Scuola Normale Superiore

PhD Thesis

October 2002

Rigorous results for mean field spin glasses: thermodynamic limit and sum rules for the free energy

Fabio Lucio Toninelli

Contents

Acknowledgments

I wish to express my deep gratitude to Francesco Guerra, who taught me to appreciate the beauty of simple ideas and from whom I learned almost all what I know on this fascinating field, and to Giovanni Jona-Lasinio, for guiding my first steps in scientific research and for his constant encouragement.

I owe particular thanks to Anton Bovier, Sergio Caracciolo, Pierluigi Contucci, Sandro Graffi, Massimiliano Gubinelli, Enzo Marinari, Andrea Montanari and Giorgio Parisi for interesting discussions and precious suggestions, and to Michel Talagrand for a stimulating correspondence.

I would also like to acknowledge the financial support of Scuola Normale Superiore and Istituto Nazionale di Fisica Nucleare, and frequent hospitality at the Department of Physics of Università degli Studi di Roma "La Sapienza".

Chapter 1

Introduction

The subject of this thesis is the rigorous study of mean field spin glass models [6], [8] and, in particular, of the Sherrington-Kirkpatrick one. This was proposed [16], [17] almost thirty years ago as a "solvable" mean field approximation of finite dimensional spin glasses. Actually, it was soon realized that the model is far from being trivial. Indeed, only many years later Parisi [18]-[20] proposed a solution for the infinite volume limit of the system, which required the introduction of completely new concepts, *e.g.*, Replica Symmetry Breaking and the ultrametric structure of the space of equilibrium states. Even if the physical picture of the model is by now well established thanks to both analytical calculations and numerical simulations, the situation is very different from the mathematical point of view, where progress has been much more difficult. The task of reducing Parisi theory, based on the so called "replica trick" and on the ultrametric *Ansatz*, to standard mathematical and statistical mechanical concepts has proved to be very tough, and has not been fully accomplished yet. Therefore, following Michel Talagrand [21] one can really say that the Sherrington-Kirkpatrick model and, in general, mean field spin glasses, are *a challenge to mathematicians*. However, the last times have witnessed remarkable progress in this field, and some long time open problems have been finally solved. As an example, let us mention here the proof of the existence of the infinite volume limit for the free energy and ground state energy per site [22], and Guerra's bounds and sum rules [23], relating the free energy to the Parisi solution with replica symmetry breaking.

The rigorous study of mean field spin glass models is important, in our opinion, at least for two reasons. First of all, many fundamental questions concerning their connection with finite dimensional spin glass models are still open. Actually, it is not clear yet whether or to what extent the main physical features of mean field models have a counterpart in more realistic ones. A full understanding of the mathematical structure underlying the physical behavior of the mean field system seems to us to be a necessary prerequisite to attack this question. The second reason is that, as we will briefly illustrate in the following, mean field spin glass models arise naturally in many different contexts, ranging from combinatorial optimization problems [24] to the theory of neural networks [7] and error correction codes [10].

In this thesis I report the results which I have obtained, in collaboration with Francesco Guerra, during my PhD studies. Our work focuses on *equilibrium properties*, and we don't deal with the (very interesting) dynamical aspects of spin glasses. In order to enable the non-expert reader to appreciate the meaning of the results I present, I have tried to put them into a more general context, introducing some of the main concepts of the theory of spin glasses, and some physical motivation for the study of these models. Moreover, I have tried as much as possible to link our results with previous works by other authors, and to compare the different methods and approaches which have been employed in the literature.

1.1 Outline of the thesis

This work is organized as follows.

- In Chapter 2, we give a very short overview on some general aspects of the theory of spin glasses and its applications. Our aim is to motivate the non-expert reader, illustrating both the physical origin of these models and their connection with other fields of physics and mathematics, especially random combinatorics and theoretical computer science.

- In Chapter 3, we introduce mean field spin glass models, and in particular the Sherrington-Kirkpatrick and p-spin models, together with the basic definitions we need in the following of this work: quenched free energy, replicas, overlaps and so on.

- In Chapter 4, before we undertake the rigorous analysis of mean field models, we recall briefly some of the results of Parisi theory of Replica Symmetry Breaking, with the purpose of making this work as self-contained as possible, and to allow the reader to fully appreciate the meaning of the results in the following chapters.

- In Chapter 5, we prove the existence of the infinite volume limit for the free energy and ground state energy per site, for a wide class of mean field spin glass models. This class includes also models with non-Gaussian disorder, with non-Ising type spin degrees of freedom and with multi-component spins. The problem of the existence of the infinite volume limit is a very basic question, which has remained open for a very long time. In this context, we introduce the very simple but powerful idea of interpolating between Hamiltonians, which will be employed many times in the following chapters.

  This chapter is based on:

  F. Guerra, F. L. Toninelli, *The Thermodynamic Limit in Mean Field Spin Glass Models*, Communications in Mathematical Physics, **230:1**, 71-79 (2002), preprint number `cond-mat/0204280`.

  F. Guerra, F. L. Toninelli, *The infinite volume limit in generalized mean field disordered models*, Markov Processes and Related Fields, to appear, preprint number `cond-mat/0208579`.

  F. Guerra, F. L. Toninelli, *The Sherrington-Kirkpatrick model with non-Gaussian disorder*, in preparation.

- In Chapter 6, we explain how to find exact sum rules for the free energy of mean field spin glasses. In the case of the Sherrington-Kirkpatrick model, this allows to rigorously prove that the free energy is bounded *below* by the Parisi solution with Replica Symmetry Breaking. The proof of these sum rules is based on a beautiful interpolation method and is due to F. Guerra [23]. We also prove that, as a consequence of Guerra's bounds, replica symmetry is necessarily broken under the Almeida-Thouless critical line. This result is based on

  F. L. Toninelli, *About the Almeida-Thouless transition line in the Sherrington-Kirkpatrick mean field spin glass model*, Europhysics Letters, to appear, preprint number `cond-mat/0207296`.

- In Chapter 7, on the other hand, we discuss how to obtain effective *upper* bounds for the free energy. The aim is to prove that the free energy is bounded *above* by Parisi solution, apart from terms which vanish as the system size goes to infinity. This task has not been fully accomplished yet, and it would provide a rigorous proof that Parisi solution actually gives the infinite volume limit for these models. In the case of the Sherrington-Kirkpatrick model, we show how this strategy can be carried on, at least in a sub-region of the expected high temperature region, where replica symmetry is not broken. Several approaches have been developed in the literature, to prove replica symmetry in some region of parameters. Here, we illustrate our "quadratic replica coupling" method, which has the advantage of being conceptually quite intuitive and technically very simple, and which was introduced in

  F. Guerra, F. L. Toninelli, *Quadratic replica coupling for the Sherrington-Kirkpatrick mean field spin glass model*, Journal of Mathematical Physics **43**, 3704 (2002), preprint number `cond-mat/0201091`.

- In Chapter 8 we prove a central limit theorem for the fluctuations of the free energy and of the overlaps, in the region where we are able to prove that replica symmetry is not broken. This result is based on

  F. Guerra, F. L. Toninelli, *Central limit theorem for fluctuations in the high temperature region of the Sherrington-Kirkpatrick spin glass model*, Journal of Mathematical Physics, to appear, preprint number `cond-mat/0201092`.

  The latter result is based on a rigorous version of the cavity method, developed by M. Talagrand.

- Finally, in Chapter 9 we summarize briefly our results, we discuss some of the main open problems, and we try to outline future research directions.

Chapter 2

A short overview on spin glasses

The simplest definition of spin glass [6, 8] is that of a spin system whose low temperature state appears as a disordered one, rather than the uniform or periodic pattern one is used to find in conventional Ising magnets. In order to produce such a state, two ingredients are necessary: the Hamiltonian must contain randomness, and there must be *frustration* [25].

Randomness means that the Hamiltonian depends not only on the configuration of the system, which we denote by σ, and on the strength of the external applied fields, like the magnetic field, but also on some random parameters (usually, the couplings among the elementary degrees of freedom), whose probability distribution is supposed to be known. We will always deal with discrete systems on a lattice, where σ denotes the configuration of the spin variables σ_i, *i.e.*, the magnetic moments of the individual atoms, in suitable units, and i is the site index. The random parameters are collectively denoted as "quenched" (or "frozen") disorder. From the physical point of view, the word "frozen" means we are modeling a disordered system whose impurities have a dynamics which is many orders of magnitude slower than the dynamics of the spin degrees of freedom. Therefore, the disorder does not thermalize, and it can be considered as fixed.

We say that frustration is present when the Hamiltonian cannot be written as the sum of many terms, all of which can be minimized by a single ground state configuration. In order to clarify this rather sloppy definition, let us begin with a *non-frustrated* system, namely, the ferromagnetic Ising model. The model is defined on the $\mathbb{Z}^d$ lattice by the Hamiltonian

$$H(\sigma, h) = -\sum_{i,j} J_{ij}\sigma_i\sigma_j, \tag{2.1}$$

where the couplings J_{ij}, for our present purposes, are only required to be *all non-negative*, and $\sigma_i = \pm 1$ are Ising spins. Of course, each spin-spin interaction term is minimized by the configurations where all spin are parallel, *i.e.*, $\sigma_i\sigma_j = +1$ for all i, j. There are two such configurations, one with all spins equal $+1$, the other with spins -1, and they are connected by the global spin-flip symmetry $\sigma_i \to -\sigma_i \,\forall i$. It is also elementary to check that any other configuration has a strictly higher energy. On the other hand, if the coupling signs are random, this is not true in general and, even worse, it is not possible to guess the ground state configuration just from symmetry considerations. Moreover, in the random sign case the ground state of the system has a high degeneracy, and the several ground state configurations are not connected one to another by elementary symmetry transformations. The reader can check this by himself, in the simple case where the system is composed of four spins, denoted by the indices $1, 2, 3, 4$, and where $J_{12} = J_{23} = J_{34} = -J_{14} = +1$, all other couplings being zero, as illustrated in Fig. 2.1. In

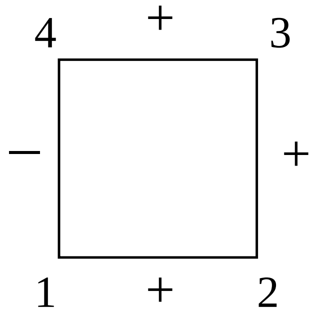

Figure 2.1: A system of four spins (a *plaquette*) with frustrated interactions. One can define a *plaquette* to be frustrated if the product of the couplings along its border is negative [8]. Notice that this is the case in the present example.

this case, there are eight ground state configurations, of course not all connected by global spin-flip. The number of ground states grows very fast as the number of spins is increased. If the absolute values of the couplings are also random, the exact degeneracy is broken, but frustration still gives rise, for large systems, to low-lying metastable configurations, which would be absent in the unfrustrated case.

Experimentally, it is very easy to find systems satisfying the requirements of being both disordered and frustrated, where spin glass behavior can be observed. The first kind of spin glass systems which was studied consisted in dilute solutions of magnetic transition metal impurities in noble metal hosts. The impurity moments produce a magnetic polarization of the host metal conduction electrons, which is positive at some distances and negative at others. For instance, in the simple approximation where the metal electrons are considered as free, one finds that the polarization induced at the point $\mathbf{r_j}$ by a spin $\vec{\sigma}(\mathbf{r_i})$ placed at the point $\mathbf{r_i}$, is proportional to

$$\frac{\cos 2(k_F r)}{r^3}\vec{\sigma}(\mathbf{r_i}) \quad \text{for } k_F r \gg 1,$$

where $r = |\mathbf{r_i} - \mathbf{r_j}|$ and k_F is the Fermi momentum. Such an interaction potential takes the name of Ruderman-Kittel-Kasuya-Yosida (or RKKY) interaction [8]. Other impurity moments then feel the local magnetic field produced by the polarized electrons and try to align themselves along it. Since the impurities are randomly placed, some of the resulting interactions are of ferromagnetic character (*i.e.*, they favor parallel alignment of the spins) and others are anti-ferromagnetic. Thus we clearly have random, competing interactions which produce frustration.

In the RKKY model we have just illustrated, randomness is in the position of the magnetic impurities. This model does not turn out to be simple enough to allow an analytical approach. Therefore, following Edwards and Anderson [26], one usually considers models where the positions of the magnetic moments are non-random and are placed on the $\mathbb{Z}^d$ lattice, and disorder is in the interactions J_{ij}, which are taken to be random variables whose distribution depends only on $(\mathbf{r_i} - \mathbf{r_j})$. A further simplification is that of considering, instead of a model of Heisenberg vector spins $\vec{\sigma}_i$, an Ising-like model where one keeps only a single spin component $\sigma_i = (\vec{\sigma}_i)_z = \pm 1$. The Hamiltonian of the model, in some magnetic field h, can be therefore written as

$$H(\sigma, h; J) = -\sum_{i,j} J_{ij}\sigma_i\sigma_j - h\sum_i \sigma_i, \tag{2.2}$$

which resembles Eq. (2.1), apart from the fundamental difference that, here, the couplings J_{ij} are random and have random sign. The J_{ij} are assumed to be independent random variables with zero mean and a variance which depends only on $|i-j|$. For instance, one can consider the Gaussian case

$$P(J_{ij}) = \frac{1}{\sqrt{2\pi\Delta_{ij}}}\exp(-J_{ij}^2/2\Delta_{ij}),$$

where $\Delta_{ij} = \Delta(\mathbf{r_i} - \mathbf{r_j})$ vanishes sufficiently fast for $|\mathbf{r_i} - \mathbf{r_j}| \to \infty$. Another often studied case is that where $J_{ij} = \pm 1$ with equal probability for $|i-j| = 1$ and $J_{ij} = 0$ otherwise (nearest neighbor interaction). The Edwards-Anderson model, though in a sense simplified with respect to the RKKY one (the disorder averages are easier to perform), retains all the essential features of spin glass systems, and is still of formidable difficulty.

The physical behavior of spin glasses is not sensitively dependent on the particular features of the interaction, like the decay of the interaction potential with distance, and this is the reason why the physically more natural random-position RKKY model can be replaced by a random-coupling one.

Spin glass behavior has also been observed in completely different physical systems, *e.g.*, ferroelectric-antiferroelectric mixtures, where the electric dipole moments take the place of the magnetic ones. We refer to [8], and references therein, for a review.

2.1 Experimental observation of spin-glass behavior

A fundamental question for a physicist is of course, how spin glass behavior can be observed experimentally. In this section, we give just a sketchy description of some basic aspects, in order to make some connection between real-life experiments and the theory we will discuss in the following of this work.

An important feature of the (low temperature) spin glass phase is that there is a local finite magnetization* $m_i \equiv \langle\sigma_i\rangle \neq 0$ even when the magnetic field is infinitesimal, while the total magnetization $m = |\Lambda|^{-1}\sum_i m_i$ vanishes. Of course, this is possible only because randomness destroys translation invariance, for a given disorder realization. Here, we are considering a system enclosed in a finite box Λ, containing $|\Lambda|$ lattice sites.

*here, the symbol $\langle . \rangle$ denotes the usual Boltzmann-Gibbs thermal average. In the following of this work (see Chapter 3) we will use a different notation, but in these introductory remarks we wish to keep notations and definitions to a minimum.

In contrast with the case of ordered antiferromagnets, any "staggered magnetization" of the form

$$m_{\mathbf{k}} \equiv \frac{1}{|\Lambda|} \sum_i e^{-i\,\mathbf{k}\,\mathbf{r}_i}\, m_i, \tag{2.3}$$

where $\mathbf{k} \in \mathbb{R}^3$ and the sum is performed on all lattice sites of position $\mathbf{r}_i$, also vanishes. This can actually be proved experimentally, since neutron scattering experiments show no magnetic Bragg peaks. In other words, there is a phase transition (non-zero local magnetization in zero external field), but no long-range order. However, the effect of the local spontaneous magnetizations can be observed experimentally, since they reduce the (static) magnetic susceptibility from the value it would otherwise have. Recall that the zero field magnetic susceptibility is defined as

$$\chi = \lim_{h \to 0^+} \frac{\partial m(h)}{\partial h},$$

where $m(h)$ denotes the magnetization in presence of the magnetic field h. For instance, if the interactions are symmetrically distributed, χ is given by the formula [8]

$$\chi = \frac{1 - |\Lambda|^{-1} \sum_i m_i^2}{T}, \tag{2.4}$$

where T is the temperature. Therefore, the reduction of the susceptibility from the Curie law $\chi = 1/T$ is a direct measure of the mean square local spontaneous magnetization. Experimentally, one finds the $1/T$ behavior for temperatures higher than the so-called freezing temperature T_f, where a cusp is observed (see Fig. 2.2).

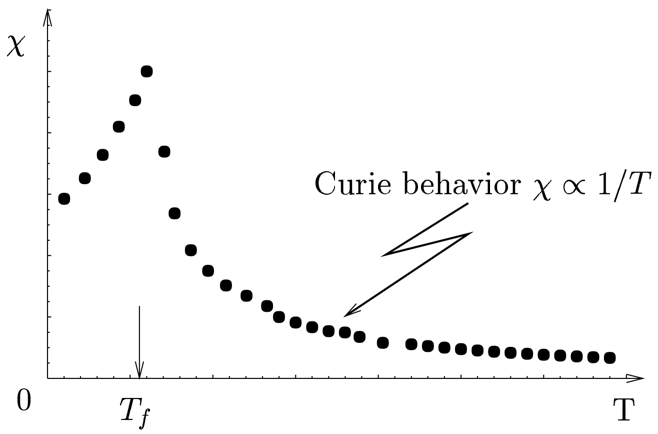

Figure 2.2: Typical behavior of the (zero-field cooled [8]) static magnetic susceptibility for a real spin glass system (in arbitrary units). T_f denotes the freezing temperature, where χ has a cusp.

The freezing temperature marks the boundary between the (low temperature) spin glass phase and the high temperature phase, which is of paramagnetic character.

Even more interesting effects are found when non-equilibrium properties of the system are investigated below T_f. For instance, the frequency-dependent susceptibility $\chi(\omega)$, defined as the response of the magnetization to an external time dependent magnetic field oscillating with frequency ω, depends strongly on the external frequency, even if ω is much smaller than the characteristic microscopic frequency of the system. This phenomenon does not occur in ordered magnets. Generally speaking, the dynamics in the spin glass phase is characterized by very long characteristic times which suggest the presence of many metastable states with high energy barriers separating them. Another important feature of spin glasses is the onset of remanence effects below the freezing temperature. For instance, the value found for the static magnetic susceptibility depends strongly on the way the experiment is performed (for instance, whether the system is cooled in zero or in non-zero external field). In other words, ergodicity of the dynamics is lost below the freezing temperature, and the experimental results depend on the initial condition of the system. The remanence effects in χ have the same origin as the frequency dependence of $\chi(\omega)$, *i.e.*, the existence of many roughly equivalent (quasi)-equilibrium states. Which state is reached when the system is prepared depends crucially on details of the experiment such as the frequency and magnitude of the applied field, the speed with which one it cools down, whether one cools in zero or finite field and so on.

2.2 Finite dimensional vs. mean field spin glasses

The Edwards-Anderson model, although somewhat simplified with respect to the actual physical situation, appeared soon too hard to be attacked analytically, and suitable approximation schemes were sought. Possibly the most important one, and the most rich of surprises, was the *mean field approximation*. In other words, it appeared natural to start from the study of a simplified model where, while maintaining the fundamental features of disorder and frustration, the geometrical structure of the lattice is disregarded, so that every magnetic moment interacts with all others, irrespective of the distance. The first model of this kind was introduced by D. Sherrington and S. Kirkpatrick [16], [17], and is defined by the Hamiltonian

$$H_N(\sigma, h; J) = -\frac{1}{\sqrt{N}} \sum_{1 \le i < j \le N} J_{ij}\sigma_i\sigma_j - h\sum_{i=1}^{N} \sigma_i. \tag{2.5}$$

While this expression is formally similar to the Hamiltonian (2.2) of the Edwards-Anderson model, here the couplings J_{ij} are independent identically distributed random variables with zero mean and *variance independent of* $|i-j|$. Notice also the presence of the normalization factor $1/\sqrt{N}$, where N is the number of spins in the system. This factor, which is absent in the Edwards-Anderson model, is necessary in the mean field case, in order to have a well defined free energy per site in the infinite volume limit (see Section 3.1 of next chapter for a detailed discussion of this point).

Of course, mean field spin glass models are not physically realistic in describing disordered magnetic alloys, since they contain no trace of the lattice structure and of its space dimensionality. However, one of the reasons why they were introduced was the hope that their physical behavior would resemble that of realistic models, at least when the space dimensionality d or the range of interactions is large. This is what happens, for instance, for non-random ferromagnetic systems, where it is well known [1] that mean field theory gives a good qualitative description of symmetry breaking and of the critical point, if d is sufficiently high. Moreover, we saw previously that in the physically more realistic RKKY model the interaction does not decay too fast ($J_{ij} \simeq |i-j|^{-3}$) with the distance, which might further justify this expectation.

After almost thirty years, the question of the connection between mean field and realistic spin glass models is still to a great extent open and under discussion. In fact, while there seems to be no doubt about the existence of a "spin glass" transition for finite range models, at least for $d \geq 3$ (see, for instance, [27], and of course the experimental observations discussed in the previous section), the nature of the spin glass - low temperature - phase is still not understood. Some authors [28] believe that the main features of the mean field picture, like the existence of an infinite number of equilibrium states and their ultrametric structure, persist in the finite dimensional case, while others [29] argue, within the droplet/scaling picture, that below the critical (freezing) temperature just two pure phases exist, connected by global spin-flip symmetry. Both pictures are supported by a mixture of numerical simulations and theoretical considerations. Recent works by Newman and Stein [30] show some of the conceptual difficulties which arise when one tries to precisely define the concept of pure thermodynamic state for disordered systems, and the authors conclude by excluding the possibility of the mean field picture for finite dimensional spin glasses. However, this rather delicate point does not seem to have been settled yet (see, for instance, the recent papers [31], [32]).

One of the few known rigorous results in this context is based on early works by Fröhlich *et al.* [33, 34], later extended by Bovier in [35], where the authors show that, at least for sufficiently high temperature and zero external magnetic field, mean field spin glasses can be actually seen as the limit of short range spin glass models, when the interaction range tends to infinity. As usual in statistical mechanics, the limit where the range of the interaction tends to infinity is performed by means of Kac asymptotics [36], which we recall briefly. Let $J(x)$ be a positive function with compact support satisfying the normalization condition

$$\int d^d x J(x) = 1,$$

and define the Kac potential J_γ as

$$J_\gamma(i,j) = \gamma^d J(\gamma(i-j)). \tag{2.6}$$

Notice that the range of the potential J_γ is of order $1/\gamma$, while its total strength is constant:

$$\sum_{j \in \mathbb{Z}^d} J_\gamma(i,j) = 1 + o(\gamma),$$

where $o(\gamma)$ vanishes as $\gamma \to 0$. In the case of spin glass systems, we define a Kac-type random potential as

$$\sqrt{J_\gamma(i,j)} J_{ij}, \tag{2.7}$$

where J_{ij} are independent identically distributed random variables with zero mean and unit variance. The potential in Eq. (2.7) has clearly a random sign, and its strength is modulated by J_γ. (The reason for the square root in (2.7) is simply that, this way, one has for the second moment of the Hamiltonian $H_\Lambda(\sigma; J)$ (see the definition below)

$$\frac{1}{|\Lambda|} E\, H_\Lambda(\sigma; J)^2 = \frac{1}{2} + o(\gamma)$$

to be compared with

$$\frac{1}{N} E\, H_N(\sigma; J)^2 = \frac{1}{2},$$

which holds for the mean field Sherrington-Kirkpatrick model (see next chapter).) The Hamiltonian of the system, enclosed in a finite box $\Lambda \in \mathbb{Z}^d$ with $|\Lambda|$ sites, is defined as

$$H_\Lambda(\sigma; J) = - \sum_{1 \le i < j \le |\Lambda|} \sqrt{J_\gamma(i,j)} J_{ij} \sigma_i \sigma_j .$$

If one lets the interaction range tend to infinity, *i.e.*, $\gamma \to 0$ after the thermodynamic limit has been performed, one finds [35] that the free energy per site approaches that of the Sherrington-Kirkpatrick model, provided that the temperature is *above* the critical temperature $T_c = 1$ of the Sherrington-Kirkpatrick model. This result is interesting but quite weak, since one would like to study the connection between finite dimensional and mean field models *below* the critical temperature, in the spin glass phase where replica symmetry and ergodicity are broken [6].

Whatever the connection between finite dimensional and mean field models may be, however, it is beyond doubt that the latter deserve a deep study. This is not only because they are interesting mathematical models by themselves, but also because, as we are going to discuss briefly in next section, they arise naturally in many contexts like, *e.g.*, combinatorial optimization problems, neural networks and so on.

2.3 Other examples of spin glass systems

One of the most interesting aspects of the theory of spin glasses is that it has applications in fields apparently very far from statistical mechanics, such as combinatorial optimization problems and neural networks. We suggest Refs. [6], [10], [7] for a beautiful introduction to these subjects.

We limit ourselves to an example, which should illustrate the connection between combinatorial optimization problems and spin glasses. Suppose there is a group of N people, which we denote by $P_1, P_2, P_3, \ldots$, who know each other. Given any couple of individuals, they can be either friends or enemies, no intermediate situation being allowed. We assume (although this is of course not realistic) that the friendship-enmity relations are assigned randomly, and independently for each couple. Of course, if (P_1, P_2) and (P_2, P_3) are two couples of friends, it is not guaranteed that P_1 and P_3 are also friends. In this sense, the system is frustrated, because P_2 has to choose between P_1 and P_3. Now, one wants to divide the N individuals into two parties, so as to minimize social discomfort, *i.e.*, one tries to group as far as possible friends together, and to separate enemies. In order to show how this problem is connected with statistical mechanics and spin glass theory, for any of the 2^N possible ways (configurations) the N people can be divided, assign to the generic individual P_i the "spin variable" $\sigma_i = -1$ if he is assigned to the first group, and $\sigma_i = 1$ in the opposite case. Moreover, given i, j set the "coupling constant" J_{ij} to -1 if i and j are enemies, and to $+1$ otherwise. In this way, it is clear that the problem to find the optimal division of the group is equivalent to find the minimum of the "cost function"

$$H_N(\sigma; J) = -\sum_{1 \leq i < j \leq N} J_{ij}\sigma_i\sigma_j, \tag{2.8}$$

over all configurations σ. Comparing the expression of the cost function with Eq. (2.5), it is clear that our minimization problem is equivalent to the search for the ground state configuration of the mean field Sherrington-Kirkpatrick model at zero magnetic field ($h = 0$). The role of disorder is played in this case by the random choice of friendship-enmity relations.

The one we have just illustrated is only a very simple, and quite academic, example. Other more complicated combinatorial optimization problems connected with spin glasses arise in theoretical computer science, and are intensively studied by information scientists, mathematicians and theoretical physicists. We refer to [37] for a nice introduction. Typically, one has N Boolean variables and looks for a configuration which satisfies certain conditions (clauses) which are generally in conflict among themselves, or for the configuration satisfying the largest possible number of clauses. These are called in general *satisfiability problems*. Also in this case, as with spin glasses, we are in presence of frustration and one cannot trivially satisfy all the conditions at the same time. The scheme is to introduce a "cost function" (which is the analog of the Hamiltonian for spin systems) which counts how many of the conditions are violated for a given Boolean configuration, and to try to minimize it. In statistical mechanics, this corresponds to search for the ground state of the system. To every Boolean variable, which can assume two different values, one assigns an Ising spin variable $\sigma = \pm 1$. The connection with spin glasses is complete when one assumes that clauses are randomly chosen from some given probability distribution, thus playing the role of quenched disorder, and when one considers the limit of N going to infinity (thermodynamic limit in statistical mechanics). This is the relevant limit, because this kind of optimization problems become interesting and hard to solve, even with the most powerful computers, when the number of Boolean variables and of clauses involved becomes large.

The spin glass models one is led to study when dealing with optimization problems are still of mean field character, but with finite connectivity [24]. This means that still there is no spatial structure or any notion of distance between lattice sites, but every spin interacts with a finite (random) number of other spins, even in the thermodynamic limit. The physical behavior of these systems is quite different from that of infinite-connectivity ones like that of Sherrington and Kirkpatrick, and a complete understanding of the corresponding replica symmetry breaking pattern has not yet been obtained, even from the heuristic point of view. A brief discussion of finite connectivity models is given in Section 3.4.

Chapter 3

Mean field spin glass models: basic definitions

In this chapter, we introduce the main definitions concerning mean field spin glass models. For clarity, we focus in the first place on the Sherrington-Kirkpatrick model, which is the one we will mostly deal with in this work. As a simple extension, we introduce Derrida's p-spin model. Finally, a brief section is dedicated to finite connectivity models, or dilute spin glasses. While a detailed study of dilute spin glasses goes well beyond the scope of the present work, it is instructive to compare their structure with that of the Gaussian models we are mostly interested in.

This chapter is not intended to give a complete list of mean field spin glass models, but rather to introduce the essential concepts and definitions we need in the following like, for instance, that of *quenched free energy* and *replicas*. As a consequence, very well known and interesting models like the Random Energy Model (REM) [38] and Generalized Random Energy Model (GREM) [39] are not presented here. These are very instructive because their exact solution is known, and the predictions of Parisi theory and the replica method can be (successfully) tested. We refer the reader interested in the REM and GREM to the many deep and detailed works which appeared in the literature (see, for instance, [40], [41], [42], [43]).

3.1 The Sherrington-Kirkpatrick model

The generic configuration of the Sherrington-Kirkpatrick model [16, 17] is determined by the N Ising variables $\sigma_i = \pm 1$, $i = 1, 2, \dots, N$. The Hamiltonian of the model, in some external magnetic field h, is

$$H_N(\sigma, h; J) = -\frac{1}{\sqrt{N}} \sum_{1 \le i < j \le N} J_{ij} \sigma_i \sigma_j - h \sum_{i=1}^{N} \sigma_i. \tag{3.1}$$

The first term in (3.1) is a long range random two body interaction, while the second represents the interaction of the spins with the magnetic field h. The external quenched disorder is given by the $N(N-1)/2$ independent and identically distributed random variables J_{ij}, defined for each couple of sites. For the sake of simplicity, we assume each J_{ij} to be a centered unit Gaussian with averages

$$E(J_{ij}) = 0, \quad E(J_{ij}^2) = 1.$$

The Gaussian choice is a matter of convenience. Indeed, as it was already noticed in [17] (see Section 5.5.2 of the present work for a sketch of the proof), any other symmetric probability distribution with finite moments could be chosen for J_{ij}, without modifying the free energy of the system, apart from error terms vanishing in the thermodynamic limit. For instance, the case $J_{ij} = \pm 1$ with equal probability 1/2 is often considered in the literature. On the other hand, independence of the J_{ij}'s is of fundamental importance.

Notice that the couplings are not only random, but have also a random sign, so that the two main features of spin glass systems, *i.e.*, disorder and frustration, are present. Moreover, the Hamiltonian we are considering clearly defines a mean field model, since there is no geometric structure in the lattice, and any spin interacts with any other.

For a given inverse temperature* β, we introduce the disorder dependent partition function $Z_N(\beta, h; J)$, the quenched average of the free energy per site $f_N(\beta, h)$, the Boltzmann-Gibbs state ω_J, and the auxiliary function $\alpha_N(\beta, h)$, according to the definitions

$$Z_N(\beta, h; J) = \sum_{\{\sigma\}} \exp(-\beta H_N(\sigma, h; J)), \tag{3.2}$$

$$-\beta f_N(\beta, h) = \frac{1}{N} E \ln Z_N(\beta, h; J) = \alpha_N(\beta, h), \tag{3.3}$$

$$\omega_J(A) = Z_N(\beta, h; J)^{-1} \sum_{\{\sigma\}} A(\sigma) \exp(-\beta H_N(\sigma, h; J)), \tag{3.4}$$

where A is a generic function of the σ's and E denotes average with respect to the quenched disorder. In the notation ω_J, we have stressed the dependence of the state on the external disorder J, but, of course, there is also a dependence on β, h and N.

The main object of interest is the quenched free energy $f_N(\beta, h)$, from which all thermodynamic quantities can be deduced. The reason why we have introduced the auxiliary function $\alpha_N(\beta, h)$ is that, in some cases, it will be more practical to deal with it, rather than with $f_N(\beta, h)$. It is very important to stress that we want to compute

$$-\frac{1}{N\beta} E \ln Z_N \tag{3.5}$$

(*quenched average* of the free energy) rather than

$$-\frac{1}{N\beta} \ln E Z_N \tag{3.6}$$

(*annealed average*). The computation of the annealed average is trivial, since the Boltzmann factor can be written as the product of statistically independent terms, one for

*here and in the following, we set the Boltzmann constant k_B equal to one, so that $\beta = 1/(k_B T) = 1/T$.

each couple of sites, so that

$$Z_N(\beta, h; J) = \sum_{\{\sigma\}} \prod_{1\le i<j\le N} \exp(\frac{\beta}{\sqrt{N}} J_{ij}\sigma_i\sigma_j) \exp(\beta h \sum_k \sigma_k)$$

and the disorder average factorizes:

$$EZ_N(\beta, h; J) = \sum_{\{\sigma\}} \exp\left(\frac{\beta^2}{2N}\frac{N(N-1)}{2}\right) \exp(\beta h \sum_k \sigma_k) = 2^N \cosh^N(\beta h) \exp\frac{\beta^2}{4}(N-1).$$

Taking the disorder average of the logarithm of the partition function corresponds to consider the J variables as a frozen (quenched) disorder, which does not take part to thermal equilibrium. The free energy is therefore computed for given J, and expectation over disorder corresponds to take the average on many different samples of the material. From the physical point of view, this means we are modeling a disordered system whose impurities have a dynamics which is many orders of magnitude slower than the dynamics of the spin degrees of freedom. In order to clarify this point, consider the β derivative of the free energy which, as it is known from thermodynamics, is connected to the internal energy. For the quenched free energy, one finds

$$-\partial_\beta \frac{1}{N} E \ln Z_N = N^{-1} E \frac{\sum_\sigma H_N(\sigma)\exp(-\beta H_N(\sigma))}{Z_N} = N^{-1} E\, \omega_J(H_N)$$

and, for the annealed one,

$$-\partial_\beta \frac{1}{N} \ln EZ_N = N^{-1} \frac{E\sum_\sigma H_N(\sigma)\exp(-\beta H_N(\sigma))}{E\sum_\sigma \exp(-\beta H_N(\sigma))}.$$

In the first case, one is first computing the thermal average of the Hamiltonian, for a fixed disorder realization, and then is averaging over the disorder distribution. This corresponds to the actual physical situation. On the other hand, in the second case the disorder and spin variables are treated on the same footing, being averaged both in the numerator and in the denominator, and this corresponds to the situation where the J variables also thermalize at the temperature $1/\beta$.

The normalization factor $1/\sqrt{N}$ in the Hamiltonian (3.1) is typical of the mean field character of the model and guarantees a good thermodynamic limit for the free energy per spin, *i.e.*, it ensures that the $N \to \infty$ limit of the free energy per site is finite and non-trivial. Recall that, in the case of the mean field (non-random) ferromagnetic Curie-Weiss model, the Hamiltonian is defined by

$$H_N^{C.W.}(\sigma, h) = -\frac{J}{N}\sum_{i,j=1}^N \sigma_i\sigma_j - h\sum_{i=1}^N \sigma_i. \tag{3.7}$$

Here, the couplings $J_{ij} \equiv J > 0$ are non-random and positive. Since the two body interaction term is the sum of N^2 terms of order 1, the correct normalization factor, which guarantees that H_N/N and the free energy density are of order 1, is $1/N$. In the case of the Sherrington-Kirkpatrick model, on the other hand, the random signs of the couplings J_{ij} produce cancellations among the many terms of H_N, and the factor $1/\sqrt{N}$ is enough to have a finite average energy per particle. This can be easily understood as follows. Suppose that $\sqrt{N}$ in (3.1) is replaced by N, as in the Curie-Weiss case, so that the Hamiltonian becomes

$$H'_N(\sigma, h; J) = -\frac{1}{N} \sum_{1 \le i < j \le N} J_{ij} \sigma_i \sigma_j - h \sum_{i=1}^{N} \sigma_i, \tag{3.8}$$

where the prime reminds us that we have changed the normalization factor. The auxiliary function $\alpha'_N(\beta, h)$, defined in analogy with (3.3), satisfies then the bounds

$$\ln 2 + \ln\cosh h + o(1) \le \alpha'_N(\beta, h) \le \frac{1}{N} \ln E Z'_N(\beta, h; J), \tag{3.9}$$

where we denote with $o(1)$ a quantity which vanishes for $N \to \infty$. The upper bound is obtained by means of the Jensen inequality, observing that the function $x \to \ln x$ is concave. As for the lower bound, it suffices to write

$$\begin{aligned} \alpha'_N(\beta, h) \;&\ge\; \frac{1}{N} E \ln \sum_{\{\sigma\}} \exp\left(-\frac{\beta}{N} \Big| \sum_{1 \le i < j \le N} J_{ij} \Big| + h \sum_i \sigma_i \right) && (3.10) \\ &=\; \ln 2 + \ln\cosh h - \frac{\beta}{N^2} E \Big| \sum_{1 \le i < j \le N} J_{ij} \Big| \\ &=\; \ln 2 + \ln\cosh h - \frac{\beta}{N^2} \sqrt{\frac{N(N-1)}{2}} E|J|, && (3.11) \end{aligned}$$

where J is a standard Gaussian random variable with zero mean and unit variance. The last term in (3.11) clearly vanishes for $N \to \infty$. The average of Z'_N is easily calculated, since it involves only Gaussian integrals, and one finds

$$E Z'_N = (2\cosh h)^N \exp\left(\frac{\beta^2}{4N^2} N(N-1)\right). \tag{3.12}$$

Then, the bounds (3.9), together with (3.11) and (3.12), imply

$$\alpha'_N(\beta, h) \to \ln 2 + \ln\cosh h$$

for any value of the temperature, so that the two-body interaction has no effect in the infinite volume limit, and the model is trivial.

3.2 Replicas and replica overlaps

Let us now introduce the important concept of replicas. Consider a generic number n of independent copies of the system, characterized by the spin variables $\sigma_i^{(1)}$, $\sigma_i^{(2)}$, ..., distributed according to the product state

$$\Omega_J = \omega_J^{(1)} \omega_J^{(2)} \dots \omega_J^{(n)},$$

where each $\omega_J^{(\alpha)}$ acts on the corresponding $\sigma_i^{(\alpha)}$ variables, and all are subject to the *same* sample J of the external disorder. Clearly, the Boltzmann factor for the replicated system is given by

$$\exp\left(-\beta(H_N(\sigma^{(1)},h;J)+H_N(\sigma^{(2)},h;J)+\cdots+H_N(\sigma^{(n)},h;J))\right). \qquad (3.13)$$

These copies of the system are usually called *real replicas*, to distinguish them from those appearing in the so called *replica trick* [6], which requires a limit to zero replicas ($n \to 0$) at some stage.

The overlap between two replicas a, b is defined according to

$$q_{ab}(\sigma^{(a)},\sigma^{(b)}) = \frac{1}{N}\sum_i \sigma_i^{(a)}\sigma_i^{(b)}, \qquad (3.14)$$

and it satisfies the obvious bounds

$$-1 \le q_{ab} \le 1.$$

For a generic smooth function F of the configuration of the n replicas, we define the $\langle\rangle$ averages as†

$$\langle F(\sigma^{(1)},\sigma^{(2)},\ldots,\sigma^{(n)})\rangle = E\,\Omega_J F\left(\sigma^{(1)},\sigma^{(2)},\ldots,\sigma^{(n)}\right), \qquad (3.15)$$

where the Boltzmann-Gibbs average Ω_J acts on the replicated σ variables, and E denotes, as usual, average with respect to the external disorder J.

Replica overlaps are quantities that can be in principle measured, at least in a numerical experiment. In a physical experiment, this would require to have two copies of the system with exactly the same realization of disorder, which is not realistic. Notice that the average over disorder introduces correlations between different groups of replicas, which would be independent under the Boltzmann-Gibbs average Ω_J. For example,

$$\Omega_J(q_{12}q_{34}) = \Omega_J(q_{12})\Omega_J(q_{34}) \qquad (3.16)$$

but

$$\langle q_{12}q_{34}\rangle \ne \langle q_{12}\rangle\langle q_{34}\rangle. \qquad (3.17)$$

On the other hand, the $\langle . \rangle$ averages are obviously invariant under permutation of replica indices. For instance,

$$\langle q_{12}q_{23}\rangle = \langle q_{24}q_{45}\rangle. \qquad (3.18)$$

Overlap distributions carry the whole physical content of the theory [6], and the averages of many physical quantities with respect to the external disorder can be expressed through $\langle . \rangle$ averages involving simple polynomials of the overlaps. As an example, for $h = 0$, the disorder average of the internal energy per spin $N^{-1}\omega_J(H_N)$ is given by [44]

$$N^{-1}E\,\omega_J(H_N) = -\frac{\beta}{2}(1-\langle q_{12}^2\rangle), \qquad (3.19)$$

†Different notations are used sometimes in the literature. For instance, the brackets $\langle . \rangle$ are often used to denote thermal average only, while (especially in the theoretical physics literature) disorder average of an observable $\mathcal{O}$ is denoted by the symbol $\overline{\mathcal{O}}$. However, we believe this should not generate any confusion, since we will stick to our notations all through this work.

while its β derivative is [44]

$$\begin{aligned} N^{-1}\partial_\beta E\,\omega_J(H_N) &= -N^{-1}E(\omega_J(H_N^2)-\omega_J^2(H_N)) \qquad (3.20)\\ &= -\frac{1}{2}(1-\langle q_{12}^2\rangle)+\frac{\beta^2}{2}N(\langle q_{12}^4\rangle-4\langle q_{12}^2q_{23}^2\rangle+3\langle q_{12}^2q_{34}^2\rangle). \end{aligned}$$

The proof of equations like (3.19) and (3.20) is very simple, since it requires only to apply the integration by parts formula

$$E(JF(J))=E\left(\frac{\partial}{\partial J}F(J)\right), \qquad (3.21)$$

which holds for a centered unit Gaussian variable J and any smooth function F.

3.3 Derrida's p-spin model

It is very natural to generalize the Sherrington-Kirkpatrick Hamiltonian, by letting the spins interact through a p-body random mean field potential, where p is an integer. The resulting Hamiltonian is

$$H_N^{(p)}(\sigma,h;J)=-\sqrt{\frac{p!}{2N^{p-1}}}\sum_{1\le i_1<\ldots<i_p\le N}J_{i_1\ldots i_p}\sigma_{i_1}\ldots\sigma_{i_p}-h\sum_{i=1}^N\sigma_i. \qquad (3.22)$$

Like for the Sherrington-Kirkpatrick model, we assume the couplings $J_{i_1\ldots i_p}$ to be independent identically distributed centered Gaussian random variables, with unit variance. One can easily show, as we did for the Sherrington-Kirkpatrick model, that the normalization factor $1/N^{(p-1)/2}$ is actually the one required to yield a good thermodynamic limit. The factor $\sqrt{p!/2}$ is introduced just to ensure that, for $p=2$, one recovers the Sherrington-Kirkpatrick model. Notice that the Gaussian random variables $H_N^{(p)}(\sigma,h;J)$ have a very simple covariance structure, which depends only on the mutual overlap of the configurations σ,σ':

$$E\left(H_N^{(p)}(\sigma)H_N^{(p)}(\sigma')\right)-E\,H_N^{(p)}(\sigma)E\,H_N^{(p)}(\sigma')=\frac{N}{2}q_{\sigma\sigma'}^p+O(1/N). \qquad (3.23)$$

This system, known as p-spin model, was first introduced by Derrida in [38], and later has been widely studied, both in the theoretical and in the mathematical physics literature (see, for instance, [45, 46, 47, 48, 23]).

The physical behavior of the p-spin model is quite different from that of the Sherrington-Kirkpatrick one. Anticipating some concepts which will be introduced in Section 4, we can say that, for $p>2$, one expects replica symmetry to be broken only at the first step (at least as long as the temperature is not too low), while for p=2 there is an infinite pattern of symmetry breaking [6]. The meaning of these statements will be clarified in the following chapter.

The limit of large p has also been considered. For $p\to\infty$, the system reduces to the so called Random Energy Model [38], which consists of 2^N statistically independent and identically distributed Gaussian energy levels $H_N(\sigma; J)$. Formally, this can be seen by letting $p \to \infty$ in (3.23), and noticing that $q^p_{\sigma\sigma'} \to 0$ whenever $\sigma \neq \sigma'$. Of course, this argument is not completely convincing, since the limit $p \to \infty$ has to be taken after the infinite volume limit (the Hamiltonian (3.22) has no meaning for $p > N$), but things can be settled [45], [47] and one can actually prove that the infinite volume free energy of the p-spin model tends to that of the REM, for $p \to \infty$. Owing to statistical independence of the energy levels, the Random Energy Model can be solved exactly, thereby confirming the predictions of Parisi theory. For p large but finite, the system approaches very fast[‡] the REM, and a very detailed analysis in this regime has been recently performed by M. Talagrand [47], [48]. We do not discuss here Talagrand's results, which rely on techniques which are rather far from those employed in this work, and we refer the reader to the original papers.

3.4 Finite connectivity models

So far, we have dealt only with infinite connectivity mean field spin glass models, *i.e.*, models where each spin interacts with all the remaining $N-1$ spins. On the other hand, of great interest are finite connectivity (or dilute) models, where the mean field character is conserved but each spin interacts only with a finite number of other spins. For instance, consider the Viana-Bray [49] model defined by the Hamiltonian

$$H_N(\sigma; J, \chi) = -\sum_{1\leq i<j\leq N} J_{ij}\chi_{ij}\sigma_i\sigma_j. \tag{3.24}$$

The J_{ij} are, as usual, independent standard Gaussian variables, while the χ_{ij}'s are random variables (independent of the J), which assume the value 0 with probability $1-\gamma/N$, for some $\gamma > 0$, and the value 1 with probability γ/N. In other words, given any two sites i and j, they interact with probability γ/N and, in this case, the interaction has a Gaussian distribution. Of course, each spin interacts on average with a finite number (of the order γ) of different spins and this is the reason why the normalization factor $1/\sqrt{N}$ is absent, like in finite dimensional models. On the other hand, there is no geometry in the system (no notion of distance among sites), so that it is still a mean field model.

From the probabilistic point of view notice that, while for infinite connectivity models the Hamiltonian $H_N(\sigma; J)$ can be seen as a correlated Gaussian process on the configuration space $\{-1, +1\}^N$, when the connectivity is finite the process is no longer Gaussian, owing to the presence of the random variables χ_{ij}, and therefore it is no more determined solely by its covariance matrix. It should not be surprising, therefore, that these models are much more complicated, even from the heuristic point of view, and that a complete analysis of Replica Symmetry Breaking, in the context of Parisi theory, has not been achieved yet.

On the other hand, the study of these models is of great interest at least for two reasons. In the first place, finite connectivity makes them more similar to models with

[‡]for p large, the difference between the p-spin and the REM free energy is of order 2^{-p}. For more precise statements, see [46], [47]

short range interactions. Secondly, such models arise very naturally in the context of combinatorial optimization problems [24], neural networks [7] and error correction codes [10].

Chapter 4

An overview of Parisi theory of Replica Symmetry Breaking

In this chapter, we present a brief and streamlined introduction to some central concepts and results of Parisi theory of mean field spin glasses. In particular, we will try to emphasize the meaning of the Replica Symmetry Breaking phenomenon, and of the so called *functional order parameter* $x(q)$. For simplicity, we will concentrate mostly on the Sherrington-Kirkpatrick model, for which the theory was first formulated.

We do not intend to give a complete or even satisfactory review of the theory, both because this would lead us away from the main scope of this work, and because beautiful and thorough reviews exist in the literature. In particular, we refer the interested reader to the original papers (see, for instance, [18], [19] and [20]), and to the book [6]. See also [50] for a different perspective. Rather, we will restrict to those concepts and results which are essential to understand the following of this work, and in particular to appreciate the relevance of the results in Chapter 6.

As a last warning, let me underline that, in this chapter, results are given without their original derivation, which would require to introduce techniques which are very different in spirit from those employ in this work. The results we report here were originally obtained by means of the so called *replica method* which, though mathematically not well understood, is of remarkable elegance and simplicity. Unfortunately, the beauty of the method gets completely lost here, and this is another reason why we warmly recommend Ref. [6] to the reader.

4.1 Replica symmetry breaking and first order phase transitions

Let us begin by recalling the concept of spontaneous symmetry breaking, and of phase coexistence, in ordinary statistical mechanics [2]. Consider for simplicity a system on a d-dimensional cubic lattice, formally defined by a Hamiltonian $H(\sigma)$, which depends on the configuration of all spins σ_i, with $i \in \mathbb{Z}^d$. In order to deal with mathematically well defined objects, one restricts the system and the Hamiltonian to a finite subset Λ of the lattice, computes the finite volume free energy per site

$$f_\Lambda(\beta) = -\frac{1}{|\Lambda|\beta} \ln Z_\Lambda(\beta), \tag{4.1}$$

where β is the inverse temperature and Z_Λ is the partition function, and then lets Λ grow to the whole $\mathbb{Z}^d$ in a suitable way. There is a certain degree of arbitrariness in this procedure because, as long as the system is finite, one has to impose boundary conditions, that is, one has to specify the configuration of the boundary spins, or their interaction with the external world. It is well known that, at least for systems with short range interactions, the free energy per site is not sensitive to boundary conditions, in the limit where $\Lambda \to \mathbb{Z}^d$. However, the equilibrium thermodynamic state of the system is not determined simply by its free energy, but also by all the correlation functions, *i.e.*, by

$$\lim_{\Lambda\to\mathbb{Z}^d} \langle \sigma_{i_1} \dots \sigma_{i_n} \rangle_\Lambda \,, \tag{4.2}$$

for all finite sets of site indices $i_1, \dots, i_n$, where $\langle . \rangle$ denotes the Boltzmann-Gibbs thermal average at temperature $1/\beta$. Unlike the free energy, the correlation functions *do* in general depend on the chosen boundary conditions, also in the infinite volume limit. When the equilibrium state is not unique, one says that a first order phase transition occurs.

Another, strictly related, way to select different equilibrium states, is to introduce auxiliary infinitesimal external fields λ_i, which are removed *only after* the thermodynamic limit $\Lambda \to \mathbb{Z}^d$ has been performed. More precisely, one adds to the Hamiltonian terms which represent the interaction of the system with the external fields, and then computes the thermodynamic limit for the free energy and for the correlation functions. For instance, in the case of an external magnetic field h the interaction term is $-h\sum_i \sigma_i$. After the infinite volume limit has been taken, one lets the λ_i go to 0, and the resulting correlation functions depend, in general, on the way the limit is performed, *e.g.*, whether $\lambda_i \to 0^+$ or $\lambda_i \to 0^-$, and more generally on the order how the various fields go to 0.

The set of all equilibrium states forms a convex set and, in addition, it is a simplex [2], so that every state can be written in an unique way as a convex linear combination of certain *extremal* states, which are called *pure states* or *pure phases*. Pure phases are characterized by the *cluster property*, or spatial decay of correlations, that is

$$\langle \sigma_{i_1} \dots \sigma_{i_n} \sigma_{j_1} \dots \sigma_{j_k} \rangle \to \langle \sigma_{i_1} \dots \sigma_{i_n} \rangle \langle \sigma_{j_1} \dots \sigma_{j_k} \rangle \,, \tag{4.3}$$

for

$$\min_{a,b} |i_a - j_b| \to \infty.$$

Usually, a first order phase transition is associated to the phenomenon of spontaneous symmetry breaking. This means that, while the Hamiltonian of the model has a certain symmetry (for instance, spin-flip or rotational symmetry), the equilibrium states possess only a smaller symmetry group. This implies that the choice of the auxiliary external fields, or equivalently of boundary conditions, which select the pure states, is suggested by the symmetry of the model. For instance, if rotation symmetry is broken, the pure states are selected by an infinitesimal uniform magnetic field which destroys space isotropy, and so on. Of course, applying the symmetry transformation to a particular symmetry-breaking phase, one obtains another (physically different) equilibrium state.

As a last comment, recall that *Gibbs' phase rule* usually holds, which states that, in order that the system has n coexisting pure phases, one has to tune $n-1$ thermodynamic parameters. For instance, in order to have coexistence of water and ice it is sufficient to fix the temperature, for a given pressure, while the triple point of vapor-liquid-solid coexistence requires to fix also the pressure.

In the case of spin glasses, the situation is much more complicated. Indeed, at low temperature one expects the pure phases to be many (actually, infinite [6]), and not related by symmetry transformations. In other words, there is phase transition but no obvious broken symmetry. Recalling the discussion above, this means that it is not clear *a priori* what should be the right boundary conditions, or external fields, which select the pure states. Moreover, Gibbs rule is clearly violated, since one has an infinite number of pure phases for a generic choice of the thermodynamic parameters.

One of the major achievements of Parisi theory of mean field spin glasses is that it showed that the study of the spin glass phase transition can be led to the framework of spontaneous symmetry breaking. However, due to the very peculiar nature of these models, the broken symmetry is also very unusual. Indeed, one has to break the symmetry under permutations for a group of n identical copies (replicas) of the system, in the limit $n \to 0$. Let me shortly give the idea of the method.

Recall that the main purpose of the theory is to compute the disorder average of the free energy, defined in the previous chapter:

$$-\frac{1}{\beta N} E \ln Z_N(\beta, h; J). \tag{4.4}$$

The replica trick is based on the observation that the average of the integer moments of Z_N is easy to compute, and that

$$E \ln Z_N = \lim_{n \to 0} \frac{E Z_N^n - 1}{n}. \tag{4.5}$$

Notice that Eq. (4.5) requires not only knowledge of the integer moments of Z_N, but also of non-integer ones. The scheme one follows is to compute $E Z_N^n$ for integer values of n, and then somehow analytically continue it to real n. Finding the right analytic continuation required long efforts and attempts at the beginning of the 80's, until the right strategy was proposed by Parisi. Computation of $E Z_N^n$ for $n \in \mathbb{N}$ is equivalent to calculate the average of the partition function of a system of n identical and non-interacting copies (replicas) of the original system, subject to the same realization of disorder:

$$E Z_N^n(\beta, h; J) = E \sum_{\{\sigma^{(1)}\}} \cdots \sum_{\{\sigma^{(n)}\}} \exp\left(-\beta \sum_{a=1}^{n} H_N(\sigma^{(a)}, h; J) \right). \tag{4.6}$$

The average over disorder can be easily performed, since it involves only Gaussian integrals, and one finds [6] an expression of the kind

$$E Z_N^n(\beta, h; J) = \int \prod_{1 \le a < b \le n} (dQ_{ab} (N\beta^2/2\pi)^{\frac{1}{2}}) \exp(-N A[Q]), \tag{4.7}$$

where Q is a $n \times n$ symmetric matrix, with zeros on the diagonal, and $A[Q]$ is a functional depending on $Q, n, \beta, \ldots$, whose explicit expression we do not report here. Equation (4.7) suggests that the n-th moment of Z_N can be computed through the saddle point method, in the limit of N going to infinity. Once the saddle point has been found, the infinite volume free energy is obtained as

$$f(\beta, h) = \lim_{n \to 0} \frac{1}{\beta n} A[Q_{sp}]. \tag{4.8}$$

Owing to equivalence among replicas, which reflects into symmetry of $A[Q]$ with respect to permutation of rows or columns of the matrix Q, the most natural idea is to look for a replica symmetric saddle point, corresponding to a matrix Q whose non-diagonal elements have all the same value q, the elements on the diagonal being zero. In this case, the integral in (4.7) reduces to an ordinary integral over the real variable q. However, when one performs the limit of zero replicas (4.8), one finds [6] an expression for the free energy (the so called replica-symmetric approximation $f_{SK}(\beta, h)$, which will be discussed in the following) which is not physically acceptable in some region of parameters, since it violates basic thermodynamic stability conditions.

Therefore, one has to look for a saddle point which breaks symmetry between replicas. Notice that the number of independent parameters of the matrix Q is $n(n-1)/2$, which becomes negative when $n \to 0$. In other words, one is looking for the relevant saddle point of an integrand which depends on a negative number of variables. In a series of remarkable papers [18]-[20], by means of a powerful *Ansatz*, Parisi proposed a form for the saddle points, which is by now widely believed to be the correct one. We will not pursue this idea here, and refer instead to Chapters 1-3 of Ref. [6] for a detailed discussion.

Rather, let us go back to our discussion about ordinary symmetry breaking, and consider the Ising model which, as it is well known, at low enough temperature and zero magnetic field has two pure phases, one with magnetization $+m(\beta)$ and the other $-m(\beta)$. If one takes two typical configurations belonging to the same phase, one finds that their overlap, as defined in Eq. (3.14), equals

$$q_{++} = q_{--} = m^2(\beta)$$

while, for two different phases,

$$q_{+-} = -m^2(\beta).$$

This holds only in the thermodynamic limit, since for finite volume systems there can be strictly speaking no symmetry breaking. Therefore, the distribution function of the overlap q_{12} between the configurations of two replicas, picked according to their Boltzmann weights, is given in the limit of infinite volume by the sum of two delta functions

$$P(q) = \frac{\delta(q - m^2(\beta)) + \delta(q + m^2(\beta))}{2}. \tag{4.9}$$

On the other hand, above the critical temperature there is just one pure phase, with zero magnetization, and in that case

$$P(q) = \delta(q). \tag{4.10}$$

This means that the analysis of $P(q)$ allows to detect the phenomenon of non-uniqueness of the state, without introducing symmetry-breaking external fields or boundary conditions. For spin glasses, where there is no obvious symmetry to be broken, and consequently no natural order parameter and associated field, the way to be pursued is to compute

$$P(q) = \lim_{N\to\infty} E\, P_J^{(N)}(q),$$

where $P_J^{(N)}(q)$ is the finite volume probability distribution of the overlap, for a given disorder realization J. If the limit distribution $P(q)$ is a single delta function (or a sum of two symmetric deltas, in absence of magnetic field) then there is just one pure state (respectively, two states related by spin-flip symmetry), and the systems is said to be replica symmetric. If, on the contrary, $P(q)$ has more than two peaks, or has a continuous part, replica symmetry is said to be broken.

4.2 The functional order parameter and Parisi solution

Parisi theory gives a precise prescription on how to compute both the infinite volume free energy per site and $P(q)$, through the replica method. In the following of the present chapter, we present the main results of the theory. To this purpose, we need some preliminary definitions. Although all of this can be found, for instance, in [6], we use a slightly different language. We refer to [51] and to the forthcoming paper [52] for a review on Parisi theory, along these lines. First of all, let us introduce the convex space $\mathcal{X}$ of the *functional order parameters* x, as non-decreasing functions of the auxiliary variable q, both x and q taking values on the interval $[0,1]$, *i.e.*

$$\mathcal{X} \ni x : [0,1] \ni q \to x(q) \in [0,1]. \tag{4.11}$$

Notice that we call x the non-decreasing function, and $x(q)$ its values. We introduce a metric on $\mathcal{X}$ through the $L^1([0,1], dq)$ norm, where dq is the Lebesgue measure.

Usually, we will consider the case of piecewise constant functional order parameters, characterized by an integer K, and two sequences $q_0, q_1, \dots, q_K$ and $m_1, m_2, \dots, m_K$ of numbers satisfying

$$0 = q_0 \le q_1 \le \dots \le q_{K-1} \le q_K = 1, \tag{4.12}$$

$$0 \le m_1 \le m_2 \le \dots \le m_K \le 1, \tag{4.13}$$

such that

$$\begin{aligned} x(q) &= m_1 \quad \text{for} \quad 0 = q_0 \le q < q_1, \\ x(q) &= m_2 \quad \text{for} \quad q_1 \le q < q_2, \\ &\dots \\ x(q) &= m_K \quad \text{for} \quad q_{K-1} \le q \le q_K. \end{aligned} \tag{4.14}$$

In the following, we will find convenient to define also $m_0 \equiv 0$, and $m_{K+1} \equiv 1$. The choice of a piecewise constant order parameter corresponds, in the frame of Parisi theory, to consider replica symmetry breaking at a finite number of steps. For instance,

$$K = 2,\ q_1 = \bar{q},\ m_1 = 0,\ m_2 = 1. \tag{4.15}$$

corresponds to the so called *replica symmetric case*. The case $K = 3$ gives the first level of replica symmetry breaking, and so on.

Let us now introduce the function f, with values $f(q, y; x, \beta)^*$, of the variables $q \in [0, 1]$, $y \in \mathbb{R}$, depending also on the functional order parameter x and on the inverse temperature β, defined as the solution of the nonlinear antiparabolic equation

$$(\partial_q f)(q, y) + \frac{1}{2}\big(f''(q, y) + x(q) f'^2(q, y)\big) = 0, \tag{4.16}$$

with final condition

$$f(1, y) = \ln\cosh(\beta y). \tag{4.17}$$

Here, we stressed only the dependence of f on q and y, and we put $f' = \partial_y f$ and $f'' = \partial_y^2 f$.

It is very simple to integrate Eq. (4.16) when x is piecewise constant. In fact, consider $x(q) = m_a$, for $q_{a-1} \leq q \leq q_a$, firstly with $m_a > 0$. Then, it is immediately seen that the correct solution of Eq. (4.16) in this interval, with the right final boundary condition at $q = q_a$, is given by

$$f(q, y) = \frac{1}{m_a} \ln \int \exp\big(m_a f(q_a, y + z\sqrt{q_a - q})\big)\, d\mu(z), \tag{4.18}$$

where $d\mu(z)$ is the centered unit Gaussian measure on the real line:

$$d\mu(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) dz.$$

On the other hand, if $m_a = 0$, then (4.16) loses the nonlinear part and the solution is given by

$$f(q, y) = \int f(q_a, y + z\sqrt{q_a - q})\, d\mu(z), \tag{4.19}$$

which can be seen also as deriving from (4.18) in the limit $m_a \to 0$. Starting from the last interval K, and using (4.18) iteratively on each interval, we easily get the solution of (4.16), (4.17), in the case of piecewise order parameter x, as in (4.14).

We refer to [52] for a detailed discussion about the properties of the solution $f(q, y; x, \beta)$ of the antiparabolic equation (4.16), with final condition (4.17), as a functional of a generic given x, as in (4.14). Here we only state the following

Theorem 1 *The function f is monotone in x, in the sense that $x(q) \leq \bar{x}(q)$, for all $0 \leq q \leq 1$, implies $f(q, y; x, \beta) \leq f(q, y; \bar{x}, \beta)$, for any $0 \leq q \leq 1$, $y \in R$. Moreover f is pointwise continuous in the $L^1([0, 1], dq)$ norm. In fact, for generic x, $\bar{x}$, we have*

$$|f(q, y; x, \beta) - f(q, y; \bar{x}, \beta)| \leq \frac{\beta^2}{2} \int_q^1 |x(q') - \bar{x}(q')|\, dq'.$$

*the reader should not confuse the function $f(q, y; x, \beta)$ with the free energy per particle $f(\beta, h)$. Both are denoted as f in the literature, and we do not wish to change conventions here.

This result is very important. In fact, any functional order parameter can be approximated in the L^1 norm through a piecewise constant one. The pointwise continuity allows us to deal mostly with piecewise constant order parameters.

Now we are ready for the following important definitions.

Definition 1 *The trial auxiliary function $\bar{\alpha}$, depending on the functional order parameter x, is defined as*

$$\bar{\alpha}(\beta, h; x) \equiv \ln 2 + f(0, h; x, \beta) - \frac{\beta^2}{2} \int_0^1 q\, x(q)\, dq. \tag{4.20}$$

Notice that in this expression the function f appears evaluated at $q = 0$, and $y = h$, where h is the value of the external magnetic field.

Definition 2 *The Parisi spontaneously broken replica symmetry solution is defined by*

$$\bar{\alpha}(\beta, h) \equiv \inf_x \bar{\alpha}(\beta, h; x), \tag{4.21}$$

where the infimum is taken with respect to all functional order parameters x.

The main prediction of Parisi theory is that, for the Sherrington-Kirkpatrick model,

$$-\beta f(\beta, h) = \lim_{N\to\infty} \frac{1}{N} E \ln Z_N(\beta, h; J) = \bar{\alpha}(\beta, h). \tag{4.22}$$

Moreover, Parisi gives the following interpretation to the functional parameter x, in correspondence of which the infimum in (4.21) is realized†:

$$x(q) = \int_0^q P(q) dq. \tag{4.23}$$

If replica symmetry holds, *i.e.*, $P(q) = \delta(q - \bar{q})$, the optimal order parameter is just a step function which equals 0 for $q < \bar{q}$, and 1 for $q > \bar{q}$. As more and more steps appear in the optimal functional order parameter $x(q)$, replica symmetry is said to be broken at more and more levels.

There is an important point which deserves to be noticed. Eqs. (4.22), (4.21) show that the infinite volume free energy $f(\beta, h)$ is obtained by *maximizing* the trial functional $-\beta^{-1}\bar{\alpha}(\beta, h; x)$ over the space of functional order parameters. This has to be compared with the usual variational principle of statistical mechanics [1], implied by the second principle of thermodynamics, which states that the true free energy is obtained through *minimization* of a suitable free energy functional on all possible trial states. More precisely, for any system at thermodynamic equilibrium (spin glasses included‡),

$$f_N(\beta, h) = \inf_\rho (u_N(\rho) - \beta^{-1} s_N(\rho)). \tag{4.24}$$

† in the particular case $h = 0$, where the system is invariant under spin-flip, $P(q)$ is symmetric and Eq. (4.23) is replaced by $x(q) = 2\int_0^q P(q)dq$. On the other hand, as soon as an arbitrarily small magnetic field h is present, one expects $P(q)$ to have support only on the interval $[0, 1]$.

‡We are always considering systems of finite volume, so that relaxation to equilibrium is always guaranteed.

Here, $\rho : \sigma \to \rho(\sigma)$ is a generic state, *i.e.*,

$$\rho(\sigma) \geq 0$$

and

$$\sum_{\{\sigma\}} \rho(\sigma) = 1,$$

and $u(\rho), s(\rho)$ are the corresponding internal energy and entropy, respectively:

$$u_N(\rho) = \sum_{\{\sigma\}} \rho(\sigma) H_N(\sigma)$$

$$s_N(\rho) = -\sum_{\{\sigma\}} \rho(\sigma) \ln \rho(\sigma).$$

As it is well known, it turns out that the infimum in (4.24) is realized in correspondence of the Boltzmann-Gibbs state

$$\rho(\sigma) = Z_N^{-1} \exp(-\beta H_N(\sigma)).$$

Together, Eqs. (4.24) and (4.22) imply that, for any order parameter x different from the optimal one, $-\beta^{-1}\bar{\alpha}(\beta, h; x)$ *cannot* be interpreted as the free energy associated to some trial state. The rather mysterious maximization procedure in (4.22) is usually justified [6] by pointing out that, in the limit where the number of replicas goes to zero, the relevant saddle point in the computation of the functional integral (4.7) is the one which *minimizes* the exponent and not the one which maximizes it since then, for $n < 1$, *the Hessian matrix has a negative number of negative eigenvalues.* As we shall see in Chapter 6, recent results by F. Guerra [23] give a more firm ground to this procedure.

4.3 The phase diagram of the Sherrington-Kirkpatrick model

It is very interesting to discuss the phase diagram of the Sherrington-Kirkpatrick model, as emerging from Parisi theory. Let us discuss first of all the high temperature, or high external field, region. To this purpose, define the Sherrington-Kirkpatrick order parameter $\bar{q}$ [16], depending on β, h, as the solution of the implicit equation

$$\bar{q} = \int \tanh^2(\beta h + \beta z\sqrt{\bar{q}})\, d\mu(z). \tag{4.25}$$

In Ref. [53] it was proven that, for any $\beta \geq 0$ and $h \neq 0$, the solution of Eq. (4.25) exists and is unique (the same result was found by R. Latala, in unpublished work). The high temperature (or replica symmetric, or ergodic) region is defined by the condition [54]

$$\beta^2 \int \frac{1}{\cosh^4(\beta h + \beta z\sqrt{\bar{q}(\beta, h)})}\, d\mu(z) \leq 1. \tag{4.26}$$

In this region, the infimum in (4.21) is obtained§ in correspondence of a $x(q)$ with the simple step form

$$x(q) = \begin{cases} 0 & 0 \leq q < \bar{q} \\ 1 & \bar{q} \leq q \leq 1. \end{cases} \tag{4.27}$$

In correspondence of this very simple expression for the functional order parameter, one finds that the Parisi solution (4.22) is given by

$$\alpha_{SK}(\beta, h) = \ln 2 + \int \ln\cosh(\beta h + \beta z\sqrt{\bar{q}})\, d\mu(z) + \frac{\beta^2}{4}(1-\bar{q})^2. \tag{4.28}$$

This is just the so called "replica symmetric solution", which was originally found in [16]. It is easy to see that the Sherrington-Kirkpatrick order parameter defined by Eq. (4.25) is just the value which minimizes the expression (4.28), considered as a function of generic $\bar{q}$, in agreement with (4.21). The authors of [16] soon realized that the replica symmetric solution cannot hold in the whole region of parameters β, h, for basic thermodynamic reasons, based on the positivity of the entropy. This is discussed in greater detail in Section 6.5. Recalling Eq. (4.23) and (4.27) we find that, in the region where the replica symmetric solution holds, the overlap does not fluctuate and its typical value is just $\bar{q}$:

$$P(q) = \delta(q - \bar{q}). \tag{4.29}$$

The critical *Almeida-Thouless line* [54]

$$\beta^2 \int \frac{1}{\cosh^4(\beta h + \beta z\sqrt{\bar{q}(\beta, h)})}\, d\mu(z) = 1 \tag{4.30}$$

is the boundary between the replica-symmetric region and the (low temperature) spin glass phase, and marks the onset of replica symmetry breaking. Below¶ this line, replica symmetry is broken at an infinite number of levels, *i.e.*, the optimal $x(q)$ is given by the limit of piecewise constant functional order parameters, where the number of steps tends to infinity, and of course their height and width go to zero. As a consequence, the overlap probability distribution has support on a whole interval $[q_m, q_M]$. A perturbative computation of $P(q)$ in the neighborhood of the Almeida-Thouless line is possible.

In the spin glass phase there exists an infinite number of thermodynamic states which, in contrast with usual ordered systems, are not connected one to another by any simple symmetry transformation. However, the set of all pure states is characterized by a very peculiar geometric structure (ultrametricity), which can be explained as follows: Let $\alpha, \beta, \ldots$ denote the different pure phases of the system and let $\sigma^{(\alpha)}$, $\sigma^{(\beta)}, \ldots$ be configurations belonging to them. Then, for almost every choice ("almost" with respect to Gibbs measure) of the configurations, the overlap

$$q_{\alpha\gamma} = \sum_i \sigma_i^{(\alpha)} \sigma_i^{(\gamma)} / N \tag{4.31}$$

§actually, this has never been rigorously proven, although it is believed to be true.

¶ When we say "below" or "above the Almeida-Thouless line", we think of the phase diagram where we put β on the x-axis, and h on the y-axis as in figure 4.1, so that the region below the critical line corresponds to low temperature and/or small magnetic field, and vice versa.

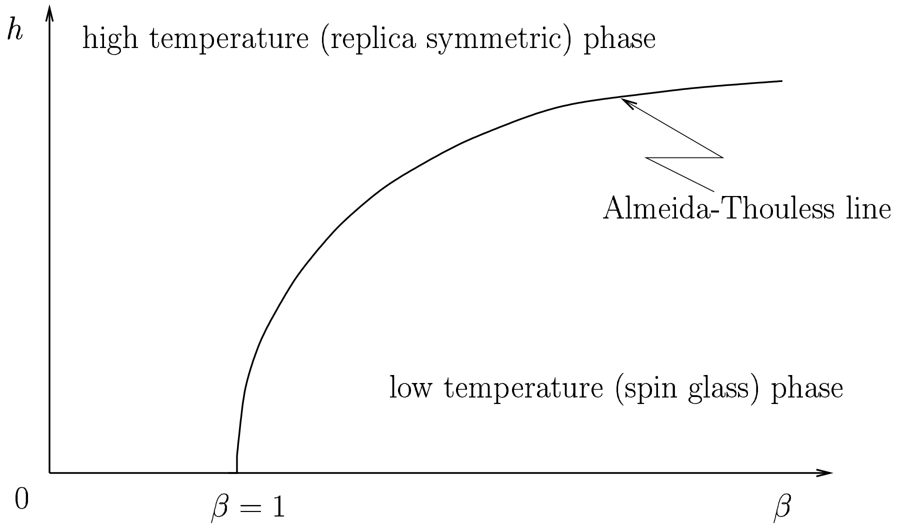

Figure 4.1: The phase diagram of the Sherrington-Kirkpatrick model and the Almeida-Thouless line.

will assume the same value

$$q_{\alpha\gamma} = \sum_i m_i^{(\alpha)} m_i^{(\gamma)}/N, \tag{4.32}$$

where $m_i^{(\alpha)}$ is the thermal average of σ_i in the state α. Now, introduce the following very natural notion of distance between two states:

$$d_{\alpha\gamma}^2 = \frac{1}{N}||m^{(\alpha)} - m^{(\gamma)}||^2 \equiv \frac{1}{N}\sum_i (m_i^{(\alpha)} - m_i^{(\gamma)})^2 \tag{4.33}$$

It follows from Parisi theory that

$$q_{\alpha\alpha} \equiv \frac{1}{N}\sum_i \left(m_i^{(\alpha)}\right)^2 = q_{EA} \tag{4.34}$$

does not depend on the state α, so that we can write

$$d_{\alpha\gamma}^2 = 2(q_{EA} - q_{\alpha\gamma}). \tag{4.35}$$

The property of ultrametricity states that, for any choice of three pure states, the resulting triangle is either equilateral or isosceles with respect to the metric $d_{\alpha\beta}$ and, in the latter case, the different side must be the smaller one. Clearly, ultrametricity is stronger than the usual triangular inequality.

It is important to stress that ultrametricity follows from the Parisi *Ansatz* for the relevant saddle point in (4.7), and therefore in a sense is assumed from the beginning. On the other hand, no one has so far proposed a different expression for the saddle point, which has no underlying ultrametric structure, and which allows to perform the limit to zero replicas as in (4.8). Even the cavity method [55], [6], which is considered as an alternative to the replica trick, and which leads to the same results, assumes ultrametricity from the beginning.

Parisi theory of Replica Symmetry breaking, proposed around twenty years ago, is by now fully accepted from the theoretical point of view and widely supported by numerical simulations. However, a mathematical justification is still lacking. On one hand, the replica trick involves not well defined objects like 0×0 matrices and functions of a negative number of variables. On the other hand, many rigorous results, among which those discussed in the next chapters, provide a strong support to Parisi theory for mean field spin glasses, and in all cases where a full rigorous analysis was possible, its predictions have been confirmed.

Chapter 5

Existence of the thermodynamic limit

5.1 Introduction

As always in statistical mechanics, also for spin glasses the main object of interest is the free energy per site, from which one can compute all equilibrium thermal averages, by performing derivatives with respect to the suitable thermodynamic parameters like temperature, magnetic field and so on. Of course, the first question one may ask is whether the free energy per site has a unique limit when the size of the systems grows to infinity, or if it depends on the particular sequence of system sizes one chooses to reach the thermodynamic limit. For ordinary non-random translation invariant systems with short range interactions, it is well known [2] that the limit is unique. The strategy of the proof consists in dividing the system into large sub-systems, and showing that the energy of interaction among the sub-systems is a surface effect which can be neglected with respect to the bulk energy. This implies that the free energy per site stays essentially the same, when the size of the system is increased. In mean field models, on the other hand, this approach does not work. The reason is that, owing to the infinite range of the interaction, what should be *surface terms are actually of the same order as the bulk terms*. For this reason, the problem of proving the existence of the thermodynamic limit for mean field spin glass models, independently of an explicit computation of the limit itself, has been considered open until very recent times. In a joint work with F. Guerra [22], we solved this problem by means of a very simple but general strategy, based on a smooth interpolation between a large system, made of N spin sites, and two similar but independent subsystems, made of N_1 and N_2 sites, respectively, with $N_1 + N_2 = N$. The main result is that, for a very large class of mean field spin glass systems, which includes most of the models introduced in Chapter 3, the limit exists and is unique. This holds both for the quenched average of the free energy per site, and for the disorder dependent one, for almost every disorder realization. It is important to emphasize that this is true not only in the high temperature region, where the limit can be explicitly computed, but for any value of the thermodynamic parameters. By simple thermodynamic considerations based on convexity of the free energy, we were able to extend these results to the thermodynamic limit of the ground state energy. Subsequently, the existence of the infinite volume limit has been extended by Contucci *et al.* [56] to include also the REM and GREM, and by Franz and Leone [57] to the finite connectivity (or diluted) spin glasses (see Section 3.4 of the present work).

We begin by recalling the usual strategy which works for short range random and non-random systems. This is very instructive, since it shows what goes wrong in the case of mean field models. Then, after discussing the self-averaging property of the free energy, we present our main results about the existence of the limit for the Sherrington-Kirkpatrick and p-spin model, together with some extensions to more general mean field spin glass systems with non-Gaussian couplings and non-Ising spin degrees of freedom, which we obtained in later work [58].

5.2 The thermodynamic limit for short range models

5.2.1 Non-random systems

In ordinary translation invariant non-random systems with short range interactions, it is simple to show that the free energy $f_N(\beta) = -1/(N\beta)\ \ln Z_N$ has a well defined limit for $N \to \infty$ [2, 3]. For simplicity, we consider the particular case of the d-dimensional Ising model with nearest-neighbor interaction. Given a finite subset Λ of the lattice $\mathbb{Z}^d$, the Hamiltonian of the system is

$$H_\Lambda(\sigma, h) = -J \sum_{|i-j|=1} \sigma_i \sigma_j - h \sum_{i=1}^{|\Lambda|} \sigma_i, \tag{5.1}$$

where the first sum is performed over all couples of neighboring spins belonging to Λ, and $|\Lambda|$ denotes the number of sites contained in the considered portion of lattice. We consider the particular case of free boundary conditions but, as we have already stated, the infinite volume free energy does not depend on the particular choice of boundary conditions. In the course of the proof we need the following inequality. Let $H_\Lambda^{(1)}, H_\Lambda^{(2)}$ be two generic Hamiltonians defined on the same finite set Λ of $\mathbb{Z}^d$, and $f_\Lambda^{(1)}(\beta), f_\Lambda^{(2)}(\beta)$ the corresponding finite volume free energies per site. Then,

$$|f_\Lambda^{(1)}(\beta) - f_\Lambda^{(2)}(\beta)| \le |\Lambda|^{-1} ||H_\Lambda^{(1)} - H_\Lambda^{(2)}||, \tag{5.2}$$

where $||H_\Lambda||$ is defined as the maximum value of $|H_\Lambda(\sigma)|$ over all spin configurations in Λ. This can be easily obtained as follows. For every configuration σ we have

$$e^{-\beta||H_\Lambda^{(1)} - H_\Lambda^{(2)}||} e^{-\beta H_\Lambda^{(2)}(\sigma)} \le e^{-\beta H_\Lambda^{(1)}(\sigma)} \le e^{\beta||H_\Lambda^{(1)} - H_\Lambda^{(2)}||} e^{-\beta H_\Lambda^{(2)}(\sigma)} \tag{5.3}$$

which, after summing over all configurations and taking the logarithm, yields (5.2). Now consider domains Λ of increasing size, and in particular hypercubes of side $L_k = 2^k$, with $k \in \mathbb{N}$. We let $H_k, f_k(\beta)$ denote the Hamiltonian and the free energy for such hypercubes, for a given inverse temperature β. Dividing the hypercube of side L_k into 2^d hypercubes of side L_{k-1}, one has

$$H_k = H_{n.i.} + H' \equiv H_{k-1}^{(1)} + \ldots + H_{k-1}^{(2^d)} + H', \tag{5.4}$$

where $H_{k-1}^{(i)}$ refers to the sub-system i, and H' contains the interaction terms between spins belonging to different sub-systems, so that $H_{n.i.}$ is the Hamiltonian of the system where the 2^d hypercubes do not interact. In absence of H', the free energy would equal $f_{n.i.}(\beta, h) = f_{k-1}(\beta, h)$ (it suffices to observe that the partition function factorizes in the product of the partition functions for the different sub-systems, and that the interaction is translation invariant). Therefore, inequality (5.2) implies

$$|f_k(\beta, h) - f_{k-1}(\beta, h)| \leq \frac{||H'||}{2^{d\,k}} \leq d|J|2^{-k}, \tag{5.5}$$

since H' is the sum of $d\,2^{k(d-1)}$ terms, each of which equals $|J|$ in absolute value. By repeatedly applying (5.5), one finds that for any positive integer m,

$$|f_{k+m}(\beta, h) - f_k(\beta, h)| \leq d|J| \sum_{j=k+1}^{\infty} 2^{-j} = d|J|2^{-k}, \tag{5.6}$$

so that $f_k(\beta, h)$ is a Cauchy sequence and admits a limit for $k \to \infty$, for any β, h. It is also easy to show that the same limit for $f_\Lambda(\beta, h)$ is obtained for an arbitrary sequence of hypercubes with edges increasing to infinity or, more generally, for domains Λ of arbitrary shape, growing to infinity in the sense of van Hove [2].

The finite range of the interaction is not really essential in the above argument, and actually it is enough to require that the potential is translation invariant and summable [2], *i.e.*,

$$\sum_{j \in \mathbb{Z}^d} |J_{ij}| < +\infty. \tag{5.7}$$

5.2.2 Random systems

The main ingredients of the above proof are the short range character of the potential, which allows to neglect, in the thermodynamic limit, the surface interaction among the sub-systems with respect to the bulk energy, and translation invariance. Both properties are lost in the case of mean field spin glass models. However, for finite-dimensional spin glasses, where the interaction has short range and the disorder *distribution* is translation-invariant, the proof essentially still works. The idea is again to divide the system into large blocks, which interact weakly owing to the short range character of the potential. The free energies of the different blocks can be therefore approximately considered as independent identically distributed random variables and the existence of the large N limit of the free energy per site follows from the strong law of large numbers [13]. This is essentially the method followed, for instance, in [59]. See also [60] for a very interesting generalization to non-summable random potentials.

For pedagogical reasons, we follow here a somewhat different strategy, which allows us to introduce some of the techniques we will employ in the following. For simplicity, we outline the proof of the existence of the thermodynamic limit in the case of the nearest-neighbor Gaussian Edwards-Anderson model [26] (see also Chapter 2 of the present work), which models a finite dimensional spin glass, and which is defined on $\mathbb{Z}^d$

by the Hamiltonian

$$H_\Lambda^{E.A.}(\sigma; J) = -\sum_{||i-j||=1} J_{ij}\sigma_i\sigma_j - h\sum_{i=1}^{|\Lambda|}\sigma_i.$$

The J_{ij} are independent identically distributed standard Gaussian variables and, as in (5.1), the sum runs over all couples of neighboring spins contained in the finite domain Λ. As regards the quenched free energy, the proof of the existence of the limit is analogous to the one we presented for non-random systems, thanks to the translation invariance of the disorder distribution: Fix a disorder realization J and divide, as before, the hypercube of side $L_k = 2^k$ into 2^d hypercubes of side 2^{k-1}. In analogy with (5.5), the estimate (5.2) implies

$$\left| f_k(\beta, h; J) - 2^{-d}\sum_{j=1}^{2^d} f_{k-1}^{(j)}(\beta, h; J)\right| \leq \frac{||H'||}{2^{d\,k}}. \tag{5.8}$$

Notice that, since the disorder realization is fixed, the system is not translation invariant, so that the free energies of the various sub-systems do not coincide. However, this problem disappears when one is interested in the disorder average of the free energy, since

$$E\, f_{k-1}^{(i)}(\beta, h; J) = f_{k-1}(\beta, h) \quad \forall i = 1, \dots, 2^d. \tag{5.9}$$

Therefore, from (5.8) one has

$$|f_k(\beta, h) - f_{k-1}(\beta, h)| \leq 2^{-d\,k} E\,||H'|| \leq d\,2^{-k} E|J_{ij}| = d\sqrt{\frac{2}{\pi}}\,2^{-k}, \tag{5.10}$$

and the proof goes on exactly like after Eq. (5.5), with the result that the limit

$$-\lim_{k\to\infty}\frac{1}{|\Lambda_k|\beta} E\,\ln Z_{\Lambda_k}(\beta, h; J) = f(\beta, h) \tag{5.11}$$

exists, for any value of β and for any sequence $\{\Lambda_k\}$ of increasing hypercubes. From Eq. (5.11), one can deduce almost sure convergence, *i.e.*,

$$-\lim_{k\to\infty}\frac{1}{|\Lambda_k|\beta}\ln Z_{\Lambda_k}(\beta, h; J) = f(\beta, h) \qquad J - a.s.\ (almost\ surely), \tag{5.12}$$

provided that one knows that the fluctuations of the disorder dependent free energy

$$f_\Lambda(\beta; J) = -\frac{1}{|\Lambda|\beta}\ln Z_\Lambda(\beta; J)$$

vanish sufficiently fast for $|\Lambda| \to \infty$ (for an alternative approach, based on the ergodic theorem, see [60]). To this purpose, we anticipate a result from next section, and we state the following

Proposition 1 *For any value of β, h and for any finite subset $\Lambda \subset \mathbb{Z}^d$, the probability of the fluctuations of the free energy per site can be estimated as follows:*

$$P\left(\left|\frac{1}{|\Lambda|\beta}\ln Z_\Lambda(\beta, h; J) - \frac{1}{|\Lambda|\beta}E\,\ln Z_\Lambda(\beta, h; J)\right| \geq u\right) \leq 2\exp\left(-\frac{|\Lambda|u^2}{2d}\right). \tag{5.13}$$

Results of this kind, known as *concentration of measure inequalities*, have been widely developed and employed by probabilists, in the more general context of Gaussian processes. We refer to the beautiful paper [61], for some early results and general motivations for concentration of measure inequalities, in the context of probability in Banach spaces. For later generalizations, we refer to [62], [63], [14], [15]. Concentration of measure techniques have been widely and successfully employed, in the framework of mean field spin glass theory, especially by M. Talagrand, as witnessed for instance in [9] and in the forthcoming book [11].

The above Proposition can be obtained as a particular case of Theorem 2 of next section. Now, we want to show how Eqs. (5.13) and (5.11), together, imply the almost sure convergence (5.12) to a non-random limit. To this purpose, we recall a very important result from probability theory, namely, Borel-Cantelli lemma [13], which states the following:

Lemma 1 (Borel-Cantelli) *Let* (Ω, Σ, μ) *be a probability space with* $\sigma-$*algebra* Σ *and probability measure* μ*, and consider a sequence of events* $A_n \in \Omega$*,* $n = 1, 2, \ldots$*, such that*

$$\sum_{n=1}^{\infty} \mu(A_n) < \infty. \tag{5.14}$$

Then,

$$\mu(\omega \in \Omega : \omega \in A_n \, for \;\; infinitely \;\; many \;\; indices \, n) = 0. \tag{5.15}$$

Going back to the problem of proving (5.12), for any arbitrary $\varepsilon >$ let k_0 be such that

$$|f_{\Lambda_k}(\beta, h) - f(\beta, h)| \leq \frac{\varepsilon}{2}$$

for $k > k_0$, where f_{Λ_k} is the finite volume average free energy corresponding to the hypercube Λ_k, so that

$$\begin{aligned} & P\left(\left|-\frac{1}{|\Lambda_k|\beta} \ln Z_{\Lambda_k}(\beta, h; J) - f(\beta, h)\right| \geq \varepsilon\right) \\ & \leq P\left(\left|-\frac{1}{|\Lambda_k|\beta} \ln Z_{\Lambda_k}(\beta, h; J) - f_{\Lambda_k}(\beta, h)\right| \geq \varepsilon/2\right) \leq 2 \exp\left(-\frac{|\Lambda_k|\varepsilon^2}{8d}\right). \end{aligned} \tag{5.16}$$

Since the right hand side is summable in k, Borel-Cantelli lemma implies that, J-almost surely,

$$\left|-\frac{1}{|\Lambda_k|\beta} \ln Z_{\Lambda_k}(\beta, h; J) - f(\beta, h)\right| \leq \varepsilon,$$

definitively in k. Now, consider a *countable* sequence $\{\varepsilon_n\}$, such that $\varepsilon_n \downarrow 0$ for $n \to \infty$. Since the intersection of a countable sequence of sets of measure one has still measure one [13], one deduces, for almost every J, the pointwise convergence (5.12). □

5.3 Self-averaging of the free energy

At this point, we abandon short range models and revert to mean field spin glasses. Even though the problem of proving the existence of the thermodynamic limit of the free energy remained open for a long time, it was nevertheless soon noticed that its disorder fluctuations vanish when the system size grows to infinity. Of course, this is not enough to show convergence, since the mean value could oscillate as N grows. When a physical quantity, such as the free energy per site, does not fluctuate in the limit of N large, it is said to be *self-averaging*. In ordinary statistical mechanics, one expects intensive quantities, such as the magnetization or the energy per site, to be self-averaging, with respect to thermal fluctuations. This holds for a generic value of the thermodynamic parameters, *i.e.*, away from phase transition points. In spin glass systems the situation is somewhat different [6] and one expects some quantities (like magnetization and internal energy) to be self-averaging, and others, in particular the overlap between the configurations of two replicas, to fluctuate even in the thermodynamic limit, at low temperature. This latter phenomenon is an indication of the occurrence of Replica Symmetry Breaking, as we explained in Chapter 4.

Self-averaging of the free energy for the Sherrington-Kirkpatrick model was first proved by Pastur and Scherbina [64] by using martingale techniques. Their result is

$$E\left(\frac{1}{N}\ln Z_N(\beta;J)\right)^2-\left(\frac{1}{N}E\,\ln Z_N(\beta;J)\right)^2\leq\frac{C}{N}+O(1/N^2), \tag{5.17}$$

for some constant C. Later, Guerra [44] made this estimate more precise, by showing that

$$C\leq\beta^2\frac{\langle q_{12}^2\rangle}{2}. \tag{5.18}$$

The self-averaging property of the free energy can also be proved in a different context, *i.e.*, by using exponential inequalities and concentration of measure arguments, as discussed briefly in the previous section. This method allows to show that fluctuations of the free energy from its disorder average are suppressed exponentially in the system size. For later convenience, we give a somewhat more general result, and we prove it by a method very similar to the one followed by Talagrand in [9].

Theorem 2 *Consider the family of Gaussian random variables $H_N(\sigma;J)$, with $\sigma\in\{-1,+1\}^N$, characterized by the mean value*

$$b_N(\sigma)=E\,H_N(\sigma;J) \tag{5.19}$$

and covariance matrix

$$c_N(\sigma,\sigma')=E\left(H_N(\sigma;J)H_N(\sigma';J)\right)-E\,H_N(\sigma;J)E\,H_N(\sigma';J), \tag{5.20}$$

and suppose that

$$|c_N(\sigma,\sigma)|\leq N\,L\quad\forall\sigma,N. \tag{5.21}$$

Then, for any β and any non-random set A_N in the configuration space $\{-1,+1\}^N$, one has

$$P\left(\left|-\frac{1}{N\beta}\ln Z_N^A(\beta;J)+\frac{1}{N\beta}E\ln Z_N^A(\beta;J)\right|\geq u\right)\leq 2\exp\left(-\frac{Nu^2}{2L}\right), \tag{5.22}$$

where

$$Z_N^A(\beta;J)=\sum_{\sigma\in A_N}\exp(-\beta H_N(\sigma;J)) \tag{5.23}$$

is the modified disorder-dependent partition function, with the sum over configurations restricted to A_N.

Of course, we will be often interested in the case where A_N coincides with the whole configuration space $\{-1,+1\}^N$.

Proof) We rewrite the Gaussian variables $H_N(\sigma;J)$ as

$$H_N(\sigma;J)=\xi_N(\sigma)+b_N(\sigma), \tag{5.24}$$

where, of course, $\xi_N(\sigma)$ is a centered Gaussian random variable, and

$$E(\xi_N(\sigma)\xi_N(\sigma'))=c_N(\sigma,\sigma').$$

Given $s\in\mathbb{R}$, we define

$$\varphi_N(t)=\ln E_1G_N(t)=\ln E_1\exp\left(s\beta^{-1}E_2\ln Z_N^A(t)\right), \tag{5.25}$$

where the interpolating parameter t varies between 0 and 1, and $Z_N^A(t)$ is the auxiliary partition function

$$\begin{aligned} Z_N^A(t) &= Z_N^A(t;J_1,J_2,\beta) \\ &= \sum_{\sigma\in A_N}\exp(-\beta\sqrt{t}\xi_N^1(\sigma)-\beta\sqrt{1-t}\xi_N^2(\sigma)-\beta b_N(\sigma)). \end{aligned} \tag{5.26}$$

Here, $\xi_N^1(\sigma)$ and $\xi_N^2(\sigma)$ are two *independent* copies of the random variable $\xi_N(\sigma)$, with the same distribution, and E_1,E_2 denote average with respect to ξ^1 and ξ^2, respectively. For simplicity of notation, we let

$$p_N(t,\sigma)=\frac{\exp(-\beta\sqrt{t}\xi_N^1(\sigma)-\beta\sqrt{1-t}\xi_N^2(\sigma)-\beta b_N(\sigma))}{Z_N^A(t)}$$

denote the modified Boltzmann weight.

It is very simple to check that

$$\varphi_N(1)-\varphi_N(0)=\ln E\exp\, s\beta^{-1}\left(\ln Z_N^A(\beta)-E\ln Z_N^A(\beta)\right). \tag{5.27}$$

Next, we compute the t derivative of $\varphi_N(t)$, and we find

$$\varphi_N'(t)=-\frac{s}{2E_1G_N(t)}E_1\left\{G_N(t)E_2\frac{\sum_{\sigma\in A_N}(\frac{1}{\sqrt{t}}\xi_N^1(\sigma)-\frac{1}{\sqrt{1-t}}\xi_N^2(\sigma))p_N(t,\sigma)}{Z_N^A(t)}\right\}.$$

An application of the integration by parts formula

$$Ex_iF(\{x\}) = \sum_j E(x_ix_j)E\partial_{x_j}F(\{x\}) \tag{5.28}$$

which holds for any family of Gaussian random variables $\{x_i\}$ and any smooth function F, gives

$$\varphi'_N(t) \;=\; \frac{s^2}{2E_1G_N(t)}E_1\left\{G_N(t)\sum_{\sigma,\sigma'\in A_N} c_N(\sigma,\sigma')E_2p_N(t,\sigma)E_2p_N(t,\sigma')\right\}.$$

Thanks to the bound (5.21), one has

$$|\varphi'_N(t)| \le \frac{s^2}{2}\max_{\sigma,\sigma'}|c_N(\sigma,\sigma')| \le \frac{Ns^2L}{2}. \tag{5.29}$$

Here, we have used the fact that, thanks to Cauchy-Schwarz inequality,

$$|c_N(\sigma,\sigma')|^2 \le c_N(\sigma,\sigma)\,c_N(\sigma',\sigma') \le N^2L^2.$$

Therefore, using Eq. (5.27) and the obvious inequality

$$e^{|x|} \le e^x + e^{-x},$$

one finds

$$E\exp\left(N|s|\left|\frac{1}{N\beta}\ln Z_N^A(\beta;J) - \frac{1}{N\beta}E\ln Z_N^A(\beta;J)\right|\right) \le 2\exp\left(\frac{s^2NL}{2}\right). \tag{5.30}$$

By Tchebyshev's inequality,

$$P\left(\left|\frac{1}{N\beta}\ln Z_N^A(\beta;J) - \frac{1}{N\beta}E\ln Z_N^A(\beta;J)\right| \ge u\right) \le 2\exp\left(-N|s|u + \frac{s^2NL}{2}\right) \tag{5.31}$$

and, choosing the optimal value $|s| = u/L$, one finally obtains the estimate (5.22). □

Theorem 2 is of very wide applicability for Gaussian models, since it relies only on the very natural hypothesis (5.21) that $c_N(\sigma,\sigma)/N$ is bounded. For instance, for the Sherrington-Kirkpatrick model defined by (3.1) one finds at once that

$$c_N(\sigma,\sigma') = N\frac{q_{\sigma\sigma'}^2}{2},$$

so that

$$L = \frac{1}{2}.$$

To conclude this section, we wish to warn the reader that the estimate of Theorem 2, though usually very useful, is not always optimal. Indeed, it states that free energy fluctuations are at most of order $1/\sqrt{N}$, irrespective of whether the system is at high or low temperature. On the other hand, in the case of zero external field and $\beta < 1$ Eq. (5.18), together with the fact that $\langle q_{12}^2\rangle = O(1/N)$ (see [6], and also Section 7.2 of the present work), shows that the free energy fluctuations are actually of order $1/N$. This was already noticed in [65] and later in [66], where the authors employed techniques of stochastic calculus, quite different from those we use here.

5.4 The interpolation method and the existence of the limit for the Sherrington-Kirkpatrick model

With this section, we begin to report the results obtained, in collaboration with F. Guerra, about the existence of the thermodynamic limit for mean field spin glass models. The results are partly contained in the two papers [22], [58], and partly are unpublished work, which will appear on a forthcoming paper [67]. For clarity, we proceed step by step, introducing our strategy first for the Sherrington-Kirkpatrick model, which is technically the simplest one, and then extending it to a wider class of mean field models.

5.4.1 A preliminary exercise: the Curie-Weiss model

For pedagogical reasons, we find it instructive to start with the exactly solvable and non-random Curie-Weiss model. This allows us to present, in a very simple context, the method of interpolation between Hamiltonians, which we introduced in [22]. Let us recall the Curie-Weiss Hamiltonian, which is given by

$$H_N^{C.W.}(\sigma, h) = -\frac{J}{N}\sum_{i,j=1}^{N} \sigma_i\sigma_j - h\sum_{i=1}^{N}\sigma_i. \tag{5.32}$$

Here, the coupling strength J is a positive constant, and h is as usual the magnetic field. For a given inverse temperature β, and defining the magnetization per spin, corresponding to the configuration σ, as

$$m(\sigma) = \frac{1}{N}\sum_{i=1}^{N}\sigma_i, \tag{5.33}$$

the partition function can be rewritten as

$$Z_N(\beta, h) = \sum_{\{\sigma\}} \exp N\beta(Jm^2(\sigma) + h\, m(\sigma)). \tag{5.34}$$

Now divide the N spin system into two subsystems of N_1 and N_2 spins each, with $N_1 + N_2 = N$. Denoting by $m_1(\sigma)$, $m_2(\sigma)$ the magnetization corresponding to the subsystems, *i.e.*,

$$m_1(\sigma) = \frac{1}{N_1}\sum_{i=1}^{N_1}\sigma_i \tag{5.35}$$

$$m_2(\sigma) = \frac{1}{N_2}\sum_{i=N_1+1}^{N}\sigma_i, \tag{5.36}$$

one sees that $m(\sigma)$ is a convex linear combination* of $m_1(\sigma)$ and $m_2(\sigma)$:

$$m(\sigma) = \frac{N_1}{N}m_1(\sigma) + \frac{N_2}{N}m_2(\sigma). \tag{5.37}$$

*recall that y is a convex combination of x_1 and x_2 if

$$y = \theta\, x_1 + (1-\theta)\, x_2,$$

Since the function $x \to x^2$ is convex, one has

$$\begin{aligned} Z_N(\beta,h) &\le \sum_{\{\sigma\}} \exp(\beta J(N_1 m_1^2(\sigma) + N_2 m_2^2(\sigma)) + \beta h(N_1 m_1(\sigma) + N_2 m_2(\sigma))) \\ &= Z_{N_1}(\beta,h) Z_{N_2}(\beta,h) \end{aligned} \tag{5.38}$$

and

$$N f_N(\beta,h) = -\frac{1}{\beta} \ln Z_N(\beta,h) \ge N_1 f_{N_1}(\beta,h) + N_2 f_{N_2}(\beta,h). \tag{5.39}$$

In other words, the free energy is superadditive in the system size, and the existence of the limit follows from standard methods [2]. For a different but equally simple method to prove the existence of the limit for the Curie-Weiss model, see Section 2 of Ref. [68].

Unfortunately, the very simple approach we illustrated above cannot be applied to the Sherrington-Kirkpatrick model, where the randomness of the couplings prevents us from exploiting subadditivity directly on the Hamiltonian H_N. However, there exists an alternative and related strategy, which allows in some sense an extension to mean field spin glass models. The main idea is to interpolate between the original systems of N spins, and two non-interacting systems, containing N_1 and N_2 spins, respectively, and to compare the corresponding free energies. To this purpose, consider the interpolating parameter $0 \le t \le 1$, and the auxiliary partition function

$$Z_N(t) = \sum_{\{\sigma\}} \exp \beta \left(N t J m^2(\sigma) + N_1 (1-t) J m_1^2(\sigma) + N_2 (1-t) J m_2^2(\sigma) + N h\, m(\sigma) \right). \tag{5.40}$$

Of course, for the boundary values $t = 0, 1$ one has

$$-\frac{1}{N\beta} \ln Z_N(1) = f_N(\beta,h) \tag{5.41}$$

$$-\frac{1}{N\beta} \ln Z_N(0) = \frac{N_1}{N} f_{N_1}(\beta,h) + \frac{N_2}{N} f_{N_2}(\beta,h) \tag{5.42}$$

and, taking the derivative with respect to t,

$$-\frac{d}{dt}\frac{1}{N\beta} \ln Z_N(t) = -J\,\Omega \left(m^2(\sigma) - \frac{N_1}{N} m_1^2(\sigma) - \frac{N_2}{N} m_2^2(\sigma) \right) \ge 0, \tag{5.43}$$

where $\Omega(.)$ denotes as usual the Boltzmann-Gibbs thermal average. Therefore, integrating in t between 0 and 1, and recalling the boundary conditions (5.41), one finds again the superadditivity property (5.39).

The interpolation method, which may look unnecessarily complicated for the Curie-Weiss model, is actually the only one working in the case of mean field spin glass systems.

---

for some $0 \le \theta \le 1$. By definition, for any convex function $x \to f(x)$, one has

$$f(\theta\, x_1 + (1-\theta)\, x_2) \le \theta\, f(x_1) + (1-\theta)\, f(x_2).$$

5.4.2 The Sherrington-Kirkpatrick model

Let us explain the main idea behind our method. As for the Curie-Weiss model, we divide the N sites into two blocks N_1, N_2 with $N_1 + N_2 = N$, and define

$$\begin{aligned} Z_N(t) &= \sum_{\{\sigma\}} \exp\left(\beta\sqrt{\frac{t}{N}} \sum_{1\le i<j\le N} J_{ij}\sigma_i\sigma_j + \beta\sqrt{\frac{1-t}{N_1}} \sum_{1\le i<j\le N_1} J'_{ij}\sigma_i\sigma_j \right. \\ &\left. +\beta\sqrt{\frac{1-t}{N_2}} \sum_{N_1<i<j\le N} J''_{ij}\sigma_i\sigma_j \right) \exp \beta h \sum_{i=1}^{N} \sigma_i, \end{aligned} \tag{5.44}$$

with $0 \le t \le 1$. The external disorder is represented by the *independent* families of unit Gaussian random variables J, J' and J''. Notice that the two subsystems are subject to a different external disorder, with respect to the original system but, of course, the probability distributions are the same. As in the case of the Curie-Weiss model to interpolate between the original N spin system at $t = 1$ and a system composed of two non-interacting parts at $t = 0$, so that

$$Z_N(1) = Z_N(\beta, h; J) \tag{5.45}$$

$$Z_N(0) = Z_{N_1}(\beta, h; J')Z_{N_2}(\beta, h; J''). \tag{5.46}$$

As a consequence, we have

$$E \ln Z_N(1) = -N\beta f_N(\beta, h) \tag{5.47}$$

$$E \ln Z_N(0) = -N_1\beta f_{N_1}(\beta, h) - N_2\beta f_{N_2}(\beta, h). \tag{5.48}$$

By taking the derivative of $-1/(N\beta)E \ln Z_N(t)$ with respect to the parameter t, we obtain

$$\begin{aligned} -\frac{d}{dt}\frac{1}{N\beta}E\ \ln Z_N(t) = -\frac{1}{2N}E\left(\frac{1}{\sqrt{tN}} \sum_{1\le i<j\le N} J_{ij}\omega_t(\sigma_i\sigma_j)\right. \\ \left. -\frac{1}{\sqrt{(1-t)N_1}} \sum_{1\le i<j\le N_1} J'_{ij}\omega_t(\sigma_i\sigma_j) - \frac{1}{\sqrt{(1-t)N_2}} \sum_{N_1<i<j\le N} J''_{ij}\omega_t(\sigma_i\sigma_j)\right) \end{aligned} \tag{5.49}$$

where $\omega_t(.)$ denotes the Gibbs state corresponding to the partition function (5.44). A standard integration by parts on the Gaussian disorder, as done for example in [6], [44], gives

$$\begin{aligned} -\frac{d}{dt}\frac{1}{N\beta}E\ \ln Z_N(t) &= -\frac{\beta}{4N^2}\sum_{i,j=1}^{N} E(1-\omega_t^2(\sigma_i\sigma_j)) + \frac{\beta}{4NN_1}\sum_{i,j=1}^{N_1} E(1-\omega_t^2(\sigma_i\sigma_j)) \\ &+\frac{\beta}{4NN_2}\sum_{i,j=N_1+1}^{N} E(1-\omega_t^2(\sigma_i\sigma_j)) \qquad (5.50) \\ &= \frac{\beta}{4}\left\langle q_{12}^2 - \frac{N_1}{N}(q_{12}^{(1)})^2 - \frac{N_2}{N}(q_{12}^{(2)})^2 \right\rangle, \qquad (5.51) \end{aligned}$$

where we have defined

$$N_1\, q_{12}^{(1)} = \sum_{i=1}^{N_1} \sigma_i^1 \sigma_i^2 \tag{5.52}$$

$$N_2\, q_{12}^{(2)} = \sum_{i=N_1+1}^{N} \sigma_i^1 \sigma_i^2 \tag{5.53}$$

as the partial two-replica overlaps, corresponding to the first (respectively, second) subsystem. Since q_{12} is a convex linear combination of $q_{12}^{(1)}$ and $q_{12}^{(2)}$ in the form

$$q_{12} = \frac{N_1}{N}\, q_{12}^{(1)} + \frac{N_2}{N}\, q_{12}^{(2)},$$

due to convexity of the function $x \to x^2$, we have the inequality

$$\left\langle q_{12}^2 - \frac{N_1}{N}(q_{12}^{(1)})^2 - \frac{N_2}{N}(q_{12}^{(2)})^2 \right\rangle \leq 0.$$

Therefore, we can state our first preliminary result.

Lemma 2 *The quenched average of the logarithm of the interpolating partition function, defined by (5.44), is increasing in t, i.e.*

$$-\frac{d}{dt}\frac{1}{N\beta} E \,\ln Z_N(t) \leq 0. \tag{5.54}$$

By integrating in t and recalling the boundary conditions (5.47), we get the first main result.

Theorem 3 *The following subadditivity property holds*

$$N\, f_N(\beta, h) \leq N_1\, f_{N_1}(\beta, h) + N_2\, f_{N_2}(\beta, h). \tag{5.55}$$

Of course, due the minus sign in (3.3), we have superadditivity for the auxiliary function $\alpha_N(\beta, h)$.

It is interesting to compare this result with the corresponding Eq. (5.39), which holds for the Curie-Weiss model. In that case, free energy is *superadditive* rather than *subadditive*.

The subadditivity property gives an immediate control on the infinite volume limit, as explained for example in [2]. In fact, we have

Theorem 4 *The infinite volume limit for $f_N(\beta, h)$ does exists and equals its* inf*:*

$$\lim_{N\to\infty} f_N(\beta, h) = \inf_N f_N(\beta, h) = f(\beta, h). \tag{5.56}$$

5.4.3 Thermodynamic limit for the ground state energy

For finite N and a given realization J of the disorder, define the ground state energy density $-e_N(J,h)$ as

$$-e_N(h;J) = \frac{1}{N} \inf_{\sigma} H_N(\sigma,h;J). \tag{5.57}$$

The minus sign in (5.57) simply guarantees that $e_N(h;J)$ is positive. Numerically, one finds that $e_N(h;J)$ approaches a well defined limit when N grows, and in particular $e_N = 0.763...$ when the magnetic field is zero [6]. Numerical computation of the ground state energy is quite hard, since any search algorithm which tries to minimize the energy by changing one or a few spins a time, tends to get stuck very soon in metastable states, which are stable under a single spin-flip, but have an energy higher than that of the ground state.

Here we show, using simple thermodynamic properties, that Eq. (5.56) of Theorem 4 implies the existence of the thermodynamic limit for $E\,e_N(h;J)$. First of all, notice that the bounds

$$e^{\beta N e_N(h;J)} \le \sum_{\{\sigma\}} e^{-\beta H_N(\sigma,h;J)} \le 2^N e^{\beta N e_N(h;J)} \tag{5.58}$$

hold for any J, N, β, h, so that

$$0 \le \frac{\ln Z_N(\beta,h;J)}{\beta N} - e_N(h;J) \le \frac{\ln 2}{\beta}. \tag{5.59}$$

The bounds (5.59), together with the obvious

$$\partial_\beta \frac{\ln Z_N(\beta,h;J)}{\beta} \le 0, \tag{5.60}$$

which is equivalent to positivity of the entropy†, imply that

$$\lim_{\beta\to\infty} \frac{\ln Z_N(\beta,h;J)}{\beta N} \downarrow e_N(h;J). \tag{5.62}$$

Of course, by taking the expectation value in (5.59) and defining

$$e_N(h) = E\,e_N(h;J),$$

one also finds

$$\lim_{\beta\to\infty} f_N(\beta,h) \uparrow -e_N(h) \tag{5.63}$$

Therefore, by taking into account the subadditivity (5.55), the inequalities (5.59), and the existence of the limit $f(\beta,h)$ for $f_N(\beta,h)$, we have from (5.63) the proof of the following

†In fact, it is easy to check that

$$\partial_\beta \frac{\ln Z_N(\beta,h;J)}{\beta} = \frac{1}{\beta^2} \sum_{\{\sigma\}} \rho(\sigma) \ln \rho(\sigma) \le 0, \tag{5.61}$$

where

$$0 \le \rho(\sigma) = Z_N^{-1} \exp(-\beta H_N(\sigma,h;J)) \le 1.$$

Theorem 5 *For the quenched average of the ground state energy we have the superadditivity property*

$$N\, e_N(h) \geq N_1\, e_{N_1}(h) + N_2\, e_{N_2}(h). \tag{5.64}$$

and the existence of the infinite volume limit

$$\lim_{N\to\infty} e_N(h) = \sup_N e_N(h) \equiv e_0(h). \tag{5.65}$$

Finally, we can write the limit $e_0(h)$ *in terms of* $f(\beta, h)$ *as*

$$\lim_{\beta\to\infty} f(\beta, h) \uparrow -e_0(h). \tag{5.66}$$

5.4.4 Almost sure convergence

After proving the existence of the thermodynamic limit for the quenched averages, we can easily extend our results to prove that convergence holds for almost every disorder realization J, as we did for the short range models of Section 5.2.2. In fact, we can state

Theorem 6 *The infinite volume limits*

$$-\lim_{N\to\infty} \frac{1}{N\beta} \ln Z_N(\beta, h; J) = f(\beta, h), \tag{5.67}$$

$$\lim_{N\to\infty} e_N(h; J) = e_0(h), \tag{5.68}$$

exist J*-almost surely.*

For the proof of (5.67), one proceeds exactly as for the random short range models, employing the asymptotic vanishing of free energy fluctuations (5.22), with $L = 1/2$, and Borel-Cantelli lemma. As regards the ground state energy one notices that, thanks to Eqs. (5.62) and (5.63), letting $\beta \to \infty$ one has

$$\begin{aligned} P\left(|e_N(h; J) - e_N(h)| \geq u\right) &= \lim_{\beta\to\infty} P\left(\left|-\frac{1}{N\beta} \ln Z_N(\beta, h; J) - f_N(\beta, h)\right| \geq u\right) \\ &\leq 2\exp(-Nu^2). \end{aligned} \tag{5.69}$$

Again, Borel-Cantelli lemma implies (5.68), and the theorem is proven. □

It is easy to realize that all the results of this section hold also in the case where on each spin σ_i acts a random magnetic field h_i, where the h_i's are independent identically distributed random variables. In fact, the one-body (random) interaction

$$-\sum_{i=1}^{N} h_i \sigma_i = -\sum_{i=1}^{N_1} h_i \sigma_i - \sum_{i=N_1+1}^{N} h_i \sigma_i$$

produces no additional interaction between the two sub-systems of size N_1 and N_2.

5.5 Generalization to other mean field spin glass models

In this section, we show how the above results on the (almost sure) existence of the thermodynamic limit for the free energy and for the ground state energy can be extended to other mean field spin glass models. In the first place, we can immediately extend the above approach to the Derrida p-spin model introduced in Section 3.3, for even p. Secondly, we can allow the couplings to be non-Gaussian random variables, provided that suitable bounds are imposed on their moments. Finally, we consider more general models, where the spin degrees of freedom are not necessarily two-valued Ising variables. The contents of this Section are based on results obtained in collaboration with F. Guerra in [22], [58].

5.5.1 The p-spin model

We turn first of all to the Derrida p-spin model introduced in Section 3.3, and we consider only the case of even p. As we did for the Sherrington-Kirkpatrick model, we define the auxiliary partition function $Z_N(t)$, in analogy with (5.44), by just replacing the two-body with a p-body interaction. By taking the t derivative, we find after integration by parts

$$\begin{aligned}-\frac{d}{dt}\frac{1}{N\beta}E\ln Z_N(t) &= \frac{\beta}{4}\left\langle q_{12}^p - \frac{N_1}{N}(q_{12}^{(1)})^p - \frac{N_2}{N}(q_{12}^{(2)})^p\right\rangle + O(1/N)\\ &\leq O(1/N),\end{aligned} \tag{5.70}$$

for p even, by the same convexity argument as before, since the function $q \to q^p$ is convex. It is easy to realize the reason for the appearance of the terms $O(1/N)$. In fact, for $p = 2$, we can write

$$\frac{2}{N^2}\sum_{1\leq i<j\leq N} E\left(1-\omega_t^2(\sigma_i\sigma_j)\right) = \frac{1}{N^2}\sum_{i,j=1}^{N} E\left(1-\omega_t^2(\sigma_i\sigma_j)\right) = (1-\langle q_{12}^2\rangle), \tag{5.71}$$

as already exploited in (5.50). On the other hand, for $p > 2$ one has

$$\begin{aligned}&\frac{p!}{N^p}\sum_{1\leq i_1<\ldots<i_p\leq N} E\left(1-\omega_t^2(\sigma_{i_1}\ldots\sigma_{i_p})\right) \qquad (5.72)\\ &= \frac{1}{N^p}\sum_{i_1,\ldots,i_p=1}^{N} E\left(1-\omega_t^2(\sigma_{i_1}\ldots\sigma_{i_p})\right) + O(1/N) = (1-\langle q_{12}^p\rangle) + O(1/N).\end{aligned}$$

From (5.70) one finds, as in the previous section, the existence of the infinite volume limits[‡]

$$\lim_{N\to\infty} f_N^{(p)}(\beta,h) = f^{(p)}(\beta,h) \tag{5.75}$$

$$\lim_{N\to\infty} e_N^{(p)}(h) \equiv e_0^{(p)}(h) = -\lim_{\beta\to\infty} f^{(p)}(\beta,h). \tag{5.76}$$

The proof of almost sure convergence, both for the free energy and for the ground state energy per site, presents no additional difficulty with respect to the case of the Sherrington-Kirkpatrick model. In fact, for the p-spin model the covariance matrix defined in (5.20) equals

$$c_N(\sigma,\sigma') = N\frac{q_{\sigma\sigma'}^p}{2} + o(1),$$

and Theorem 2 holds with $L = 1/2$.

Unfortunately, our proof does not extend directly to the case of odd p, since in this case the function $q \to q^p$ is not convex, for negative values of q. However one expects that the overlap between two configuration is non-negative with probability approaching one, when N tends to infinity. This is in contrast with the case of p even, where the system is invariant under global spin-flip symmetry, and the distribution of the overlap is symmetric. It is interesting that M. Talagrand proved [48] that, for odd p large enough, the probability distribution of the overlap has support in $[0,1]$, asymptotically for N large, and the probability that a couple of configurations σ^1, σ^2 has overlap $q_{12} < 0$ tends to zero. Moreover, in a very recent still unpublished work [11] he generalized the result to a generic odd value of p. For a more precise statement, see the forthcoming book [11]. As a consequence, modulo some minor technical modifications, the existence of the thermodynamic limit for the free energy and ground state energy of the p-spin model follows from (5.70), also for odd p.

‡it is easy to see that the term $O(1/N)$ in (5.70) does not spoil the argument leading to the existence of the thermodynamic limit. More generally, if

$$f_N \le \frac{N_1}{N} f_{N_1} + \frac{N_2}{N} f_{N_2} + C/N^\gamma, \tag{5.73}$$

with $C, \gamma > 0$, then the limit exists. Indeed, fix $x_{1,2} = N_{1,2}/N$, and rewrite C/N^γ as

$$\frac{C}{N^\gamma} = \frac{N_1}{N}\frac{a}{N_1^\gamma} + \frac{N_2}{N}\frac{a}{N_2^\gamma} - \frac{a}{N^\gamma},$$

where

$$a = C\left(x_1^{1-\gamma} + x_2^{1-\gamma} - 1\right)^{-1}.$$

In this way, Eq. (5.73) gives

$$f_N + a/N^\gamma \le \frac{N_1}{N}(f_{N_1} + a/N_1^\gamma) + \frac{N_2}{N}(f_{N_2} + a/N_2^\gamma). \tag{5.74}$$

The existence of the limit for $f_N + a/N^\gamma$, and therefore also for f_N, follows from the usual subadditivity argument.

5.5.2 Non-Gaussian couplings

In Section 3.4, a comparison was made between infinite and finite connectivity mean field spin glass models, and it was stated that they can be characterized, in probabilistic terms, as Gaussian and non-Gaussian processes on $\{-1,+1\}^N$, respectively. In the present section, on the other hand, we deal with infinite connectivity models (essentially, the Sherrington-Kirkpatrick and p-spin models), where the couplings between spins are *non-Gaussian* random variables. In this case, of course, the Hamiltonian $H_N(\sigma;J)$ is also non-Gaussian. This, however, does not contradict the above statement. In fact, for N large, $H_N(\sigma;J)$ is the sum of many (for the Sherrington-Kirkpatrick model, $N(N-1)/2$) independent identically distributed random variables, each having a small weight (of the order $1/\sqrt{N}$, for the Sherrington-Kirkpatrick model), so that $H_N/\sqrt{N}$ has a Gaussian distribution, for N going to infinity, as guaranteed by the central limit theorem[§] [13]. As we shall see, the precise probability distribution of the couplings is inessential for N large, provided that certain conditions of symmetry and finiteness of the moments are satisfied.

That the probability distribution of the couplings J_{ij} should not influence the thermodynamic limit of the model was already clear to the authors of [17], and was rigorously proved by Talagrand [69], in the particular case of two-valued quenched variables J_{ij}, which assume the values ± 1 with equal probability $1/2$. Of course, the probability distribution of the couplings J_{ij} is expected to have an effect on the disorder fluctuations of the various physical quantities, and on finite size corrections to the thermodynamic limit. Here, we extend Talagrand's result to more general conditions, thereby extending to these models the proof of the existence of the limit. While this presents no difficulty for the quenched free energy, the proof of almost sure convergence requires an extension of inequality (5.22), which expresses exponential self-averaging of the free energy and which, the way it is stated, holds only for Gaussian models.

Consider the p-spin model (which, for $p=2$, reduces to the Sherrington-Kirkpatrick model) defined by the Hamiltonian

$$\tilde{H}_N^{(p)}(\sigma,h;\eta) = -\sqrt{\frac{p!}{2N^{p-1}}} \sum_{1\le i_1<\ldots<i_p\le N} \eta_{i_1\ldots i_p}\sigma_{i_1}\ldots\sigma_{i_p} - h\sum_{i=1}^N \sigma_i. \tag{5.78}$$

and suppose that the couplings $\eta_{i_1\ldots i_p}$ are independent identically distributed random variables, with a symmetric distribution and finite fourth moment, *i.e.*,

$$P(\eta_{i_1\ldots i_p}) = P(-\eta_{i_1\ldots i_p}) \tag{5.79}$$

$$E\eta^4_{i_1\ldots i_p} < \infty. \tag{5.80}$$

[§] for instance, one can check directly that

$$\frac{EH_N^4(\sigma,h;J) - 3(EH_N^2(\sigma,h;J))^2}{N^2} = O(1/N), \tag{5.77}$$

where $H_N(\sigma,h;J)$ is the Sherrington-Kirkpatrick Hamiltonian, and the couplings J_{ij} are independent identically distributed symmetric random variables with finite moments. Similar relations hold for the higher order moments of $H_N/\sqrt{N}$, which therefore tend to the moments of a Gaussian variable with variance $EH_N^2(\sigma,h;J)/N = 1/2$.

Moreover, we assume the disorder variables to have unit variance,

$$E(\eta_{i_1\ldots i_p}^2) = 1, \tag{5.81}$$

as in the Gaussian case. This last assumption is just a matter of convention, since a change in the variance is equivalent to a change in the scale of temperatures. Conditions similar to (5.79), (5.80) have been also exploited for the study of the Sherrington-Kirkpatrick model, at zero external field and high temperature, for example in [65] and [33].

As we have shown in the previous sections, integration by parts on the Gaussian disorder (3.21) is a simple but essential tool. In the present case, Eq. (3.21) no longer holds, and its role is played instead by the formula

$$E\,\eta F(\eta) = E\,\eta^2 F'(\eta) - \frac{1}{4} E\,|\eta| \int_{-|\eta|}^{|\eta|} (\eta^2 - u^2) F'''(u)\,du, \tag{5.82}$$

which holds for any symmetric random variable η and for sufficiently regular functions F, as a simple direct calculation shows. A similar expression has been exploited in [69] by Talagrand, for two-valued random variables $\eta = \pm 1$.

Denote as $f_N^{(p)}(\beta, h)$ and $\tilde{f}_N^{(p)}(\beta, h)$ the finite volume quenched free energies of the Gaussian and non-Gaussian p-spin model, respectively. Then, the following result holds:

Theorem 7 *If conditions (5.79)-(5.81) are satisfied,*

$$|f_N^{(p)}(\beta, h) - \tilde{f}_N^{(p)}(\beta, h)| = O(N^{1-p}). \tag{5.83}$$

In particular, for p even this implies that the limit

$$\lim_{N\to\infty} \tilde{f}_N^{(p)}(\beta, h) = \lim_{N\to\infty} f^{(p)}(\beta, h) = f^{(p)}(\beta, h) \tag{5.84}$$

exists for any β, h. A similar statement holds for the ground state energy:

$$\lim_{N\to\infty} \tilde{e}_N^{(p)}(h) = \lim_{N\to\infty} e_N^{(p)}(h) = e_0^{(p)}(h) = -\lim_{\beta\to\infty} f^{(p)}(\beta, h). \tag{5.85}$$

Proof) Just for simplicity, we consider the case $h = 0$. It should be clear, by now, that in order to compare the Gaussian and the non-Gaussian free energies, we need to suitably interpolate between them. In this case, the interpolating auxiliary partition function is given by

$$Z_N(t) = \sum_{\{\sigma\}} \exp\left(-\beta\sqrt{t}\tilde{H}_N^{(p)}(\sigma;\eta) - \beta\sqrt{1-t}H_N^{(p)}(\sigma;J)\right), \tag{5.86}$$

such that

$$-\frac{1}{N\beta} E \ln Z_N(0) = f_N^{(p)}(\beta) \tag{5.87}$$

$$-\frac{1}{N\beta} E \ln Z_N(1) = \tilde{f}_N^{(p)}(\beta). \tag{5.88}$$

As regards the t derivative, one easily finds

$$
\begin{aligned}
-\frac{d}{dt}\frac{1}{N\beta}E\ln Z_N(t) &= -\frac{1}{2N}\sqrt{\frac{p!}{2N^{p-1}t}}\sum_{1\le i_1<\ldots<i_p\le N} E\,\eta_{i_1\ldots i_p}\omega_t(\sigma_{i_1}\ldots\sigma_{i_p}) \\
&+\frac{\beta p!}{4N^p}\sum_{1\le i_1<\ldots<i_p\le N}(1-E\,\omega_t^2(\sigma_{i_1}\ldots\sigma_{i_p})), \qquad (5.89)
\end{aligned}
$$

where we used Gaussian integration by parts on the variables $J_{i_1\ldots i_p}$ in the second term. Applying formula (5.82), one can see that the first term in the right hand side cancels the second, apart from terms of order $O(N^{1-p})$. More precisely, (5.82) gives

$$
E\,\eta_{i_1\ldots i_p}\omega_t(\sigma_{i_1}\ldots\sigma_{i_p}) = \beta\sqrt{\frac{p!t}{2N^{p-1}}}(1-E\,\eta^2_{i_1\ldots i_p}\omega_t^2(\sigma_{i_1}\ldots\sigma_{i_p}))+O(N^{-3(p-1)/2}), \quad (5.90)
$$

where the error term arises from the third derivative F''' in (5.82), since the dependence of the Boltzmann-Gibbs state on the single variable $\eta_{i_1\ldots i_p}$ is of order $N^{-(p-1)/2}$. One has also to make use of assumption (5.80). Moreover, employing the simple identity

$$
\begin{aligned}
E\,\eta^2_{i_1\ldots i_p}\omega_t^2(\sigma_{i_1}\ldots\sigma_{i_p}) &= E\,\omega_t^2(\sigma_{i_1}\ldots\sigma_{i_p}) \\
&+E\,(\eta^2_{i_1\ldots i_p}-1)\int_0^{\eta_{i_1\ldots i_p}} du\,\partial_u\,\omega_t^2(\sigma_{i_1}\ldots\sigma_{i_p})\Big|_{\eta_{i_1\ldots i_p}=u},
\end{aligned}
$$

one has

$$
E\,\eta^2_{i_1\ldots i_p}\omega_t^2(\sigma_{i_1}\ldots\sigma_{i_p}) = E\,\omega_t^2(\sigma_{i_1}\ldots\sigma_{i_p})+O(N^{-(p-1)}). \quad (5.91)
$$

Putting Eqs. (5.89), (5.90) and (5.91) together, one finally obtains (5.83). Statements (5.84) and (5.85) are obvious consequences of Eqs. (5.83) and (5.75)-(5.76), referring to the Gaussian p-spin model. □.

As we explained in Section 5.4.4, for models with Gaussian couplings almost sure convergence of the free energy and ground state energy per site follows from the convergence of the quenched average and from Borel-Cantelli lemma, provided that an estimate of the type (5.22) holds. The extension of (5.22) to the non-Gaussian case is not entirely trivial. For instance, Theorem 2 of [61] allows to extend it to the case where the probability distribution of the variables η_{ij} vanishes for $|\eta_{ij}|\to\infty$ at least as fast as in the Gaussian case (for instance, this includes the case of bounded random variables), but we are not aware of much more general results concerning concentration of measure inequalities for functions of many non-Gaussian random variables. Here, we prove the following

Theorem 8 *[67] If the quenched variables satisfy conditions (5.79), (5.80), then for the $p-$spin model free energy*

$$
P\left(\left|\frac{1}{\beta N}\ln Z_N(\beta,h;\eta)-\frac{1}{\beta N}E\,\ln Z_N(\beta,h;\eta)\right|\ge u\right)\le C(\beta)\,e^{-N\,C(\beta)u^2}, \quad (5.92)
$$

where $C(\beta)$ is some constant, which depends only on the temperature and on the disorder distribution.

Proof) We give just a brief sketch of the proof, since the strategy is the same we followed to prove Theorem 2, apart from some technical estimates, which have already been employed to prove Theorem 7.

We define, in analogy with (5.26), an auxiliary partition function $Z_N(t)$, depending on the interpolating parameter $0 \le t \le 1$ and on two independent families $\eta^{(1)}, \eta^{(2)}$ of independent identically distributed random variables, with the same distribution as the original disorder η:

$$Z_N(t) = \sum_{\{\sigma\}} \exp\left(\beta\sqrt{\frac{p!}{2N^{p-1}}}\sum(\sqrt{t}\,\eta^{(1)}_{i_1\dots i_p} + \sqrt{1-t}\,\eta^{(2)}_{i_1\dots i_p})\sigma_{i_1}\dots\sigma_{i_p} + \beta h\sum_{i=1}^{N}\sigma_i\right).$$

Exactly as in Eq. (5.25), one defines

$$\varphi_N(t) = \ln\, E_1\, \exp\frac{s}{\beta}\, E_2\, \ln Z_N(t),$$

such that

$$\varphi_N(1) - \varphi_N(0) = \ln\, E\, \exp\frac{s}{\beta}\left(\ln Z_N^{(p)}(\beta,h,\eta) - E\ln Z_N^{(p)}(\beta,h,\eta)\right). \tag{5.93}$$

Next, one computes as usual the t derivative of $\varphi_N(t)$ and, after some algebra, one finds

$$|\varphi'_N(t)| \le C(\beta)(Ns^2+1), \tag{5.94}$$

in analogy with (5.29). Here and in the following, $C(\beta)$ denotes some positive constant, which needs not be the same at each occurrence. The detailed proof of Eq. (5.94) is rather lengthy, but conceptually straightforward, since it relies only on formula (5.82), and on estimates of the kind (5.90), (5.91), which simply express the weak dependence of the Gibbs state from any individual coupling variable. For further details, see the forthcoming paper [67]. Notice that in the present case, with respect to (5.29), there is also dependence on the temperature in the bound (5.94). Recalling Eq. (5.93), we find

$$E\,\exp\left(N|s|\left|\frac{1}{N\beta}Z_N^{(p)}(\beta,h;\eta) - \frac{1}{N\beta}E\,Z_N^{(p)}(\beta,h;\eta)\right|\right) \le C(\beta)\,e^{NC(\beta)s^2}.$$

At this point, one proceeds like after Eq. (5.30) and the theorem is proved. □

In conclusion, as a simple consequence of Theorem 8, we have

Theorem 9 *For p even, under assumptions (5.79)-(5.81), there exist the limits*

$$-\lim_{N\to\infty}\frac{1}{N\beta}\ln Z_N^{(p)}(\beta,h;\eta) = f^{(p)}(\beta,h) \quad \eta - a.e., \tag{5.95}$$

and

$$\lim_{N\to\infty} e_N^{(p)}(h;\eta) = e_0^{(p)}(h) \quad \eta - a.e. \tag{5.96}$$

Proof) The first statement simply follows from Borel-Cantelli lemma and (5.92).

As for the ground state energy recall that, in the Gaussian case, the result simply follows from the fact that $e_N(h;J)$ is the zero-temperature limit of the disorder dependent free energy per site (apart from a minus sign), and from the lack of β-dependence of the right hand side of the exponential estimate (5.22). In fact, the limit $\beta \to \infty$ in the exponential inequality can be then taken without problems, as in (5.69). In the present case, due to the β dependence of the right hand side of (5.92), the proof is not so direct, and rather relies on thermodynamic considerations, based on convexity properties of the free energy. Indeed, the inequality (5.60), together with (5.62) implies, for any η, N, β, h,

$$e_N^{(p)}(h;\eta) \leq \frac{\ln Z_N^{(p)}(\beta,h;\eta)}{N\beta}$$

and, thanks to the almost sure existence of the limit for the free energy (5.95),

$$\limsup_{N\to\infty} e_N^{(p)}(h;\eta) \leq -f^{(p)}(\beta,h) \quad \eta - a.s.$$

Choosing a sequence $\{\beta^{(i)}\}$ with $\lim_{i\to\infty}\beta^{(i)} = \infty$ and recalling Eq. (5.85), we get

$$\limsup_{N\to\infty} e_N^{(p)}(h;\eta) \leq e_0^{(p)}(h) \quad \eta - a.s. \tag{5.97}$$

On the other side, since the function $\beta \to \ln Z_N(\beta)$ is convex in β, for any $\bar\beta > \beta$ we have

$$\frac{\ln Z_N^{(p)}(\bar\beta,h;\eta)}{\bar\beta} \geq \frac{\ln Z_N^{(p)}(\beta,h;\eta)}{\bar\beta} + \frac{(\bar\beta-\beta)}{\bar\beta}\partial_\beta \ln Z_N^{(p)}(\beta,h;\eta)$$

so that, letting $\bar\beta \to \infty$ with β fixed,

$$e_N^{(p)}(h;\eta) \geq \frac{\partial_\beta \ln Z_N^{(p)}(\beta,h;\eta)}{N} \quad \forall \beta, N, \eta.$$

Choosing a sequence $\{\beta^{(i)}\}$ which goes to infinity and such that $\partial_\beta f^{(p)}(\beta^{(i)},h)$ exists, we have¶, in the infinite volume limit

$$\liminf_{N\to\infty} e_N^{(p)}(h;\eta) \geq \partial_\beta \alpha^{(p)}(\beta^{(i)},h) \quad \eta - a.s. \tag{5.98}$$

¶Recall that the function $\beta \to -\beta f^{(p)}(\beta,h)$ is convex in β, and that a convex function is differentiable almost everywhere [70], so that it is not a problem to find the required sequence $\{\beta^{(i)}\}$.

Moreover, it is known [70] that, given a sequence $\{f_N(x)\}$ of convex functions which converges to a limit function

$$f(x) = \lim_{N\to\infty} f_N(x),$$

for any x such that $f'(x)$ exists, then also

$$f'_N(x) \to f'(x).$$

This fact is used in Eq. (5.98), when the infinite volume limit is taken in the right hand side.

Convexity of $\alpha^{(p)}(\beta,h)$ with respect to β implies that

$$\frac{\alpha^{(p)}(\beta,h)-\alpha^{(p)}(0,h)}{\beta}$$

is increasing in β. This, together with the fact that $\alpha^{(p)}(\beta,h)/\beta$ decreases with β (see Eq. (5.60)) implies that

$$-\frac{\alpha^{(p)}(0,h)}{\beta^2}\leq\partial_\beta\left(\frac{\alpha^{(p)}(\beta,h)}{\beta}\right)\leq 0.$$

Therefore,

$$\lim_{i\to\infty}\partial_\beta\alpha^{(p)}(\beta^{(i)},h)\geq\lim_{i\to\infty}\partial_\beta(\beta^{(i)}\frac{\alpha^{(p)}(\beta^{(i)},h)}{\beta^{(i)}})=\lim_{i\to\infty}\frac{\alpha^{(p)}(\beta^{(i)},h)}{\beta^{(i)}}=e_0^{(p)}(h), \tag{5.99}$$

and

$$\liminf_{N\to\infty}e_N^{(p)}(h;\eta)\geq e_0^{(p)}(h)\quad \eta-a.s, \tag{5.100}$$

which, together with (5.97), completes the proof of the theorem. □

5.5.3 Mean field disordered models with non-Ising type spins

In all the models we have considered so far, the basic degrees of freedom are two-valued Ising variables $\sigma_i=\pm1$. In this section, we consider a more general class of mean field spin glass models, whose spin degrees of freedom may, in general, have many components, taking arbitrary values in $\mathbb{R}$. This generalization is conceptually quite important, at least for two reasons. First of all, as we explained in Chapter 2, one of the main purposes of spin glass theory is to describe disordered magnetic alloys, and the magnetic moments of the impurities have of course three components. Secondly, we will see in the following of this work that it is very useful to consider identical coupled replicas of the mean field spin glass system, in order to obtain upper bounds for the free energy per site of the one-replica system. While the approach we outlined in the previous sections does not allow directly to prove the existence of the thermodynamic limit for the coupled replicas, the generalization we present in this section covers also this case. Indeed, we will see that this model can be seen as a system whose spin degrees of freedom have more than one Ising-like component.

The generic configuration σ of the system is again defined by N spin degrees of freedom $\sigma_1,\sigma_2,\ldots$. We suppose each σ_i to belong to a set $\mathcal{S}\in\mathbb{R}^n$, $n\in\mathbb{N}$, equipped with an *a priori* measure ν. For instance, the case $\mathcal{S}=\{-1,+1\}$ and $\nu=1/2(\delta_{-1}+\delta_{+1})$ corresponds to the usual Ising two-valued variables. The Hamiltonian of the model, $H_N(\sigma;J)$, depends on the spin configuration, on the system size N and on some quenched disorder, which we denote as J. Of course, the Hamiltonian can also depend on some additional external fields, *e.g.*, on the magnetic field h. We do not indicate this dependence explicitly. The mean field character of the model consists in the condition that, if two configurations σ and σ' are related by a permutation of the site indices, the random variables $H_N(\sigma;J)$ and $H_N(\sigma';J)$ have the same distribution.

In analogy with definitions (3.2)-(3.4), we introduce the disorder dependent partition function $Z_N(\beta; J)$, the quenched free energy per site $f_N(\beta)$, and the Boltzmann-Gibbs state ω_J, according to the definitions

$$Z_N(\beta; J) = \int_{\mathcal{S}^N} d\nu(\sigma_1)\dots d\nu(\sigma_N) \exp(-\beta H_N(\sigma; J)), \tag{5.101}$$

$$-\beta f_N(\beta) = N^{-1} E \ln Z_N(\beta; J), \tag{5.102}$$

$$\omega_J(A) = Z_N(\beta; J)^{-1} \int_{\mathcal{S}^N} d\nu(\sigma_1)\dots d\nu(\sigma_N) A(\sigma) \exp(-\beta H_N(\sigma; J)), \tag{5.103}$$

where A is a generic function of the σ's.

We restrict our analysis to the case of Gaussian models, *i.e.*, to models for which the $H_N(\sigma; J)$ are (correlated) Gaussian random variables. Of course, these random variables are fully characterized by their mean values $b_N(\sigma)$ and covariance matrix $c_N(\sigma, \sigma')$, defined in (5.19), (5.20). In order to prove the existence of the thermodynamic limit, we suppose that the following conditions are satisfied. First of all, we require that

$$\frac{b_N(\sigma)}{N} = g_1(m_N^{(1)}(\sigma), \dots, m_N^{(k)}(\sigma)) + O(N^{-1}). \tag{5.104}$$

Here, $k \in \mathbb{N}$, g_1 is a smooth function of class C^1 and the $m_N^{(i)}(\sigma)$ are bounded functions, with $N m_N^{(i)}$ additive in the system size. In other words,

$$|m_N^{(i)}(\sigma)| \leq M \qquad \forall i, N, \sigma \tag{5.105}$$

and

$$N m_N^{(i)}(\sigma) = N_1 m_{N_1}^{(i)}(\sigma^{(1)}) + N_2 m_{N_2}^{(i)}(\sigma^{(2)}), \tag{5.106}$$

if $N = N_1 + N_2$ and if the configuration σ can be decomposed as

$$\sigma = (\sigma_1^{(1)}, \dots, \sigma_{N_1}^{(1)}, \sigma_1^{(2)}, \dots, \sigma_{N_2}^{(2)}).$$

As regards the covariance matrix, we assume that

$$\frac{c_N(\sigma, \sigma')}{N} = g_2(Q_N^{(1)}(\sigma, \sigma'), \dots, Q_N^{(k)}(\sigma, \sigma')) + O(N^{-1}), \tag{5.107}$$

where g_2 is a convex function with continuous derivatives. The variables $Q_N^{(i)}$ must satisfy properties analogous to (5.105)-(5.106), *i.e.*,

$$|Q_N^{(i)}(\sigma, \sigma)| \leq M \qquad \forall i, N, \sigma \tag{5.108}$$

and

$$N\, Q_N^{(i)}(\sigma, \sigma') = N_1\, Q_{N_1}^{(i)}(\sigma^{(1)}, \sigma'^{(1)}) + N_2\, Q_{N_2}^{(i)}(\sigma^{(2)}, \sigma'^{(2)}). \tag{5.109}$$

It is interesting to observe that the Ising spin glass models of Chapter 3 have the additional properties that $c_N(\sigma, \sigma)$ does not depend on the configuration σ, and that g_1 is a linear function.

Now, we can state our result:

Theorem 10 *[58] If conditions (5.104) to (5.109) are satisfied, then the thermodynamic limit of the quenched free energy exists:*

$$\lim_{N\to\infty} -\frac{1}{N\beta} E \ln Z_N(\beta; J) = f(\beta). \tag{5.110}$$

Moreover, the disorder dependent free energy converges almost surely, with respect to the disorder realization:

$$\lim_{N\to\infty} -\frac{1}{N\beta} \ln Z_N(\beta; J) = f(\beta) \qquad J-\text{almost surely}, \tag{5.111}$$

and its disorder fluctuations can be estimated as

$$P\left(\left|-\frac{1}{N\beta} \ln Z_N(\beta; J) - f_N(\beta)\right| \geq u\right) \leq 2\exp\left(-\frac{Nu^2}{2L}\right), \tag{5.112}$$

where

$$L = \max_{|x_i|\leq M\ \forall i} |g_2(x_1, \ldots, x_k)|, \tag{5.113}$$

and M is the same constant as in (5.108).

Remark As we already showed for the Ising type models, from Eqs. (5.110)-(5.112) follows also the convergence, both under quenched average and J-almost surely, of the ground state energy per site.

Before we turn to the proof of the theorem, we give a few examples of physically meaningful systems to which it applies.

1. The Sherrington-Kirkpatrick model with non-Ising type spins, defined as

$$H_N(\sigma, h; J) = -\frac{1}{\sqrt{N}} \sum_{1\leq i<j\leq N} J_{ij}\sigma_i\sigma_j - h\sum_{i=1}^{N} \sigma_i. \tag{5.114}$$

As for the spin degrees of freedom, we suppose that $\sigma_i \in \mathcal{S} = [-a, a]$, while the measure ν on $\mathcal{S}$, which appears in the definition of the partition function, is arbitrary. In this case,

$$\frac{b_N(\sigma)}{N} = -h\, m_N(\sigma) = -\frac{h}{N}\sum_{i=1}^{N} \sigma_i$$

and conditions (5.104) to (5.106) are clearly satisfied, since $|m_N(\sigma)| \leq a$ and the total magnetization $\sum_i \sigma_i$ is linear in the system size. Of course, the function g_1 in (5.104) is just $g_1(x) = -h\,x$. As regards the covariance matrix, one finds easily

$$\frac{c_N(\sigma, \sigma')}{N} = \frac{q^2_{\sigma\sigma'}}{2} + O(N^{-1}),$$

where

$$q_{\sigma\sigma'} = \frac{1}{N}\sum_{i=1}^{N} \sigma_i\sigma'_i, \qquad |q_{\sigma\sigma'}| \leq a^2,$$

is as usual the overlap of the two configurations. Since $Nq_{\sigma\sigma'}$ is additive and $g_2(x) = x^2/2$ is convex, conditions (5.20) to (5.109) are also satisfied.

2. The Sherrington-Kirkpatrick model with an additional Curie-Weiss interaction, defined as

$$H_N(\sigma, h; J) = -\frac{1}{\sqrt{N}} \sum_{1\le i<j\le N} J_{ij}\sigma_i\sigma_j - \frac{J_0}{N}\sum_{i,j=1}^{N}\sigma_i\sigma_j - h\sum_{i=1}^{N}\sigma_i,$$

where the non-random constant J_0 is the strength of the Curie-Weiss mean field ferromagnetic interaction and, again, $\sigma_i \in [-a, a]$. This model can be obtained from the previous one, if one supposes that the Gaussian variables J_{ij} in (5.114) have mean value $2J_0/\sqrt{N}$. This case can be dealt with in analogy with the previous one, with the only difference that

$$\frac{b_N(\sigma)}{N} = -J_0\, m_N(\sigma)^2 - h\, m_N(\sigma),$$

so that $g_1(x) = -h\,x - J_0\,x^2$.

3. The Sherrington-Kirkpatrick model with Heisenberg type interaction, defined by the Hamiltonian

$$H_N(\sigma, \vec{h}; J) = -\frac{1}{\sqrt{N}} \sum_{1\le i<j\le N} J_{ij}\vec{\sigma}_i\,\vec{\sigma}_j - \sum_{i=1}^{N}\vec{h}\,\vec{\sigma}_i, \tag{5.115}$$

where each $\vec{\sigma}_i$ has n bounded components $\sigma_i^{(1)}, \ldots, \sigma_i^{(n)}$, and $\vec{u}\,\vec{v}$ denotes scalar product in $\mathbb{R}^n$. In this case,

$$\frac{c_N(\sigma, \sigma')}{N} = \frac{1}{2}\sum_{a,b=1}^{n}(q^{ab}_{\sigma\sigma'})^2 + O(N^{-1}), \tag{5.116}$$

where

$$q^{ab}_{\sigma\sigma'} = \frac{1}{N}\sum_{i=1}^{N}\sigma_i^{(a)}\sigma_i'^{(b)}.$$

The check of properties (5.104) to (5.109) is trivial, and is left to the reader.

4. The multi-replica Sherrington-Kirkpatrick model, with coupled replicas. In this case, the Hamiltonian depends on the configurations $\sigma^{(1)}, \ldots, \sigma^{(n)}$ of the n replicas of the usual Sherrington-Kirkpatrick model defined in Chapter 3, which interact through a term depending on the mutual overlaps:

$$H_N(\sigma^{(1)}, \ldots, \sigma^{(n)}; J) = -\frac{1}{\sqrt{N}} \sum_{1\le i<j\le N} J_{ij}(\sigma_i^{(1)}\sigma_j^{(1)} + \ldots + \sigma_i^{(n)}\sigma_j^{(n)}) + Ng_1(\{q_{ab}\}),$$

where g_1 is a smooth C^1 function of all the overlaps. This model also fits our general scheme, as it is easily realized by noticing that the generic configuration of the model can be seen as a collection of variables $\tilde{\sigma}_i$, $i = 1, 2, \ldots, N$, each with n Ising components $\tilde{\sigma}_i^a = \sigma_i^{(a)} = \pm 1$. Therefore, the present model can be seen as a particular case of the Sherrington-Kirkpatrick model with Heisenberg-type interaction discussed at point 3.

It is instructive to verify explicitly that, for these models, the method introduced in [22] does not work, and requires an extension. Indeed, let us try to apply naively the approach of Section 5.4.2 to the model 1. When one integrates by parts on the Gaussian random variables in (5.49) one finds, instead of Eq. (5.50),

$$-\frac{d}{dt}\frac{1}{N\beta}E\ln Z_N(t) = -\frac{\beta}{4N^2}\sum_{i,j=1}^{N} E(\omega_t(\sigma_i^2\sigma_j^2) - \omega_t^2(\sigma_i\sigma_j))$$
$$+\frac{\beta}{4NN_1}\sum_{i,j=1}^{N_1} E(\omega_t(\sigma_i^2\sigma_j^2) - \omega_t^2(\sigma_i\sigma_j)) + \frac{\beta}{4NN_2}\sum_{i,j=N_1+1}^{N} E(\omega_t(\sigma_i^2\sigma_j^2) - \omega_t^2(\sigma_i\sigma_j)) + O(1/N)$$
$$= \frac{\beta}{4}\left\langle q_{12}^2 - \frac{N_1}{N}(q_{12}^{(1)})^2 - \frac{N_2}{N}(q_{12}^{(2)})^2\right\rangle - \frac{\beta}{4}\left\langle q_{11}^2 - \frac{N_1}{N}(q_{11}^{(1)})^2 - \frac{N_2}{N}(q_{11}^{(2)})^2\right\rangle + O(1/N). \tag{5.117}$$

While in the Ising spin case

$$q_{11} = q_{11}^{(1)} = q_{11}^{(2)} = 1$$

identically and the second term in Eq. (5.117) vanishes, for generic spin variables this is no more the case, and the two terms in the right hand side of (5.117) give contributions of opposite sign. Therefore, we cannot conclude immediately that the t derivative is non-positive and find subadditivity for the free energy.

Proof of Theorem 10) Eq. (5.112) is a trivial extension of Theorem 2 we discussed above. Therefore, we turn to the main statements of the theorem, concerning the existence of the thermodynamic limit. For simplicity, we assume that

$$N^{-1}b_N(\sigma) = g_1(m_N(\sigma)) + O(N^{-1}),$$

and

$$N^{-1}c_N(\sigma,\sigma') = g_2(Q_N(\sigma,\sigma')) + O(N^{-1}),$$

corresponding to the case $k = 1$ in Eqs. (5.104), (5.107), and we assume that $L = 1$ in (5.113) (the general case can be obtained as a simple extension).

First of all, we prove the existence of the limit along sequences of the type $\{N_K\} = \{N_0 n^K\}$, with $n, N_0 \in \mathbb{N}$. As in [22], the idea is to find a suitable interpolation between the original system, of size N_K, and a system composed of n non-interacting subsystems, of size N_{K-1} each. However, in the present case, it is also necessary to divide the configuration space into sets, such that $m_N(\sigma)$ and $Q_N(\sigma,\sigma)$ are approximately constant within each set. This idea was introduced by Talagrand [11], and developed in an important series of applications. For any $0 < \varepsilon < 1$, we can rewrite the partition function (5.101) as

$$Z_{N_K}(\beta;J) = \sum_{i,j=0}^{[1/\varepsilon]} Z_{N_K}^{(ij)}(\beta;J) \equiv \sum_{i,j=0}^{[1/\varepsilon]}\int_{C_{ij}} d\tilde{\nu}(\sigma)\exp(-\beta H_{N_K}(\sigma;J)), \tag{5.118}$$

where

$$C_{ij} = \{\sigma \in \mathcal{S}^{N_K} : i\varepsilon \le Q_{N_K}(\sigma,\sigma), < (i+1)\varepsilon, j\varepsilon \le m_{N_K}(\sigma), < (j+1)\varepsilon\}, \tag{5.119}$$

$[x]$ denotes the integer part of x and, for simplicity of notations, we put

$$d\tilde{\nu}(\sigma) = d\nu(\sigma_1)\dots d\nu(\sigma_N).$$

Since $N_K = nN_{K-1}$, we can divide the system into n subsystems of N_{K-1} spins each, and we denote the configuration of the ℓ-th subsystem as $\sigma^{(\ell)}$, with $\ell = 1, 2, \dots, n$. Of course, the following inequality holds

$$Z^{(ij)}_{N_K}(\beta; J) \geq \tilde{Z}^{(ij)}_{N_K}(\beta; J) = \int_{\tilde{C}_{ij}} d\tilde{\nu}(\sigma) \exp(-\beta H_{N_K}(\sigma; J)), \tag{5.120}$$

where

$$\begin{aligned} C_{ij} \supseteq \tilde{C}_{ij} \; &= \; \{\sigma \in \mathcal{S}^{N_K} : i\varepsilon \leq Q_{N_{K-1}}(\sigma^{(\ell)}, \sigma^{(\ell)}) < (i+1)\varepsilon, \\ & \qquad j\varepsilon \leq m_{N_{K-1}}(\sigma^{(\ell)}) < (j+1)\varepsilon, \ \forall \ell\}. \end{aligned} \tag{5.121}$$

Now, we introduce an interpolating parameter $0 \leq t \leq 1$, and the auxiliary partition function

$$\begin{aligned} \tilde{Z}^{(ij)}_{N_K}(t, \beta) \; &= \; \int_{\tilde{C}_{ij}} d\tilde{\nu}(\sigma) \exp \beta \left(-\sqrt{t}\xi_{N_K}(\sigma) - tb_{N_K}(\sigma) - \sqrt{1-t} \sum_{\ell=1}^{n} \xi^{\ell}_{N_{K-1}}(\sigma^{(\ell)}) \right. \\ & \left. -(1-t) \sum_{\ell=1}^{n} b_{N_{K-1}}(\sigma^{(\ell)}) \right), \end{aligned}$$

where $H_N(\sigma; J)$ has been rewritten as in (5.24), and $\xi^{\ell}_N(\sigma)$ are n *independent* copies of the random variable $\xi_N(\sigma)$. Clearly, for the boundary values of the parameter t one has

$$-\frac{1}{N_K \beta} E \ln \tilde{Z}^{(ij)}_{N_K}(0, \beta) = -\frac{1}{N_{K-1}\beta} E \ln Z^{(ij)}_{N_{K-1}}(\beta; J) \tag{5.122}$$

and

$$-\frac{1}{N_K \beta} E \ln \tilde{Z}^{(ij)}_{N_K}(1, \beta) = -\frac{1}{N_K \beta} E \ln \tilde{Z}^{(ij)}_{N_K}(\beta; J) \geq -\frac{1}{N_K \beta} E \ln Z^{(ij)}_{N_K}(\beta; J). \tag{5.123}$$

As regards the t derivative, we apply the integration by parts formula (5.28) and, recalling that the random variables $\xi^{\ell}_N(\sigma)$ are statistically independent for different ℓ, we find after some straightforward computations,

$$-\frac{d}{dt}\frac{1}{N_K \beta} E \ln \tilde{Z}^{(ij)}_{N_K}(t, \beta) =$$

$$-\frac{\beta}{2} \left\langle g_2(Q_{N_K}(\sigma, \sigma)) - \frac{1}{n} \sum_{\ell=1}^{n} g_2(Q_{N_{K-1}}(\sigma^{(\ell)}, \sigma^{(\ell)})) \right\rangle \tag{5.124}$$

$$+\frac{\beta}{2} \left\langle g_2(Q_{N_K}(\sigma, \sigma')) - \frac{1}{n} \sum_{\ell=1}^{n} g_2(Q_{N_{K-1}}(\sigma^{(\ell)}, \sigma'^{(\ell)})) \right\rangle \tag{5.125}$$

$$+ \left\langle g_1(m_{N_K}(\sigma)) - \frac{1}{n} \sum_{\ell=1}^{n} g_1(m_{N_{K-1}}(\sigma^{(\ell)})) \right\rangle + O\left(\frac{1}{N_K}\right), \tag{5.126}$$

where the averages are, of course, restricted to configurations belonging to $\tilde{C}_{ij}$. Since Q_{N_K} is a convex combination of the $Q_{N_{K-1}}$ and g_2 is a convex function, the term (5.125) is non-positive. On the other hand, since g_2 is a function of class C^1, and $Q_{N_{K-1}}(\sigma^{(\ell)}, \sigma^{(\ell)})$ and are constrained to belong to the interval $[i\varepsilon, (i+1)\varepsilon)$ for each ℓ, the term (5.124) is of order ε. The same holds for the term (5.126). This implies that, for K large enough,

$$-\frac{d}{dt}\frac{1}{N_K\beta}E\ln\tilde{Z}^{(ij)}_{N_K}(t,\beta) \leq C\varepsilon, \tag{5.127}$$

for some positive constant C independent of N. Recalling Eqs. (5.122), (5.123), this means that

$$-\frac{1}{N_K\beta}E\ln Z^{(ij)}_{N_K}(\beta;J) + \frac{1}{N_{K-1}\beta}E\ln Z^{(ij)}_{N_{K-1}}(\beta;J) \leq C\varepsilon. \tag{5.128}$$

Now, we want to turn this inequality, which involves disorder averages, into an inequality valid $J-$almost everywhere. To this purpose, we choose $\varepsilon = N_K^{-1/4}$ and we observe that, thanks to the estimate (5.22),

$$P\left(-\frac{1}{N_K\beta}\ln Z^{(ij)}_{N_K}(\beta;J) \geq -\frac{1}{N_K\beta}E\ln Z^{(ij)}_{N_K}(\beta;J) + C\varepsilon\right) \leq 2\exp\left(-\frac{\sqrt{N_K}C^2}{2}\right)$$

and

$$P\left(-\frac{1}{N_{K-1}\beta}\ln Z^{(ij)}_{N_{K-1}}(\beta;J) \leq -\frac{1}{N_{K-1}\beta}E\ln Z^{(ij)}_{N_{K-1}}(\beta;J) - C\varepsilon\right) \leq 2\exp\left(-\frac{\sqrt{N_K}C^2}{2n}\right).$$

Therefore, with probability $P \geq 1 - 4\sqrt{N_K}\exp\left(-\frac{\sqrt{N_K}C^2}{2n}\right)$, one has

$$\begin{aligned}&-\frac{1}{N_K\beta}\ln Z^{(ij)}_{N_K}(\beta;J) \leq -\frac{1}{N_{K-1}\beta}\ln Z^{(ij)}_{N_{K-1}}(\beta;J) + 3CN_K^{-1/4}\\ &\forall i,j = 0,\dots,[N_K^{1/4}].\end{aligned} \tag{5.129}$$

Since the probability of the complementary event is summable in K, it follows from Borel-Cantelli lemma that inequality (5.129) holds J-almost surely, for K large enough. As a consequence, one obtains

$$\begin{aligned} Z_{N_K}(\beta;J) &= \sum_{i,j=0}^{[N_K^{1/4}]} Z^{(ij)}_{N_K}(\beta;J) \geq e^{-3\beta CN_K^{3/4}}\sum_{i,j=0}^{[N_K^{1/4}]}\left(Z^{(ij)}_{N_{K-1}}(\beta;J)\right)^n \\ &\geq e^{-3\beta CN_K^{3/4}}N_K^{(1-n)/2}\left(\sum_{i,j=0}^{[N_K^{1/4}]} Z^{(ij)}_{N_{K-1}}(\beta;J)\right)^n \\ &= \frac{e^{-3\beta CN_K^{3/4}}}{N_K^{(n-1)/2}}Z^n_{N_{K-1}}(\beta;J).\end{aligned} \tag{5.130}$$

Here, we have used the property

$$\sum_{i=1}^k x_i^n \geq k^{1-n}\left(\sum_{i=1}^k x_i\right)^n, \tag{5.131}$$

which holds if $x_i \geq 0$, thanks to the convexity of the function $x \to x^n$. Taking the logarithm and dropping terms of lower order in N_K, one has

$$-\frac{1}{N_K\beta}\ln Z_{N_K}(\beta;J) \leq -\frac{1}{N_{K-1}\beta}\ln Z_{N_{K-1}}(\beta;J) + 3CN_K^{-1/4} \quad J-a.s., \tag{5.132}$$

for K large enough. Notice that, with respect to (5.128), the above inequality involves the original free energy, where the sum over configurations has no restrictions. From (5.132), it follows that the thermodynamic limit exists, J-almost surely by subadditivity, the term $N_K^{-1/4}$ being inessential, as discussed in the footnote on page 48. On the other hand, the exponential estimate (5.112) implies that the limit has a non random value $f(\beta)$, for almost every disorder realization J.

Once the almost sure convergence is proved, the convergence of the quenched average can be obtained easily, provided that the probability that $1/N \ln Z_N$ assumes large values is sufficiently small. For instance, one has the following criterion [12]: given random variables X_K and X, if $X_K \longrightarrow X$ almost surely for $K \to \infty$, and if

$$\lim_{\lambda\to\infty}\sup_K E\left(|X_K|\, 1_{\{|X_K|\geq\lambda\}}\right) = 0, \tag{5.133}$$

where 1_A denotes the characteristic function of the set A, then

$$EX_K \to EX.$$

In the present case, $X_K = -1/(N_K\beta)\ln Z_{N_K}(\beta;J)$, $X = f(\beta)$, and the condition (5.133) can be easily checked, by employing the exponential bound (5.112).

In conclusion, we have proved almost sure convergence for the free energy, and convergence of its quenched average, for any subsequence of the form $\{N_0 n^K\}$. It is not difficult to show, by standard methods, that this implies convergence along any increasing subsequence $\{N_K\}$, and the uniqueness of the limit. □

Chapter 6

Sum rules and lower bounds for the free energy

The previous chapter was entirely devoted to the proof of the existence of the thermodynamic limit for the free energy per site, and to the estimation of its disorder fluctuations, for a wide class of mean field spin glass models. Even if this is a very important conceptual problem, the mere existence of the limit does not tell us much about the physics of the system. Of course, one would like to compute the infinite volume free energy, in order to verify the predictions of Parisi theory. This task has proven to be very hard, and actually it has been accomplished so far only in very particular cases.

Of course, we have a few exactly solvable models, *e.g.*, Derrida's Random Energy Model and Generalized Random Energy Model. These models are very instructive, but they look even more artificial than the mean field spin glass models introduced in Chapter 3. For more "realistic" ones, like that of Sherrington and Kirkpatrick or the p-spin, the control of the limit has been so far achieved essentially only for high temperature, where replica symmetry is not broken and the system is considered as trivial in the theoretical physics literature. A remarkable exception is the case of the p-spin model for p large, where M. Talagrand has been able to give [47], [48] a complete description of the phase where replica symmetry breaking occurs, provided that the temperature is not too low*. Also in this case, the physical picture given by Parisi theory has been fully confirmed.

The status of knowledge has been evolving quite fast in the last times, and a general strategy to approach the problem seems to have emerged. First of all, given a mean field spin glass model one proves that its free energy per site is *bounded from below* by the corresponding Parisi solution

$$-\beta^{-1}\bar{\alpha}(\beta,h),$$

which was defined in (4.22) of Chapter 4 in the particular case of the Sherrington-Kirkpatrick model. This result was proved recently by F. Guerra [23], for the Sherrington-Kirkpatrick and p-spin models, in the whole range of thermodynamic parameters β, h, and in particular in the low temperature region. The bounds derive from a *sum rule*,

*essentially, the analysis by Talagrand covers the range of temperature where one-step replica symmetry breaking occurs. One expects continuous replica symmetry breaking to occur at even lower temperatures.

expressing the true free energy of the system as the sum of the Parisi solution plus a positive term, which involves the fluctuation of the overlap in suitable states (see Section 6.2 in this chapter for details). The next necessary step to control the infinite volume limit is to prove that the opposite bound also holds, *i.e.*, to prove that $-\beta^{-1}\bar{\alpha}(\beta, h)$ is also an *upper bound* for the free energy, apart from error terms which vanish in the thermodynamic limit. To this purposes one considers two replicas of the system, suitably coupled. The idea is to use, as coupling interaction, just the fluctuation terms which appear in the above mentioned sum rule, in order to show that they vanish in the infinite volume limit. This second step has been performed so far essentially only for high enough temperature, in situations where replica symmetry holds, as reported in detail in Chapter 7 of the present work.

In this chapter, we illustrate Guerra's *broken replica symmetry bounds* for the free energy. The conceptual importance of the result can be hardly overestimated. Indeed, it shows how the functional parameter $x(q)$ introduced in Chapter 4 and Parisi solution arise in a very natural way in the theory, without recourse to the replica trick, and it gives strong support to Parisi's prediction (4.22) for the infinite volume limit of the free energy. In the second place, Guerra's result gives a firm ground to the maximization procedure in Eq. (4.22) which, as we already noticed, is quite unclear from the physical point of view. Guerra's bounds have been later extended by Franz *et al.* [57] to finite-connectivity models.

As a byproduct we also prove that, in the whole region of parameters below the Almeida-Thouless line, replica symmetry cannot hold for the Sherrington-Kirkpatrick model.

It is quite surprising that the proof of the broken replica symmetry bounds turns out to be technically rather easy, since it relies on an interpolation scheme somewhat similar to the one we employed in the previous chapter to prove the existence of the infinite volume limit.

6.1 An example: the replica-symmetric bound

As a preliminary exercise, we show that, for the Sherrington-Kirkpatrick model, the following lower bound holds for the free energy:

$$f_N(\beta, h) \geq f_{SK}(\beta, h) \tag{6.1}$$

for any value of N, β, h. Here, $f_{SK}(\beta, h)$ is the so called "replica symmetric free energy"

$$f_{SK}(\beta, h) = -\beta^{-1}\alpha_{SK}(\beta, h), \tag{6.2}$$

where α_{SK} was defined in (4.28). Inequality (6.1) was obtained by Guerra in [53].

As we discussed in Chapter 4, f_{SK} arises from the functional integral (4.7) when one looks for a saddle point symmetric under permutation of replicas or, equivalently, from the trial functional (4.20), when one considers a functional order parameter of the simple one step form (4.27). Recall that f_{SK} is expected to be the infinite volume of the free energy per site, in the region above the Almeida-Thouless line, *i.e.*, when condition (4.26) holds.

The essential idea to prove Eq. (6.1) is to compare, by means of a suitable interpolation, the Sherrington-Kirkpatrick model and an exactly solvable model, where a random site dependent external field takes the place of the random two-body potential. To this purpose, fix β, h and consider the auxiliary interpolating partition function

$$\tilde{Z}_N(t) = \sum_{\{\sigma\}} \exp\left(\beta\sqrt{\frac{t}{N}} \sum_{1\leq i<j\leq N} J_{ij}\sigma_i\sigma_j + \beta\sqrt{\bar{q}(1-t)}\sum_i J_i\sigma_i + \beta h \sum_i \sigma_i\right), \qquad (6.3)$$

depending on the interpolating parameter $0 \leq t \leq 1$, where J_i are independent identically distributed standard unit Gaussian variables, independent of the disorder J_{ij}. For $t = 0$ all sites decouple due to the disappearance of the two-body interaction, so that

$$\tilde{Z}_N(0) = \sum_{\{\sigma\}} e^{\beta\sum_i(\sqrt{\bar{q}}J_i+h)\sigma_i} = 2^N \prod_i \cosh(\beta h + \beta\sqrt{\bar{q}}J_i), \qquad (6.4)$$

and $1/NE\ln\tilde{Z}_N$ can be computed explicitly, with the result

$$\frac{1}{N}E\ln\tilde{Z}_N(0) = \ln 2 + \int d\mu(z)\ln\cosh(\beta h + \beta z\sqrt{\bar{q}}). \qquad (6.5)$$

The similarity between (6.5) and the expression for the replica symmetric solution (4.28) is evident. On the other hand,

$$\frac{1}{N}E\ln\tilde{Z}_N(1) = \frac{1}{N}E\ln Z_N(\beta,h) = -\beta f_N(\beta,h). \qquad (6.6)$$

As for the t derivative of $1/NE\ln\tilde{Z}$, the computation is not very different from that of (5.49). Indeed, one finds

$$\frac{d}{dt}\frac{1}{N}E\ln\tilde{Z}_N(t) = \frac{\beta}{2N\sqrt{Nt}}E\sum_{1\leq i<j\leq N} J_{ij}\omega_t(\sigma_i\sigma_j) - \frac{\beta\sqrt{\bar{q}}}{2N\sqrt{1-t}}\sum_i EJ_i\omega_t(\sigma_i), \qquad (6.7)$$

where the Boltzmann-Gibbs average ω_t refers to the t-dependent partition function (6.3). After integration by parts on the Gaussian disorder, and grouping terms in order to reconstruct squares, one is left with

$$\frac{d}{dt}\frac{1}{N}E\ln\tilde{Z}_N(t) = \frac{\beta^2}{4}(1-\bar{q})^2 - \frac{\beta^2}{4}\left\langle(q_{12}-\bar{q})^2\right\rangle_t. \qquad (6.8)$$

Therefore, integrating along t, and taking into account the boundary conditions, one finds the *sum rule*

$$f_N(\beta,h) = f_{SK}(\beta,h) + \frac{\beta}{4}\int_0^1 \langle(q_{12}-\bar{q})^2\rangle_{t'}\,dt' \geq f_{SK}(\beta,h), \qquad (6.9)$$

and (6.1) is proven. □

We call Eq. (6.9) a “sum rule” because it expresses the true free energy as the sum of two positive terms, one of which is the replica symmetric solution, and the other is the integral of the "source" $\langle (q_{12}-\bar{q})^2\rangle_{t'}$. From Eq. (6.9) it appears clear that, as one expects from the physical point of view, the replica symmetric approximation for the free energy is equivalent to the assumption that the overlap q_{12} is self-averaging. This was already proved, in a different context, by Pastur and Scherbina in [64].

It is very important that the "source" term has a very natural interpretation, as the mean square fluctuation of the overlap around the replica symmetric value $\bar{q}$, in the *auxiliary* state indexed by the parameter t. As we will explain in Chapter 7, the expression for the source suggests in a very natural way how to couple two replicas of the system, in order to obtain an *upper bound* for the free energy, and to show that replica symmetry holds, in a suitable region of parameters.

6.2 Guerra's Broken Replica Bounds

Here, we start by proving Guerra's bounds for the Sherrington-Kirkpatrick model. The present section, as well as Sections 6.3 and 6.4, is based on Ref. [23] by Guerra.

In Chapter 4 we introduced the structure of the space of functional order parameters and of Parisi trial solutions with replica symmetry breaking. The main motivation for the introduction of these definitions is the following expected tentative Theorem:

Theorem 11 (**expected**) *In the thermodynamic limit, for the Sherrington-Kirkpatrick model, we have*

$$\lim_{N\to\infty} N^{-1} E \ln Z_N(\beta,h;J) = \bar{\alpha}(\beta,h).$$

for any value of the temperature and magnetic field.

Of course, the present technology is far from being able to give a complete rigorous proof. However, in this section we prove that $\bar{\alpha}(\beta,h)$ is at least a rigorous upper bound for $N^{-1}E\ln Z_N(\beta,h;J)$, uniformly in N.

The main results of this section are summarized in the following

Theorem 12 *[23] For all values of the inverse temperature β, and the external magnetic field h, and for any functional order parameter x, the following bound holds*

$$-\frac{1}{N\beta}E\ln Z_N(\beta,h;J) \geq -\frac{1}{\beta}\bar{\alpha}(\beta,h;x),$$

uniformly in N, where $\bar{\alpha}(\beta,h;x)$ is defined in (4.20). Consequently, we have also

$$-\frac{1}{N\beta}E\ln Z_N(\beta,h;J) \geq -\frac{1}{\beta}\bar{\alpha}(\beta,h),$$

uniformly in N, where $\bar{\alpha}(\beta,h)$ is the infimum defined in (4.21). Moreover, for the thermodynamic limit, we have

$$-\lim_{N\to\infty}\frac{1}{N\beta}E\ln Z_N(\beta,h;J) \equiv f(\beta,h) \geq -\frac{1}{\beta}\bar{\alpha}(\beta,h),$$

and

$$-\lim_{N\to\infty}\frac{1}{N\beta}\ln Z_N(\beta,h;J) = f(\beta,h) \geq -\frac{1}{\beta}\bar{\alpha}(\beta,h),$$

J-almost surely.

The proof is long, and will be split in a series of intermediate results. Consider a generic piecewise constant functional order parameter x, as in (4.14), and define the auxiliary partition function $\tilde{Z}$ as follows

$$\begin{aligned}\tilde{Z}_N(\beta,h;t;x;J) \equiv & \sum_{\{\sigma\}} \exp\left(\beta\sqrt{\frac{t}{N}} \sum_{1\le i<j\le N} J_{ij}\sigma_i\sigma_j \right. \\ & \left. +\beta h \sum_i \sigma_i + \beta\sqrt{1-t}\sum_{a=1}^{K} \sqrt{q_a - q_{a-1}} \sum_i J_i^a \sigma_i \right). \qquad (6.10)\end{aligned}$$

Here, we have introduced additional independent centered unit Gaussian variables J_i^a, $a = 1,\dots,K$, $i = 1,\dots,N$. The interpolating parameter t runs in the interval $[0,1]$.

For $a = 1,\dots,K$, let us call E_a the average with respect to all random variables J_i^a, $i = 1,\dots,N$. Analogously, we call E_0 the average with respect to all J_{ij}, and denote by E averages with respect to all J random variables.

Now we define recursively the random variables $Z_0, Z_1, \dots, Z_K$

$$Z_K = \tilde{Z}_N(\beta,h;t;x;J),\ Z_{K-1}^{m_K} = E_K Z_K^{m_K}, \dots,\ Z_0^{m_1} = E_1 Z_1^{m_1}, \qquad (6.11)$$

and the auxiliary function $\tilde{\alpha}_N(t)$

$$\tilde{\alpha}_N(t) = \frac{1}{N} E_0 \ln Z_0. \qquad (6.12)$$

Notice that, due to the partial integrations, any Z_a depends only on the J_{ij}, and on the J_i^b with $b \le a$, while in $\tilde{\alpha}$ all J variables have been completely averaged out.

The basic motivation for the introduction of $\tilde{\alpha}$ is given by

Lemma 3 *At the extreme values of the interpolating parameter t we have*

$$\begin{aligned}\tilde{\alpha}_N(1) &= \frac{1}{N} E \ln Z_N(\beta,h;J) = \alpha_N(\beta,h), & (6.13)\\ \tilde{\alpha}_N(0) &= \ln 2 + f(0,h;x,\beta), & (6.14)\end{aligned}$$

where $\alpha_N(\beta,h)$ is the auxiliary function defined in Eq. (3.3), and f is the solution of the antiparabolic equation (4.16) introduced in Chapter 4.

The proof is simple. In fact, at $t = 1$, the J_i^a disappear, and $\tilde{Z}$ reduces to the usual partition function Z in (3.2). On the other hand, at $t = 0$, the two site couplings J_{ij} disappear, while all effects of the J_i^a factorize with respect to the sites i. Therefore, we are essentially reduced to a one site problem, and it is clear that we can simply take $N = 1$:

$$\tilde{\alpha}_N(0) = \tilde{\alpha}_1(0).$$

Now,

$$\tilde{Z}_1(\beta,h;0;x,J) = 2\cosh(\beta h + \beta \sum_{a=1}^{K} \sqrt{q_a - q_{a-1}} J^a)$$

and it is immediate to recognize that the averages involved in the definition of $\ln Z_0$ (see Eq. (6.11)) reduce to the Gaussian averages necessary for the computation of the solution of the antiparabolic problem (4.16), (4.17), as given by the repeated application of (4.18), with the f function evaluated at $q = 0$, and $y = h$. □

It is clear that now we have to proceed to the calculation of the t derivative of $\tilde{\alpha}_N(t)$. To this purpose, we need some additional definitions. Introduce the random variables f_a, $a = 1, \dots, K$,

$$f_a = \frac{Z_a^{m_a}}{E_a(Z_a^{m_a})}, \tag{6.15}$$

and notice that they depend only on the J_i^b with $b \le a$, and are normalized, $E(f_a) = 1$. Moreover, we consider the t-dependent state ω_J (we omit the subscript t for simplicity) associated to the Boltzmann factor in (6.10), and its replicated Ω_J. A very important role is played by the following states $\tilde{\omega}_a$, and their replicated ones $\tilde{\Omega}_a$, $a = 0, \dots, K$, defined as

$$\tilde{\omega}_K(.) = \omega(.),\ \ \tilde{\omega}_a(.) = E_{a+1} \dots E_K(f_{a+1} \dots f_K \omega(.)). \tag{6.16}$$

Finally, we define the $\langle . \rangle_a$ averages, through a generalization of (3.15),

$$\langle . \rangle_a = E(f_1 \dots f_a \tilde{\Omega}_a(.)). \tag{6.17}$$

As it will be clear in the following, the $\langle . \rangle_a$ averages are able, in a sense, to concentrate the overlap fluctuations around the value q_a, if q_a refers to the optimal functional order parameter which realizes the infimum in (4.21).

Now, we have all definitions in order to be able to state the following important results.

Theorem 13 *The t derivative of $\tilde{\alpha}_N(t)$ in (6.12) is given by*

$$\begin{aligned} \frac{d}{dt}\tilde{\alpha}_N(t) &= -\frac{\beta^2}{4}\big(1 - \sum_{a=0}^{K}(m_{a+1} - m_a)q_a^2\big) \\ &\quad -\frac{\beta^2}{4}\sum_{a=0}^{K}(m_{a+1} - m_a)\langle (q_{12} - q_a)^2 \rangle_a. \end{aligned} \tag{6.18}$$

Theorem 14 *For any functional order parameter x, of the type given in (4.14), the following sum rule holds*

$$\bar{\alpha}(\beta, h; x) = \frac{1}{N} E \ln Z_N(\beta, h; J) + \frac{\beta^2}{4}\sum_{a=0}^{K}(m_{a+1} - m_a)\int_0^1 \langle (q_{12} - q_a)^2 \rangle_a(t)\, dt. \tag{6.19}$$

Clearly, Theorem 14 follows from the previous Theorem 13, by integrating with respect to t, taking into account the boundary values in Lemma 3, and the definition of $\bar{\alpha}(\beta, h; x)$ given in Chapter 4. Moreover, one should use also the obvious identity

$$\frac{1}{2}\big(1 - \sum_{a=0}^{K}(m_{a+1} - m_a)q_a^2\big) = \int_0^1 q\, x(q)\, dq, \tag{6.20}$$

which is evident from the definition (4.14) of the piecewise constant functional order parameter. By taking into account that all terms in the sum rule are non-negative, since $m_{a+1} \geq m_a$, we can immediately establish the validity of Theorem 12.

Now we must attack Theorem 13. The proof is straightforward, and involves integration by parts with respect to the Gaussian quenched disorder. We sketch only the main points. Let us begin with

Lemma 4 *The t derivative of $\tilde{\alpha}_N(t)$ in (6.12) is given by*

$$\frac{d}{dt}\tilde{\alpha}_N(t) = \frac{1}{N}E(f_1 f_2 \ldots f_K Z_K^{-1}\partial_t Z_K),$$

where

$$\begin{aligned} Z_K^{-1}\partial_t Z_K &= \tilde{Z}_N^{-1}\partial_t \tilde{Z}_N \\ &= \frac{\beta}{2\sqrt{tN}}\sum_{1\leq i<j\leq N} J_{ij}\omega(\sigma_i\sigma_j) - \frac{\beta}{2\sqrt{1-t}}\sum_{a=1}^{K}\sqrt{q_a - q_{a-1}}\sum_i J_i^a \omega(\sigma_i). \end{aligned}$$

The proof is very simple. In fact, from the definitions in (6.11), we have, for $a = 0, 1, \ldots, K-1$,

$$Z_a^{-1}\partial_t Z_a = E_{a+1}(f_{a+1}Z_{a+1}^{-1}\partial_t Z_{a+1}).$$

The rest follows from iteration of this formula, and simple calculations.

Clearly, now we have to evaluate

$$\begin{aligned} E(J_{ij}f_1 f_2 \ldots f_K \omega(\sigma_i\sigma_j)) &= \sum_{a=1}^{K} E(\ldots \partial_{J_{ij}} f_a \ldots \omega(\sigma_i\sigma_j)) + E(f_1 \ldots f_K \partial_{J_{ij}}\omega(\sigma_i\sigma_j)), \\ E(J_i^a f_1 f_2 \ldots f_K \omega(\sigma_i)) &= \sum_{b=1}^{K} E(\ldots \partial_{J_i^a} f_b \ldots \omega(\sigma_i)) + E(f_1 \ldots f_K \partial_{J_i^a}\omega(\sigma_i)), \end{aligned}$$

where we have exploited standard integration by parts on the Gaussian J variables.

The following lemma gives all additional information necessary for the proof of Theorem 13.

Lemma 5 *For the J-derivatives we have*

$$\partial_{J_{ij}}\omega(\sigma_i\sigma_j) = \beta\sqrt{\frac{t}{N}}(1-\omega^2(\sigma_i\sigma_j)), \tag{6.23}$$

$$\partial_{J_i^a}\omega(\sigma_i) = \beta\sqrt{1-t}\sqrt{q_a - q_{a-1}}(1-\omega^2(\sigma_i)), \tag{6.24}$$

$$\partial_{J_{ij}} f_a = \beta\sqrt{\frac{t}{N}} m_a f_a(\tilde{\omega}_a(\sigma_i\sigma_j) - \tilde{\omega}_{a-1}(\sigma_i\sigma_j)), \tag{6.25}$$

$$\partial_{J_i^a} f_b = 0 \quad if \quad b < a, \tag{6.26}$$

$$\partial_{J_i^a} f_a = \beta\sqrt{1-t}\sqrt{q_a - q_{a-1}} m_a f_a \tilde{\omega}_a(\sigma_i), \tag{6.27}$$

$$\partial_{J_i^a} f_b = \beta\sqrt{1-t}\sqrt{q_a - q_{a-1}} m_b f_b(\tilde{\omega}_b(\sigma_i) - \tilde{\omega}_{b-1}(\sigma_i)) \; if \; b > a. \tag{6.28}$$

The proof of (6.23) and (6.24) is a standard calculation. On the other hand, Eq. (6.25) follows from the definition (6.15) and the easily established equalities

$$\begin{aligned}
&\partial_{J_{ij}} Z_a^{m_a} = m_a Z_a^{m_a} Z_a^{-1} \partial_{J_{ij}} Z_a,\\
&Z_a^{-1}\partial_{J_{ij}} Z_a = E_{a+1}(f_{a+1} Z_{a+1}^{-1} \partial_{J_{ij}} Z_{a+1}),\ a = 1, \dots, K-1,\\
&Z_K^{-1}\partial_{J_{ij}} Z_K = \tilde{Z}_N^{-1} \partial_{J_{ij}} \tilde{Z}_N = \beta\sqrt{\frac{t}{N}}\omega(\sigma_i\sigma_j),\\
&Z_a^{-1}\partial_{J_{ij}} Z_a = \beta\sqrt{\frac{t}{N}} E_{a+1}(f_{a+1}\dots f_K \omega(\sigma_i\sigma_j)) = \beta\sqrt{\frac{t}{N}}\tilde{\omega}_a(\sigma_i\sigma_j).
\end{aligned}$$

In the same way, we can establish (6.26), (6.27), (6.28). However, here we have to take into account that Z_b does not depend on J_i^a if $b < a$.

A careful combination of all information given by Lemma 4 and Lemma 5, finally leads to the proof of Theorem 13. On the other hand, the main Theorem 12 follows easily from Theorem 14, and from the results of Chapter 5 about the existence of the thermodynamic limit. □

6.3 Broken replica symmetry bounds for the ground state energy

The lower bounds for the free energy, given by Theorem 12, can be turned into lower bounds for the ground state energy, once a suitable $\beta \to \infty$ limit is performed. Let us consider the ground state energy density of the Sherrington-Kirkpatrick model $-e_N(h; J)$, defined as

$$-e_N(h; J) \equiv \frac{1}{N} \inf_{\sigma} H_N(\sigma, h; J) = -\lim_{\beta\to\infty} \frac{\ln Z_N(\beta, h; J)}{\beta N},$$

where we employed Eq. (5.62) in the last equality. By taking the expectation values we also have

$$e_N(h) \equiv E(e_N(h; J)) = \lim_{\beta\to\infty} \frac{E \ln Z_N(\beta, h; J)}{\beta N}. \tag{6.30}$$

From Theorem 12 of the previous section we have, for any functional order parameter x,

$$\frac{E \ln Z_N(\beta, h; J)}{\beta N} \le \beta^{-1} \bar{\alpha}(\beta, h; x), \tag{6.31}$$

uniformly in N.

When $\beta \to \infty$, the overlap concentrates at the value $q_{12} = +1$[†]. Thanks to Eq. (4.23), this means that the optimal functional order parameter $x(q)$ tends to zero for $q < 1$. Therefore, in order to obtain meaningful broken replica symmetry bounds for the

[†]In fact, recall that

$$\partial_\beta \frac{1}{N} E \ln Z_N(\beta, h; J) = \frac{\beta}{2}(1 - \langle q_{12}^2 \rangle).$$

Since $\frac{1}{N} E \ln Z_N$ cannot grow faster than linearly for $\beta \to \infty$, it is clear that $\langle q_{12}^2 \rangle$ must tend to 1 in this limit.

ground state energy analogous to those of Theorem 12 for the free energy, it is necessary to rescale $x(q)$ with β in a suitable way. To this purpose, let us now introduce an arbitrary sequence

$$0 \leq \bar{m}_1 \leq \bar{m}_2 \leq \cdots \leq \bar{m}_K, \tag{6.32}$$

and the corresponding order parameter $\bar{x}$, defined as in (4.14), but with all m_a replaced by $\bar{m}_a$. Notice that there is no upper bound equal to 1 for $\bar{m}_K$, and consequently for $\bar{x}$. However, for sufficiently large β, we definitely have $\bar{m}_K \leq \beta$. Therefore, we can take in (6.31) the order parameter x defined by $x(q) = \bar{x}(q)/\beta$, with $0 \leq x(q) \leq 1$. Then one can easily establish the following Lemma.

Lemma 6 *[23] In the limit $\beta \to \infty$ we have*

$$\lim_{\beta\to\infty} \beta^{-1}\bar{\alpha}(\beta, h; x) = \hat{\alpha}(h; \bar{x}) \equiv \bar{f}(0, h; \bar{x}) - \frac{1}{2}\int_0^1 q\,\bar{x}(q)\,dq, \tag{6.33}$$

where the function $\bar{f}$, with values $\bar{f}(q, y; \bar{x})$ satisfies the antiparabolic equation

$$(\partial_q \bar{f})(q, y) + \frac{1}{2}\big(\bar{f}''(q, y) + \bar{x}(q)\bar{f}'^2(q, y)\big) = 0, \tag{6.34}$$

with final condition

$$\bar{f}(1, y) = |y|. \tag{6.35}$$

The proof follows easily from the definition (4.20) of $\bar{\alpha}(\beta, h; x)$. In fact, the recursive solution for f coming from (4.18), allows to prove immediately

$$\lim_{\beta\to\infty} \beta^{-1} f(q, y; \bar{x}/\beta) = \bar{f}(q, y; \bar{x}),$$

by taking into account the elementary property $\lim_{\beta\to\infty} \beta^{-1} \ln\cosh(\beta y) = |y|$. The identification of the second term in $\hat{\alpha}(h; \bar{x})$ is obvious from the relation $\bar{x}(q) = \beta\, x(q)$. $\square$

Therefore we have established

Theorem 15 *[23] The following inequalities hold*

$$\begin{aligned} e_N(h) &\leq \hat{\alpha}(h; \bar{x}), & (6.37)\\ e_N(h) &\leq \hat{\alpha}(h) \equiv \inf_{\bar{x}} \hat{\alpha}(h; \bar{x}), & (6.38)\\ \lim_{N\to\infty} e_N(h) \equiv e_0(h) &\leq \hat{\alpha}(h; \bar{x}), & (6.39)\\ e_0(h) &\leq \hat{\alpha}(h). & (6.40) \end{aligned}$$

A detailed study of the numerical information coming from the variational bound of Theorem 15 will be presented in a forthcoming paper [71].

6.4 Broken replica symmetry bounds in the p-spin model

The bounds developed in the previous sections for the Sherrington-Kirkpatrick model can be easily extended [23] to the Derrida p-spin model introduced in Chapter 3, and defined by the Hamiltonian

$$H_N(\sigma, h; J) = -\sqrt{\frac{p!}{2N^{p-1}}} \sum_{1\le i_1<\cdots<i_p\le N} J_{i_1\ldots i_p}\sigma_{i_1}\ldots\sigma_{i_p} - h\sum_i \sigma_i. \tag{6.41}$$

In order to motivate the following results, we give a very short sketch of Parisi theory for the p-spin model, along the same lines we followed in Chapter 4 for the Sherrington-Kirkpatrick one. For a more satisfactory treatment, we refer especially to Refs. [45] and [46]. The general structure of the theory is closely analogous to that we illustrated for the Sherrington-Kirkpatrick model, but the predicted phase diagram is quite different.

Piecewise constant order parameters are introduced as in (4.12), (4.14), but now we assume $q_K = p/2$. We still introduce the function $f(q, y; x, \beta)$, defined by the antiparabolic equation (4.16), but in this case q ranges between 0 and $p/2$, and the final condition is

$$f(p/2, y) = \ln\cosh(\beta y). \tag{6.42}$$

We also introduce the change of variables $q \to \bar{q}$, defined by $2q = p\,\bar{q}^{p-1}$, so that, in particular, $\bar{q}_K \le 1$. The definitions (4.20) and (4.21) must be modified as follows.

Definition 3 *The trial auxiliary function, associated to a given p-spin mean field spin glass system, depending on the functional order parameter x, is defined as*

$$\bar{\alpha}^{(p)}(\beta, h; x) \equiv \ln 2 + f(0, h; x, \beta) - \frac{\beta^2}{2}\int_0^{\frac{p}{2}} \bar{q}(q)\, x(q)\, dq. \tag{6.43}$$

Definition 4 *The spontaneously broken replica symmetry solution for the p-spin model is defined by*

$$\bar{\alpha}^{(p)}(\beta, h) \equiv \inf_x \bar{\alpha}^{(p)}(\beta, h; x), \tag{6.44}$$

where the infimum is taken with respect to all functional order parameters x.

The prediction of Parisi theory, in this case, is that

$$-\beta f^{(p)}(\beta, h) = \lim_{N\to\infty} \frac{1}{N} E \ln Z_N^{(p)}(\beta, h; J) = \bar{\alpha}^{(p)}(\beta, h). \tag{6.45}$$

Moreover, the optimal order parameter where the infimum in (6.44) is realized is related to the infinite volume overlap probability distribution by‡

$$x(q(\bar{q})) = \int_0^{\bar{q}} P(q')dq'. \tag{6.47}$$

‡When the magnetic field is zero and p is even, the system is spin-flip invariant and $P(q)$ is symmetric and Eq. (6.47) is replaced by

$$x(q(\bar{q})) = 2\int_0^{\bar{q}} P(q')dq'. \tag{6.46}$$

See also the note on page 27.

Notwithstanding the strict analogy with the Sherrington-Kirkpatrick case, the physical behavior and the phase diagram of the p-spin model are quite different. While we refer to the beautiful paper [46] for all the details, we want to outline here the main results. For simplicity, we suppose that the magnetic filed h is infinitesimal, but strong enough to break spin flip symmetry when p is even. For any integer $p \geq 3$, there are two relevant critical temperatures $1/\beta_c^{(1)}(p) > 1/\beta_c^{(2)}(p)$.

- For $\beta < \beta_c^{(1)}(p)$, the annealed approximation holds, *i.e.*,

$$\lim_{N\to\infty} \frac{1}{N} E \ln Z_N^{(p)}(\beta; J) = \lim_{N\to\infty} \frac{1}{N} \ln E Z_N^{(p)}(\beta; J) = \ln 2 + \frac{\beta^2}{4}. \tag{6.48}$$

  In this case, replica symmetry is not broken and the infimum in (6.44) is assumed in correspondence of a functional order parameter of the simplest form

$$x(q) = 1 \ \text{ for } \ 0 \leq q \leq \frac{p}{2},$$

  so that the overlap does not fluctuate,

$$P(q) = \delta(q)$$

  as it is easily seen from (6.47).

- for $\beta_c^{(1)}(p) < \beta < \beta_c^{(2)}(p)$, the overlap assumes *two* different values, and one says that replica symmetry is broken at one step. In other words, the optimizing functional order parameter has the form

$$\begin{aligned} x(q) &= m \quad \text{for} \quad 0 \leq q \leq q_1 \\ x(q) &= 1 \quad \text{for} \quad q_1 < q \leq p/2, \end{aligned} \tag{6.49}$$

  with $0 < m < 1$. As a consequence,

$$P(q) = m\delta(q) + (1-m)\delta(q - \bar{q}_1),$$

  where $2q_1 = p\,\bar{q}_1^{p-1}$. The values of q_1 and m are determined by two coupled equations, arising from the minimization in (6.44).

- for $\beta > \beta_c^{(2)}(p)$, replica symmetry is broken at infinite steps, and $P(q)$ has support on a whole interval, as it happens for the Sherrington-Kirkpatrick model, below the Almeida-Thouless line.

A similar picture holds in presence of a non-vanishing magnetic field. In this case, however, for $\beta < \beta_c^{(1)}(h, p)$ annealing does not hold, even if the system is replica symmetric.

For the Sherrington-Kirkpatrick model ($p = 2$) the two critical temperatures coincide, and there is no region where replica symmetry is broken just at one step. On the other hand, one can obtain an expansion of the critical temperatures for p going to infinity, with the result that

$$\beta_c^{(1)}(p) \to 2\sqrt{\ln 2}$$

and

$$\beta_c^{(2)}(p) \to \infty.$$

In other words, the phase with continuous replica symmetry breaking disappears for p going to infinity and one-step breaking holds in the whole low-temperature region. Asymptotic expansions for p large may be found in [46]. It is interesting to verify that, in this limit, one recovers the behavior of the Derrida Random Energy Model [38], where one has just one critical value for β, which equals $2\sqrt{\ln 2}$.

Now, we go back to the problem of giving rigorous lower bounds to the free energy per site of the p-spin model, involving the Parisi solution (6.44). In the following we consider only the case of even p, for reasons that will be clarified in the following. With the same procedure as described in Section 6.2, we arrive at the sum rule given by

Theorem 16 *[23] In the p-spin model, for any functional order parameter, the following sum rule holds*

$$\begin{aligned}\bar{\alpha}^{(p)}(\beta, h; x) &= \frac{1}{N} E \ln Z_N^{(p)}(\beta, h; J) \\ &+ \frac{\beta^2}{4} \sum_{a=0}^{K} (m_{a+1} - m_a) \int_0^1 \langle q_{12}^p - p\, q_{12} \bar{q}_a^{p-1} + (p-1) \bar{q}_a^p \rangle_a(t)\, dt \\ &+ O(1/N), \end{aligned} \tag{6.50}$$

where $\bar{\alpha}^{(p)}(\beta, h; x)$ is defined in (6.43).

Notice that the terms in the sum are still positive, for p even, since the minimum over $\mathbb{R}$ of the function $x \to x^p - p\, x\, \bar{q}^{p-1}$ is just $(1-p)\bar{q}^p$. The $O(1/N)$ correction in (6.50), which is absent in the Sherrington-Kirkpatrick counterpart (6.19), is typical of the p-spin model.

From the sum rule we have also

Theorem 17 *[23] In the p-spin model with p even, for any functional order parameter x, the following bound holds*

$$-\frac{1}{N\beta} E \ln Z_N^{(p)}(\beta, h; J) \geq -\frac{1}{\beta} \bar{\alpha}^{(p)}(\beta, h; x) + O(1/N),$$

where $\bar{\alpha}^{(p)}(\beta, h; x)$ is defined in (6.43). Consequently, we have also

$$-\frac{1}{N\beta} E \ln Z_N^{(p)}(\beta, h; J) \geq -\frac{1}{\beta} \bar{\alpha}^{(p)}(\beta, h) + O(1/N),$$

where $\bar{\alpha}^{(p)}(\beta, h)$ is defined in (6.44). Moreover, for the thermodynamic limit, we have

$$-\lim_{N\to\infty} \frac{1}{N\beta} E \ln Z_N^{(p)}(\beta, h; J) \equiv f^{(p)}(\beta, h) \geq -\frac{1}{\beta} \bar{\alpha}^{(p)}(\beta, h),$$

and

$$-\lim_{N\to\infty} \frac{1}{N\beta} \ln Z_N^{(p)}(\beta, h; J) = f^{(p)}(\beta, h) \geq -\frac{1}{\beta} \bar{\alpha}^{(p)}(\beta, h),$$

J-almost surely.

In the case of odd p, the sum rule (6.50) still holds, but it is not obvious that the source terms are positive. A sufficient condition for this to be true is that

$$-1 + p\bar{q}_a^{p-1} + (p-1)\bar{q}_a^p \geq 0 \quad \forall a, \tag{6.51}$$

where $\bar{q}_a$ corresponds to the infimum in (6.44). Checking this property, while conceptually straightforward, is in practice difficult, since the Parisi solution is not given in an explicit form, but rather as the solution of a variational problem. We refer to the forthcoming paper [72] for a more detailed treatment.

6.5 Replica Symmetry Breaking below the Almeida-Thouless line

As a very simple yet important consequence of Guerra's broken replica symmetry bounds for the Sherrington-Kirkpatrick model, expressed by Theorem 12, we prove that the replica symmetric solution f_{SK} does not hold in the whole region below the Almeida-Thouless line. That the replica symmetric free energy cannot be the infinite volume free energy of the system, for very low temperature, was already proved by Sherrington and Kirkpatrick in [16]. Indeed, consider the average entropy per site

$$s_N(\beta, h) = \beta^2 \frac{\partial}{\partial \beta} f_N(\beta, h), \tag{6.52}$$

which for a lattice system is non-negative by definition. In fact, it is easily seen that

$$s_N(\beta, h) = -\frac{1}{N} E \sum_{\{\sigma\}} \rho_N(\sigma, \beta, h; J) \ln \rho_N(\sigma, \beta, h; J), \tag{6.53}$$

where $\rho_N(\sigma, \beta, h; J)$ is the Boltzmann weight

$$0 \leq \rho_N(\sigma, \beta, h; J) = Z_N(\beta, h)^{-1} \exp(-\beta H_N(\sigma, \beta, h; J)) \leq 1, \tag{6.54}$$

and

$$-x \ln x \geq 0$$

for $0 \leq x \leq 1$. On the other hand, if one computes the entropy per site of the replica symmetric free energy f_{SK}, using Eq. (6.52), one easily finds out that $s(\beta, h)$ is negative, for β sufficiently large. However, this does not happen in the whole region below the Almeida-Thouless line, while one expects replica symmetry to be broken for all thermodynamic parameters satisfying

$$\beta^2 \int \frac{1}{\cosh^4(\beta h + \beta z \sqrt{\bar{q}(\beta, h)})} d\mu(z) > 1. \tag{6.55}$$

This belief is based on an analysis of the stability of the replica symmetric saddle point in the functional integral (4.7), performed by de Almeida and Thouless in [54]. In the present section we prove the following result:

Theorem 18 *[73] For any value of β and h satisfying the condition (6.55),*

$$f(\beta,h) \equiv \lim_{N\to\infty} f_N(\beta,h) > f_{SK}(\beta,h), \tag{6.56}$$

where $f_{SK}(\beta,h) = -\beta^{-1}\alpha_{SK}(\beta,h)$ is the Sherrington-Kirkpatrick replica symmetric solution, defined in Eq. (4.28).

Proof) Recall first of all Guerra' bounds

$$f_N(\beta,h) \geq -\beta^{-1} \inf_x \bar{\alpha}(\beta,h;x), \tag{6.57}$$

where $\bar{\alpha}(\beta,h;x)$ is defined as

$$\bar{\alpha}(\beta,h;x) = \ln 2 + f(0,h;x,\beta) - \frac{\beta^2}{2}\int_0^1 q\, x(q)\, dq. \tag{6.58}$$

The infimum is taken over the space of functional order parameters x, and $f(q,y;x,\beta)$ satisfies the antiparabolic equation

$$(\partial_q f)(q,y) + \frac{1}{2}\big(f''(q,y) + x(q) f'^2(q,y)\big) = 0, \tag{6.59}$$

with final condition

$$f(1,y) = \ln\cosh(\beta y). \tag{6.60}$$

In order to prove the Theorem 18, it suffices to show that, if (6.55) holds, there exists a functional order parameter $\tilde{x}$ such that $\bar{\alpha}(\beta,h;\tilde{x})$ is *strictly smaller* than $\alpha_{SK}(\beta,h)$. As discussed in Section 4.3, if one takes

$$\begin{cases} x(q) = 0 & q \in [0,\bar{q}] \\ x(q) = 1 & q \in (\bar{q},1], \end{cases} \tag{6.61}$$

one finds that $\bar{\alpha}(\beta,h;x)$ is just $\alpha_{SK}(\beta,h)$. Therefore, we slightly deform the simple one step functional order parameter (6.61), and choose

$$\begin{cases} \tilde{x}(q') = 0 & q' \in [0,\bar{q}] \\ \tilde{x}(q') = m & q' \in (\bar{q},q] \\ \tilde{x}(q') = 1 & q' \in (q,1], \end{cases} \tag{6.62}$$

where $0 \leq m \leq 1$ and $\bar{q} \leq q \leq 1$. We denote with $\bar{\alpha}(\beta,h;m,q)$ the corresponding Parisi function $\bar{\alpha}(\beta,h;\tilde{x})$. Of course, since $\bar{\alpha}(\beta,h;1,q) = \alpha_{SK}(\beta,h)$, it is sufficient to prove that

$$\partial_m \bar{\alpha}(\beta,h;m,q)|_{m=1} > 0,$$

for some $q > \bar{q}$. First of all, $\bar{\alpha}(\beta,h;m,q)$ is easily found to be

$$\bar{\alpha}(\beta,h;m,q) = \ln 2 + \frac{\beta^2}{2}(1-q) - \frac{\beta^2}{4}(1-q^2+m(q^2-\bar{q}^2)) + \tag{6.63}$$

$$+\frac{1}{m}\int d\mu(z')\ln\int d\mu(z)\cosh^m(\beta h + \beta z\sqrt{q-\bar{q}} + \beta z'\sqrt{\bar{q}}). \tag{6.64}$$

In fact, it is easy to integrate iteratively the antiparabolic equation with the help of (4.18), the order parameter being piecewise constant. Next, we compute the derivative with respect to m, keeping q fixed, and we find

$$\partial_m \bar{\alpha}(\beta, h; m, q)|_{m=1} = K(\beta, h; q) \equiv \tag{6.65}$$
$$\equiv -\frac{\beta^2}{4}(q^2 - \bar{q}^2) - \int d\mu(z') \ln \int d\mu(z) \cosh(\beta h + \beta z\sqrt{q-\bar{q}} + \beta z'\sqrt{\bar{q}})$$
$$+ \int d\mu(z') \frac{\int d\mu(z) \cosh(\beta h + \beta z\sqrt{q-\bar{q}} + \beta z'\sqrt{\bar{q}}) \ln\cosh(\beta h + \beta z\sqrt{q-\bar{q}} + \beta z'\sqrt{\bar{q}})}{\int d\mu(z) \cosh(\beta h + \beta z\sqrt{q-\bar{q}} + \beta z'\sqrt{\bar{q}})}.$$

It is clear that, for $q \downarrow \bar{q}$, the integration over z disappears, and

$$K(\beta, h; \bar{q}) = 0.$$

Therefore, in order to check the sign of $K(\beta, h; \bar{q})$, we have to expand around $q = \bar{q}$. By performing the first two derivatives with respect to q, one finds

$$\partial_q K(\beta, h; q)|_{q=\bar{q}} = 0 \tag{6.66}$$

and

$$\partial_q^2 K(\beta, h; q)\big|_{q=\bar{q}} = -\frac{\beta^2}{2}\left(1 - \beta^2 \int d\mu(z) \frac{1}{\cosh^4(z\beta\sqrt{\bar{q}} + \beta h)}\right). \tag{6.67}$$

This computation requires a simple integration by parts on a Gaussian variable, and is sketched below. It is clear that, when condition (6.55) holds, $\partial_q^2 K(\beta, h; q)\big|_{q=\bar{q}} > 0$, so that

$$\partial_m \bar{\alpha}(\beta, h; m, q)|_{m=1} > 0,$$

at least for $q - \bar{q}$ small. This completes the proof of the result (6.56), *i.e.*, of the instability of the replica symmetric solution, in the whole region below the Almeida-Thouless line.

For completeness, we compute here the first two derivatives of $K(\beta, h; q)$ with respect to q, and we prove Eqs. (6.66), (6.67). Defining for simplicity

$$\phi_{z,z'} = \beta h + \beta z\sqrt{q-\bar{q}} + \beta z'\sqrt{\bar{q}},$$

one finds

$$\begin{aligned}\partial_q K(\beta, h; q) \quad = \quad & -\frac{\beta^2}{2}q + \frac{\beta}{2\sqrt{q-\bar{q}}} \int d\mu(z') \frac{\int d\mu(z) z \sinh(\phi_{z,z'}) \ln\cosh(\phi_{z,z'})}{\int d\mu(z) \cosh(\phi_{z,z'})} \\ & -\frac{\beta}{2\sqrt{q-\bar{q}}} \int d\mu(z') \frac{\int d\mu(z) \cosh(\phi_{z,z'}) \ln\cosh(\phi_{z,z'}) \int d\mu(z) z \sinh(\phi_{z,z'})}{(\int d\mu(z) \cosh(\phi_{z,z'}))^2}.\end{aligned}$$

At this point, one employs the usual integration by parts formula

$$\int d\mu(z) z\, g(z) = \int d\mu(z) g'(z),$$

which holds for any centered unit Gaussian variable z and smooth function g, obtaining

$$\partial_q K(\beta, h; q) = -\frac{\beta^2}{2}q + \frac{\beta^2}{2} \int d\mu(z') \frac{\int d\mu(z) \sinh(\phi_{z,z'}) \tanh(\phi_{z,z'})}{\int d\mu(z) \cosh(\phi_{z,z'})}. \tag{6.68}$$

Recalling the definitions of $\phi_{z,z'}$ and $\bar{q}$, Eq. (6.66) is then easily proved. As for the second derivative of $K(\beta,h;q)$, one proceeds in a similar way, finding

$$\begin{aligned}\partial_q^2 K(\beta,h;q) &= -\frac{\beta^2}{2} - \frac{\beta^3}{4\sqrt{q-\bar{q}}}\int d\mu(z')\frac{\int d\mu(z)\sinh(\phi_{z,z'})\tanh(\phi_{z,z'})\int d\mu(z) z\sinh(\phi_{z,z'})}{(\int d\mu(z)\cosh(\phi_{z,z'}))^2}\\ &+\frac{\beta^3}{4\sqrt{q-\bar{q}}}\int d\mu(z')\frac{\int d\mu(z) z\sinh(\phi_{z,z'})(1+\cosh^{-2}(\phi_{z,z'}))}{\int d\mu(z)\cosh(\phi_{z,z'})}\end{aligned}$$

which, after integration by parts, gives Eq. (6.67).

□

Strictly speaking, Theorem 18 shows only that, below the Almeida-Thouless line, the true infinite volume free energy is higher than its replica-symmetric approximation, and not that replica symmetry is broken there, in the sense that the mean square fluctuations of the overlap are finite in the thermodynamic limit. However, this can be easily deduced from Eq. (6.56) by means of the cavity method, as proven originally by Pastur and Scherbina [64], and later in a simpler way by M. Talagrand (see [9] and the forthcoming book [11]). The proof proceeds by contradiction, and one shows that, if the overlap self-averages, then necessarily $f_N(\beta,h) \to f_{SK}(\beta,h)$, which is not possible thanks to inequality (6.56).

Chapter 7

Quadratic replica coupling and the high temperature region

7.1 Introduction

In the previous chapter, by means of rigorous sum rules involving overlap fluctuations we showed that the free energy per site is bounded below by the Parisi replica symmetry breaking solution $-\beta^{-1}\bar{\alpha}(\beta,h)$ defined in (4.21), for a wide class of mean field spin glass models. Of course, since one expects Parisi solution to be the true infinite volume limit for the free energy, one would like to prove also *upper bounds* of the type

$$f_N(\beta,h) \leq -\beta^{-1}\bar{\alpha}(\beta,h) + o(1),$$

where $o(1)$ vanishes in the thermodynamic limit. So far, this task has been achieved only for high enough temperature, in regions where replica symmetry is not broken. In the present chapter we are going to report recent progress on this point.

In the replica symmetric region, where the overlaps self-average, the system is considered as physically trivial (although it is mathematically quite hard), since the infinite volume equilibrium state is a product state, which means that spins on different sites are uncorrelated random variables. The same happens for instance also for the Curie-Weiss model in the high temperature region. The factorization of the correlation functions is a direct consequence of self-averaging of the overlaps. Indeed, one has

$$\lim_{N\to\infty} \left\langle (q_{13}-q_{14})^2 (q_{23}-q_{24})^2 \right\rangle = 0 \tag{7.1}$$

and, exploiting symmetry between sites, this is easily seen [6], [9] to be equivalent to

$$\lim_{N\to\infty} E\left(\omega_J(\sigma_i\sigma_j) - \omega_J(\sigma_i)\omega_J(\sigma_j)\right)^2 = 0 \tag{7.2}$$

for any $i \neq j$.

In the following of this chapter, we restrict ourselves for simplicity to the Sherrington-Kirkpatrick model. As we already stated in Chapter 4, one expects the replica symmetric region to be determined by the condition

$$\beta^2 \int \frac{1}{\cosh^4(\beta h + \beta z\sqrt{\bar{q}(\beta,h)})}\, d\mu(z) \leq 1, \tag{7.3}$$

whose boundary is just the Almeida-Thouless line (4.30) we discussed above. In this region, the fluctuations of the overlap should vanish in the infinite volume limit, and the free energy should converge to the replica symmetric expression, originally found by Sherrington and Kirkpatrick in their seminal paper [16] and defined in Eq. (4.28):

$$-\beta f_N(\beta, h) \to \alpha_{SK}(\beta, h) \ \text{ for } \ N \to \infty. \tag{7.4}$$

In correspondence of the Almeida-Thouless line a phase transition should occur and one expects it to be of second order [6]. The probability distribution $P(q)$ of the overlap, which is a delta function peaked at $q = \bar{q}$ above* the critical line, has a finite width below it. By "second order transition" we mean the following: when the Almeida-Thouless line is approached from the low temperature phase, $P(q)$ becomes more and more peaked in correspondence of the replica symmetric value $\bar{q}$, with a finite but small dispersion around it. This is in contrast with the case of the p-spin model, where a first order transition occurs and, as soon as a suitably defined critical line is crossed, $P(q)$ *is the sum of two delta functions located at finite distance one from the other [6].* The computation of de Almeida and Thouless is based on the observation that, when inequality (7.3) does not hold, the replica-symmetric saddle point in the functional integral (4.7) becomes unstable, *i.e.*, some of the eigenvalues of the corresponding stability matrix have the wrong sign. It is clear from their approach why replica symmetry breaking is expected to be of second order. Indeed, the eigenvalues of the Hessian matrix are continuous with respect to the thermodynamic parameters, and change sign just at the critical line. Therefore, for T slightly below the critical temperature, the saddle point which gives the dominant contribution to the functional integral is close to the replica-symmetric one and moves away smoothly from it, as temperature is lowered. This is in contrast with the situation one encounters in the case of first order phase transition where, changing for instance the value of temperature, the new relevant saddle point suddenly appears far from the old one, which is still locally stable.

From statistical mechanics we expect, in correspondence of a second order phase transition, some derivative of the free energy to be singular, denoting divergence of physical quantities like specific heat or magnetic susceptibility. However, except for $h = 0$, the replica symmetric solution is analytic at the transition line, as it is easy to verify from the definitions (4.25), (4.28) of $\bar{q}$ and α_{SK}. As we shall see in Chapter 8, the forecast of the replica symmetry breaking phase transition from the side of the replica symmetric region is the divergence of the suitably rescaled overlap fluctuations, as the Almeida-Thouless line is approached. This can be equivalently seen as the divergence of the susceptibility with respect to a field coupling two replicas of the system.

As we have already stated, from the mathematical point of view the problem of proving that replica symmetry holds is far from being trivial and Eq. (7.4) has been established so far only for β, h belonging to a subset of the expected region of validity (7.3). Before we illustrate our results, we give a brief summary of previous work by other authors. The best studied case is that of the Sherrington-Kirkpatrick model at zero magnetic field and $\beta < 1$. In this case, not only replica symmetry holds, but the quenched free energy coincides with the annealed one, in the infinite volume limit. This

* "above" and "below" refer to the phase diagram in the (β, h) plane, as in see Fig. 4.1

was first proven (among other more refined results, like a central limit theorem for free energy fluctuations) by Aizenman, Lebowitz and Ruelle in Ref. [65], employing cluster expansion techniques. Later, their results were extended by Comets and Neveu [66], using stochastic calculus, and by Guerra [74], who exploited essentially only convexity of the free energy and positivity properties of fluctuations. Other recent results are those of Talagrand [21], by means of the so called *second moment method*. All these approaches give a quite satisfactory picture of the high temperature phase of the model when the magnetic field is absent, but unfortunately they can not be immediately extended, as far as we know, to the general case where $h \neq 0$. The case with external field was also studied in the literature, for instance by Scherbina [75] and Talagrand (see [9] and references therein), but the control of the entire region (7.3) was not obtained.

A very innovative idea was introduced recently by M. Talagrand [76], [77], based on the very sound physical idea that the spontaneous replica symmetry breaking phenomenon can be understood by exploring the properties of the model, under the application of auxiliary interactions, which explicitly break permutation symmetry among replicas. This is somewhat similar to the idea of studying ordinary spontaneous symmetry breaking by means of external fields which break the symmetry, as discussed at the beginning of Chapter 4. In particular, Talagrand considered two replicas of the system, coupled by a term depending linearly on the mutual overlap, and was able to prove Eq. (7.4) in a subset of the entire region above the Almeida-Thouless line. The importance of the result in [77] is not just that it allows to prove replica symmetry in a larger region with respect to previous works, but rather that it shows the power of the replica coupling approach in obtaining a strong and rigorous control of the system. It must be emphasized, however, that the idea of coupling replicas had been already employed for a long time in the theoretical physics literature (see, for instance, [78]), even if it seems to have been somewhat overlooked by mathematical physicists, until [76] appeared.

In a joint work with F. Guerra [79], we extended Talagrand's idea by coupling two replicas of the Sherrington-Kirkpatrick model by means of an interaction term which is just the square of the deviation of the overlap of the two replicas from its replica symmetric value $\bar{q}$. We call this method *quadratic replica coupling*. We believe that our method is quite instructive and paradigmatic, since it is technically very simple and, above all, is connected in a clear way with the sum rules and lower bounds for the free energy we illustrated in the previous chapter. However, it must be said that even our result does not allow to control the whole region above the Almeida-Thouless line. This is still an open problem from the rigorous point of view.

For pedagogical reasons, we start by illustrating our quadratic coupling method in the technically simpler case of zero external field, and then we proceed to the more general situation.

7.2 Quadratic coupling for zero external field

The high temperature region ($\beta < 1$) of the zero external field Sherrington-Kirkpatrick model is a very particular case where everything can be computed. As it is well known, in this case annealing is exact in the infinite volume limit, *i.e.*, the quenched and the annealed free energies coincide for $N \to \infty$[†]. In fact, for $\beta < 1$ we have

$$\lim_{N\to\infty} \frac{1}{N} E \ln Z_N(\beta; J) = \alpha_{SK}(\beta, 0) = \ln 2 + \frac{\beta^2}{4} = \lim_{N\to\infty} \frac{1}{N} \ln E Z_N(\beta; J). \tag{7.6}$$

In this section we give a new proof of Eq. (7.6), based on sum rules for the free energy. Our method is very simple and can be easily extended to the case of nontrivial external field, as shown in the next section.

When $h = 0$ and $\beta < 1$, the only solution of the equation (4.25) defining the Sherrington-Kirkpatrick order parameter is $\bar{q} = 0$. Consider the Boltzmann-Gibbs state defined by the partition function

$$\tilde{Z}_N(\beta, t; J) = \sum_{\{\sigma\}} \exp\left(\beta \sqrt{\frac{t}{N}} \sum_{1\le i<j\le N} J_{ij}\sigma_i\sigma_j\right), \tag{7.7}$$

and define also the auxiliary function[‡]

$$\alpha_N(\beta, t) = \frac{1}{N} E \ln \tilde{Z}_N(\beta, t; J), \tag{7.8}$$

in analogy with Eq. (3.3), related to the free energy $f_N(\beta)$ by

$$\alpha_N(\beta, 1) = \frac{1}{N} E \ln Z_N(\beta; J) = -\beta f_N(\beta). \tag{7.9}$$

As in the previous sections, $0 \le t \le 1$ plays the role of an interpolating parameter. By successive derivation and integration of $\alpha_N(\beta, t)$ with respect to t, one finds the sum rule, closely related to (6.9)

$$\alpha_N(\beta, t) = \ln 2 + \frac{\beta^2}{4} t - \frac{\beta^2}{4} \int_0^t \langle q_{12}^2 \rangle_{t'} \, dt'. \tag{7.10}$$

[†]Recall the definitions (3.5), (3.6) of the quenched and annealed free energies. We say that annealing holds, in the infinite volume limit, if

$$\lim_{N\to\infty} \left(\frac{1}{N} E \ln Z_N - \frac{1}{N} \ln E Z_N \right) = 0. \tag{7.5}$$

Notice that it is only for $h = 0$ and $\beta < 1$ that the annealed and the replica symmetric approximations coincide, since in this case $\bar{q} = 0$:

$$\alpha_{SK}(\beta, 0) = \ln 2 + \frac{\beta^2}{4} = \lim_{N\to\infty} \frac{1}{N} \ln E Z_N.$$

For $h \neq 0$, one has

$$\alpha_{SK}(\beta, h) < \lim_{N\to\infty} \frac{1}{N} \ln E Z_N = \ln 2 + \frac{\beta^2}{4} + \ln\cosh(\beta h).$$

[‡]notice that $\alpha_N(\beta, h)$ in Chapter 3 has the inverse temperature and the magnetic field as arguments, while the α function we define here depends on β and on the interpolating parameter t. To be precise, we should introduce a different symbol for this new auxiliary function, but we think that our choice will not generate confusion, also because $\alpha_N(\beta, t)$ appears only in this section.

The presence of $\langle q_{12}^2 \rangle_{t'}$ in Eq. (7.10), as source of the difference between the replica symmetric (or annealed) free energy and the true one, suggests to couple two replicas with a term proportional to the square of the overlap, the corresponding partition function being

$$\tilde{Z}_N(\beta,\lambda,t;J) = \sum_{\{\sigma^1,\sigma^2\}} \exp\left(\beta\sqrt{\frac{t}{N}} \sum_{1\leq i<j\leq N} J_{ij}(\sigma_i^1\sigma_j^2 + \sigma_i^2\sigma_j^2) + \frac{\lambda}{2}N\beta^2 q_{12}^2\right), \qquad (7.11)$$

with $\lambda \geq 0$. The effect of the added term is to give a larger weight to the configurations having $q_{12} \neq 0$, thus favoring non-selfaveraging of the overlap. Of course, the system is spin-flip invariant also for $\lambda \neq 0$, so that $\langle q_{12} \rangle = 0$. Therefore, if $\langle q_{12}^2 \rangle \neq 0$ then the overlap is not self-averaging. It is important to realize that now, when we replicate the system, we have to take copies of the two coupled replicas. We will denote these copies as $(1,2)$, $(3,4)$, $(5,6)$ and so on, where replica 1 is quadratically coupled with replica 2, replica 3 with 4, and so on but, for instance, 1 is not coupled with 3. The situation is pictorially illustrated in Fig. 7.1. Notice that now full permutation symmetry among replicas is explicitly broken so that, for instance, $\langle q_{12} \rangle \neq \langle q_{13} \rangle$.

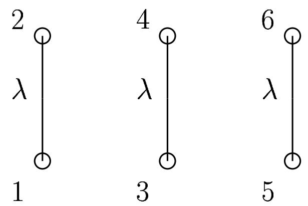

Figure 7.1: Quadratic replica coupling. Circles denote replicas, and replicas joined by a line are coupled.

The basic idea of our method is to show that, as long as $\beta < 1$ and λ is small enough, the term $\lambda N q_{12}^2$ does not change the value of the free energy in the thermodynamic limit. Therefore, "most" configuration must have $q_{12}{=}0$ and the overlap must be self-averaging. In order to implement this intuitive idea, one introduces the λ dependent auxiliary function

$$\tilde{\alpha}_N(\beta,\lambda,t) = \frac{1}{2N} E \ln \tilde{Z}_N(\beta,\lambda,t;J),$$

where the normalization factor 1/2 is chosen so that $\tilde{\alpha}_N(\beta,0,t) = \alpha_N(\beta,t)$. It is interesting to observe that, thanks to the results of Section 5.5.3, the infinite volume limit of $\tilde{\alpha}_N$ exists, for any β,λ,t. Through a simple explicit calculation, we can easily find the t derivative in the form

$$\partial_t \tilde{\alpha}_N(\beta,\lambda,t) = \frac{\beta^2}{4}\left(1 + \langle q_{12}^2 \rangle_{\lambda,t} - 2\langle q_{13}^2 \rangle_{\lambda,t}\right), \qquad (7.12)$$

where now all averages $\langle . \rangle$ involve the λ-dependent state with Boltzmann factor given in agreement with (7.11). Moreover, it is obvious that

$$\partial_\lambda \tilde{\alpha}_N(\beta,\lambda,t) = \frac{\beta^2}{4}\langle q_{12}^2 \rangle_{\lambda,t}.$$

Starting from some point $\lambda_0 > 0$, consider the linear trajectory $\lambda(t) = \lambda_0 - t$, with obvious invertibility in the form $\lambda_0 = \lambda(t) + t$. Let us take the t derivative of $\tilde{\alpha}_N$ along this trajectory

$$\frac{d}{dt}\tilde{\alpha}_N(\beta, \lambda(t), t) = (\partial_t - \partial_\lambda)\,\tilde{\alpha}_N = \frac{\beta^2}{4} - \frac{\beta^2}{2}\langle q_{13}^2\rangle_{\lambda(t),t}.$$

Notice that the term containing $\langle q_{12}^2\rangle$, which would give a contribution with the "wrong sign" in the sum rule, has disappeared. By integration on t we get the sum rule and the inequality

$$\tilde{\alpha}_N(\beta, \lambda, t) = \tilde{\alpha}_N(\beta, \lambda_0, 0) + \frac{\beta^2}{4}t - \frac{\beta^2}{2}\int_0^t \langle q_{13}^2\rangle_{\lambda(t'),t'}\,dt' \leq \tilde{\alpha}_N(\beta, \lambda_0, 0) + \frac{\beta^2}{4}t. \quad (7.13)$$

Next, we compute the initial condition $\tilde{\alpha}_N(\beta, \lambda_0, 0)$. We introduce an auxiliary unit Gaussian random variable z, and perform simple rescaling, in order to obtain

$$\begin{aligned}\tilde{\alpha}_N(\beta, \lambda_0, 0) &= \frac{1}{2N}\ln\sum_{\{\sigma,\sigma'\}} e^{\frac{1}{2}\lambda_0\beta^2 N q_{12}^2} = \frac{1}{2N}\ln\sum_{\{\sigma,\sigma'\}}\int e^{\beta\sqrt{\lambda_0 N}q_{12}z}\,d\mu(z) &(7.14)\\ &= \ln 2 + \frac{1}{2N}\ln\int\left(\cosh z\beta\sqrt{\frac{\lambda_0}{N}}\right)^N d\mu(z)\\ &= \ln 2 + \frac{1}{2N}\ln\int dy\sqrt{\frac{N\lambda_0}{2\pi}}\exp N\left(-\lambda_0\frac{y^2}{2} + \ln\cosh(y\beta\lambda_0)\right), &(7.15)\end{aligned}$$

where we used the identity

$$e^{a^2/2} = \int e^{a\,z}\,d\mu(z) \quad (7.16)$$

in the first step, and we performed the change of variables $z \to y\sqrt{N\lambda_0}$ in the last one. It is immediately recognized that the integral in (7.15) appears in the ordinary treatment of the well known ferromagnetic mean field Curie-Weiss model. The saddle point method

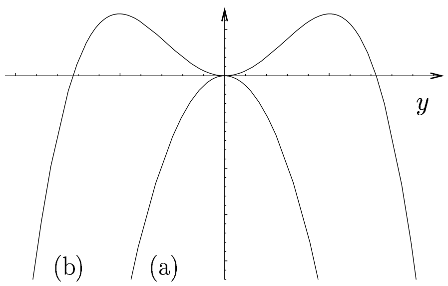

Figure 7.2: Plot of the function at the exponent of Eq. (7.15), for λ_0 smaller (a) or larger (b) than the critical value $\lambda_c = \beta^{-2}$.

gives immediately

$$\lim_{N\to\infty}\tilde{\alpha}_N(\beta, \lambda_0, 0) = \ln 2 + \frac{1}{2}\max_y\left(-\lambda_0\frac{y^2}{2} + \ln\cosh(y\beta\lambda_0)\right). \quad (7.17)$$

The situation is illustrated in Fig. 7.2. The critical value for λ_0 is $\lambda_c = 1/\beta^2$. For $\beta^2\lambda_0 > 1$ we have

$$\lim_{N\to\infty} \tilde{\alpha}_N(\beta, \lambda_0, 0) > \ln 2,$$

while for $\beta^2\lambda_0 < 1$, one can use the elementary property $2\ln\cosh x \le x^2$ to find

$$\ln 2 \le \tilde{\alpha}_N(\beta, \lambda_0, 0) \le \ln 2 + \frac{1}{4N}\ln\frac{1}{1-\lambda_0\beta^2}. \tag{7.18}$$

Notice that, when $\beta^2\lambda_0$ approaches the value 1^-, the term of order $1/N$ diverges, since Gaussian fluctuations around the saddle point become larger and larger.

Thanks to (7.18), the inequality in (7.13) becomes

$$\tilde{\alpha}_N(\beta, \lambda, t) \le \ln 2 + \frac{\beta^2}{4}t + \frac{1}{4N}\ln\frac{1}{1-\lambda_0\beta^2},$$

which holds for $0 \le \beta^2\lambda_0 < 1$, *i.e.*, for $0 \le \beta^2(t+\lambda) < 1$.

Next, we use convexity of $\tilde{\alpha}_N(\beta, \lambda, t)$ with respect to λ and the fact that

$$\partial_\lambda \tilde{\alpha}_N(\beta, \lambda, t)|_{\lambda=0} = \frac{\beta^2}{4}\langle q_{12}^2\rangle_t$$

to write

$$\alpha_N(\beta, t) + \frac{\lambda\beta^2}{4}\langle q_{12}^2\rangle_t \le \tilde{\alpha}_N(\beta, \lambda, t) \le \ln 2 + \frac{\beta^2}{4}t + \frac{1}{4N}\ln\frac{1}{1-(t+\lambda)\beta^2}.$$

Notice that, here, $\langle q_{12}^2\rangle_t$ refers to the system defined by the partition function (7.11), without the interaction between the two replicas. For $0 \le \beta \le \bar{\beta} < 1$ and $0 \le t \le 1$, choose

$$\lambda = \frac{1}{2}(\bar{\beta}^{-2} - 1) > 0,$$

so that

$$\beta^2(t+\lambda) \le \bar{\lambda}_0 \equiv \frac{1}{2}(1+\bar{\beta}^2) < 1$$

and

$$\frac{\beta^2}{4}\langle q_{12}^2\rangle_t \le \frac{1}{\lambda}\left((\ln 2 + \frac{\beta^2}{4}t) - \alpha_N(\beta, t)\right) + \frac{1}{4N\lambda}\ln\frac{1}{1-\bar{\lambda}_0}. \tag{7.19}$$

Recalling Eq. (7.10), one has

$$\frac{d}{dt}\left((\ln 2 + \frac{\beta^2}{4}t) - \alpha_N(\beta, t)\right) = \frac{\beta^2}{4}\langle q_{12}^2\rangle_t \le \frac{1}{\lambda}\left((\ln 2 + \frac{\beta^2}{4}t) - \alpha_N(\beta, t)\right) + \frac{1}{4N\lambda}\ln\frac{1}{1-\bar{\lambda}_0}. \tag{7.20}$$

For $t = 0$, it is easily checked that

$$\alpha_N(\beta, 0) = \ln 2,$$

since in that case the interaction in (7.11) disappears and one is left with a system of N free Ising spins with partition function $Z_N = 2^N$. Then, Eq. (7.20) soon implies

$$\alpha_N(\beta, t) = \ln 2 + \frac{\beta^2}{4}t + O(1/N)$$

uniformly for $0 \leq t \leq 1$ and $0 \leq \beta \leq \bar{\beta} < 1$. In particular, choosing $t = 1$ one has

$$\frac{1}{N} E \ln Z_N(\beta; J) = \alpha_N(\beta, 1) = \ln 2 + \frac{\beta^2}{4} + O(1/N), \tag{7.21}$$

for $\beta < 1$, which proves Eq. (7.6). Of course, from Eq. (7.13) and convexity of $\tilde{\alpha}_N$ one also has

$$\begin{aligned} &\tilde{\alpha}_N(\beta, \lambda, t) = \ln 2 + \frac{\beta^2}{4} t + O(1/N), \\ &\langle q_{13}^2 \rangle_{\lambda,t} = O(1/N) \\ &\langle q_{12}^2 \rangle_{\lambda,t} = O(1/N), \end{aligned}$$

for

$$0 \leq \beta^2 (t + \lambda) \leq \bar{\lambda}_0 < 1. \tag{7.22}$$

As we have stated above, what happens is that for λ sufficiently small the infinite volume free energy does not feel the quadratic coupling, and this implies self-averaging to zero of the overlaps.

We have gained a complete control of the system in the region (7.22). Notice that we have not only proved Eq. (7.6) but we also shown that the leading correction to annealing is of order at most $1/N$.

7.3 The general case

The method we follow to prove replica symmetry in the general case, where an external magnetic field h is present, is a direct generalization of the one explained in the previous section. In the zero external field case, the way how to couple the two replicas was suggested by the sum rule (7.10). In the same way, in the present case we consider the sum rule

$$\alpha_N(\beta, h, t) = \alpha_{SK}(\beta, h, t) - \frac{\beta^2}{4} \int_0^t \left\langle (q_{12} - \bar{q}(\beta, h))^2 \right\rangle_{t'} dt', \tag{7.23}$$

which is soon derived from (6.8) by integration over t. Here, we define

$$\alpha_{SK}(\beta, h, t) = \ln 2 + \frac{\beta^2}{4} t (1 - \bar{q})^2 + \int d\mu(z) \ln \cosh(\beta h + \beta z \sqrt{\bar{q}}). \tag{7.24}$$

Notice that $\alpha_{SK}(\beta, h, 1)$ is just the usual replica symmetric solution $\alpha_{SK}(\beta, h)$, defined in (4.28). Recall that $\alpha_N(\beta, h, t)$ and the average $\langle . \rangle_t$ refer to the auxiliary partition function (6.3), depending on the interpolating parameter $0 \leq t \leq 1$, which corresponds to the situation where the two-body coupling has strength t and the one-body random field has strength $\bar{q}(1-t)$. Moreover, for $t = 1$, α_N is related to the quenched free energy f_N by

$$\alpha_N(\beta, h, 1) = -\beta f_N(\beta, h),$$

so that in this case (7.23) reduces to the sum rule (6.9). In analogy with the zero magnetic field case discussed in the previous section, we are led to introduce the auxiliary function

$$\tilde{\alpha}_N(\beta, h, \lambda, t) = \frac{1}{2N} E \ln \tilde{Z}_N(\beta, h, \lambda, t; J),$$

where $\tilde{Z}_N$ is the partition function for a system of two replicas of the system defined by the partition function (6.3), coupled by the term

$$\frac{\beta^2\lambda}{2}N(q_{12}-\bar{q}(\beta,h))^2,$$

with $\lambda \geq 0$. Now the t derivative is given by

$$\partial_t\tilde{\alpha}_N = \frac{\beta^2}{4}\left(1+\langle q_{12}^2\rangle_{\lambda,t} - 2\langle q_{13}^2\rangle_{\lambda,t}\right) - \frac{\beta^2\bar{q}}{2}(1+\langle q_{12}\rangle_{\lambda,t} - 2\langle q_{13}\rangle_{\lambda,t}),$$

while the λ derivative appears as

$$\partial_\lambda\tilde{\alpha}_N = \frac{\beta^2}{4}\left\langle (q_{12}-\bar{q})^2\right\rangle_{\lambda,t}. \tag{7.25}$$

Starting from points $\lambda_0 > 0$, consider the linear trajectories $\lambda(t) = \lambda_0 - t$, with obvious invertibility as explained before. As for the total t derivative of $\tilde{\alpha}_N$, one finds

$$\frac{d}{dt}\tilde{\alpha}_N(\beta,h,\lambda(t),t) = (\partial_t - \partial_\lambda)\,\tilde{\alpha}_N = \frac{\beta^2}{4}(1-\bar{q})^2 - \frac{\beta^2}{2}\left\langle (q_{13}-\bar{q})^2\right\rangle_{\lambda,t}.$$

Notice that in this case the term containing $\langle (q_{12}-\bar{q})^2\rangle$ disappeared.

By integration, we get the sum rule and the inequality

$$\begin{aligned}\tilde{\alpha}_N(\beta,h,\lambda,t) &= \tilde{\alpha}_N(\beta,h,\lambda_0,0) + \frac{\beta^2}{4}(1-\bar{q})^2 t - \frac{\beta^2}{2}\int_0^t \left\langle (q_{13}-\bar{q})^2\right\rangle_{\lambda(t'),t'}\,dt' \\ &\leq \tilde{\alpha}_N(\beta,h,\lambda_0,0) + \frac{\beta^2}{4}(1-\bar{q})^2 t,\end{aligned}$$

where $\langle (q_{13}-\bar{q})^2\rangle_{\lambda(t'),t'}$ refers to

$$\lambda(t') = \lambda_0 - t' = \lambda + t - t'.$$

If Ω_J is the product state for two replicas with the original Boltzmann factor associated to (6.3), then we can write

$$\tilde{\alpha}_N(\beta,h,\lambda,t) - \alpha_N(\beta,h,t) = \frac{1}{2N}E\ln\Omega_J\left(\exp\frac{\beta^2\lambda}{2}N(q_{12}-\bar{q})^2\right).$$

Therefore, by exploiting Jensen inequality, we have, for $\lambda \geq 0$,

$$\frac{\beta^2\lambda}{4}\left\langle (q_{12}-\bar{q})^2\right\rangle_{\lambda=0,t} \leq \tilde{\alpha}_N(\beta,h,\lambda,t) - \alpha_N(\beta,h,t).$$

Let us also define, for given values of β, h,

$$0 \leq \Delta_N(\lambda_0) \equiv \tilde{\alpha}_N(\beta,h,\lambda_0,0) - \alpha_N(\beta,h,0) = \frac{1}{2N}E\ln\Omega_J^0\left(\exp\frac{\beta^2\lambda_0}{2}N(q_{12}-\bar{q})^2\right), \tag{7.26}$$

where we have introduced the state Ω_J^0 for two replicas, corresponding to $t = 0$ in (6.3). Notice that Ω_J^0 is a factor state over the sites i, since the two body interaction J_{ij} disappears for $t = 0$. By the definition (6.3), one has

$$\alpha_N(\beta, h, 0) = \ln 2 + \int d\mu(z) \ln \cosh(\beta h + \beta z \sqrt{\bar{q}}).$$

By collecting all our definitions and inequalities we have

$$\frac{\beta^2 \lambda}{4} \left\langle (q_{12} - \bar{q})^2 \right\rangle_t \leq \Delta_N(\lambda_0) + \alpha_{SK}(\beta, h, t) - \alpha_N(\beta, h, t).$$

Let us now introduce the critical value λ_c such that, for any $\lambda_0 \leq \lambda_c$, one has

$$\lim_{N \to \infty} \Delta_N(\lambda_0) = 0.$$

Of course, λ_c depends on β and h, since Δ_N does. Then, by the same reasoning already exploited starting from (7.19), and taking into account (7.23), we obtain the proof of the following

Theorem 19 *For any $0 \leq \lambda + t \leq \lambda_c(\beta, h)$, with $\lambda \geq 0$ (see Fig. 7.3), we have the convergence*

$$\lim_{N \to \infty} \tilde{\alpha}_N(\beta, h, \lambda, t) = \alpha_{SK}(\beta, h, t) \tag{7.27}$$

and

$$\left\langle (q_{12} - \bar{q})^2 \right\rangle_{\lambda,t} \leq O(1/N) \tag{7.28}$$

$$\left\langle (q_{13} - \bar{q})^2 \right\rangle_{\lambda,t} \leq O(1/N). \tag{7.29}$$

In particular, if $\lambda_c(\beta, h) \geq 1$, then we can reach the point $t = 1, \lambda = 0$ and

$$\lim_{N \to \infty} \alpha_N(\beta, h) = \lim_{N \to \infty} \frac{1}{N} E \ln Z_N(\beta, h; J) = \alpha_{SK}(\beta, h). \tag{7.30}$$

For the specification of λ_c, we can easily establish the complete characterization of the Δ_N limit. In fact, the following holds:

Theorem 20 *The infinite volume limit of Δ_N is given by*

$$\lim_{N \to \infty} \Delta_N(\lambda_0) = \Delta(\lambda_0).$$

Here, $\Delta(\lambda_0)$ is defined through the variational expression

$$\Delta(\lambda_0) \equiv \frac{1}{2} \max_{\nu} \left(\int \ln(\cosh \nu + \tanh^2(\beta h + \beta z \sqrt{\bar{q}}) \sinh \nu) d\mu(z) - \nu \bar{q} - \frac{\nu^2}{2 \lambda_0 \beta^2} \right), \tag{7.31}$$

where $d\mu(z)$ is the centered unit Gaussian measure.

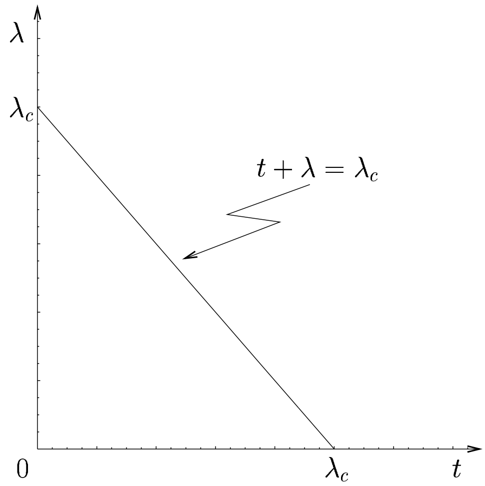

Figure 7.3: The triangular region in the (t, λ) plane, where we prove replica symmetry for the two-replica system. The value of λ_c depends on β and h.

It is easy to check that the expression (7.31) is in agreement with (7.17), when there are no external fields.

The proof of the Theorem 20 is easy. First of all let us establish the elementary bound, uniform in N,

$$\Delta_N(\lambda_0) \geq \Delta(\lambda_0). \tag{7.32}$$

In fact, starting from the definition of $\Delta_N(x_0, \lambda_0)$ given in (7.26), we can write, for $\lambda_0 \neq 0$ and any ν,

$$(q_{12} - \bar{q})^2 \geq 2\frac{\nu}{\beta^2\lambda_0}(q_{12} - \bar{q}) - \left(\frac{\nu}{\beta^2\lambda_0}\right)^2,$$

and conclude that

$$\Delta_N(\lambda_0) \geq \alpha_0(\nu) - \frac{\nu^2}{4\beta^2\lambda_0}, \tag{7.33}$$

where we have defined

$$\begin{aligned} \alpha_0(\nu) &\equiv \frac{1}{2N} E \ln \Omega_J^0 \left(\exp \nu N(q_{12} - \bar{q})\right) \\ &= \frac{1}{2}\int \ln\left(\cosh\nu + \tanh^2(\beta h + \beta z\sqrt{\bar{q}})\sinh\nu\right)\, d\mu(z) - \frac{1}{2}\nu\bar{q}. \end{aligned}$$

Of course, it is convenient to take the maximum over ν in the right hand side of (7.33), so that the bound in (7.32) is established. The proof that the bound is in effect the limit, as $N \to \infty$, can be obtained in a very simple way by using a Gaussian transformation on (7.26), as it was done in (7.14). In fact, now we have

$$\frac{1}{2N} E \ln \Omega_J^0 \left(\exp \frac{\beta^2\lambda_0}{2} N(q_{12} - \bar{q})^2\right) = \frac{1}{2N} E \ln \int \Omega_J^0 \left(\exp \beta\sqrt{\lambda_0 N}(q_{12} - \bar{q})z\right)\, d\mu(z).$$

Therefore, by exploiting the fact that also Ω^0_J factorizes with respect to the sites i, we can write

$$\begin{aligned}\Delta_N(\lambda_0) &= \frac{1}{2N} E \ln \int \prod_i \left(\cosh \beta \sqrt{\frac{\lambda_0}{N}} z + \tanh^2(\beta h + \beta J_i \sqrt{\bar{q}}) \sinh \beta \sqrt{\frac{\lambda_0}{N}} z \right) \\ &\quad \times \exp\left(-\beta \sqrt{\lambda_0 N} \bar{q} z \right) \, d\mu(z). \end{aligned} \tag{7.34}$$

Now, we find convenient to introduce a small $\epsilon > 0$, so that

$$\frac{1}{\lambda_0} = \frac{1}{\lambda'_0} + \epsilon. \tag{7.35}$$

Notice that $\lambda_0 < \lambda'_0$. We also introduce the auxiliary (random) function

$$\phi_N(y, \lambda'_0) \equiv \frac{1}{N} \sum_i \ln \left(\cosh y + \tanh^2(\beta h + \beta J_i \sqrt{\bar{q}}) \sinh y \right) - \bar{q} y - \frac{1}{2} \frac{y^2}{\beta^2 \lambda'_0}.$$

By the strong law of large numbers [13], as $N \to \infty$, for any y, we have the J almost sure convergence of $\phi_N(y, \lambda'_0)$ to $\phi(y, \lambda'_0)$ defined by

$$\begin{aligned}\phi(y, \lambda'_0) &\equiv \int \ln \left(\cosh y + \tanh^2(\beta h + \beta z \sqrt{\bar{q}}) \sinh y \right) d\mu(z) - \bar{q} y - \frac{1}{2} \frac{y^2}{\beta^2 \lambda'_0} \\ &= E \phi_N(y, \lambda'_0).\end{aligned}$$

Here $d\mu(z)$ performs the averages with respect to the J_i variables. Let us also remark that the convergence is J almost surely uniform for any finite number of values of the variable y. Now we can go back to (7.34), write explicitly the unit Gaussian measure $d\mu(z)$, perform the change of variables $y = z\beta\sqrt{\lambda_0 N^{-1}}$, make the transformation (7.35), take the $\sup_y$ for the ϕ_N, and perform the residual Gaussian integration over y. We end up with the estimate

$$\Delta_N(x_0, \lambda_0) \leq \frac{1}{2} E \sup_y \phi_N(y, \lambda'_0) + \frac{1}{2N} \ln \frac{1}{\sqrt{\lambda_0 \epsilon}}. \tag{7.36}$$

Since the J-dependent $\sup_y$ is reached in some finite interval, for any fixed λ'_0, and the function ϕ_N is continuous with respect to y, with bounded derivatives, we can perform the $\sup_y$ with y running over a finite discrete mesh of values, by tolerating a small error, which becomes smaller and smaller as the mesh interval is made smaller. But in this case the strong law of large numbers allows us to substitute ϕ_N with ϕ, in the infinite volume limit $N \to \infty$. On the other hand, the second term in the right hand side of (7.36) vanishes in the limit. Therefore, we conclude that

$$\limsup_{N \to \infty} \Delta_N(\lambda_0) \leq \frac{1}{2} \sup_y \phi(y, \lambda'_0).$$

From continuity with respect to λ_0 and arbitrariness of ϵ, we can let λ'_0 approach λ_0, and the theorem is proven. □

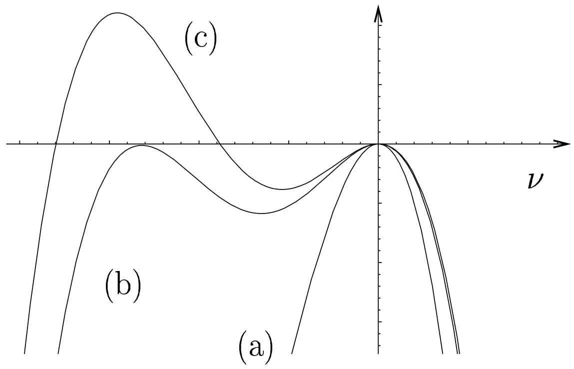

Figure 7.4: Plot of the function whose maximum over ν defines $\Delta(\lambda_0)$, through the variational principle (7.31). The function is asymmetric, for $h \neq 0$. The three curves correspond to: (a) $\lambda_0 < \lambda_c$, (b) $\lambda_0 = \lambda_c$, (c) $\lambda_0 > \lambda_c$. Notice the typical first order character of the transition.

It is easy to realize that the region in the (β, h) plane, where Theorem 19 allows to prove replica symmetry for the Sherrington-Kirkpatrick model, is strictly smaller than the expected high temperature or high external field region defined by the condition (7.3), which arises from the stability analysis of de Almeida and Thouless [54]. This can be seen either studying numerically the maximum in the definition (7.31) of $\Delta(\lambda_0)$, or computing it perturbatively as a power series in β, h in the neighborhood of the point $\beta = 1, h = 0$, where the expression for $\Delta(\lambda_0)$ can be written down explicitly. On the contrary, in the zero external field case our quadratic coupling method allows us to reach the point $\beta = 1$, which is known to be the true critical point of the theory.

Before we conclude this chapter, we wish to discuss the possible physical meaning of the line in the (β, h) plane, which marks the boundary of the region where Theorem 19 gives a proof of replica symmetry. To this purpose, let us summarize the situation in the two quite different situations of zero and non-zero magnetic field, anticipating also some of the results of next chapter.

- For $h = 0$ the result of Theorem 19 is optimal, that is, we prove replica symmetry up to the point $\beta = 1$, beyond which the overlap cannot be self-averaging (see for instance Section 6.5). Moreover, for $\beta < 1$ one finds that the rescaled overlap $\sqrt{N} q_{12}$ behaves as a Gaussian random variable, for N going to infinity, and its variance diverges for $\beta \to 1^-$. This is typical of second-order phase transitions, which are forecasted by the divergence of the rescaled fluctuations of the relevant order parameter (in this case, the overlap).

- On the other hand, for $h \neq 0$ we can prove replica symmetry only up to a critical inverse temperature $\beta_c(h)$ which is *strictly smaller* than the one predicted by de Almeida and Thouless. For $\beta < \beta_c(h)$, the rescaled fluctuations of the overlap around $\bar{q}$, defined as $\sqrt{N}(q_{12} - \bar{q})$, tend also to Gaussian variables, but *their variance does not diverge for β approaching $\beta_c(h)^-$*. If the $\beta_c(h)$ we find is the real transition point of the model, after which replica symmetry is broken (which we do *not* prove), then the transition must be one of first order type, which is not forecasted by a divergence of the fluctuations. It has to be emphasized that the prediction of Almeida and Thouless is based on the analysis of the *local* stability of the replica symmetric saddle point in the functional integral (4.7), and therefore it would not be able, as far as we understand, to detect the occurrence of a discontinuous (first order) transition. However, even if the possibility of a first order phase transition above the Almeida-Thouless line cannot be in principle excluded, all numerical indications and theoretical arguments within Parisi theory seem to show that the true critical line is the Almeida-Thouless one, where overlap fluctuations diverge and the transition is of second order. This would mean that the appearance of $\beta_c(h)$ is just an artifact of our method of proof, without a precise physical meaning.

Chapter 8

Central limit theorems for the fluctuations at high temperature

8.1 Introduction

In the previous chapter, we focused on the Sherrington-Kirkpatrick model and we showed that, in a region of high temperature or high magnetic field, replica symmetry is not broken. In other words, the overlap q_{12} between two different replicas is self-averaging and its typical value is the Sherrington-Kirkpatrick order parameter $\bar{q}$. However, this result gives no sharp information on how fast overlap fluctuations decay with the system size, and about their asymptotic (for large N) probability distribution. A similar question may be posed about the free energy: we know it converges almost surely to the replica symmetric solution, but we would like to investigate more precisely the behavior of the fluctuations around the limit.

Here, we prove that the fluctuations of the free energy and of the overlaps are Gaussian, on the scale $1/\sqrt{N}$, for N large. Namely, we show first of all that the suitably rescaled fluctuations of the disorder dependent free energy around the replica symmetric solution,

$$\hat{f}_N(\beta, h; J) = \sqrt{N}\left(-\frac{1}{N\beta}\ln Z_N(\beta, h; J) - f_{SK}(\beta, h)\right), \tag{8.1}$$

tend to a Gaussian random variable in the thermodynamic limit. The proof is quite simple, and it relies on an interpolation scheme which is similar, in many aspects, to the one we used to prove the existence of the thermodynamic limit and Guerra's bounds in the previous chapters. As a byproduct, we also recover a well known result by Aizenman, Lebowitz, Ruelle [65] and by Comets, Neveu [66] for the free energy fluctuations in the high temperature region of the zero-external field Sherrington-Kirkpatrick model.

As a second result, we show that, defining

$$\xi_{ab}^N = \sqrt{N}(q_{ab} - \bar{q}), \tag{8.2}$$

the random variables ξ_{ab} behave as correlated centered Gaussian variables in the thermodynamic limit, and we compute explicitly the covariance structure. Recall that

$$\xi_{ab}^N = \frac{\sum_{i=1}^N (\sigma_i^a \sigma_i^b - \bar{q})}{\sqrt{N}}, \tag{8.3}$$

and that, owing to the two body interaction $J_{ij}\sigma_i\sigma_j$, the different terms in ξ_{ab}^N are not statistically independent. Therefore, what we are proving is a central limit theorem for the sum of weakly dependent identically distributed random variables. The reason why we speak of "weak dependence" is that, since replica symmetry is not broken, the "cluster property" (7.2) holds. The central limit theorem for the overlap fluctuations ξ_{ab}^N, in the high temperature region, has been independently proved also by M. Talagrand [11], with a somewhat different method. In any case, our proof is based essentially on concentration of measure inequalities for the free energy, and on a rigorous version of the cavity method, previously developed by M. Talagrand himself [76], [9], which is particularly appropriate for the study of the high temperature region.

The Gaussian character of the fluctuations and the scaling $\sqrt{N}$ are of course not surprising. Indeed, from statistical mechanics [1] we expect that, *away from critical points* , intensive physical quantities like magnetization and energy per site have Gaussian fluctuations on the scale $1/\sqrt{|\Lambda|}$, where $|\Lambda|$ is the volume of the system. Moreover, the variance is just given by the susceptibility of the system, with respect to a suitable external field. For instance, for an Ising ferromagnet in a uniform magnetic field h, the "block variable"

$$|\Lambda|^{-1/2} \sum_{i\in\Lambda} (\sigma_i - \langle\sigma_i\rangle)$$

tends to a centered Gaussian random variable, with variance given by the magnetic susceptibility

$$\chi(h) = \frac{\partial m(h)}{\partial h} = \frac{\partial \langle\sigma_1\rangle}{\partial h}.$$

Here, $m(h)$ is the average magnetization, Λ is a finite region of $\mathbb{Z}^d$, and the limit we consider is $\Lambda \to \mathbb{Z}^d$. This has been proved in the past years for a wide class of models, see for instance Refs. [80], [81]. While the methods introduced in those papers are very beautiful and elegant, they don't seem to be applicable to our case, where one has neither informations on the position of the Lee-Yang zeroes of the partition function [82], as required by [80], nor any form of the FKG ferromagnetic correlation inequalities [83], as required in [81].

The present chapter is organized as follows. In Section 8.2 we state our results and comment briefly on them. In Section 8.3 we prove the central limit theorem for the rescaled free energy fluctuations. Then, in Section 8.4 we turn to the central limit theorem for the overlap fluctuations: We begin by illustrating, in Sections 8.4.1 and 8.4.2, the phenomenon of exponential suppression of overlap fluctuations and (Talagrand's version of) the cavity method, respectively, and, in Sections 8.4.3-8.4.4 we prove the limit theorem for the fluctuations of the overlaps. The proof is quite long and technical, so we divide it into intermediate steps, with the hope to make it more readable.

8.2 The main results

First of all, it is convenient to recall shortly some notations and results of Chapter 7. We denote by $\alpha_N(\beta,h,t)$ the auxiliary function

$$\alpha_N(\beta,h,t)=\frac{1}{N}E\ln\tilde{Z}_N(\beta,h,t;J),$$

where $\tilde{Z}_N$ is defined in Eq. (6.3), for $0\le t\le 1$. Recall that

$$\alpha_N(\beta,h,1)=\alpha_N(\beta,h), \tag{8.4}$$

where $\alpha_N(\beta,h)$ was defined in (3.3). With Theorem 19 we proved that, given β and h, there exists a critical value $\lambda_c(\beta,h)$ and a constant C such that, for $t<\lambda_c$,

$$\left\langle (q_{12}-\bar{q})^2\right\rangle_t \le \frac{C}{N} \tag{8.5}$$

and

$$\begin{aligned}\alpha_N(\beta,h,t) &= \alpha_{SK}(\beta,h,t)+O(1/N) \\ &= \ln 2+\frac{\beta^2}{4}t(1-\bar{q})^2+\int d\mu(z)\ln\cosh(\beta h+\beta z\sqrt{\bar{q}})+O(1/N).\end{aligned} \tag{8.6}$$

Let us consider free energy fluctuations first. Aizenman, Lebowitz and Ruelle [65] proved that, in the case of zero external field and $\beta<1$, the variable

$$\ln Z_N(\beta;J)-\ln EZ_N(\beta;J) \tag{8.7}$$

tends in distribution to a non-zero mean Gaussian random variable whose variance diverges at $\beta=1$. In the general case the situation is quite different. Indeed, the following result holds, which refines the convergence (8.6) of the free energy per site:

Theorem 21 *[84] Define the rescaled fluctuations of the disorder dependent free energy as*

$$\hat{f}_N(\beta,h,t;J)\equiv-\frac{\sqrt{N}}{\beta}\left(\frac{\ln\tilde{Z}_N(\beta,h,t;J)}{N}-\alpha_{SK}(\beta,h,t)\right).$$

If $t<\lambda_c(\beta,h)$ then

$$\hat{f}_N(\beta,h,t;J)\xrightarrow{d}\mathcal{N}(0,\sigma^2(\beta,h,t)),$$

where convergence holds in distribution and

$$\sigma^2(\beta,h,t)=\frac{1}{\beta^2}Var\left(\ln\cosh(\beta z\sqrt{\bar{q}}+\beta h)\right)-\frac{\bar{q}^2t}{2}. \tag{8.8}$$

Here, $Var(.)$ denotes the variance of a random variable and z is a standard unit Gaussian variable $\mathcal{N}(0,1)$.

Notice that *fluctuations of the extensive free energy* $\ln Z_N$ *are of order* 1 *at zero external field* (notice the rescaling in (8.7)) *and of order* $\sqrt{N}$ *otherwise.*

Analysis of free energy fluctuations in mean field spin glass models has been performed also by several other authors. For instance, Bovier *et al.* [85] considered the high temperature region of the p-spin model and of its $p \to \infty$ limit, the Random Energy Model. For finite p, they also found a Gaussian behavior for the fluctuations. For the REM, the picture is not so simple and one finds that two different regimes exist. At very high temperature, one has again Gaussian fluctuations, while at lower temperature (but always above the transition temperature where annealing breaks down) there are non-Gaussian fluctuations driven by the Poisson process of the extreme values of the random energies. We refer to [85] for details and further results.

As for the overlap fluctuations, the following theorem* refines the result (8.5):

Theorem 22 *[84] If* $t < \lambda_c(\beta, h)$*, the rescaled overlaps* ξ_{ab}^N *defined in (8.2) tend in distribution, for* $N \to \infty$*, to jointly Gaussian variables* ξ_{ab}*, with covariances*

$$\begin{aligned}\langle \xi_{ab}^2 \rangle &= A(\beta, h, t)\\ \langle \xi_{ab}\xi_{ac} \rangle &= B(\beta, h, t)\\ \langle \xi_{ab}\xi_{cd} \rangle &= C(\beta, h, t),\end{aligned}$$

where the indices a, b, c, d *are all distinct.* A, B *and* C *are explicitly given by*

$$A(\beta, h, t) = (1 + 2R + 4R^2)Y + c_0 R^2 \tag{8.9}$$

$$B(\beta, h, t) = (1 + 4R)RY + c_0 R^2 \tag{8.10}$$

$$C(\beta, h, t) = 4R^2 Y + c_0 R^2, \tag{8.11}$$

where

$$Y(\beta, h, t) = \frac{1}{Y_0^{-1} - \beta^2 t} \tag{8.12}$$

$$R(\beta, h, t) = \frac{d_0}{Y_0^{-1} + 2d_0 - \beta^2 t}$$

and $Y_0(\beta, h), c_0(\beta, h)$ *and* $d_0(\beta, h)$ *are chosen in such a way that* A, B, C *satisfy the initial conditions*

$$\begin{aligned}A(\beta, h, 0) &= 1 - \bar{q}^2\\ B(\beta, h, 0) &= \bar{q} - \bar{q}^2\\ C(\beta, h, 0) &= \int \tanh^4(\beta z\sqrt{\bar{q}} + \beta h)\, d\mu(z) - \bar{q}^2.\end{aligned}$$

In particular, one has

$$Y_0 = \int \cosh^{-4}(\beta z\sqrt{\bar{q}} + \beta h)\, d\mu(z). \tag{8.13}$$

*as we already stated in the introduction to the present chapter, a similar result has been recently proved by M. Talagrand, and will appear in the forthcoming book [11]. Both results are based on a rigorous version of the cavity method, previously developed by Talagrand himself in [76], [9]. In the following, we follow our method of proof.

For $h = 0$, one has the initial conditions $A(\beta,0,0) = 1$, $B(\beta,0,0) = C(\beta,0,0) = 0$, which imply $R = 0$. As a consequence, $\langle \xi_{ab}\xi_{bc}\rangle$ and $\langle \xi_{ab}\xi_{cd}\rangle$ vanish for any t, as it is clear from Eqs. (8.10)-(8.11), and the limit variables ξ_{ab} are independent Gaussian variables with variance

$$\langle \xi_{ab}^2\rangle = Y(\beta,0,t) = 1/(1-t\beta^2), \tag{8.14}$$

which is well known to hold† [65], [66].

The expressions for A, B and C were first given by Guerra in Ref. [53] where he showed that, *provided that* the variables ξ_{ab} are Gaussian in the infinite volume limit, their covariance structure is given by Eqs. (8.9)-(8.11). It is instructive to sketch the proof here. First of all one shows, through a direct long computation, that the averages of the rescaled overlaps satisfy the streaming equation

$$\frac{d}{dt}\langle F\rangle_t = \frac{\beta^2}{2}\left\langle F\left(\sum_{1\le a<b\le s}(\xi_{ab}^N)^2 - s\sum_{a=1}^{s}(\xi_{a,s+1}^N)^2 + \frac{s(s+1)}{2}(\xi_{s+1,s+2}^N)^2\right)\right\rangle_t, \tag{8.15}$$

where F is any smooth function of the ξ_{ab}^N variables, with $a, b \le s$, and the t dependent thermal averages refer, as usual, to the system defined by the partition function (6.3). Notice that the equations (8.15) are in themselves independent of N. Of course, the initial conditions at $t = 0$ do depend on N. It is very easy to control the limit $N \to \infty$ at $t = 0$, by an elementary application of the central limit theorem, since the interaction there factorizes, and the spin variables at each site become independent. Indeed, one finds that at $t = 0$ the rescaled overlaps tend in the thermodynamic limit to centered Gaussian variables with covariances given by

$$\langle \xi_{ab}^2\rangle = A(\beta,h,0) \tag{8.16}$$
$$\langle \xi_{ab}\xi_{ac}\rangle = B(\beta,h,0) \tag{8.17}$$
$$\langle \xi_{ab}\xi_{cd}\rangle = C(\beta,h,0), \tag{8.18}$$

in accord with Theorem 22. Moreover, through a direct calculation, it is possible to show that the centered Gaussian variables ξ_{ab} with covariances defined by (8.9)-(8.11) satisfy both the streaming equations (8.15) and the initial conditions (8.16). Of course, this does not prove that the variables ξ_{ab}^N tend to ξ_{ab} as $N \to \infty$, for $t > 0$. For this to be true, we should prove that the limit $N \to \infty$ and evolution in t commute.

If one assumes that the validity of Theorem 22 can be extended to the entire high temperature region defined by (4.26), then one has a clear characterization of the Almeida-Thouless critical line, as the line where the fluctuations of the rescaled overlap fluctuations diverge. In fact, $Y(\beta,h,t)$ as defined in Eq. (8.12) is singular when $t\beta^2 = Y_0^{-1}$, *i.e.*,

$$t\beta^2\int \cosh^{-4}(\beta z\sqrt{\bar q} + \beta h)\, d\mu(z) = 1, \tag{8.19}$$

†Actually, in order to recover the known result, one has to make the identification $\beta\sqrt{t} \to \beta$ in (7.11). In this case, one finds that the variance of ξ_{ab}^N tends to $1/(1-\beta^2)$ for $\beta < 1$, which was found in Refs. [65], [66].

which is nothing but the definition of the Almeida-Thouless critical line‡. It is interesting to notice that, on the contrary, free energy fluctuations on the scale $1/\sqrt{N}$, as given by Eq. (8.8) in Theorem 21, show no singularity at the same point. Again, this shows the qualitative difference between the Sherrington-Kirkpatrick model with external field and the zero-field one, where the fluctuation of the rescaled overlap and of the (suitably rescaled) free energy (8.7) diverge at the same critical point $\beta = 1$.

One can also easily check from the expressions for A, B, C that, when the Almeida-Thouless line is approached, the following relations hold [53]:

$$\lim \frac{B}{A} = \lim \frac{\langle \xi_{12}^N \xi_{23}^N \rangle}{\langle (\xi_{12}^N)^2 \rangle} = \frac{1}{2} \tag{8.20}$$
$$\lim \frac{C}{A} = \lim \frac{\langle \xi_{12}^N \xi_{34}^N \rangle}{\langle (\xi_{12}^N)^2 \rangle} = \frac{1}{3}.$$

These relations seem to be the forecast, from the side of the replica symmetric region, of the ultrametric equalities [6]

$$\left\langle (q_{12} - \langle q_{12} \rangle)(q_{23} - \langle q_{23} \rangle) \right\rangle = \frac{1}{2} \left\langle (q_{12} - \langle q_{12} \rangle)^2 \right\rangle \tag{8.21}$$
$$\left\langle (q_{12} - \langle q_{12} \rangle)(q_{34} - \langle q_{34} \rangle) \right\rangle = \frac{1}{3} \left\langle (q_{12} - \langle q_{12} \rangle)^2 \right\rangle,$$

which have been proved to hold also in the low temperature region [44, 86, 87].

As a last remark, notice that the estimate (8.5) we proved in Theorem 19 of the previous chapter gives a uniform bound only for the second moment of the ξ variables,

$$\left\langle (\xi_{ab}^N)^2 \right\rangle \leq C.$$

In order to transfer this information into bounds for higher order moments, we need to employ the cavity method, as illustrated in Section 8.4.2.

8.3 Free energy fluctuations

In this section, we prove Theorem 21. As it is well known [13], in order to prove convergence in distribution it is enough to show that the characteristic function of $\hat{f}_N$ converges to that of $\mathcal{N}(0, \sigma^2(\beta, h, t))$, *i.e.*,

$$\lim_{N \to \infty} E\, e^{i\,u\,\hat{f}_N(\beta,h,t;J)} = \exp\left(-\frac{u^2}{2} \sigma^2(\beta, h, t) \right). \tag{8.22}$$

Define for simplicity

$$\hat{f}_N(t) = \hat{f}_N(\beta, h, t; J)$$
$$\alpha_{SK}(t) = \alpha_{SK}(\beta, h, t)$$
$$\zeta_N(t) = \frac{1}{N} \ln Z_N(\beta, h, t; J).$$

‡In fact, recall that the the original Sherrington-Kirkpatrick model defined in Chapter 3, without the one-body random interaction J_i, is recovered when $t = 1$.

The values of β and h will be held fixed in the course of the proof. The characteristic function of $\hat{f}_N$ can be written as

$$E\, e^{i\,u\hat{f}_N(\beta,h,t;J)} \;=\; E\, e^{i\,u\hat{f}_N(\beta,h,0;J)} + i\,u\,E\int_0^t e^{i\,u\hat{f}_N(t')}\,\frac{d}{dt'}\hat{f}_N(t')\,dt'. \qquad (8.23)$$

From definitions (7.24), (6.3) one finds

$$\frac{d}{dt}\alpha_{SK}(t) = \frac{\beta^2}{4}(1-\bar{q})^2$$
$$\frac{d}{dt}\zeta_N(t) = \frac{\beta}{2\sqrt{t}N^{3/2}}\sum_{1\le i<j\le N} J_{ij}\omega_t(\sigma_i\sigma_j) - \frac{\beta}{2N}\sqrt{\frac{\bar{q}}{1-t}}\sum_i J_i\omega_t(\sigma_i),$$

and, through integration by parts on the Gaussian disorder,

$$E\left\{e^{-i\,u\frac{\sqrt{N}}{\beta}\zeta_N(t)}\frac{d}{dt}\zeta_N(t)\right\} \;=\; \frac{\beta^2}{4}E\left\{e^{-i\,u\frac{\sqrt{N}}{\beta}\zeta_N(t)}[1-\Omega_J(q_{12}^2)-2\bar{q}\,(1-\Omega_J(q_{12}))]\right\} \qquad (8.24)$$
$$-\frac{i\,u\beta}{4\sqrt{N}}E\left\{e^{-i\,u\frac{\sqrt{N}}{\beta}\,\zeta_N(t)}[\Omega_J(q_{12}^2)-2\bar{q}\,\Omega_J(q_{12})-N^{-1}]\right\}.$$

By using (8.24) in Eq. (8.23), one finds

$$E\, e^{i\,u\hat{f}_N(\beta,h,t;J)} \;=\; E\, e^{i\,u\hat{f}_N(\beta,h,0;J)} + \frac{u^2\bar{q}^2}{4}E\int_0^t e^{i\,u\hat{f}_N(t')}\,dt' \qquad (8.25)$$
$$-\frac{u^2}{4N}\int_0^t E\,e^{i\,u\hat{f}_N(t')}\left(\Omega_J((\xi_{12}^N)^2)-1\right)\,dt'$$
$$+\frac{i\,u\beta}{4\sqrt{N}}\int_0^t E\,e^{i\,u\hat{f}_N(t')}\Omega_J((\xi_{12}^N)^2)\,dt'.$$

At $t=0$, all sites are decoupled and the central limit theorem for independent identically distributed random variables implies that

$$\hat{f}_N(\beta,h,0;J) \xrightarrow{d} \mathcal{N}(0,\sigma^2(\beta,h,0)), \qquad (8.26)$$

so that

$$E\, e^{i\,u\hat{f}_N(\beta,h,0;J)} \to \exp\left(-\frac{u^2}{2}\sigma^2(\beta,h,0)\right). \qquad (8.27)$$

The last two terms in Eq. (8.25) clearly vanish for $N\to\infty$. In fact, for instance,

$$N^{-1/2}\left|E\,e^{i\,u\hat{f}_N(t)}\Omega_J((\xi_{12}^N)^2)\right| \le N^{-1/2}E\,\Omega_J((\xi_{12}^N)^2) = N^{-1/2}\left\langle(\xi_{12}^N)^2\right\rangle_t = O(N^{-1/2}),$$

thanks to (8.5). Therefore, Eq. (8.25) yields the following linear integral equation for the characteristic function:

$$E\, e^{i\,u\hat{f}_N(\beta,h,t;J)} = E\, e^{i\,u\hat{f}_N(\beta,h,0;J)} + \frac{u^2\bar{q}^2}{4}E\int_0^t e^{i\,u\hat{f}_N(\beta,h,t';J)}\,dt' + o(1),$$

whose unique solution is, keeping into account the initial condition (8.27) and the definition (8.8) of σ^2,

$$E\, e^{i\,u \hat{f}_N(\beta,h,t;J)} = e^{-\frac{u^2}{2}\sigma^2(\beta,h,t)} + o(1).$$

□

Before we conclude this section, we wish to note that from Eq. (8.25) one can also obtain in a very simple way a well known [65, 66] result for free energy fluctuations at zero external field and $\beta < 1$, *i.e.*,

$$\eta_N(\beta; J) \equiv \ln Z_N(\beta; J) - N\left(\ln 2 + \frac{\beta^2}{4}\right) \xrightarrow{d} \hat{Y}_\beta - \frac{1}{4}\ln\frac{1}{1-\beta^2}, \tag{8.28}$$

where $\hat{Y}_\beta$ is a centered Gaussian random variable of variance

$$\frac{1}{2}\left(\ln\frac{1}{1-\beta^2} - \beta^2\right).$$

First of all notice that, for $0 \le t \le 1$,

$$\eta_N(\beta\sqrt{t}; J) \stackrel{d}{=} -\sqrt{N}\beta \hat{f}_N(\beta, 0, t; J),$$

where equality holds in law, and that

$$\eta_N(0; J) \equiv 0,$$

since for $\beta = 0$ the interaction disappears and we are left with a system of free spins. Therefore, setting $u = -s\sqrt{N}\beta$ and $h = 0$ in Eq. (8.25), one obtains the equation

$$\begin{aligned} E\, e^{i\,s\,\eta_N(\beta\sqrt{t};J)} &= 1 - \frac{s^2\beta^2}{4}\int_0^t E\, e^{i\,s\,\eta_N(\beta\sqrt{t'})}\left(\Omega_J((\xi_{12}^N)^2) - 1\right) dt' \\ &\quad - \frac{is\beta^2}{4}\int_0^t E\, e^{i\,s\,\eta_N(\beta\sqrt{t'})}\Omega_J((\xi_{12}^N)^2)\, dt'. \end{aligned} \tag{8.29}$$

Theorem 22 implies, for vanishing external field and $t\beta^2 < 1$, that the limit variables ξ_{ab} are independent if the indices are different, *i.e.*, $\langle \xi_{ab}\xi_{bc}\rangle = \langle \xi_{ab}\xi_{cd}\rangle = 0$, so that

$$E\left(\Omega_J((\xi_{12}^N)^2) - \langle(\xi_{12}^N)^2\rangle_t\right)^2 = \langle(\xi_{12}^N)^2\,(\xi_{34}^N)^2\rangle_t - \langle(\xi_{12}^N)^2\rangle_t^2 = o(1),$$

thanks to Wick's theorem. Then, $\Omega_J((\xi_{12}^N)^2)$ can be replaced by $\langle \xi_{ab}^2\rangle_t = (1 - t\beta^2)^{-1}$ (see Eq. (8.14)), and Eq. (8.29) yields

$$\begin{aligned} E\, e^{i\,s\,\eta_N(\beta\sqrt{t};J)} &= 1 - \frac{s^2\beta^2}{4}\int_0^t E\, e^{i\,s\,\eta_N(\beta\sqrt{t'};J)}\left(\frac{1}{1-t'\beta^2} - 1\right) dt' \\ &\quad - \frac{is\beta^2}{4}\int_0^t E\, e^{i\,s\,\eta_N(\beta\sqrt{t'};J)}\frac{1}{1-t'\beta^2}\, dt' + o(1). \end{aligned} \tag{8.30}$$

Derivation with respect to t reduces the last relation to a linear differential equation, and the result (8.28) easily follows. □

8.4 Overlap fluctuations

8.4.1 Exponential suppression of overlap fluctuations

In Chapter 7 we proved that, in a certain region of the (β, h) plane, the overlap fluctuations around the Sherrington-Kirkpatrick order parameter $\bar{q}$ are of order $1/\sqrt{N}$ (the mean square fluctuations are of order $1/N$). Here, we prove that the results of Theorem 19, together with the concentration of measure arguments of Section 5.3, allow to get a bound, exponentially small in N, on the probability that the difference between the overlap and $\bar{q}$ is greater than some given value. We learned this nice argument in [77]. To this purpose, recall Theorem 2 on the self-averaging of the free energy, which in particular implies that

$$P\left(\left|\frac{1}{2N}\ln\Omega_J\left(e^{N\frac{\beta^2\lambda}{2}(q_{12}-\bar{q})^2}\right)-(\tilde{\alpha}_N(\beta,h,\lambda,t)-\alpha_N(\beta,h,t))\right|\geq u\right)\leq\exp(-NKu^2)\quad(8.31)$$

where K is a number which depends on β, t, h but not of N, and Ω_J is the t-dependent Boltzmann-Gibbs average corresponding to the partition function (6.3). The bound (8.31) can be derived from Eq. (5.22), observing that $(2N)^{-1}\ln\Omega_J(\dots)$ can be written as the difference of two disorder dependent free energies,

$$\frac{1}{2N}\ln\Omega_J\exp\left(N\frac{\beta^2\lambda}{2}(q_{12}-\bar{q})^2\right)=\frac{1}{2N}\ln\tilde{Z}_N(\beta,h,\lambda,t;J)-\frac{1}{2N}\ln\tilde{Z}_N(\beta,h,0,t;J),\quad(8.32)$$

and that

$$\tilde{\alpha}_N(\beta,h,\lambda,t)-\alpha_N(\beta,h,t)$$

is just its disorder average. From Theorem 19 we know that, for fixed values of β and h, in the triangular region[§]

$$0\leq t+\lambda\leq\bar{\lambda}<\lambda_c(\beta,h)\quad(8.33)$$

the following bounds hold uniformly:

$$0\leq\tilde{\alpha}_N(\beta,h,\lambda,t)-\alpha_N(\beta,h,t)\leq\frac{C}{N}\quad(8.34)$$

for some constant C. Therefore, with probability at least $1-\exp(-NKu^2)$ one has

$$\Omega_J\left(\exp\frac{\beta^2\lambda}{2}N(q_{12}-\bar{q})^2\right)\leq\exp(2Nu+2C).\quad(8.35)$$

By employing Tchebyshev's inequality one finds

$$\Omega_J\left(1_{\{|q_{12}-\bar{q}|\geq v\}}\right)\leq e^{-\frac{\beta^2\bar{\lambda}}{2}Nv^2}\Omega_J\left(\exp\frac{\beta^2\bar{\lambda}}{2}N(q_{12}-\bar{q})^2\right)\leq e^{2N(u-\frac{\beta^2\bar{\lambda}}{4}v^2)+2C},\quad(8.36)$$

so that, choosing

$$u=\frac{\beta^2\bar{\lambda}}{8}v^2\quad(8.37)$$

[§]since t and λ are restricted to be non-negative, Eq. (8.33) defines a triangle in the (t, λ) plane.

one has

$$\Omega_J\left(1_{\{|q_{12}-\bar{q}|\geq v\}}\right)\leq \exp\left(-N\frac{v^2\beta^2\bar{\lambda}}{4}+2C\right), \tag{8.38}$$

with probability at least $1-\exp(-NK'v^4)$, where

$$K'=K\frac{\beta^4\bar{\lambda}^2}{64}$$

is a positive constant. Finally, the estimate we are looking for easily follows:

$$\left\langle 1_{\{|q_{12}-\bar{q}|\geq v\}}\right\rangle = E\,\Omega_J\left(1_{\{|q_{12}-\bar{q}|\geq v\}}\right)\leq e^{-N\frac{v^2\bar{\lambda}}{4}+2C}+e^{-Nv^4K'}. \tag{8.39}$$

Of course, this is much more than just self-averaging of the overlaps. This estimate is going to be important in controlling the infinite volume limit of the probability distribution of the fluctuations by means of the cavity equations.

8.4.2 The cavity method in the replica symmetric region

The cavity method [6] allows to express thermal averages of physical quantities defined on the N-spin system as functions of averages on a system of $N-1$ spins, at a slightly higher temperature. In a sense, it is induction on N. The cavity method was introduced in [55] as an alternative to the replica trick and, while the mysterious $n\to 0$ limit does not appear here, the results it gives and the *Ansatz* on which it is based are equivalent. The cavity method has been widely applied not only in the theoretical physics literature but also in the mathematical physics one (see, for instance, Refs. [74], [88], [64], [11]).

Here, we discuss a version of the cavity method developed by M. Talagrand [76], [11], where a careful control on error terms is obtained. This method has been employed with success by Talagrand to obtain detailed results on the replica symmetric region of the Sherrington-Kirkpatrick model, and we will follow a similar strategy to prove Theorem 22.

Consider a generic number k of replicas of the N-spin system, and single out the $N-$th spin, letting

$$\sigma^a=(\eta^a,\epsilon^a), \qquad a=1,2,\dots,k \tag{8.40}$$

where

$$\eta^a\in\{-1,1\}^{N-1} \tag{8.41}$$

is the configuration of a system with the last spin deleted, and

$$\epsilon^a=\sigma^a_N=\pm 1.$$

Then, the Hamiltonian corresponding to the partition function (6.3) can be rewritten as

$$H_N(\sigma^a,t,h;J) \;=\; H'_{N-1}(\eta^a,t,h;J)-\epsilon^a\left(h+\sqrt{\bar{q}(1-t)}g_0+\sqrt{t/N}g\eta^a\right), \tag{8.42}$$

where H'_{N-1} is the Hamiltonian of the $N-1$-spin system, *with the two-body coupling* J_{ij} *replaced by* $J_{ij}\sqrt{1-1/N}$. The random variables

$$g_0 = J_N \tag{8.43}$$
$$g_i = J_{N\,i}, \quad i = 1, 2, \ldots, N-1, \tag{8.44}$$

are independent identically distributed standard Gaussians, independent also of the random variables appearing in H_{N-1}. Here and in the following, $g\eta^a$ denotes the $(N-1)$-dimensional scalar product

$$g\eta^a = \sum_{i=1}^{N-1} g_i \eta_i^a.$$

The weakening of the the couplings J_{ij}, when going from $N-1$ to N sites, corresponds in a sense to lowering β, or raising the temperature[¶]. It is interesting to notice that only the two body interaction term is modified in H'_N, while the external (random and non-random) magnetic field is left unchanged.

If we let $\Omega'_J(.)$ denote the Boltzmann-Gibbs average for a system of $N-1$ spins, with Hamiltonian given by H'_{N-1}, Eq. (8.42) implies

$$\Omega_J\left(F(\sigma^1, \ldots, \sigma^k)\right) = \frac{\Omega'_J\left(Av\, F(\eta^1, \epsilon^1, \ldots, \eta^k, \epsilon^k)\Psi^{(k)}\right)}{\Omega'_J\left(Av\, \Psi^{(k)}\right)}, \tag{8.45}$$

where F is a generic function of the configuration of the k replicas, Av denotes the average over the spin variables $\epsilon^a = \pm 1$ and

$$\Psi^{(k)} \equiv \exp \beta \sum_{a=1}^{k} \epsilon^a \left(h\sqrt{t/N} g\eta^a + \sqrt{\bar{q}(1-t)} g_0 + h\right). \tag{8.46}$$

In the region where the overlap self-averages, the cavity equations (8.45) can be simplified, expanding $\Psi^{(k)}$ both in the numerator and in the denominator for η^a close to its thermal average $\Omega'_J(\eta^a)$. This idea was introduced by M. Talagrand in [76], and further employed in [9], [11]. In order to state Talagrand's result, we need some additional preliminary definitions. First of all, we define

$$b = \Omega'_J(\eta^a) \tag{8.47}$$

to be the thermal average of η^a (of course, it does not depend on the replica index a), with respect to the Boltzmann-Gibbs state Ω'_J. Moreover, we let

$$\dot{\eta}^a = \eta^a - b \tag{8.48}$$
$$X = \beta(\sqrt{t/N} g\, b + \sqrt{\bar{q}(1-t)} g_0 + h)$$
$$\Psi_0^{(k)} = \exp(X \sum_{a=1}^{k} \epsilon^a)$$

[¶]To be precise, this analogy is correct only for vanishing external field, since a change in the temperature would reflect also on the one-body interaction.

and we rewrite

$$F(\sigma^1,\ldots,\sigma^k)=F(\eta^1,\epsilon^1,\ldots,\eta^k,\epsilon^k), \tag{8.49}$$

with a slight abuse of notation. Then, the following holds:

Theorem 23 *[76] For any value of the thermodynamic parameters, and for any smooth function F,*

$$\langle F(\sigma^1,\ldots,\sigma^k)\rangle = E\frac{1}{\cosh^k X}\Omega'_J\left(Av\,F\,\Psi_0^{(k)}\right) \tag{8.50}$$

$$+t\beta^2\,E\frac{1}{\cosh^k X}\Omega'_J\left(Av\,F\,\Psi_0^{(k)}\sum_{1\le a<c\le k}\epsilon^a\epsilon^c\frac{\dot\eta^a\dot\eta^c}{N}\right) \tag{8.51}$$

$$+t\beta^2\,E\frac{1}{\cosh^k X}\Omega'_J\left(Av\,F\,\Psi_0^{(k)}\sum_{1\le a\neq c\le k}\epsilon^a\epsilon^c\frac{\dot\eta^c b}{N}\right) \tag{8.52}$$

$$-k\,t\beta^2\,E\frac{\tanh X}{\cosh^k X}\Omega'_J\left(Av\,F\,\Psi_0^{(k)}\sum_{a=1}^{k}\epsilon^a\frac{\dot\eta^a b}{N}\right)+S \tag{8.53}$$

and the "error term" S can be estimated as

$$|S|\le w_k(\beta,h,t)E\,\Omega'_J\left(Av\,|F|\left(\sum_{a=1}^{k+1}\left(\frac{\dot\eta^a b}{N}\right)^2+\sum_{1\le a<c\le k+2}\left(\frac{\dot\eta^a\dot\eta^c}{N}\right)^2\right)\right), \tag{8.54}$$

where w_k is a smooth function of its arguments, independent of N.

With respect to Theorem 3.2 of [76], the last sum in the right hand side of (8.54) is performed on $a<c$ instead of $a\le c$. However, the proof of Theorem 23 proceeds exactly as in [76].

Theorem 23 is a sort of Taylor expansion of the cavity equations around $\eta^a=b$. Increasing orders in the expansion carry increasing powers in $\sqrt{t/N}$ and, owing to symmetry of the distribution of the variables g_i, only integer powers of t appear. The real "expansion parameter", however, is $\dot\eta_a b/N$, which is small in the high temperature region, since η^a fluctuates very little, as shown by Eq. (8.39). Moreover, in this region S vanishes as $N\to\infty$, so what the theorem says is that the only terms that matter in the expansion are the zeroth and first order ones. This will be explained more precisely in the following.

8.4.3 A preparatory example: zero external field

For simplicity, we discuss separately the cases of zero and non-zero external field. The former is much simpler and is of course a particular case of the latter. The reason why we treat them separately is simply of pedagogical character, since when $h=0$ one can easily get an idea of how the proof works. In the general case, the strategy is the same but the proof is technically more complicated.

For the Sherrington-Kirkpatrick model with $h = 0$, Theorem 22 states that, for $t\beta^2 < 1$, the rescaled overlaps ξ_{ab}^N converge in distribution to statistically independent centered Gaussian variables whose variance

$$A(\beta, h = 0, t) = 1/(1 - t\beta^2)$$

diverges at the critical point $t\beta^2 = 1$. In order to prove this, it is convenient to work with the characteristic function of the random vector $\xi^N = (\xi_{12}^N, \xi_{13}^N, \ldots)$. What we have to show [13] is that it tends to the characteristic function of the product of independent Gaussian variables $\mathcal{N}(0, A) \otimes \mathcal{N}(0, A) \otimes \ldots$. An alternative (equivalent) approach is that of showing convergence for all polynomials of the form

$$\langle (\xi_{12}^N)^{n_{12}} (\xi_{13}^N)^{n_{13}} \ldots \ldots \rangle \tag{8.55}$$

to the expression given by Wick's theorem. The result is therefore the following:

Theorem 24 *Let $\phi_N(u), u = (u_{12}, \ldots, u_{s-1,s})$ be the characteristic function of the rescaled overlaps ξ_{ab}^N for $a, b < s$:*

$$\phi_N(u) = \left\langle \exp i \sum_{1 \le a < b \le s} u_{ab}\, \xi_{ab}^N \right\rangle. \tag{8.56}$$

Then, for $t\beta^2 < 1$

$$\lim_{N \to \infty} \phi_N(u) = \prod_{1 \le a < b \le s} \exp - \frac{u_{ab}^2}{2(1 - t\beta^2)}. \tag{8.57}$$

Remark This result was also obtained by Aizenman *et al.* in [65], and by Comets and Neveu in [66], where the authors studied the fluctuations of free energy, energy and overlaps in the high temperature regime for the zero field Sherrington-Kirkpatrick model, using techniques of stochastic calculus.

Proof) We give here only the main ideas about the proof and skip some technicalities, which are analyzed in greater detail in next section, where the more general case $h \neq 0$ is considered. The strategy we follow is to employ the cavity equations and self-averaging of the overlaps, which we know to hold for $t\beta^2 < 1$ (see Section 7.2), in order to obtain a linear differential equation for the characteristic function. The linear equation can be then solved explicitly, giving the wanted result.

First of all, owing to symmetry among spins one can write

$$\partial_{u_{ab}} \phi_N^t(u) \;=\; i\sqrt{N} \left\langle \sigma_N^a \sigma_N^b \, e^{iu\,\xi_{12}^N} \right\rangle. \tag{8.58}$$

Next we notice that, in the particular case of zero external field, the thermal average of each spin variable σ_i vanishes so that $b = 0$ in (8.47), and Theorem 23 reduces to

Proposition 2 *For any smooth function F,*

$$\begin{aligned} \left\langle F(\sigma^1, \ldots, \sigma^k) \right\rangle \;&=\; E\,\Omega'_J \left(Av\, F(\eta^1, \epsilon^1, \ldots, \eta^k, \epsilon^k) \right) \\ &+ t\beta^2 E\,\Omega'_J \left(Av\, F(\eta^1, \epsilon^1, \ldots, \eta^k, \epsilon^k) \sum_{1 \le a < c \le k} \epsilon^a \epsilon^c \frac{\eta^a \eta^c}{N} \right) + S, \end{aligned} \tag{8.59}$$

where Av denotes average over $\epsilon^a = \pm 1$, *and*

$$|S| \leq w_k(\beta, t) E\, \Omega'_J \left(Av\, |F| \sum_{1 \leq a < c \leq k+2} \left(\frac{\eta^a \eta^c}{N} \right)^2 \right). \tag{8.60}$$

Here, w_k *is a smooth function of its arguments, independent of* N.

By applying Proposition 2 to Eq. (8.58) and using the fact that

$$\exp(iu/\sqrt{N}) = 1 + iu/\sqrt{N} + O(1/N), \tag{8.61}$$

one finds

$$\partial_{u_{ab}} \phi_N(u) = -u_{ab} \left\langle \exp i \sum_{1 \leq c < d \leq s} u_{cd} \frac{\eta^c \eta^d}{\sqrt{N}} \right\rangle' \tag{8.62}$$

$$+ it\beta^2 \left\langle \frac{\eta^a \eta^b}{\sqrt{N}} \exp i \sum_{1 \leq c < d \leq s} u_{cd} \frac{\eta^c \eta^d}{\sqrt{N}} \right\rangle' + S + o(1)$$

$$= -u_{ab}\, \phi_N(u) + t\beta^2\, \partial_{u_{ab}} \phi_N(u) + S + o(1) \tag{8.63}$$

where, of course, $\langle . \rangle'$ denotes the double average $E\, \Omega'_J(.)$. In Eq. (8.63) we were quite sloppy, and we performed substitutions of the kind

$$\left\langle \exp i \sum_{1 \leq c < d \leq s} u_{cd} \frac{\eta^c \eta^d}{\sqrt{N}} \right\rangle' \to \phi_N(u) + o(1) \tag{8.64}$$

and similar, which are not self-evident. The replacement (8.64) can be rigorously justified by applying once more Proposition 2, but here we skip details for simplicity, since analogous estimates will be proved in next section. Finally, the error term S can be estimated by means of inequality (8.60), which gives

$$|S| \leq C N^{-3/2} \left\langle (\eta^1 \eta^2)^2 \right\rangle'. \tag{8.65}$$

Using formula (8.5), *i.e.*, the uniform bound on the second moment of ξ, one easily sees that

$$|S| = O(N^{-1/2}). \tag{8.66}$$

Therefore, apart from terms which are negligible in the thermodynamic limit, $\phi_N(u)$ satisfies the set of *linear differential equations*

$$(1 - \beta^2 t) \partial_{u_{ab}} \phi_N(u) = -u_{ab}\, \phi_N(u) \tag{8.67}$$

with the initial condition

$$\phi_N(0) = 1. \tag{8.68}$$

Condition (8.68) is obvious if one recalls the definition of $\phi_N(u)$ as characteristic function of a random variable. Clearly, this implies that $\phi_N(u)$ factorizes as a product of functions of the different u_{ab}'s and the solution is easily found to be

$$\lim_{N \to \infty} \phi_N(u) = \prod_{1 \leq a < b \leq s} \exp - \frac{u_{ab}^2}{2(1 - \beta^2 t)}. \tag{8.69}$$

□

8.4.4 The general case

The case with non-vanishing external field is similar to that considered in the previous section, but it is technically more involved. The additional complication arises from the fact that, while for vanishing field the distribution of ξ_{ab}^N is symmetric about $\xi_{ab}^N = 0$, in the general case there is no reason why this should hold for finite N. Indeed, symmetry holds only in the ergodic region for $N \to \infty$. This implies that some of the terms appearing in the expansion in Theorem 23 are no more identically zero and have to be kept into account, therefore increasing the algebraic complexity of the equations. This is also reflected in the fact that the limit Gaussian variables are characterized by a non-diagonal correlation matrix.

To prove Theorem 22, it suffices to show that for any integer s, the characteristic function

$$\phi_N(u) = \langle \exp i\, u\, \xi^N \rangle = \left\langle \exp i \sum_{1 \le a < b \le s} u_{ab}\, \xi_{ab}^N \right\rangle$$

converges for $N \to \infty$ to

$$\phi(u) = \exp\left\{ -\frac{1}{2} (\hat{L}u, u) \right\}, \tag{8.70}$$

where $(.\,,.)$ denotes scalar product and $\hat{L}$ is the $s(s-1)/2 \times s(s-1)/2$ dimensional matrix of elements

$$\begin{aligned} L_{(ab),(ab)} &= A(\beta, h, t) \\ L_{(ab),(ac)} &= B(\beta, h, t) \\ L_{(ab),(cd)} &= C(\beta, h, t). \end{aligned}$$

The idea of the proof is to obtain a set of closed linear differential equations for $\phi_N(u)$, which determine uniquely the solution as (8.70), for $N \to \infty$. Some of the calculations involved in the proof are quite long, although straightforward, and are therefore just sketched.

In order to prove Theorem 22, we first exploit symmetry among site indices to write

$$\partial_{u_{rr'}} \phi_N(u) = i \left\langle \xi_{rr'}^N\, e^{i\,u\,\xi^N} \right\rangle = i\sqrt{N} \left\langle (\sigma_N^r \sigma_N^{r'} - \bar{q})\, e^{i\,u\,\xi^N} \right\rangle \tag{8.71}$$

$$\varphi_N^{(a)}(u) \equiv i \left\langle \xi_{a,s+1}^N\, e^{i\,u\,\xi^N} \right\rangle = i\sqrt{N} \left\langle (\sigma_N^a \sigma_N^{s+1} - \bar{q})\, e^{i\,u\,\xi^N} \right\rangle \tag{8.72}$$

$$\psi_N(u) \equiv i \left\langle \xi_{s+1,s+2}^N\, e^{i\,u\,\xi^N} \right\rangle = i\sqrt{N} \left\langle (\sigma_N^{s+1} \sigma_N^{s+2} - \bar{q})\, e^{i\,u\,\xi^N} \right\rangle, \tag{8.73}$$

and then employ the cavity equations to express these quantities as functions of ϕ, φ, ψ themselves. For instance, apply Theorem 23 to the right hand side of Eq. (8.71) and consider the term arising from (8.50). After averaging on the dichotomic variables ϵ, one is left with

$$i\,(\sqrt{N}-i\sum_{1\le a<b\le s}u_{ab}\,\bar q)E\left\{(\tanh^2 X-\bar q)\Omega'_J\exp\left(i\,u'\,\xi^{N-1}\right)\right\}+ \tag{8.74}$$
$$-\,u_{rr'}E\left\{(1-\bar q\tanh^2 X)\Omega'_J\exp\left(i\,u'\,\xi^{N-1}\right)\right\}$$
$$-(1-\bar q)\sum_{a\neq r,r'}(u_{ar}+u_{ar'})E\left\{\tanh^2 X\,\Omega'_J\exp\left(i\,u'\,\xi^{N-1}\right)\right\}$$
$$-\sum_{1\le c<d\le s\,c,d\neq r,r'}u_{cd}\,E\left\{\tanh^2 X\left(\tanh^2 X-\bar q\right)\Omega'_J\exp\left(i\,u'\,\xi^{N-1}\right)\right\}+o(1),$$

where $u'=u\sqrt{1-1/N}$. The term $o(1)$ arises when $\exp\left(iu/\sqrt{N}\right)$ is expanded around u=0 and the terms of order u^2 or higher are neglected. Indeed, one has

$$\left|\sqrt{N}\left\langle(\sigma_N^r\sigma_N^{r'}-\bar q)\,e^{iu'\xi^{N-1}}\left(e^{i\sum_{1\le a<b\le s}\frac{u_{ab}}{\sqrt N}(\sigma_N^a\sigma_N^b-\bar q)}-1-i\sum_{1\le a<b\le s}\frac{u_{ab}}{\sqrt N}(\sigma_N^a\sigma_N^b-\bar q)\right)\right\rangle\right|$$
$$\le 2\sqrt{N}\left|e^{i\sum_{1\le a<b\le s}\frac{u_{ab}}{\sqrt N}(\sigma_N^a\sigma_N^b-\bar q)}-1-i\sum_{1\le a<b\le s}\frac{u_{ab}}{\sqrt N}(\sigma_N^a\sigma_N^b-\bar q)\right|=O(N^{-1/2}).$$

Now, rewrite E as $E\,E_g$, where E_g denotes the average only with respect to the random variables g_0 and $g_i, i=1,\dots,N-1$, and notice that the thermal average Ω'_J does not depend on g_0,g_i. Computation of $E_g(\dots)$ would be simpler if, instead of X, there were

$$\bar X=\beta(\sqrt{t/N}g\,\bar b+\sqrt{\bar q(1-t)}g_0+h),$$

where

$$\bar b\equiv\frac{b}{||b||}\sqrt{N\bar q}$$

and $\bar q$ is, as usual, the Sherrington-Kirkpatrick order parameter defined by Eq. (4.25). Of course, one has

$$\bar X\stackrel{d}{=}\beta z\sqrt{\bar q}+\beta h,$$

where z is a standard unit Gaussian variable and equality holds in distribution so that, for instance,

$$E_g\tanh^2\bar X=\bar q.$$

The idea is, therefore, to expand in $X-\bar X$. As a preliminary fact, notice that the second moment of the random variable $(b-\bar b)$ is bounded uniformly in N. Indeed,

$$E||b-\bar b||^2=E(||b||-\sqrt{N\bar q})^2\le\frac{1}{\bar qN}E(||b||^2-N\bar q)^2 \tag{8.75}$$
$$=\frac{1}{\bar q}\langle\xi_{12}^{N-1}\xi_{34}^{N-1}\rangle+O(1/N)=O(1), \tag{8.76}$$

thanks to Eq. (8.5). As an example, let us examine in detail the first term in (8.74), that is,

$$i\sqrt{N}E\,E_g(\tanh^2 X)\,\Omega'_J\exp\left(i\,u'\xi^{N-1}\right)-i\,\bar q\sqrt{N}E\,\Omega'_J\exp\left(i\,u'\xi^{N-1}\right). \tag{8.77}$$

By a simple second order Taylor expansion in $X-\bar X$ and an integration by parts on the Gaussian quenched disorder g, one finds

$$E_g\,\tanh^2 X = E_g\,\tanh^2\bar X + \frac{t}{N}(b-\bar b)\bar b\, E_g\;\partial_x^2\tanh^2 x\big|_{x=\bar X} \tag{8.78}$$

$$+\frac{t\beta^2}{2N}E_g\,\partial_x^2\tanh^2 x\big|_{x=\bar X+\theta(X-\bar X)}\left(g(b-\bar b)\right)^2 \tag{8.79}$$

$$=\bar q+\frac{t\beta^2}{2N}(b-\bar b)(b+\bar b)E_g\,\partial_x^2\tanh^2\bar X \tag{8.80}$$

$$+\frac{t\beta^2}{2N}||b-\bar b||^2E_g\left(\partial_x^2\tanh^2 x\big|_{x=\bar X+\theta(X-\bar X)}-\partial_x^2\tanh^2\bar X\right) \tag{8.81}$$

$$+\frac{t^2\beta^4}{2N^2}E_g\;\partial_x^4\tanh^2 x\big|_{x=\bar X+\theta(X-\bar X)}\left[(b-\bar b)(\bar b+\theta(b-\bar b))\right]^2 \tag{8.82}$$

where $0\le\theta\le 1$. Let analyse each term separately. Recalling the definitions of b and $\bar b$, the second term in (8.80) equals

$$\frac{t\beta^2}{2N}\Omega'_J(\eta^{s+1}\eta^{s+2}-N\bar q)\int\partial_x^2\tanh^2(\beta h+\beta z\sqrt{\bar q})d\mu(z) \tag{8.83}$$

$$=\frac{\beta^2 t}{\sqrt N}\Omega'_J(\xi^{N-1}_{s+1,s+2})(3Y_0+2\bar q-2)+O(1/N), \tag{8.84}$$

where Y_0 was defined in (8.13). Another application of Taylor expansion and integration by parts, together with Cauchy-Schwarz inequality and the fact that the derivatives of the function $x\to\tanh^2 x$ are bounded, shows that the terms (8.81) and (8.82) can be bounded by

$$\frac{k}{N}||b-\bar b||^2.$$

Therefore, using the estimate (8.76), the expression (8.77) reduces to

$$i\,\beta^2 t(3Y_0+2\bar q-2)E\,\Omega'_J\left[\xi_{s+1,s+2}\exp\left(i\,u'\,\xi^{N-1}\right)\right]+O(N^{-1/2}),$$

and

$$i\sqrt N E\left\{(\tanh^2 X-\bar q)\Omega'_J\exp\left(i\,u'\xi^{N-1}\right)\right\}\;=\;\beta^2 t(2\bar q-2+3Y_0)\psi'+o(1),$$

where

$$\psi'\equiv i\left\langle\xi^N_{s+1,s+2}\,e^{i\,u'\,\xi^{N-1}}\right\rangle'. \tag{8.85}$$

Recall that $u'=u\sqrt{1-1/N}$. The other terms in (8.74) are much simpler than (8.77), and can be dealt with in the same way. Finally, the whole expression (8.74) can be rewritten as

$$\beta^2 t(2\bar q-2+3Y_0)\,\psi'-u_{rr'}(1-\bar q^2)\,\phi' \tag{8.86}$$
$$-(\bar q-\bar q^2)\sum_{a\neq r,r'}(u_{ar}+u_{ar'})\phi'-(Y_0-(1-\bar q)^2)\sum_{1\le c<d\le s:c,d\neq r,r'}u_{cd}\,\phi'+o(1),$$

where ϕ' is defined in analogy with ψ', as

$$\phi'\equiv\left\langle e^{i\,u'\,\xi^{N-1}}\right\rangle'.$$

The steps leading to expression (8.86) can be repeated with minor changes for the remaining terms (8.51) to (8.53). These terms, although they look more complicated than (8.50) at first sight, are actually simpler to treat, since a first (instead of second) order Taylor expansion in $X - \bar{X}$ is sufficient. This is due to the presence of terms like $\dot{\eta}_a \dot{\eta}_b / N$ or $\dot{\eta}_a b/N$, which are small with large probability, thanks to (8.5). Also in this case, one finds that terms (8.51) to (8.53) give quantities linear in $\phi', \partial\phi', \varphi', \psi'$, apart from terms of order $o(1)$. As for the "error term" S which appears in Theorem 23, one can easily check that it vanishes in the thermodynamic limit. This is a consequence of the exponential decay of overlap fluctuations, as expressed by (8.39).

Next, we show that ϕ', φ' and ψ' can be substituted by ϕ, φ and ψ, apart from negligible error terms. Indeed, for instance,

$$\begin{aligned}\phi_N(u) &= \left\langle \exp\left(i\, u' \,\xi^{N-1} + i\, u\, (\sigma_N^1 \sigma_N^2 - \bar{q})/\sqrt{N} \right) \right\rangle \\ &= \left\langle \exp i\, u' \,\xi^{N-1} \right\rangle (1 + o(1)) = \phi' + o(1).\end{aligned}$$

In the last step, we used Theorem 23. Therefore, Eq. (8.71) reduces to a linear relation between ϕ, φ and ψ, apart from a remainder which becomes irrelevant in the thermodynamic limit. In the same way, one sees that also Eqs. (8.72), (8.73) yield linear equations for ϕ, φ, ψ. Putting everything together, in the thermodynamic limit one has a set of coupled linear differential equations of the form

$$\mathbf{\Phi}(u) = \phi(u)\, \mathbf{v}(u) + t\beta^2\, \hat{M}\, \mathbf{\Phi}(u) \tag{8.87}$$

where $\mathbf{\Phi}(u)$ is the vector

$$\mathbf{\Phi}(u) = (\partial_{u_{12}} \phi(u), \ldots, \partial_{u_{s-1,s}} \phi(u), \varphi^{(1)}(u), \ldots, \varphi^{(s)}(u), \psi(u)).$$

$\mathbf{v}(u)$ is a vector whose components are homogeneous linear functions of the variables u, while $\hat{M}$ is a real square matrix with elements depending on $\bar{q}, Y_0$ alone. We do not report here the explicit expressions of $\mathbf{v}(u)$ and $\hat{M}$, which are quite complicated. However, it is instructive to check that, for instance, the term (8.86) is in agreement with this structure. In fact, the coefficient of ϕ' is a homogeneous linear function of the u variables, while the coefficient of ψ' is linear in $\beta^2 t$ and depends only on Y_0 and $\bar{q}$. As will be clear in the following, only the structure (8.87), and not the specific form of $\mathbf{v}$ and $\hat{M}$, are needed to conclude the proof of the theorem.

Assume at first that the matrix $(1 - \beta^2 t\, \hat{M})$ is invertible, which in principle can fail only for a finite number of values of t (for fixed β, h) since $\hat{M}$ is finite dimensional. In this case, Eq. (8.87) can be reduced to a first order differential system in normal form:

$$\mathbf{\Phi}(u) = \phi(u)(1 - \beta^2 t\, \hat{M})^{-1} \mathbf{v}(u), \tag{8.88}$$

which can be easily integrated. The most general solution for $\phi(u)$, compatible with the initial condition

$$\phi(0) = 1,$$

is of the form

$$\phi(u) = \exp\left\{ -\frac{1}{2}(\hat{K}u, u) + (p, u) \right\}, \tag{8.89}$$

where p is some $s(s-1)/2$ dimensional u-independent vector, and $\hat{K}$ is a $s(s-1)/2 \times s(s-1)/2$ real symmetric positive definite matrix. The symmetry and non negativity of $\hat{K}$ derive from the obvious property of permutation symmetry among replicas, and from the bound

$$|\phi(u)| \leq 1,$$

which holds for any characteristic function. The quadratic dependence on u of the exponent of $\phi(u)$ stems from the linear dependence of the components of $\mathbf{v}(u)$. Clearly, Eq. (8.89) means that the random variables $\{\xi_{ab}^N\}$ converge to some Gaussian process $\{\xi_{ab}\}$. Moreover, it turns out that the identification

$$p = 0$$

and

$$\hat{K} = \hat{L}$$

are straightforward. Indeed, as we discussed in Section 8.2, it was shown by Guerra in [53] that, if the limit process is Gaussian, then it is centered and its covariance function is exactly $\hat{L}$.

In order to conclude the proof, it remains to show convergence of the characteristic function for those possible values $\tilde{t}$ where $(1 - \beta^2 t \hat{M})$ is singular. For any $\delta > 0$ one can write

$$\phi_N\big|_{t=\tilde{t}} = \phi_N\big|_{t=\tilde{t}-\delta} + \delta\, \partial_t\, \phi_N\big|_{t=\tilde{t}-\theta_N \delta}\, ,$$

where $0 < \theta_N < 1$. After a straightforward computation one finds that

$$\partial_t \phi_N = \frac{\beta^2}{2} \left\langle e^{i\, u \xi^N} \left(\sum_{1 \leq a < b \leq s} (\xi_{ab}^N)^2 - s \sum_{a=1}^{s} (\xi_{a,s+1}^N)^2 + \frac{s(s+1)}{2} (\xi_{s+1,s+2}^N)^2 \right) \right\rangle .$$

By exploiting the uniform bound (8.5) and the arbitrariness of δ, one finds therefore that the theorem holds also for $t = \tilde{t}$. □

Chapter 9

Conclusions

In the present work I have reported some new results obtained, mostly in collaboration with Francesco Guerra, in the context of the rigorous study of mean field spin glass models. Our first main result is the existence of the infinite volume free energy and ground state energy per site for a wide class of models, which includes the well known Sherrington-Kirkpatrick model and Derrida's p-spin. Many generalizations have been considered, for instance to models with non-Gaussian random interactions, with non-Ising spin degrees of freedom, and to systems composed of several interacting replicas of the original mean field spin glass model. In the second place, we concentrated on the study of the high temperature region of the Sherrington-Kirkpatrick model, where we were able to obtain a quite detailed picture, proving replica symmetry and central limit theorems for the rescaled fluctuations of overlaps and free energy. As regards the low temperature region, as a consequence of Guerra's "broken replica symmetry bounds" for the free energy, we proved that replica symmetry is broken in the whole region of parameters below the Almeida-Thouless line.

Many of the above results can be expressed as comparisons, or sum rules, involving the free energies of two different systems. For instance, to prove subadditivity of the free energy one compares the free energy of the N spin system with that of two subsystems of size N_1 and N_2, with $N = N_1 + N_2$. As a second example, the extension to non-Gaussian couplings is performed by showing that the difference between the free energy of the Gaussian model and that of its non-Gaussian counterpart vanishes in the thermodynamic limit. Again, in order to prove replica symmetry at high temperature we compare the free energy of the system with that of an auxiliary exactly solvable system of two replicas, where the two-body disordered interaction is replaced by a random one-body interaction plus a non-random quadratic coupling between the two replicas. In all these cases, the idea we employed is that of suitably interpolating between the Hamiltonians of the systems we wish to compare, and of estimating the derivative of the free energy as the interpolating parameter is varied. Even the proof of Guerra's bounds, which we illustrated in Chapter 6, is based on an interpolation technique of this kind, though more complicated.

Notwithstanding the many recent developments in the rigorous study of mean field spin glasses, many fundamental questions remain open. For instance, one would like to extend the proof of replica symmetry for the Sherrington-Kirkpatrick model up to the Almeida-Thouless line, and to understand what is the infinite volume behavior of the fluctuations on the critical line.

Of course, the ultimate goal is to compute the infinite volume free energy in the whole range of thermodynamic parameters, and possibly to prove that Parisi solution holds. To this purpose, one should prove that the auxiliary states $\langle . \rangle_a$, which appear in Guerra's sum rules discussed in Chapter 6, are able to concentrate the overlap around the non-random values $\bar{q}_a$ given by the optimal Parisi functional order parameter. A difficulty one encounters in this task is that these states do not appear in Parisi theory, and therefore one has little physical intuition about their properties. In the case of the high temperature region, we saw that the computation of the infinite volume free energy can be actually performed, as a consequence of lower bounds for the free energy of a system of two quadratically coupled replicas. We believe that this fact should guide us when trying to attack the more general problem of showing that Parisi solution holds for all values of thermodynamic parameters, and in particular in the low temperature region. However, it is not clear yet how to couple replicas in this case, and how to obtain lower bounds for the free energy of the coupled system. This will be one of the main subjects of our future research, and we hope to report on this soon.

It would also be very interesting to extend the techniques we illustrated in this work, to the more general situation of finite connectivity spin glasses and neural networks. A first step was performed in [57], but a lot of work remains to be done.

Bibliography

General references

Statistical mechanics

[1] L. D. Landau, E. M. Lifshitz, *Statistical physics*, Pergamon Press, London (1958).

[2] D. Ruelle, *Statistical mechanics. Rigorous results*, W.A. Benjamin Inc., New York (1969).

[3] R. B. Griffiths, *Rigorous results and theorems*, in *Phase transitions and critical phenomena* vol. 1, C. Domb and M. S. Green eds., Academic Press, New York (1972).

[4] R. B. Israel, *Convexity in the Theory of Lattice Gases*, Princeton University Press, Princeton (1979).

[5] A. Bovier, *Statistical Mechanics of Disordered Systems*, lecture notes, available at the web page `http://www.wias-berlin.de/~bovier`.

Spin glass theory and related fields

[6] M. Mézard, G. Parisi and M. A. Virasoro, *Spin glass theory and beyond*, World Scientific, Singapore (1987).

[7] D. J. Amit, *Modeling brain function*, Cambridge University Press, Cambridge (1989).

[8] K. H. Fischer, J. A. Hertz, *Spin glasses*, Cambridge University Press, Cambridge (1991).

[9] M. Talagrand, *Mean field models for spin glasses: a first course*, to appear in the Proceedings of the 2000 Saint Flour Summer School in probability.

[10] H. Nishimori, *Statistical physics of spin glasses and information processing: an introduction*, Oxford University Press, Oxford (2001).

[11] M. Talagrand, *Spin glasses: a challenge for mathematicians. Mean field models and cavity method*, Springer Verlag, to appear.

Probability theory and concentration of measure inequalities

[12] Y. S. Chow, H. Teicher, *Probability theory*, Springer Verlag, Berlin (1978).

[13] A. N. Shiryaev, *Probability*, Springer Verlag, Berlin (1989).

[14] M. Ledoux, M. Talagrand, *Probability in Banach spaces. Isoperimetry and processes*, Springer, Berlin (1991).

[15] M. Ledoux, *The concentration of measure phenomenon*, Mathematical Surveys and Monographs **89**, Am. Math. Soc., Providence, RI (2001).

Specific references

[16] D. Sherrington, S. Kirkpatrick, *Solvable model of a spin-glass*, Phys. Rev. Lett. **35** 1792 (1975).

[17] S. Kirkpatrick, D. Sherrington, *Infinite-ranged models of spin-glasses*, Phys. Rev. B **17**, 4384 (1978).

[18] G. Parisi, *Toward a mean field theory for spin glasses*, Phys. Lett. A **73**, 203 (1979).

[19] G. Parisi, *A sequence of approximated solutions to the S-K model for spin glasses*, J. Phys. A **13**, L-115 (1980).

[20] G. Parisi, *The order parameter for spin glasses: a function on the interval* $0-1$, J. Phys. A **13**, 1101 (1980).

[21] M. Talagrand, *The Sherrington Kirkpatrick model: a challenge for mathematicians*, Probab. Rel. Fields **110**, 109-176 (1998).

[22] F. Guerra, F. L. Toninelli, *The Thermodynamic Limit in Mean Field Spin Glass Models*, Commun. Math. Phys. **230:1**, 71-79 (2002), `cond-mat/0204280`.

[23] F. Guerra, *Broken Replica Symmetry Bounds in the Mean Field Spin Glass Model*, Commun, Math. Phys., to appear, `cond-mat/0205123`.

[24] O. Martin, R. Monasson, R. Zecchina. *Statistical mechanics methods and phase transitions in optimization problems*, Theoret. Comput. Sci. **265**, 3 (2001).

[25] G. Toulouse, *Theory of Frustration Effect in Spin Glasses*, Commun. Phys. **2**, 115 (1977).

[26] S. F. Edwards, P. W. Anderson, *Theory of spin glasses*, J. Phys F **5**, 1965 (1975).

[27] A. Bovier, J. Fröhlich, *A heuristic theory of the spin class phase*, J. Stat. Phys. **44**, 347 (1986).

[28] see, for instance:

G. Parisi, *Some recent developments in the theory of replica symmetry breaking*, Physica A **185**, 316 (1992);

G. Parisi, *Spin glasses, complexity and all that*, Physica A **194**, 28 (1993);

E. Marinari, G. Parisi, J. J. Ruiz-Lorenzo, F. Ritort, *Numerical Evidence for Spontaneously Broken Replica Symmetry in 3D Spin Glasses*, Phys. Rev. Lett. **76**, 843 (1996).

[29] see, for instance:

D. S. Fisher, D. A. Huse, *Ordered Phase of Short-Range Ising Spin-Glasses*, Phys. Rev. Lett. **56**, 1601 (1986);

D. S. Fisher, D. A. Huse, *Equilibrium behavior of the spin-glass ordered phase*, Phys. Rev. B **38**, 386 (1988).

[30] see, for instance:

C. M. Newman, D. L. Stein, *Simplicity of State and Overlap Structure in Finite-Volume Realistic Spin Glasses*, Phys. Rev. E, **57**, 1356 (1998);

C. M. Newman, D. L. Stein, *The Metastate approach to thermodynamic chaos*, Phys. Rev. E **55**, 5194 (1997).

[31] C. M. Newman, D. L. Stein, *The State(s) of Replica Symmetry Breaking: Mean Field Theories vs. Short-Ranged Spin Glasses*, `cond-mat/0105282`.

[32] E. Marinari, G. Parisi, F. Ricci-Tersenghi, J. J. Ruiz-Lorenzo, F. Zuliani, *Replica Symmetry Breaking in Short Range Spin Glasses: A Review of the Theoretical Foundations and of the Numerical Evidence*, J. Stat. Phys **98**, 973 (2000).

[33] J. Fröhlich, B. Zegarlinski, *Some comments on the Sherrington-Kirkpatrick model of spin glasses*, Commun. Math. Phys. **112**, 553 (1987).

[34] J. Fröhlich, B. Zegarlinski, *The high temperature phase of long-range spin glasses*, Commun. Math. Phys. **110**, 121 (1987).

[35] A. Bovier, *The Kac version of the Sherrington-Kirkpatrick model at high temperatures*, J. Stat. Phys. **91**, 459 (1998).

[36] M. Kac, G. E. Uhlenbeck, P. C. Hemmer, *On the van der Waals Theory of the Vapor-Liquid Equilibrium. I. Discussion of a One-Dimensional Model*, J. Math. Phys. **4**, 216 (1963).

[37] B. Hayes, *Can't get no satisfaction*, Am. Sci. **85**, 108 (1997).

[38] B. Derrida, *Random energy model: An exactly solvable model of disordered systems*, Phys. Rev. B **24**, 2613 (1981).

[39] B. Derrida, E. Gardner, *Solution of the generalized random energy model*, J. Phys. C **19**, 2253 (1986).

[40] E. Olivieri, P. Picco, *On the existence of thermodynamics for the random energy model*, Commun. Math. Phys. **96**, 125 (1984).

[41] D. Capocaccia, M. Cassandro, P. Picco, *On the existence of thermodynamics for the generalized random energy model*, J. Stat. Phys. **46**, 493 (1987).

[42] D. Ruelle, *A mathematical reformulation of Derrida's REM and GREM*, Commun. Math. Phys. **108**, 225 (1987).

[43] A. Bovier, I. Kurkova, *Rigorous results on some simple spin glass models*, to appear, `cond-mat/0206562`.

[44] F. Guerra, *About the overlap distribution in mean field spin glass models*, Int. Jou. Mod. Phys. B **10**, 1675-1684 (1996).

[45] D. J. Gross, M. Mézard, *The simplest spin glass*, Nucl. Phys. B **240**, 431 (1984).

[46] E. Gardner, *Spin glasses with p-spin interactions*, Nucl. Phys. B **257**, 747 (1985).

[47] M. Talagrand, *Rigorous low temperature results for the p-spin mean field spin glass model*, Prob. Theory and Rel. Fields **117** 303 (2000).

[48] M. Talagrand, *On the p-spin interaction model at low temperature*, C. R. A. S., to appear.

[49] L. Viana, A. J. Bray, *Phase diagrams for dilute spin-glasses*, J. Phys. C **18**, 3037 (1985).

[50] G. Parisi, *On the probabilistic formulation of the replica approach to spin glasses*, cond-mat/9801081.

[51] F. Baffioni and F. Rosati, *On the Ultrametric Overlap Distribution for Mean Field Spin Glass Models (I)*, The European Physical Journal B **17**, 439-447 (2000).

[52] F. Guerra, *On the mean field spin glass model*, in preparation.

[53] F. Guerra, *Sum rules for the free energy in the mean field spin glass model*, in *Mathematical Physics in Mathematics and Physics: Quantum and Operator Algebraic Aspects*, Fields Institute Communications **30**, American Mathematical Society (2001).

[54] J. R. L. de Almeida, D. J. Thouless, *Stability of the Sherrington-Kirkpatrick solution of a spin glass model*, J. Phys. A **11** 983 (1978).

[55] M. Mézard, G. Parisi, M. A. Virasoro, *SK model: the Replica solution without Replicas*, Europhys. Lett. **1**, 77 (1985).

[56] P. Contucci, M. Degli Esposti, C. Giardinà, S. Graffi, *Thermodynamical limit for correlated Gaussian random energy models*, to appear, `cond-mat/0206007`.

[57] S. Franz, M. Leone, *Replica bounds for optimization problems and diluted spin systems*, to appear, `cond-mat/0208280`.

[58] F. Guerra, F. L. Toninelli, *The infinite volume limit in generalized mean field disordered models*, Markov Processes and Rel. Fields, to appear, `cond-mat/0208579`.

[59] P. A. Vuillermot, *Thermodynamics of quenched random spin systems, and application to the problem of phase transitions in magnetic (spin) glasses*, J. Phys. A **10**, 1319 (1977).

[60] A. C. D. van Enter, J. L. van Hemmen, *The Thermodynamic Limit for Long-Range Random Systems*, J. Stat. Phys. **32**, 141 (1983).

[61] B. S. Cirel'son, I. A. Ibragimov, V. N. Sudakov, *Norms of Gaussian sample functions*, Proceedings of the third Japan-USSR symposium on probability theory, Lect. Notes Math. **550**, 20-41 (1976).

[62] G. Pisier, *Probabilistic methods in the geometry of Banach spaces*, Probability and analysis (Varenna, 1985), Lect. Notes Math. **1206**, 167-241 (1986).

[63] M. Talagrand *A new look at independence*, Ann. Probab. **24**, 1 (1996).

[64] L. Pastur, M. Shcherbina, *The absence of self-averaging of the order parameter in the Sherrington-Kirkpatrick model*, J. Stat. Phys. **62**, 1 (1991).

[65] M. Aizenman, J. Lebowitz and D. Ruelle, *Some rigorous results on the Sherrington-Kirkpatrick spin glass model*, Commun. Math. Phys. **112**, 3 (1987).

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

[67] F. Guerra, F. L. Toninelli, *The Sherrington-Kirkpatrick model with non-Gaussian disorder*, in preparation.

[68] P. Contucci, S. Graffi, S. Isola, *Mean field behaviour of spin systems with orthogonal interaction matrix*, J. Stat. Phys., **106**, 895-914 (2002).

[69] M. Talagrand, *Gaussian averages, Bernoulli averages and Gibbs' measure*, Random structures and algorithms, to appear.

[70] R. A. Wayne, D. E. Varberg, *Convex functions*, Academic Press, New York (1973).

[71] F. Guerra and F. L. Toninelli, *About the ground state energy in the mean field spin glass model*, in preparation.

[72] F. Guerra and F. L. Toninelli, *Spontaneous Replica Symmetry Breaking in the p-spin Model*, in preparation.

[73] F. L. Toninelli, *About the Almeida-Thouless transition line in the Sherrington-Kirkpatrick mean field spin glass model*, Europhys. Lett., to appear, `cond-mat/0207296`.

[74] F. Guerra, *Fluctuations and thermodynamic variables in mean field spin glass models*, in "Stochastic provesses, physics and geometry, II", S. Albeverio *et al.* eds., Singapore (1995).

[75] M. Shcherbina, *On the replica symmetric solution for the Sherrington-Kirkpatrick model*, Helv. Phys. Acta **70**, 838 (1997).

[76] M. Talagrand, *Replica symmetry breaking and exponential inequalities for the Sherrington Kirkpatrick model*, Ann. Probab. **28**, 1018 (2000).

[77] M. Talagrand, *On the high temperature region of the Sherrington-Kirkpatrick model*, Ann. Probab., to appear.

[78] S. Caracciolo, G. Parisi, S. Paternello, N. Sourlas, *Low temperature behaviour of 3-D spin glasses in a magnetic field*, J. Phys. France **51**, 1877 (1990).

[79] F. Guerra, F. L. Toninelli, *Quadratic replica coupling for the Sherrington-Kirkpatrick mean field spin glass model*, J. Math. Phys. **43**, 3704 (2002), `cond-mat/0201091`.

[80] D. Iagolnitzer, B. Souillard, *Lee-Yang theory and normal fluctuations*, Phys. Rev. B **19:3**, 1515 (1979).

[81] C. M. Newman, *A General Central Limit Theorem for FKG Systems*, Commun. Math. Phys. **91**, 75 (1983).

[82] C. N. Yang, T. D. Lee, *Statistical Theory of Equations of State and Phase Transitions. I. Theory of condensation*, Phys. Rev. **87**, 404 (1952).

T. D. Lee, C. N. Yang, *Statistical Theory of Equations of State and Phase Transitions. II. Lattice Gas and Ising Model*, Phys. Rev. **87**, 410 (1952).

[83] C. M. Fortuin, P. W. Kasteleyn, J. Ginibre, *Correlation Inequalities on Some Partially Ordered Sets*, Commun. Math. Phys. **22**, 89 (1971).

[84] F. Guerra, F. L. Toninelli, *Central limit theorem for fluctuations in the high temperature region of the Sherrington-Kirkpatrick spin glass model*, J. Math. Phys., to appear, `cond-mat/0201092`.

[85] A. Bovier, I. Kurkova, M. Löwe, *Fluctuations of the free energy in the REM and the p-spin SK models*, to appear on Ann. Probab.

[86] S. Ghirlanda, F. Guerra, *General properties of overlap distributions in disordered spin systems. Towards Parisi ultrametricity*, J. Phys. A, **31** 9149-9155 (1998).

[87] M. Aizenman, P. Contucci, *On the stability of the quenched state in mean field spin glass models*, J. Stat. Phys. **92**, 765 (1998).

[88] F. Guerra, *The cavity method in the mean field spin glass model. Functional reperesentations of thermodynamic variables*, in "Advances in dynamical systems and quantum physics", S. Albeverio *et al.* eds., Singapore (1995).