k}{2}} = \frac{N!}{\left(\frac{N+k}{2}\right)!\left(\frac{N-k}{2}\right)!}\\
&\ge \frac{1}{L}\frac{\sqrt{N}}{\sqrt{(N-k)(N+k)}}\frac{N^N}{\left(\frac{N+k}{2}\right)^{(N+k)/2}\left(\frac{N-k}{2}\right)^{(N-k)/2}}\\
&\ge \frac{2^N}{L\sqrt{N}}\frac{1}{\left(1+\frac{k}{N}\right)^{(N+k)/2}\left(1-\frac{k}{N}\right)^{(N-k)/2}}\\
&= \frac{2^N}{L\sqrt{N}}\exp\left(-N\mathcal{I}\left(\frac{k}{N}\right)\right) .\end{aligned} \tag{A.25}$$

This reverses the inequality (A.24) within the factor $L\sqrt{N}$.

Since by Lemma 4.3.5 the function $t\mapsto \log\mathrm{ch}\sqrt{t}$ is concave, it follows from (A.15) that

$$\mathsf{P}\Big(\sum_{i\le N}a_i\eta_i \ge t\sqrt{N}\Big) \le \exp N\Big(-\lambda t + \log\mathrm{ch}\,\lambda\sqrt{\sum_{i\le N}a_i^2}\Big)$$

and, using (A.21)

$$\mathsf{P}\Big(\sum_{i\le N}a_i\eta_i \ge t\sqrt{N}\Big) \le \exp\Bigg(-N\mathcal{I}\Bigg(\frac{t}{\sqrt{\sum_{i\le N}a_i^2}}\Bigg)\Bigg) . \tag{A.26}$$

A.5 How to Use Tail Estimates

It will often occur that for a r.v. X, we know an upper bound for the probabilities $\mathsf{P}(X\ge t)$, and that we want to deduce an upper bound for $\mathsf{E}F(X)$ for a certain function F. For example, if Y is a r.v., $Y\ge 0$, then

$$\mathsf{E}Y = \int_0^\infty \mathsf{P}(Y\ge t)\,\mathrm{d}t , \tag{A.27}$$

using Fubini theorem to compute the "area under the graph of Y".

More generally, if $X \geq 0$ and F is a continuously differentiable non-decreasing function on $\mathbb{R}^+$ we have

$$F(X) = F(0) + \int_0^X F'(t)\mathrm{d}t = F(0) + \int_{\{t \leq X\}} F'(t)\mathrm{d}t \ .$$

Taking expectation, and using Fubini's theorem to exchange the integral in t and the expectation, we get that

$$\mathsf{E}F(X) = F(0) + \int_0^\infty F'(t)\mathsf{P}(X \geq t)\mathrm{d}t \ . \tag{A.28}$$

For a typical application of (A.28) let us assume that X satisfies the following tail condition:

$$\forall t \geq 0 \ , \ \mathsf{P}(|X| \geq t) \leq 2\exp\left(-\frac{t^2}{2A^2}\right) \ , \tag{A.29}$$

where A is a certain number. Then, using (A.28) for $F(x) = x^k$ and $|X|$ instead of X we get

$$\mathsf{E}|X|^k \leq 2k\int_0^\infty t^{k-1}\exp\left(-\frac{t^2}{2A^2}\right)\mathrm{d}t \ .$$

The right-hand side can be recursively computed by integration by parts. If $k \geq 3$,

$$\int_0^\infty t^{k-1}\exp\left(-\frac{t^2}{2A^2}\right)\mathrm{d}t = (k-2)A^2\int_0^\infty t^{k-3}\exp\left(-\frac{t^2}{2A^2}\right)\mathrm{d}t.$$

In this manner one obtains e.g.

$$\mathsf{E}X^{2k} \leq 2^{k+1}k!A^{2k} \ .$$

This shows in particular that "the moments of order k of X grow at most like $\sqrt{k}$." Indeed, using the crude inequality $k! \leq k^k$ we obtain

$$(\mathsf{E}|X|^k)^{1/k} \leq (\mathsf{E}X^{2k})^{1/2k} \leq 2A\sqrt{k} \ . \tag{A.30}$$

Suppose, conversely, that for a given r.v. X we know that for a certain number B and any $k \geq 1$ we have $\mathsf{E}X^{2k} \leq B^{2k}k^k$ (i.e. an inequality of the type (A.30) for even moments). Then, using the power expansion $\exp x^2 = \sum_{k\geq 0} x^{2k}/k!$, for any number C we have

$$\mathsf{E}\exp\frac{X^2}{C^2} = \sum_{k\geq 0}\frac{\mathsf{E}X^{2k}}{C^{2k}k!} \leq \sum_{k\geq 0}\frac{B^{2k}k^k}{C^{2k}k!} \ .$$

Now, by Stirling's formula, there is a constant L_0 such that $k^k \leq L_0^k k!$, and therefore there is a number L (e.g. $L = 2L_0$) such that

$$\mathsf{E} \exp \frac{X^2}{LB^2} \le 2 \ .$$

This implies in turn that

$$\mathsf{P}(X \ge t) \le 2 \exp\left(-\frac{t^2}{LB^2}\right) .$$

Many r.v.s considered in this work satisfy the condition (A.29). The previous considerations explain why, when convenient, we control these r.v.s through their moments.

If F is a continuously differentiable non-decreasing function on $\mathbb{R}$, $F \ge 0$, $F(-\infty) = 0$, we have

$$F(X) = \int_{-\infty}^{X} F'(t)\mathrm{d}t = \int_{\{t \le X\}} F'(t)\mathrm{d}t \ .$$

Taking expectation, and using again Fubini's theorem to exchange the integral in t and the expectation, we get now that

$$\mathsf{E}\, F(X) = \int_{-\infty}^{\infty} F'(t)\, \mathsf{P}(X \ge t)\, \mathrm{d}t \ . \tag{A.31}$$

This no longer assumes that $X \ge 0$. Considering now $a < b$ we have

$$\mathsf{E}(F(\min(X,b))\mathbf{1}_{\{X \ge a\}}) = F(a)\mathsf{P}(X \ge a) + \int_a^b F'(t)\mathsf{P}(X \ge t)\mathrm{d}t \ . \tag{A.32}$$

This is seen by using (A.31) for the conditional probability that $X \ge a$, and for the r.v. $\min(X,b)$ instead of X.

A.6 Bernstein's Inequality

Theorem A.6.1. *Consider a r.v. X with $\mathsf{E}X = 0$ and an independent sequence $(X_i)_{i \le N}$ distributed like X. Assume that, for a certain number A, we have*

$$\mathsf{E} \exp \frac{|X|}{A} \le 2 \ . \tag{A.33}$$

Then, for all $t > 0$ we have

$$\mathsf{P}\left(\sum_{i \le N} X_i \ge t\right) \le \exp\left(-\min\left(\frac{t^2}{4NA^2}, \frac{t}{2A}\right)\right) \tag{A.34}$$

$$\mathsf{P}\left(\sum_{i \le N} X_i \ge t\right) \le \exp\left(-\frac{t^2}{2N\mathsf{E}X^2}\left(1 - \frac{4A^3 t}{N(\mathsf{E}X^2)^2}\right)\right) \ . \tag{A.35}$$

Proof. From (A.14) we obtain

$$\mathsf{P}\Big(\sum_{i\le N} X_i \ge t\Big) \le \exp(-\lambda t + N\log \mathsf{E}\exp \lambda X)\ . \tag{A.36}$$

We have

$$\mathsf{E}\exp\lambda X = 1 + \mathsf{E}\,\varphi(\lambda X) \tag{A.37}$$

where $\varphi(x) = e^x - x - 1$. We observe that $\mathsf{E}\varphi(|X|/A) \le \mathsf{E}\exp(|X|/A) - 1 = 2 - 1 = 1$. Now power series expansion yields that $\varphi(x) \le \varphi(|x|)$ and that for $x > 0$, the function $\lambda \to \varphi(\lambda x)/\lambda^2$ increases. Thus, for $\lambda \le 1/A$, we have

$$\mathsf{E}\,\varphi(\lambda X) \le \lambda^2 A^2 \mathsf{E}\,\varphi(|X|/A) \le \lambda^2 A^2\ .$$

Combining (A.37) with the inequality $\log(1+x) \le x$, we obtain $\log\mathsf{E}\exp\lambda X \le \lambda^2 A^2$. Consequently (A.36) implies

$$\mathsf{P}\Big(\sum_{i\le N} X_i \ge t\Big) \le \exp(-\lambda t + N\lambda^2 A^2)\ .$$

We choose $\lambda = t/2NA^2$ if $t \le 2NA$ (so that $\lambda \le 1/A$). When $t \ge 2NA$, we choose $\lambda = 1/A$, and then

$$-\lambda t + N\lambda^2 A^2 = -\frac{t}{A} + N \le -\frac{t}{2A}\ .$$

This proves (A.34). To prove (A.35) we replace (A.37) by

$$\mathsf{E}\exp\lambda X = 1 + \frac{\lambda^2 \mathsf{E} X^2}{2} + \mathsf{E}\,\varphi_1(\lambda X)$$

where $\varphi_1(x) = e^x - x^2/2 - x - 1$. We observe that $\mathsf{E}\varphi_1(|X|/A) \le \mathsf{E}\varphi(|X|/A) \le 1$. Using again power series expansion yields $\varphi_1(x) \le \varphi_1(|x|)$ and that for $x > 0$ the function $\lambda \mapsto \varphi_1(\lambda x)/\lambda^3$ increases. Thus, if $\lambda \le 1/A$, we get

$$\mathsf{E}\,\varphi_1(\lambda X) \le \lambda^3 A^3 \mathsf{E}\,\varphi_1(|X|/A) \le \lambda^3 A^3$$

so that $\log\mathsf{E}\exp\lambda X \le \lambda^2\mathsf{E}X^2/2 + \lambda^3 A^3$ and we choose $\lambda = t/N\mathsf{E}X^2$ to obtain (A.35) when $t \le N\mathsf{E}X^2/A$. When $t \ge N\mathsf{E}X^2/A$, then

$$\frac{4A^3 t}{N(\mathsf{E}X^2)^2} \ge \frac{4A^2}{\mathsf{E}X^2} \ge 1$$

because $\mathsf{E}X^2/2A^2 \le \mathsf{E}\exp|X|/A \le 2$. Thus (A.35) is automatically satisfied in that case since the right-hand side is ≥ 1. □

Another important version of Bernstein's inequality assumes that

$$|X| \le A\ . \tag{A.38}$$

In that case for $p \geq 2$ we have $\mathsf{E} X^p \leq A^{p-2} \mathsf{E} X^2$, so that when $\lambda \leq 1$, and since $\sum_{p \geq 2} 1/p! = e - 2 \leq 1$,

$$\mathsf{E}\, \varphi(\lambda X) = \sum_{p \geq 2} \frac{\lambda^p}{p!} \mathsf{E}\, X^p \leq \lambda^2\, \mathsf{E}\, X^2 \sum_{p \geq 2} \frac{(\lambda A)^{p-2}}{p!} \leq \lambda^2 \mathsf{E}\, X^2 \,.$$

Proceeding as before, and taking now $\lambda = \min(t/\mathsf{E}\, X^2, 1/A)$, we get

$$\mathsf{P}\Big(\sum_{i \leq N} X_i \geq t\Big) \leq \exp\left(-\min\left(\frac{t^2}{4N\mathsf{E}\, X^2}, \frac{t}{2A}\right)\right) \,. \tag{A.39}$$

We will also need a version of (A.34) for martingale difference sequences. Assume that we are given an increasing sequence $(\Xi_i)_{0 \leq i \leq N}$ of σ-algebras. A sequence $(X_i)_{1 \leq i \leq N}$ is called a martingale difference sequence if X_i is Ξ_i-measurable and $\mathsf{E}_{i-1}(X_i) = 0$, where E_{i-1} denotes conditional expectation given Ξ_{i-1}. Let us assume that for a certain number A we have

$$\forall\, i \leq N\,,\ \mathsf{E}_{i-1} \exp \frac{|X_i|}{A} \leq 2\,. \tag{A.40}$$

Exactly as before, this implies that for $|\lambda|\, A \leq 1$ we have $\mathsf{E}_{i-1} \exp \lambda X_i \leq \exp \lambda^2 A^2$. Thus

$$\begin{aligned}\mathsf{E}_{k-1} \exp \lambda \sum_{i \leq k} X_i &= \exp\Big(\lambda \sum_{i \leq k-1} X_i\Big)\, \mathsf{E}_k \exp \lambda\, X_k \\ &\leq \exp\Big(\lambda \sum_{i \leq k-1} X_i + \lambda^2 A^2\Big)\,.\end{aligned}$$

By decreasing induction over k, this shows that for each k we have

$$\mathsf{E}_{k-1} \exp \lambda \sum_{i \leq N} X_i \leq \exp\Big(\lambda \sum_{i \leq k-1} X_i + (N - k + 1)\, \lambda^2 A^2\Big)\,.$$

Using this for $k = 1$ and taking expectation yields $\mathsf{E} \exp \lambda \sum_{i \leq N} X_i \leq \exp N \lambda^2 A^2$. Use of Chebyshev inequality as before gives

$$\mathsf{P}\Big(\sum_{i \leq N} X_i \geq t\Big) \leq \exp\left(-\min\left(\frac{t^2}{4NA^2}, \frac{t}{2A}\right)\right) \,. \tag{A.41}$$

A.7 ε-Nets

A *ball* of $\mathbb{R}^M$ is a convex balanced set with non-empty interior. The convex hull of a set A is denoted by conv$A$.

Proposition A.7.1. *Given a ball B of $\mathbb{R}^M$, we can find a subset A of B such that*

$$\operatorname{card} A \le \left(1 + \frac{1}{\varepsilon}\right)^M \tag{A.42}$$

$$\forall\, x \in B\,,\;\; A \cap (x + 2\,\varepsilon B) \ne \emptyset \tag{A.43}$$

$$\operatorname{conv} A \supset (1 - 2\,\varepsilon)B\;. \tag{A.44}$$

Moreover, given a linear functional φ on $\mathbb{R}^M$, we have

$$\sup_{x\in A} \varphi(x) \ge (1 - 2\,\varepsilon) \sup_{x\in B} \varphi(x)\;. \tag{A.45}$$

As a corollary, we can find a subset A of $(1-2\varepsilon)^{-1}B$ such that $\operatorname{card}A \le (1+\varepsilon^{-1})^M$ and $B \subset \operatorname{conv}A$. The case $\varepsilon = 1/4$ is of interest: $\operatorname{card}A \le 5^M$ and $\sup_{x\in A}\varphi(x) \ge (1/2)\sup_{x\in B}\varphi(x)$.

Proof. We simply take for A a maximal subset of B such that the sets $x+\varepsilon B$ are disjoint for $x \in A$. These sets are of volume $\varepsilon^M \mathrm{Vol}B$, and are entirely contained in the set $(1+\varepsilon)B$, which is of volume $(1+\varepsilon)^M \mathrm{Vol}B$. This proves (A.42).

Given x in B, we can find y in A with $(x+\varepsilon B)\cap(y+\varepsilon B) \ne \emptyset$, for otherwise this would contradict the maximality of A. Thus $y \in (x + 2\varepsilon B) \cap A$. This proves (A.43).

Using (A.43), given x in B, we can find y_0 in A with $x - y_0 \in 2\,\varepsilon B$. Applying this to $(x - y_0)/2\varepsilon$, we find y_1 in A with $x - y_0 - 2\varepsilon y_1 \in (2\varepsilon)^2 B$, and in this manner we find a sequence (y_i) in A with

$$y = \sum_{i\ge 0} (2\varepsilon)^i y_i \in (1 - 2\varepsilon)^{-1}\operatorname{conv}A\;,$$

since A is finite. This proves (A.44), of which (A.45) is an immediate consequence. □

A.8 Random Matrices

In this section we get some control of the norm of certain random matrices. Much more detailed (and difficult) results are known.

Lemma A.8.1. *If $(g_{ij})_{1\le i<j\le N}$ are independent standard Gaussian r.v.s, then, with probability at least $1 - L\exp(-N)$ we have*

$$\forall\,(x_i)_{i\le N}\,, \forall\,(y_i)_{i\le N}\,,\;\; \left|\sum_{i<j} g_{ij}\, x_i\, y_j\right| \le L\sqrt{N}\left(\sum_{i\le N} x_i^2 \sum_{i\le N} y_i^2\right)^{1/2}\;. \tag{A.46}$$

Proof. Let us denote by B the Euclidean ball of $\mathbb{R}^N$, and by A a subset of $2B$ with $\operatorname{card} A \le 5^N$ and $\operatorname{conv} A \supset B$, as provided by Proposition A.7.1. If $(x_i)_{i\le N}$ and $(y_i)_{i\le N}$ belong to A, then $\mathsf{E}\big(\sum_{i<j} g_{ij}\, x_i\, y_j\big)^2 \le \sum_{i\le N} x_i^2 \sum_{j\le N} y_j^2 \le 16$ and (A.4) implies

$$\mathsf{P}\left(\left|\sum_{i<j} g_{ij}\, x_i\, y_j\right| \ge t\right) \le 2\exp\left(-\frac{t^2}{32}\right) ,$$

so that with probability at least $1 - 2(25)^N \exp(-64N)$ it holds that

$$\forall\, (x_i)_{i\le N}\,,\ \forall\, (y_i)_{i\le N} \in A\,,\ \left|\sum_{i<j} g_{ij}\, x_i\, y_j\right| \le 32\sqrt{N}\,,$$

and hence

$$\forall\, (x_i)_{i\le N}\,,\ \forall\, (y_i)_{i\le N} \in B\,,\ \left|\sum_{i<j} g_{ij}\, x_i\, y_j\right| \le 32\sqrt{N}\,,$$

and this implies (A.46). □

We consider independent Bernoulli r.v.s $(\eta_{i,k})_{i\le N, k\le M}$, that is, $\mathsf{P}(\eta_{i,k} = \pm 1) = 1/2$.

Lemma A.8.2. *Consider numbers* $(\alpha_{k,k'})_{k,k'\le M}$ *with* $\sum \alpha_{k,k'}^2 \le 1$. *Then, for* $t > 0$ *we have*

$$\mathsf{P}\left(\sum_{k\ne k'} \sum_{i\le N} \alpha_{k,k'}\, \eta_{i,k}\, \eta_{i,k'} \ge t\right) \le \exp\left(-\min\left(\frac{t^2}{NL}, \frac{t}{L}\right)\right) \tag{A.47}$$

$$\mathsf{P}\left(\sum_{k\ne k'} \sum_{i\le N} \alpha_{k,k'}\, \eta_{i,k}\, \eta_{i,k'} \ge t\right) \le \exp\left(-\frac{t^2}{2N}\left(1 - \frac{Lt}{N}\right)\right) . \tag{A.48}$$

Proof. The r.v.s $X_i = \sum_{k\ne k'} \alpha_{k,k'}\, \eta_{i,k}\, \eta_{i,k'}$ are i.i.d., and obviously $\mathsf{E} X_i = 0$, $\mathsf{E} X_i^2 = \sum \alpha_{k,k'}^2 \le 1$. An important result of C. Borell [20] implies that then $\mathsf{E}\exp(|X_i|/L) \le 2$ so that (A.47) is a consequence of (A.34) and (A.48) is a consequence of (A.35). □

Proposition A.8.3. *Consider a number* $0 < a \le 1$ *and* $n \le M$. *If* $n\log(eM/n) \le Na^2$, *the following event occurs with probability at least* $1 - \exp(-a^2 N)$. *Given any subset* I *of* $\{1, \dots, M\}$ *with* $\operatorname{card} I = n$, *and any sequences* $(x_k)_{k\le M}$, $(y_k)_{k\le M}$, *we have*

$$\sum_{i\le N}\left(\sum_{k\in I} x_k\, \eta_{i,k}\right)\left(\sum_{k\in I} y_k\, \eta_{i,k}\right) \le N\sum_{k\in I} x_k\, y_k + NLa\left(\sum_{k\in I} x_k^2\right)^{1/2}\left(\sum_{k\in I} y_k^2\right)^{1/2} . \tag{A.49}$$

Corollary A.8.4. *If $a \le 1$ and $M \le Na^2$, then with probability at least $1 - \exp(-a^2 N/L)$, for any sequences $(x_k)_{k\le M}$ and $(y_k)_{k\le M}$ we have*

$$\sum_{i\le N}\Big(\sum_{k\le M} x_k\,\eta_{i,k}\Big)\Big(\sum_{k\le M} y_k\,\eta_{i,k}\Big) \le N\sum_{k\le M} x_k\,y_k + NLa\Big(\sum_{k\le M} x_k^2\Big)^{1/2}\Big(\sum_{k\le M} y_k^2\Big)^{1/2}, \tag{A.50}$$

and

$$\sum_{i\le N}\Big(\sum_{k\le M} x_k\,\eta_{i,k}\Big)^2 \le N(1+La)\Big(\sum_{k\le M} x_k^2\Big). \tag{A.51}$$

Proof. The case $n = M$ of (A.49) is (A.50) and the case $y_k = x_k$ of (A.50) is (A.51). □

Proof of Proposition A.8.3. We rewrite (A.49) as

$$\sum_{i\le N}\ \sum_{k\ne k',\,k,\,k'\in I} x_k\,y_{k'}\,\eta_{i,k}\,\eta_{i,k'} \le LNa\Big(\sum_{k\in I} x_k^2\Big)^{1/2}\Big(\sum_{k\in I} y_k^2\Big)^{1/2}. \tag{A.52}$$

Consider a subset A of $\mathbb{R}^n$, with $\operatorname{card} A \le 5^n$, $A \subset 2B$ and $\operatorname{conv} A \supset B$, where B is the Euclidean ball $\sum_{k\le n} x_k^2 = 1$. To ensure (A.52) it suffices that

$$\sum_{i\le N}\ \sum_{k\ne k',k,k'\in I} x_k\,y_{k'}\,\eta_{i,k}\,\eta_{i,k'} \le LNa \tag{A.53}$$

whenever $(x_k)_{k\in I} \in A$ and $(y_k)_{k\in I} \in A$. Now, given any such sequences (A.47) implies

$$\mathsf{P}\Big(\sum_{i\le N}\ \sum_{k\ne k',k,k'\in I} x_k\,y_{k'}\,\eta_{i,k}\,\eta_{i,k'} \ge Nu\Big) \le \exp\left(-\frac{N}{L}\min(u^2,u)\right). \tag{A.54}$$

Since $n \le M$ and $n\log(eM/n) \le Na^2$ it holds that $n \le Na^2$. We observe also that $25 \le e^4$. Thus the number of possible choices for I and the sequences $(x_k)_{k\in I}$, $(y_k)_{k\in I}$ is at most

$$\binom{M}{n}(\operatorname{card} A)^2 \le \left(\frac{eM}{n}\right)^n 25^n = 25^n \exp\left(n\log\left(\frac{eM}{n}\right)\right) \le \exp 5Na^2$$

so that taking $u = L'a$ where L' large enough, all the events (A.53) simultaneously occur with a probability at least $1 - \exp(-Na^2)$. □

Our next result resembles Proposition A.8.3, but rather than restricting the range of k we now restrict the range of i.

Proposition A.8.5. *Consider a number* $0 < a < 1$. *Consider a number* $N_0 \le N$ *such that* $N_0 \log(eN/N_0) \le a^2 N$, *and assume that* $M \le a^2 N$. *Then the following event occurs with probability at least* $1 - \exp(-a^2 N)$*: Given any subset* J *of* $\{1, \ldots, N\}$ *with* $\mathrm{card} J \le N_0$, *and any sequence* $(x_k)_{k \le M}$, *we have*

$$\sum_{i \in J} \Big(\sum_{k \le M} x_k \eta_{i,k} \Big)^2 \le N_0 \sum_{k \le M} x_k^2 + L \max(Na^2, \sqrt{NN_0}a) \Big(\sum_{k \le M} x_k^2 \Big) . \tag{A.55}$$

Proof. The proof is very similar to the proof of Proposition A.8.3. It suffices to prove that for all choices of (x_k) and (y_k) we have

$$\sum_{i \in J} \sum_{k \ne k'} x_k y_{k'} \eta_{i,k} \eta_{i,k'} \le L \max(Na^2, \sqrt{NN_0}a) \Big(\sum_{k \le M} x_k^2 \Big)^{1/2} \Big(\sum_{k \le M} y_k^2 \Big)^{1/2} . \tag{A.56}$$

Consider a subset A of $\mathbb{R}^M$, with $\mathrm{card} A \le 5^M$, $A \subset 2B$, $B \subset \mathrm{conv} A$, where B is the Euclidean ball $\sum_{k \le M} x_k^2 \le 1$. To ensure (A.56) it suffices that

$$\sum_{i \in J} \sum_{k \ne k'} x_k y_{k'} \eta_{i,k} \eta_{i,k'} \le L \max(Na^2, \sqrt{NN_0}a)$$

whenever $\mathrm{card} J \le N_0$, $(x_k)_{k \le M}$, $(y_k)_{k \le M} \in A$. It follows from (A.47) that for $v > 0$,

$$\mathsf{P}\Big(\sum_{i \in J} \sum_{k \ne k'} x_k y_k \eta_{i,k} \eta_{i,k'} \ge v \mathrm{card} J \Big) \le \exp\left(- \frac{\mathrm{card} J}{L} \min(v^2, v) \right) ,$$

and using this for $v = uN_0/\mathrm{card} J \ge u$ entails

$$\mathsf{P}\Big(\sum_{i \in J} \sum_{k \ne k'} x_k y_k \eta_{i,k} \eta_{i,k'} \ge N_0 u \Big) \le \exp\left(- \frac{N_0}{L} \min(v^2, u) \right) . \tag{A.57}$$

The number of possible choices for J and the sequences $(x_k)_{k \le M}$, $(y_k)_{k \le M}$ is at most

$$\sum_{n \le N_0} \binom{N}{n} (\mathrm{card} A)^2 \le \left(\frac{eN}{N_0} \right)^{N_0} 25^M \le \exp 5Na^2 ,$$

so that by taking $u = L' \max(a^2 N/N_0, a\sqrt{N/N_0})$ where L' is large enough, all the events (A.56) simultaneously occur with a probability at least $1 - \exp(-Na^2)$. □

Here is another nice consequence of Lemma A.8.2.

Lemma A.8.6. *If* $\varepsilon > 0$ *we have*

$$\mathsf{P}\Big(\sum_{1\le k<k'\le M} NR^2_{k,k'} \ge (1-2\,\varepsilon)^{-2}\,u\Big) \le \left(1+\frac{1}{\varepsilon}\right)^{M^2} \exp\left(-\frac{u}{2}\left(1-L\sqrt{\frac{u}{N}}\right)\right)$$

where $R_{k,k'} = N^{-1}\sum_{i\le N}\eta_{i,k}\,\eta_{i,k'}$.

Proof. We start the proof by observing that

$$\Big(\sum_{k<k'} R^2_{k,k'}\Big)^{1/2} = \sup \sum_{k<k'} \alpha_{k,k'}\,R_{k,k'}$$

where the supremum is taken over the subset B of $\mathbb{R}^{M(M-1)/2}$ of sequences $\alpha_{k,k'}$ with $\sum_{k<k'}\alpha^2_{k,k'} \le 1$. We use Proposition A.7.1 to find a subset A of B with $\operatorname{card} A \le (1+\varepsilon^{-1})^{M^2}$ such that

$$\sup_A \sum \alpha_{k,k'}\,R_{k,k'} \ge (1-2\,\varepsilon)\Big(\sum_{k<k'} R^2_{k,k'}\Big)^{1/2}.$$

Thus

$$\begin{aligned}
&\mathsf{P}\Big(\sum_{k<k'} NR^2_{k,k'} \ge (1-2\,\varepsilon)^{-2}\,u\Big) \\
&= \mathsf{P}\bigg(\Big(\sum_{k<k'} R^2_{k,k'}\Big)^{1/2} \ge (1-2\,\varepsilon)^{-1}\sqrt{\frac{u}{N}}\bigg) \\
&\le \mathsf{P}\Big(\sup_A \sum_{k<k'} \alpha_{k,k'}\,R_{k,k'} \ge \sqrt{\frac{u}{N}}\Big) \\
&\le \left(1+\frac{1}{\varepsilon}\right)^{M^2} \exp\left(-\frac{u}{2}\left(1-L\sqrt{\frac{u}{N}}\right)\right),
\end{aligned}$$

where we use (A.48) for $t=\sqrt{uN}$ in the last line. □

Corollary A.8.7. *We have*

$$2^{-Nn}\operatorname{card}\Big\{(\boldsymbol{\sigma}^1,\ldots,\boldsymbol{\sigma}^n)\,;\ \sum_{1\le\ell<\ell'\le n} NR^2_{\ell,\ell'} \ge (1-2\,\varepsilon)^{-2}\,u\Big\}$$
$$\le \left(1+\frac{1}{\varepsilon}\right)^{n^2} \exp\left(-\frac{u}{2}\left(1-L\sqrt{\frac{u}{N}}\right)\right)$$

where $R_{\ell,\ell'} = N^{-1}\sum_{i\le N}\sigma_i^\ell\,\sigma_i^{\ell'}$.

Proof. This is another way to formulate Lemma A.8.6 when $M=n$. □

A.9 Poisson Random Variables and Point Processes

A Poisson random variable X of expectation a is an integer-valued r.v. such that, for $k = 0, 1, \ldots$

$$\mathsf{P}(X = k) = \frac{a^k}{k!}\, e^{-a}$$

so that

$$\mathsf{E} \exp \lambda X = \sum_{k \geq 0} \frac{a^k}{k!}\, e^{\lambda k - a} = \exp a(e^{\lambda} - 1)\,. \tag{A.58}$$

Differentiating 1, 2, or 3 times this relation in λ and setting $\lambda = 0$ we see that

$$\mathsf{E}X = a\,;\ \mathsf{E}X^2 = a + a^2\,;\ \mathsf{E}X^3 = a + 3a^2 + a^3\,. \tag{A.59}$$

Using from (A.3) that for $\lambda > 0$ and a r.v. Y we have $\mathsf{P}(Y \geq t) \leq e^{-\lambda t}\mathsf{E} \exp \lambda Y$ and $\mathsf{P}(Y \leq t) \leq e^{\lambda t}\mathsf{E} \exp(-\lambda Y)$, and optimizing over λ we get that for $t > 1$ we have

$$\mathsf{P}(X \geq at) \leq \exp(-a(t \log t - t - 1))$$

and

$$\mathsf{P}(X \leq a/t) \leq \exp(-a(t \log t - t - 1))\,.$$

In particular we have

$$\mathsf{P}(|X - a| \geq a/2) \leq \exp\left(-\frac{a}{L}\right)\,. \tag{A.60}$$

Of course, such an inequality holds for any constant instead of 1/2.

If X_1, X_2 are independent Poisson r.v.s, $X_1 + X_2$ is a Poisson r.v. The following lemma prove a less known remarkable property of these variables.

Lemma A.9.1. *Consider a Poisson r.v X and i.i.d. r.v.s $(\delta_i)_{i \geq 1}$ such that $\mathsf{P}(\delta_i = 1) = \delta, \mathsf{P}(\delta_i = 0) = 1 - \delta$ for a certain number δ. Then the r.v.s*

$$X_1 = \sum_{i \leq X} \delta_i\ ;\ X_2 = \sum_{i \leq X} (1 - \delta_i)$$

are independent Poisson r.v.s, of expectation respectively $\delta \mathsf{E}X$ and $(1 - \delta)\mathsf{E}X$.

In this lemma we "split X in two pieces". In a similar manner, we can split X in any number of pieces.

Proof. We compute

$$\begin{aligned}\mathsf{E}\exp(\lambda X_1+\mu X_2) &= \mathsf{E}\exp\Big(\lambda\sum_{i\le X}\delta_i+\mu\sum_{i\le X}(1-\delta_i)\Big)\\ &= \sum_{k\ge 0}\frac{a^k}{k!}e^{-a}\mathsf{E}\exp\Big(\lambda\sum_{i\le k}\delta_i+\mu\sum_{i\le k}(1-\delta_i)\Big)\\ &= \sum_{k\ge 0}\frac{a^k}{k!}e^{-a}\big(\mathsf{E}\exp(\lambda\delta_i+\mu(1-\delta_1))\big)^k\\ &= \sum_{k\ge 0}\frac{a^k}{k!}e^{-a}(\delta e^{\lambda}+(1-\delta)e^{\mu})^k\\ &= \exp a(\delta e^{\lambda}+(1-\delta)e^{\mu}-1)\\ &= \exp a\delta(e^{-\lambda}-1)\exp a(1-\delta)(e^{-\mu}-1)\\ &= \mathsf{E}\exp(\lambda Y_1+\mu Y_2)\ ,\end{aligned}$$

where Y_1 and Y_2 are independent Poisson r.v.s with expectation respectively δ and $1-\delta$. □

Consider a positive measure μ of finite total mass $|\mu|$ (say on $\mathbb{R}^3$), and assume for simplicity that μ has no atoms. A Poisson point process of intensity measure μ is a random finite subset $\Pi = \Pi_\mu$ with the following properties:

1. $\operatorname{card}\Pi$ is a Poisson r.v. of expectation $|\mu|$.
2. Given that $\operatorname{card}\Pi = k$, Π is distributed like the set $\{X_1,\ldots,X_k\}$ where $X_1,\ldots,X_k$ are i.i.d. r.v.s of law $\mu/|\mu|$.

(Some inessential complications occur when μ has atoms, and one has to count points of the Poisson point process "with their order of multiplicity".) We list without proof some of the main properties of Poisson point processes. (The proofs are all very easy.)

Given two disjoint Borel sets, A, B, $\Pi\cap A$ and $\Pi\cap B$ are independent Poisson point processes.

Given two finite measures μ_1, μ_2, if Π_{μ_1} and Π_{μ_2} are independent Poisson point processes of intensity measure μ_1 and μ_2 respectively, then $\Pi_{\mu_1}\cup\Pi_{\mu_2}$ is a Poisson point process of intensity measure $\mu_1+\mu_2$.

Given a (continuous) map φ, $\varphi(\Pi)$ is a Poisson point process of intensity measure $\varphi(\mu)$, the image measure of the intensity measure μ of Π by φ.

Consider a positive measure μ and a Poisson point process Π_μ of intensity measure μ. If ν is a probability (say on $\mathbb{R}^3$), and $(U_\alpha)_{\alpha\ge 1}$ are i.i.d. r.v.s of law ν, we can construct a Poisson point process of intensity measure $\mu\otimes\nu$ as follows. We number in a random order the points of Π as $x_1,\ldots,x_k$, and we consider the couples $(x_1,U_1),\ldots,(x_k,U_k)$.

Consider now a positive measure μ on $\mathbb{R}^+$. We do not assume that μ is finite, but we assume that $\mu([a,\infty))$ is finite for each $a\ge 0$. We denote by μ_0 the restriction of μ to $[1,\infty)$, by μ_k its restriction to $[2^{-k},2^{-k+1}[$, $k\ge 1$. Consider for $k\ge 0$ a Poisson point process Π_k of intensity measure μ_k, and assume that these are independent. We can define a Poisson point process of intensity measure μ as $\Pi = \cup_{k\geq 0}\Pi_k$. Then for each a, $\Pi \cap [a,\infty)$ is a Poisson point process, the intensity measure of which is the restriction of μ to $[a,\infty)$.

A.10 The Paley-Zygmund Inequality

This simple (yet important) argument is also known as the second moment method. It goes back to the work of Paley and Zygmund on trigonometric series.

Proposition A.10.1. *Consider a r.v. $X \geq 0$. Then*

$$\mathsf{P}\left(X \geq \frac{1}{2}\,\mathsf{E}X\right) \geq \frac{1}{4}\,\frac{(\mathsf{E}X)^2}{\mathsf{E}X^2}\,. \tag{A.61}$$

Proof. If $A = \{X \geq \mathsf{E}X/2\}$, then, since $X \leq \mathsf{E}X/2$ on the complement A^c of A, we have

$$\mathsf{E}X = \mathsf{E}(X\mathbf{1}_A) + \mathsf{E}(X\mathbf{1}_{A^c}) \leq \mathsf{E}(X\mathbf{1}_A) + \frac{1}{2}\,\mathsf{E}X\,.$$

Thus, using the Cauchy-Schwarz inequality,

$$\frac{1}{2}\,\mathsf{E}X \leq \mathsf{E}(X\mathbf{1}_A) \leq (\mathsf{E}X^2)^{1/2}\,\mathsf{P}(A)^{1/2}\,. \qquad \square$$

A.11 Differential Inequalities

We will often meet simple differential inequalities, and it is worth to learn how to handle them. The following is a form of the classical Gronwall's lemma.

Lemma A.11.1. *If a function $\varphi \geq 0$ satisfies*

$$|\varphi'_r(t)| \leq c_1\varphi(t) + c_2$$

for $0 < t < 1$, where $c_1, c_2 \geq 0$ and where φ'_r is the right-derivative of φ, then

$$\varphi(t) \leq \exp(c_1 t)\left(\varphi(0) + \frac{c_2}{c_1}\right)\,. \tag{A.62}$$

Proof. We note that

$$\left|\left(\varphi(t) + \frac{c_2}{c_1}\right)'_r\right| \leq c_1\left(\varphi(t) + \frac{c_2}{c_1}\right)\,,$$

so that

$$\varphi(t) + \frac{c_2}{c_1} \leq \exp(c_1 t)\left(\varphi(0) + \frac{c_2}{c_1}\right)\,. \qquad \square$$

References

1. Arguin L-P. (2008) A remark on the infinite-volume Gibbs measures of spin glasses. *J. Math. Phys.* **49**, no. 12, 125204, 8 pp.
2. Arguin L-P. (2008) Competing particle systems and the Ghirlanda-Guerra identities. *Electron. J. Probab.* **13**, no. 69, pp. 2101–2117.
3. Arguin L-P. (2007) A dynamical characterization of Poisson-Dirichlet distributions. *Electron. Comm. Probab.* **12**, pp. 283–290.
4. Arguin L-P. (2007) Spin glass computations and Ruelle's probability cascades. *J. Stat. Phys.* **126**, no. 4–5, pp. 951–976.
5. Arguin L-P., Aizenman M. (2009) On the structure of quasi-stationary competing particle systems. *Ann. Probab.* **37**, no. 3, pp. 1080–1113.
6. Arguin L-P., Kistler N. (2009) Small perturbations of a spin glass system. *J. Stat. Phys.* **135**, no. 5–6, pp. 1167–1180.
7. Arguin L-P., Chatterjee S. (2010) Random Overlap Structures: Properties and Applications to Spin Glasses. arXiv:1011.1823
8. Aizenman M., Contucci P. (1998) On the stability of the quenched state in mean-field spin-glass models. *J. Statist. Phys.* **92**, no. 5–6, pp. 765–783.
9. Aizenman M., Lebowitz J.L., Ruelle D. (1987) Some rigorous results on the Sherrington-Kirkpatrick model. *Comm. Math. Phys.* **112**, pp. 3–20.
10. Ruzmaikina A., Aizenman M., Ruzmaikina A. (2005) Characterization of invariant measures at the leading edge for competing particle systems. *Ann. Probab.* **33**, no. 1, pp. 82–113.
11. Aizenman M., Sims R., Starr S. (2003) Extended variational principle for the Sherrington-Kirkpatrick spin-glass model. *Phys. Rev. B* **68**, 214403.
12. Aizenman M., Sims R., Starr S., Shannon L. (2007) Mean-field spin glass models from the cavity-ROSt perspective. Prospects in mathematical physics, *Contemp. Math.*, **437**, pp. 1–30, Amer. Math. Soc., Providence, RI.
13. Almeida J.R.L., Thouless D.J. (1978) Stability of the Sherrington-Kirkpatrick solution of a spin glass model. *J. Phys. A: Math. Gen.* **II**, pp. 983–990.
14. Amit D.J., Gutfreund H., Sompolinsky H. (1987) Statistical mechanics of neural networks near saturation. *Annals of Physics* **173**, pp. 30–67.
15. Baffioni F., Rosati F. (2000) On the Ultrametric Overlap Distribution for Mean Field Spin Glass Models (I). *The European Physical Journal B* **17**, pp. 439–447.
16. Bolthausen E., Sznitman A.-S. (1998) On Ruelle's probability cascades and an abstract cavity method. *Comm. Math. Phys.* **197**, no. 2, pp. 247–276.
17. Bolthausen E., Kistler N. (2006) On a nonhierarchical version of the Generalized Random Energy Model. *Ann. Appl. Prob.* **16**, no. 1, pp. 1–14.
18. Bolthausen E., Kistler N. (2008) Universal structures in some mean field spin glasses, and an application. *J. Math. Phys.* **49**, 125205, pp. 1–17.
19. Bolthausen E., Kistler N. (2009) On a nonhierarchical version of the Generalized Random Energy Model. II. Ultrametricity. *Stoch. Proc. Appl.* **119**, pp. 2357–2386.

M. Talagrand, *Mean Field Models for Spin Glasses*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys in Mathematics 55, DOI 10.1007/978-3-642-22253-5,

20. Borell C. (1984) On polynomial chaos and integrability. *Probab. Math. Statist.* **3**, no. 2, pp. 191–203.
21. Bouchaud J.-P., Mézard M. (1997) Universality classes for extreme values statistics. *J. Phys. A. Math. Gen.* **30**, pp. 7997–8015.
22. Bovier A. (1994) Self-averaging in a class of generalized Hopfield models. *J. Phys. A* **27**, no. 21, pp. 7069–7077.
23. Bovier A., van Enter A.C.D., Niederhauser B. (1999) Stochastic symmetry-breaking in a Gaussian Hopfield model. *J. Statist. Phys.* **95**, no. 1–2, pp. 181–213.
24. Bovier A., Gayrard V. (1997) An almost sure central limit theorem for the Hopfield model. *Markov Process. Related Fields* **3**, no. 2, pp. 151–173.
25. Bovier A., Klimovsky A. (2009) The Aizenman-Sims-Starr and Guerra's schemes for the SK model with multidimensional spins. *Electron. J. Probab.* **14**, no. 8, pp. 161–241.
26. Bovier A., Kurkova I. (2004) Derrida's generalised random energy models. I. Models with finitely many hierarchies. *Ann. Inst. H. Poincaré Probab. Statist.* **40**, no. 4, pp. 439–480.
27. Bovier A., Kurkova I. (2007) Local energy statistics in spin glasses. *J. Stat. Phys.* **126**, no. 4–5, pp. 933–949.
28. Bovier A., Kurkova I., Loewe M. (2002) Fluctuations of the free energy in the REM and the p-spin SK model. *Ann. Probab.* **30**, pp. 605–651.
29. Bovier A., Niederhauser B. (2001) The spin-glass phase-transition in the Hopfield model with p-spin interactions. *Adv. Theor. Math. Phys.* **5**, no. 6, pp. 1001–1046.
30. Chatterjee S. (2007) Estimation in spin glasses: a first step. *Ann. Statist.* **35**, no. 5, pp. 1931–1946.
31. ——— Solution to the chaos in disorder problem. Manuscript, 2009.
32. ——— (2010) Spin glasses and Stein's method. *Probab. Theory Related Fields* **148**, no. 3–4, pp. 567–600.
33. Contucci P., Giardinà C. (2005) Spin-glass stochastic stability: a rigorous proof. *Ann. Henri Poincaré* **6**, no. 5, pp. 915–923.
34. Contucci P. (2003) Towards a classification of stochastically stable quenched measures. *Markov Process. Related Fields* **9**, no. 2, pp. 167–176.
35. Crisanti A., Rizzo T. (2003) Chaos in temperature in the Sherrington-Kirkpatrick model. *Phys. Rev. Lett.* **90**, 137201.
36. Derrida B. (1981) Random energy model: An exactly solvable model of disordered systems. *Phys. Rev. B* **24**, no. 5, pp. 2613–2626.
37. Derrida B., Gardner E. (1988) Optimal storage properties of neural network models. *J. Phys. A* **21**, pp. 271–284.
38. ——— (1989) Three unfinished works on the optimal storage capacity of networks. Special issue in memory of Elizabeth Gardner, 1957–1988. *J. Phys. A* **22**, no. 12, pp. 1983–1994.
39. Fischer K.H., Hertz J. (1991) Spin glasses. *Cambridge Studies in Magnetism*, **1**, Cambridge University Press, Cambridge, x+408 pp.
40. Gardner E. (1985) Spin glasses with p-spin interactions. *Nuclear Phys. B* **257**, no. 6, pp. 747–765.
41. ——— (1988) The space of interactions in neural network models. *J. Phys. A* **21**, pp. 257–270.
42. Ghirlanda S., Guerra F. (1998) General properties of overlap probability distributions in disordered spin systems. Towards Parisi ultrametricity. *J. Phys. A* **31**, no. 46, pp. 9149–9155.
43. Gross D., Mézard M. (1984) The simplest spin glass. *Nuclear Phys. B* **240**, pp. 431–452.

44. Guerra F. (1995) Fluctuations and thermodynamic variables in mean field spin glass models. In S. Albeverio et al., editors, *Stochastic Processes, Physics and Geometry*, World Scientific, Singapore.
45. ______ (1996) About the overlap distribution in mean field spin glass models. *Int. J. Modern Physics B* **10**, no. 13–14, pp. 1675–1684.
46. ______ (1995) Functional order parameters for the quenched free energy in mean field spin glass models. In S. De Lillo, P. Sodano, F.C. Khanna, G.W. Semenov, editors, *Field Theory and Collective Phenomena*, World Scientific, Singapore.
47. ______ (1996) About the overlap distribution in mean field spin glass models. *International Journal of Modern Physics B* **10**, pp. 1675–1684.
48. ______ (2001) Sum rules for the free energy in the mean field spin glass model. *Fields Institute Communications* **30**, pp. 161–170.
49. ______ (2003) Replica broken bounds in the mean field spin glass model. *Communications Math. Phys.* **233**, pp. 1–12.
50. ______ (2005) Mathematical aspects of mean field spin glass theory. European Congress of Mathematics, pp. 719–732, Eur. Math. Soc., Zürich, 2005.
51. Guerra F., Toninelli F.L. (2002) Quadratic replica coupling for the Sherrington-Kirkpatrick mean field spin glass model. *J. Math. Phys.* **43**, no. 7, pp. 3704–3716.
52. ______ (2003) Infinite volume limit and spontaneous replica symmetry breaking in mean field spin glass models. *Ann. Henri Poincaré* **4**, suppl. 1, pp. S441–S444.
53. Hertz J., Krogh A., Palmer R.C. (1991) *Introduction to the theory of neural computation* Addison Wesley.
54. Kingman J.F.C. (1970) The representation of partition structures. *J. London Math Soc.* **18**, pp. 374–380.
55. ______ (1975) Random discrete distributions. *Jour. Royal Stat. Soc. Ser B* **37**, pp. 1–22.
56. ______ (1982) Exchangeability and the evolution of large partitions. In G. Koch and F. Spizzichino, editors, *Exchangeability in Probability and Statistics*, (Rome, 1981) pp. 97–112, North Holland, Amsterdam.
57. Krauth W., Mézard M. (1989) Storage capacity of memory network with binary couplings. *J. Phys.* **50**, pp. 3057–3066.
58. Ledoux M., Talagrand M. (1991) *Probability in Banach Spaces*, Springer-Verlag, Berlin.
59. Mézard M. (1988) The space of interactions in neural networks: Gardner's computation with the cavity method. *J. Phys. A* **22**, pp. 2181–2190.
60. Mézard M., Parisi G. (1985) Replicas and optimization. *J. Physique. Lett.* **46**, L771.
61. ______ (1986) Mean field equations for the matching and the traveling salesman problem. *Europhys. Lett.* **916**.
62. Mézard M., Parisi G., Virasoro M. (1987) *Spin glass theory and beyond*, World Scientific, Singapore.
63. Nishimori, H. (2001) Statistical physics of spin glasses and information processing. An introduction. *International Series of Monographs on Physics*, **111**, Oxford University Press, New York, xii+243 pp.
64. Panchenko D. (2005) A note on the free energy of the coupled system in the Sherrington-Kirkpatrick model. *Markov Process. Related Fields* **11**, no. 1, pp. 19–36.
65. ______ (2005) Free energy in the generalized Sherrington-Kirkpatrick mean field model. *Rev. Math. Phys.* **17**, no. 7, pp. 793–857.

66. ——— (2005) A question about the Parisi functional. *Electron. Comm. Probab.* **10**, pp. 155–166.
67. ——— (2007) A note on Talagrand's positivity principle. *Electron. Comm. Probab.* **12**, pp. 401–410.
68. ——— (2008) Exponential control of overlap in the replica method for p-spin Sherrington-Kirkpatrick model. *J. Stat. Phys.* **130**, no. 5, pp. 831–842.
69. ——— (2008) On differentiability of the Parisi formula. *Electron. Commun. Probab.* **13**, pp. 241–247.
70. ——— (2009) Cavity method in the spherical SK model. *Ann. Inst. Henri Poincaré Probab. Stat.* **45**, no. 4, pp. 1020–1047.
71. ——— (2010) A connection between Ghirlanda-Guerra indentities and ultrametricity. *Ann. Proba.* **38**, no. 1, pp. 327–347.
72. ——— (2010) On the Dobysh-Sudakov representation theorem. *Electron. Commun. Probab.* **15**, pp. 330–338.
73. ——— (2010) The Ghuirlanda-Guerra identities for the mixed p-spin model. *Comptes Rendus Acad. Sci.* **348**, no. 3–4, pp. 189–192.
74. ——— Spin glass models from the point of view of spin distributions. Manuscript, 2010.
75. ——— (2011) A deletion-invariance property for random measures satisfying the Ghirlanda-Guerra identities. *Comptes Rendus Acad. Sci.*, to appear.
76. Panchenko D., Talagrand M. (2004) Bounds for diluted mean-fields spin glass models. *Probab. Theory Related Fields* **130**, no. 3, pp. 319–336.
77. ——— (2007) On the overlap in the multiple spherical SK model. *Ann. Probab.* **35**, no. 6, pp. 2321–2355.
78. Parisi G. (1992) Field theory, disorder, simulation. *World Scientific Lecture Notes in Physics*, **45**, World Scientific, Singapore.
79. Parisi G. (2009) The mean field theory of spin glasses: the heuristic replica approach and recent rigorous results. *Lett. Math. Phys.* **88**, no. 1–3, pp. 255–269.
80. Parisi G., Rizzo T. (2008) Large deviations in the free energy of mean-field spin glasses. *Phys. Rev. Lett.* **101**, 117205.
81. Parisi G., Talagrand M. (2004) On the distribution of the overlaps at given disorder. *C. R. Math. Acad. Sci. Paris* **339**, no. 4, pp. 303–306.
82. Pitman J. (2006) Combinatorial Stochastic Processes, Lecture from the 32 nd St. Flour Summer School, July 2002, *Lecture Notes in Math.*, **1875**, Springer Verlag, x+256 pp.
83. Pitman J., Yor M. (1997) The two-parameter Poisson-Dirichlet distribution derived from a stable subordinator. *Ann. Probab.* **25**, pp. 855–900.
84. Rizzo T. (2001) Against chaos in temperature in mean field models. *J. Phys. A: Math. Gen.* **34**, no. 27, 5531.
85. Ruelle D. (1987) A mathematical reformulation of Derrida's REM and GREM. *Comm. Math. Phys.* **108**, pp. 225–239.
86. Shcherbina M.V., Tirozzi B. (2003) On the rigorous solution of Gardner's problem. *Comm. Math. Phys.* **234**, no. 3, pp. 383–422.
87. ——— (2003) Central limit theorems for order parameters of the Gardner problem. *Markov Process. Related Fields* **9**, no. 4, pp. 803–828.
88. ——— (2005) Central limit theorems for the free energy of the modified Gardner model. *Markov Process. Related Fields* **11**, no. 1, pp. 133–144.
89. Sherrington D., Kirkpatrick S. (1972) Solvable model of a spin glass. *Phys. Rev. Lett.* **35**, pp. 1792–1796.
90. Talagrand M. (1995) Concentration of measure and isoperimetric inequalities in product spaces. *Publ. Math. I.H.E.S.* **81**, pp. 73–205.
91. ——— (1996) A new look at independence. *Ann. Probab.* **24**, pp. 1–34.

92. ——— (1998) The Sherrington-Kirkpatrick model: a challenge to mathematicians. *Prob. Th. Relat. Fields* **110**, pp. 109–176.
93. ——— (1998) Rigorous results for the Hopfield model with many patterns. *Prob. Th. Relat. Fields* **110**, pp. 177–276.
94. ——— (1998) Huge random structures and mean field models for spin glasses, Proceedings of the Berlin International Congress of mathematicians. *Documenta Math.*, Extra Vol. I, pp. 507–536.
95. ——— (2000) Rigorous low temperature results for the mean field p-spin interaction model. *Probab. Theor. Relat. Fields* **117**, pp. 303–360.
96. ——— (2000) Multiple levels of symmetry breaking. *Probab. Theor. Relat. Fields* **117**, pp. 449–466.
97. ——— (2000) Intersecting random half-spaces: towards the Derrida-Gardner formula. *Ann. Probab.* **28**, pp. 725–758.
98. ——— (2000) Exponential inequalities and replica-symmetry breaking for the Sherrington-Kirkpatrick model. *Ann. Probab.* **28**, pp. 1018–1062.
99. ——— (2000) Exponential inequalities and convergence of moments in the replica-symmetric phase of the Hopfield model. *Ann. Probab.* **28**, pp. 1393–1469.
100. ——— (2000) Large deviation principles and generalized Sherrington-Kirkpatrick models. *Ann. Fac. Sci. Toulouse Math.* **9**, pp. 203–244.
101. ——— (2000) Spin glasses: a new direction for probability theory? Mathematics towards the third millennium (Rome, 1999). *Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei* **9**, Mat. Appl. Special Issue, pp. 127–146.
102. ——— (2002) On the high temperature phase of the Sherrington-Kirkpatrick model. *Ann. Probab.* **30**, pp. 364–381.
103. ——— (2003) Spin glasses: a challenge for mathematicians. Cavity and mean field models. *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics*, **46**, Springer-Verlag, Berlin, x+586 pp.
104. ——— (2003) Self organization in the low temperature phase of a spin glass model. *Review Math. Phys.* **14**, pp. 1–78.
105. ——— (2003) On Guerra's broken replica-symmetry bound. *C. R. Math. Acad. Sci. Paris* **337**, no. 7, pp. 477–480.
106. ——— (2004) Magnetization of the pure states in the p-spin interaction model. *Internat. J. Modern Phys. B* **18**, no. 4–5, pp. 773–784.
107. ——— (2005) *The Generic Chaining*, Springer Verlag, viii+222 pp.
108. ——— (2006) The Parisi formula. *Ann. of Math. (2)* **163**, no. 1, pp. 221–263.
109. ——— (2006) Parisi measures. *J. Funct. Anal.* **231**, no. 2, pp. 269–286.
110. ——— (2006) Free energy of the spherical mean field model. *Probab. Theory Related Fields* **134**, no. 3, pp. 339–382.
111. ——— (2007) Large deviations, Guerra's and A.S.S. Schemes, and the Parisi hypothesis. *J. Stat. Phys.* **126**, no. 4–5, pp. 837–894.
112. ——— (2007) Mean field models for spin glasses: some obnoxious problems. *Lecture Notes in Math.*, **1900**, Springer, Berlin, pp. 63–80.
113. ——— (2010) Construction of pure states in mean field models for spin glasses. *Probab. Teo. Relat. Fields* **148**, no. 3–4, pp. 60–643.
114. Toninelli F. (2002) About the Almeida-Thouless transition line in the Sherrington-Kirkpatrick mean field spin glass model. *Europhysics Letters* **60**, no. 5, pp. 764–767.

Index

M. Talagrand, *Mean Field Models for Spin Glasses*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys in Mathematics 55, DOI 10.1007/978-3-642-22253-5,