Advanced Statistical Physics:
Quenched disordered systems

Leticia F. Cugliandolo
`leticia@lpthe.jussieu.fr`
Sorbonne Université
Laboratoire de Physique Théorique et Hautes Energies

December 3, 2018

Contents

1 Random fields, random interactions

No material is perfectly homogeneous: impurities of different kinds are distributed randomly throughout the samples. In ultra-cold atom systems, so much studied nowadays, disorder can be realised, for example, using speckle laser light. It is quite natural to expect that disorder will change the equilibrium and dynamical properties of the systems on which it acts.

A natural effect of disorder should be to lower the critical temperature of a macroscopic interacting system. Much attention has been payed to the effect of *weak disorder* on phase transitions, that is to say, situations in which the nature of the ordered and disordered phases is not modified by the impurities but the critical phenomenon is. On the one hand, the critical exponents of second order phase transitions might be modified by disorder, on the other hand, disorder may smooth out the discontinuities of first order phase transitions rendering them of second order. *Strong disorder* instead changes the nature of the low-temperature phase and before discussing the critical phenomenon one needs to understand how to characterise the new ordered 'glassy' phase.

In this Section we shall discuss several types of *quenched disorder* and models that account for it. We shall also overview some of the theoretical methods used to deal with the static properties of models with quenched disorder, namely, scaling arguments and the droplet theory, mean-field equations, and the replica method.

1.1 Quenched and annealed disorder

Imagine that one mixes some random impurities in a melt and then very slowly cools it down in such a way that the impurities and the host remain in thermal equilibrium. If the interest sets on the statistical properties of the full system, one has to compute the full partition function in which a sum over all configurations of the host components and the impurities has to be performed. This is called *annealed disorder*.

In the opposite case in which upon cooling the impurities do not equilibrate with the host nor with the environment but remain blocked in random fixed positions, one talks about *quenched disorder*. Basically, the relaxation time associated with the diffusion of the impurities in the sample is so long that these remain trapped:

$$t_{\rm obs} \sim 10^4 \text{ sec} \ll t_{\rm diff} \; . \tag{1.1}$$

As concerns the host variables, they have their own, typically microscopic time scale that is much smaller than the observational time scale and therefore fluctuate. For example, in magnetic system, this time scale is the typical time-scale needed to reverse a spin, $\tau_o \sim 10^{-12} - 10^{-15}$ sec $\ll t_{\rm obs}$.

The annealed case is easier to treat analytically but it brings in less theoretical novelties and, in many cases of interest, is of less physical relevance. The quenched one is the one that leads to new phenomena and ideas that we shall discuss next.

Quenched disorder is static. Instead, in annealed disorder the impurities are in thermal equilibrium in the experimental time-scales, and they can simply be included in the statistical mechanic description of the problem, by summing over their degrees of freedom in the partition function.

1.2 Properties

Let us list a number of properties of systems with frozen-in randomness.

1.2.1 Lack of homogeneity

It is clear that the presence of quenched disorder, in the form of random interactions, fields, dilution, *etc.* breaks spatial homogeneity and renders single samples heterogeneous. Homogeneity is recovered though, if one performs an average of all possible realisations of disorder, each weighted with its own probability.

1.2.2 Frustration

We already discussed *frustration* in the context of magnetic models without disorder. It is quite clear that disorder will also introduce frustration in magnetic (and other) systems.

An example of an Ising model with four sites and four links is shown in Fig. 1.1-left, where we took three positive exchanges and one negative one all, for simplicity, with the same absolute value, J. Four configurations minimise the energy, $E_f = -2J$, but none of them satisfies the lower link. One can easily check that any closed loop such that the product of the interactions takes a negative sign is frustrated. Frustration naturally leads to a *higher energy* and a *larger degeneracy* of the number of ground states. This is again easy to grasp by comparing the number of ground states of the frustrated plaquette in Fig. 1.1-left to its unfrustrated counterpart shown on the central panel. Indeed, the energy and degeneracy of the ground state of the unfrustrated plaquette are $E_u = -4J$ and $n_u = 2$, respectively.

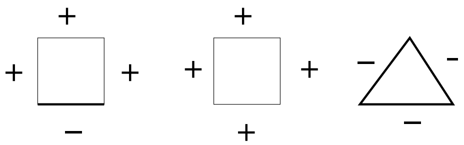

Figure 1.1: A frustrated (left) and an unfrustrated (center) square plaquette for an Ising model with nearest-neighbour interactions. A frustrated triangular plaquette (right).

Frustration may also be due to pure geometrical constraints. The canonical example is an antiferromagnet on a triangular lattice in which each plaquette is frustrated, see Fig. 1.1-right.

In short, frustration arises when the geometry of the lattice and/or the nature of the interactions make impossible the simultaneous minimisation of the energy of all pair couplings between the spins. Any loop of connected spins is said to be frustrated if the product of the signs of connecting bonds is negative. In general, energy and entropy of the ground states increase due to frustration.

Later in this Section, in Eq. (1.16), we will introduce the interaction energy between any pair of spins in a spin-glass sample. Depending on the value of the distance r_{ij} the numerator in this expression can be positive or negative implying that both ferromagnetic and antiferromagnetic interactions exist. This leads to frustration, which means that in any configuration some two-body interactions remain unsatisfied. In other words, there is no spin configuration that minimises all terms in the Hamiltonian.

1.2.3 Random parameters

Each given sample has its own peculiar realisation of the quenched disorder (the interactions between the fluctuating and the frozen-in variables) that is determined by the way in which the sample was prepared. It would be illusory, and quite impossible, to know all details about it. The idea put forward by several theoreticians is to consider that quenched disorder will be typical and hence modelled by random exchanges, fields or potentials, taken from time-independent *probability distribution functions*. What one has to determine first are the characteristics of these distributions (functional form, mean, variance, momenta). Having this modellisation in mind, one then talks about *quenched randomness*.

Exercise 1.1 This exercise provides a useful example of the distinction between *typical and average* values of random variables. Consider a random variable x that takes only two values $x_1 = e^{\alpha\sqrt{N}}$ and $x_2 = e^{\beta N}$, with α and β two positive and finite numbers with α unconstrained and $\beta > 1$. The probabilities of the two events are $p_1 = 1 - e^{-N}$ and $p_2 = e^{-N}$. First, confirm that these probabilities are normalised. Second, compute the average $\langle x \rangle$, where the angular brackets indicate average with the probabilities p_1, p_2, and evaluate it in the limit $N \to \infty$. Third, calculate the most probable value taken by x, that we call $x_{\rm typ}$, for typical (indeed, if we were to draw the variable we would typically get this value). Compare and conclude. Now, let us study the behaviour of the quantity $\ln x$ that is also a random variable. Compute its average. By which value of x is it determined? Does $\langle \ln x \rangle = (\ln x)_{\rm typ}$ in the large N limit? Is $\langle \ln x \rangle = \ln\langle x \rangle$? The last result demonstrate the difference between what are called *quenched and annealed* averages. Which value is larger? Does the comparison comply with Jenssen's inequality? (See App. A for its definition.)

1.2.4 Self-averageness

Say that the quenched randomness is given by random exchanges (*i.e.*, random J_{ij} in an Ising model). If each sample is characterised by its own realisation of the exchanges, should one expect a totally different behaviour from sample to sample? Fortunately, many generic static and dynamic properties of spin-glasses (and other systems with quenched disorder) do not depend on the specific realisation of the random couplings and are *self-averaging*.

Owing to the fact that each disorder configuration has a probability of occurrence, each physical quantity A has a probability distribution $P(A)$ given by $P(A) = \sum_J p(J)\delta(A - A(J))$ where we denoted J a generic disorder realisation. When the size of the system increases one expects (and even prove in some cases) that the distribution $P(A)$ becomes narrower and narrower. Therefore the only quantity which can be observed in the thermodynamic limit is the most probable, or typical value, A_J^{typ}, the value around which most of the distribution is concentrated.

For some quantities the typical value is very close to the average over the disorder and

$$A_J^{\text{typ}} = [\, A_J \,] \tag{1.2}$$

in the thermodynamic limit. Henceforth, we use square brackets to indicate the average over the random couplings. More precisely, in self-averaging quantities *sample-to-sample fluctuations* with respect to the mean value are expected to be $\mathrm{O}(N^{-a})$ with N the number of variables in the system and $a > 0$. Roughly, observables that involve summing over the entire volume of the system are expected to be self-averaging. In particular, the free-energy density of models with short-ranged interactions is expected to be self-averaging in the infinite size limit. There can be, though, in the same system quantities for which $B_J^{\text{typ}} \neq [\, B_J \,]$ even in the thermodynamic limit. We will show examples of both below.

An example: the disordered Ising chain

The meaning of this property can be grasped from the solution of the random bond Ising chain defined by the energy function $H_J[\{s_i\}] = -\sum_i J_i s_i s_{i+1}$ with spin variables $s_i = \pm$, for $i = 1, \ldots, N$ and random bonds J_i independently taken from a probability distribution $P(J_i)$. For simplicity, we consider periodic boundary conditions. The disorder-dependent partition function reads

$$Z_J = \sum_{\{s_i = \pm 1\}} e^{\beta \sum_i J_i s_i s_{i+1}} \tag{1.3}$$

and this can be readily computed introducing the change of variables $\sigma_i \equiv s_i s_{i+1}$. (Note that these new variables are not independent, since they are constrained to satisfy $\prod_i \sigma_i = 1$ and one should take it into account to perform the sum. However, this constraint becomes irrelevant in the thermodynamic limit and one can simply ignore it.) One finds

$$Z_J = \prod_i 2\cosh(\beta J_i) \qquad \Rightarrow \qquad -\beta F_J = \sum_i \ln\cosh(\beta J_i) + N \ln 2 \; . \tag{1.4}$$

The partition function is a *product* of *i.i.d.* random variables and it is itself a random variable with a log-normal distribution. The free-energy density instead is a *sum* of *i.i.d.* random variables and, using the central limit theorem, in the large N limit becomes a Gaussian random variable narrowly peaked at its maximum. The typical value, given by the maximum of the Gaussian distribution, coincides with the average, $\lim_{N\to\infty}(f_J^{\rm typ} - [\,f_J\,]) = 0$ with $f_J = F_J/N$.

Exercise 1.2 Take a one dimensional Ising model with a Gaussian probability distribution of the interaction strengths J_i, with zero mean and variance J^2. Draw histograms of the partition function Z_J, the total free-energy F_J, and the free-energy density f_J. Study these for increasing value of N. Conclude. Repeat the analysis for different probability distributions of the interaction strengths. In particular, consider distribution functions with fat tails, that is to say, with power law decays. What is the difference?

General argument

A simple argument justifies the self-averageness of the free-energy density in generic finite dimensional systems with short-range interactions. Let us divide a, say, cubic system of volume $V = L^d$ in n subsystems, say also cubes, of volume $v = \ell^d$ with $V = nv$. If the interactions are short-ranged, the total free-energy is the sum of two terms, a contribution from the bulk of the subsystems and a contribution from the interfaces between the subsystems: $-\beta F_J = \ln Z_J = \ln \sum_{\rm conf} e^{-\beta H_J({\rm conf})} \approx \ln \sum_{\rm conf} e^{-\beta H_J({\rm bulk}) - \beta {\rm H}_J({\rm surf})} = \ln \sum_{\rm bulk} e^{-\beta H_J({\rm bulk})} + \ln \sum_{\rm surf} e^{-\beta H_J({\rm surf})} = -\beta F_J^{\rm bulk} - \beta F_J^{\rm surf}$ (we neglected the contributions from the interaction between surface and bulk). If the interaction extends over a short distance σ and the linear size of the boxes is $\ell \gg \sigma$, the surface energy is negligible with respect to the bulk one and $-\beta F_J \approx \ln \sum_{\rm bulk} e^{-\beta H_J({\rm bulk})}$. In the thermodynamic limit, the disorder dependent free-energy is then a sum of $n = (L/\ell)^d$ random numbers, each one being the disorder dependent free-energy of the bulk of each subsystem: $-\beta F_J \approx \sum_{k=1}^n \ln \sum_{{\rm bulk}_k} e^{-\beta H_J({\rm bulk}_k)}$. In the limit of a very large number of subsystems ($L \gg \ell$ or $n \gg 1$) the central limit theorem (see App. B) implies that the free-energy density is Gaussian distributed with the maximum reached at a value $f_J^{\rm typ}$ that coincides with the average over all realisations of the randomness $[\,f_J\,]$. Moreover, the dispersion about the typical value of the total free-energy vanishes in the large n limit, $\sigma_{F_J}/[\,F_J\,] \propto \sqrt{n}/n = n^{-1/2} \to 0$ and the one of the free-energy density, or intensive free-energy, $f_J = F_J/N$, as well, $\sigma_{f_J}/[f_J] = O(n^{-1/2})$. In a sufficiently large system the typical f_J is then very close to the averaged $[\,f_J\,]$ and one can compute the latter to understand the static properties of typical systems. This is very convenient from a calculation point of view.

Lack of self-averageness in the correlation functions

Once one has $[f_J]$, one derives all disordered average thermal averages by taking derivatives of the disordered averaged free-energy with respect to sources introduced in the partition function. For example,

$$[\langle s_i \rangle] = -\left.\frac{\partial [F_J]}{\partial h_i}\right|_{h_i=0} , \tag{1.5}$$

$$[\langle s_i s_j \rangle - \langle s_i \rangle \langle s_j \rangle] = -T \left.\frac{\partial^2 [F_J]}{\partial h_i h_j}\right|_{h_i=0} , \tag{1.6}$$

with $H_J \to H_J - \sum_i h_i s_i$. Connected correlation functions, though, *are not* self-averaging quantities. This can be seen, again, studying the random bond Ising chain. Take $i < j$. One can easily check that

$$\langle s_i s_j \rangle_J - \langle s_i \rangle_J \langle s_j \rangle_J = Z_J^{-1} \frac{\partial}{\partial \beta J_{j-1}} \dots \frac{\partial}{\partial \beta J_i} Z_J = x \tanh(\beta J_i) \dots \tanh(\beta J_j) , \tag{1.7}$$

where we used $\langle s_i \rangle = 0$ (valid for a distribution of random bonds with zero mean) and the dots indicate all sites on the chain between the ending points i and j, *i.e.* $i+1 \leq k \leq j-1$. The last expression is a product of random variables and it is not equal to its average (1.6) – not even in the large separation limit $|\vec{r}_i - \vec{r}_j| \to \infty$.

See the TD on correlation functions for more details on them.

1.2.5 Annealed disorder

The thermodynamics of a system with annealed disorder is obtained by averaging the partition function over the impurity degrees of freedom,

$$Z = [Z_J] \tag{1.8}$$

since one needs to do the partition sum over the disorder degrees of freedom as well. In general, this calculation does not present any particular difficulty. In some cases, the annealed average gives a good description of the high temperature phases of problems with quenched randomness but it fails to predict the phase transition and the low temperature properties correctly.

In general, one can prove

$$f^{\text{quenched}} \geq f^{\text{annealed}} . \tag{1.9}$$

Exercise 1.3 Prove this inequality using Jenssen's inequality (see App. A).

1.3 Percolation (geometry)

The understanding of *fluid flow in porous media* needs, as a first step, the understanding of the static geometry of the connected pores. The typical example, that gave the name to the problem, is coffee percolation, where a solvent (water) filter or trickle through the permeable substance that is the coffee grounds and in passing picks up soluble constituents (the chemical compounds that give coffee its color, taste, and aroma).

Another problem that needs the comprehension of a static random structure is the one of *conduction across a disordered sample*. Imagine that one mixes randomly a set of conducting and insulating islands. Whether the mix can conduct an electric current from one end to the other of the container is the question posed, and the answer depends on the structure formed by the conducting islands.

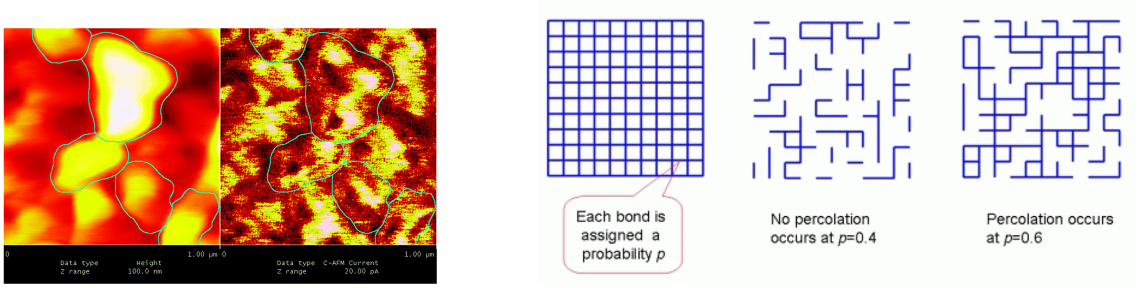

Figure 1.2: Left: a measurement of the topography (left) and local current (right) in an inhomogeneous mixture of good and bad conducting polymers. The brighter the zone the more current passing through it. Several grains are contoured in the left image. Right: an example of bond percolation.

Percolation [3, 4, 5, 6] is a simple *geometric* problem with a *critical threshold*. It is very helpful since it allows one to become familiar with important concepts of *critical phenomena* such as fractals, scaling, and renormalisation group theory in a very intuitive way. Moreover, it is not just a mathematical model, since it is at the basis of the understanding of the two physical problems mentioned above among many others.

Site dilute lattices with missing vertices are intimately related to the site percolation problem. Imagine that one builds a lattice by occupying a site with probability p (and not occupying it with probability $1-p$). For $p=0$ the lattice will be completely empty while for $p=1$ is will be totally full. For intermediate values of p, on average, order pL^d sites will be occupied, with L the linear size of the lattice. *Site percolation theory* is about the geometric and statistical properties of the structures thus formed. In particular, it deals with the behaviour of the clusters of nearest neighbour occupied sites.

Similarly, one can construct bond dilute lattices and compare them to the *bond percolation* problem.

The site percolation problem describes, for example, a binary alloy or dilute ferromagnetic crystal, also called a doped ferromagnet. The question in this context is how much dilution is needed to destroy the ferromagnetic order in the sample at a given temperature. The bond percolation problem corresponds to a randomly blocked maze through which the percolation of a fluid can occur. Many other physical problems can be set in terms of percolation: the distribution of grain size in sand and photographic emulsions, the vulcanisation of rubber and the formation of cross-linked gels, the propagation of an infection, etc.

The main interest lies on characterising the statistical and geometric properties of the *clusters* on a lattice of linear size L as a function of the probability p. The clusters are connected ensembles of nearest neighbour sites. Their easiest geometric property is their *size*, defined as the number of sites that compose them. Other geometric properties are also interesting and we will define them below.

The percolation problem is specially interesting since it has a threshold phenomenon, with a critical value p_c at which a first spanning cluster that goes from one end of the lattice to the opposite in at least one of the Cartesian directions appears. For $p < p_c$ there are only finite clusters, for $p > p_c$ there is a *spanning cluster* as well as finite clusters.

The first natural question is whether the value p_c depends on the particular sample studied or not, that is to say, whether it suffers from *sample-to-sample fluctuations*. All samples are different as the sites erased or the links cut are not the same. The threshold value is therefore a random variable and it does not take the same value for different samples. The 'surprise' is that the mean-square deviations of p_c from its mean value vanish as a power law with the system size,

$$\delta^2_{p_c}(N) \equiv \frac{1}{\mathcal{N}} \sum_{k=1}^{\mathcal{N}} (p_c^{(k)} - \overline{p}_c)^2 \simeq C^2 N^{-\nu} \;, \qquad \overline{p}_c \equiv \frac{1}{\mathcal{N}} \sum_{k=1}^{\mathcal{N}} p_c^{(k)} \;, \tag{1.10}$$

with k labelling different measurements and $\mathcal{N}$ counting its total number. N the number of sites in the sample. (C turns out to be 0.54 and $\nu = 1.3$ in $d = 2$.) In the infinite system size limit, p_c does not fluctuate from sample to sample.

One can then count the number of sites belonging to the largest cluster and compare this number to the total number of sites in the sample:

$$r_L(p) \equiv \frac{N_{\max}(p)}{N} \;. \tag{1.11}$$

This is, again, a fluctuating quantity that, in the infinite system size limit does no longer fluctuate and defines

$$r_\infty \equiv \lim_{L\to\infty} r_L(p) \;. \tag{1.12}$$

The precise definition of the *critical threshold* p_c involves the infinite size limit and it can be given by

$$r_\infty(p) = \lim_{L\to\infty} r_L(p) = \begin{cases} 0 & \text{for} \quad p < p_c \\ > 0 & \text{for} \quad p > p_c \end{cases} \tag{1.13}$$

where $r_\infty(p)$ denotes the fraction of sites belonging to the largest cluster in the finite lattice with linear size L. In the magnetic application of percolation, this means that the magnetisation vanishes for $p < p_c$ and it takes the value that the magnetisation takes on the largest cluster for $p > p_c$ (as in both cases the magnetisation on the finite clusters is independent and averages to zero).

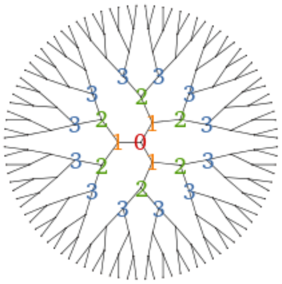

Figure 1.3: A Bethe lattice with coordination number $z = 3$. The root is labeled 0 and the generations 1, 2 and 3 are shown with different colours. Figure taken from Wikipedia.

An equivalent definition of the critical threshold p_c is given by

$$P_\infty(p) = \lim_{L\to\infty} P_L(p) = \begin{cases} 0 & \text{for} \quad p < p_c \\ 1 & \text{for} \quad p > p_c \end{cases} \tag{1.14}$$

where $P_L(p)$ denotes the probability of there being a percolating cluster in the finite lattice with linear size L.

The percolation threshold p_c depends on the lattice geometry and its dimensionality. Moreover, it is not the same for bond percolation and site percolation. Exact results are known for special lattices as the Cayley tree. Examples of how these results are found are given in [3]. Numerical data for finite dimensional lattices are complemented by rigorous upper and lower bounds and the outcome of series expansions for the mean cluster value. Harris showed that $p_c \geq 1/2$ for the bond percolation problem on a planar square lattice and the numerics suggests $p_c = 1/2$. Fisher put several bounds on p_c on various $2d$ lattices for the site and bond problem. In particular, $p_c \geq 1/2$ for site percolation on planar regular lattices with no crossings.

1.4 Bethe lattices and random graphs

The *Bethe lattice* is a tree, in which each site has z neighbours and each branch gives rise to $z-1$ new branches, see Fig. 1.3. Two important properties of these lattices are:
- there are no closed loops.
- the number of sites on the border is of the same order of magnitude as the total number of sites on the lattice.
- It is a *rooted tree*, with all other nodes arranged in shells around the root node, also called the origin of the lattice.

Exercise 1.4 Show that the total number of sites on the Bethe lattice with $z = 3$ and g generations (or the distance from the site designed as the central one) is $n_{\text{tot}} = 3\ 2^{g-1}$. Prove that the surface to volume ratio tends to $1/2$.

Due to its distinctive topological structure, the statistical mechanics of lattice models on this graph are often exactly solvable.

Exercise 1.5 Take a hypercubic lattice in d dimensions and estimate the surface to volume ratio. Show that this ratio tends to a finite value only if $d \to \infty$.

A *random graph* is obtained by starting with a set of n isolated vertices and adding successive edges between them at random. A popular ensemble is the one denoted $G(n,p)$, in which every possible edge occurs independently with probability $0 < p < 1$. Random graphs with fixed connectivity are also commonly used.

Random graphs are used in social sciences modeling (nodes representing individuals and edges the friendship relationship), technology (interconnections of routers in the Internet, pages of the WWW, or production centers in an electrical network), biology (interactions of genes in a regulatory network) [1, 2]. Disordered systems are usually defined on random graphs, especially the ones motivated by combinatorial optimisation.

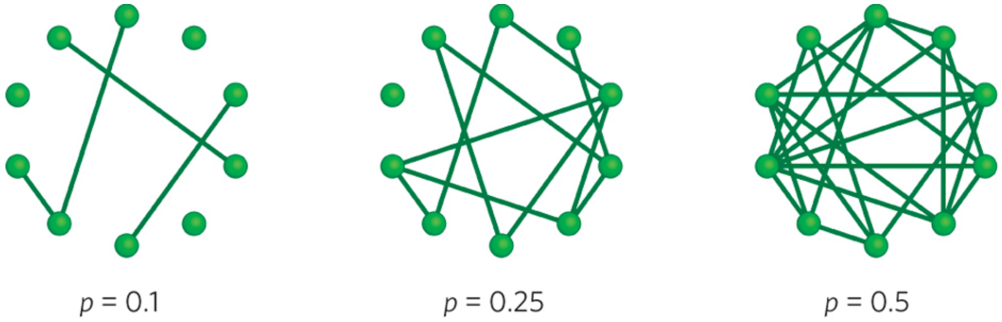

Figure 1.4: Random graphs with $N = 10$ and different probabilities p of joining two nodes.

1.5 Energetic models

Let us briefly describe here some representative models with quenched randomness.

1.5.1 Dilute spin models

Lattice models with site or link dilution are

$$H_J^{\text{site dil}} = -J \sum_{\langle ij \rangle} s_i s_j \epsilon_i \epsilon_j \; , \qquad H_J^{\text{link dil}} = -J \sum_{\langle ij \rangle} s_i s_j \epsilon_{ij} \; , \tag{1.15}$$

with $P(\epsilon_i = 1, 0) = p, 1-p$ in the first case and $P(\epsilon_{ij} = 1, 0) = p, 1-p$ in the second. These models are intimately related to Percolation theory [3, 4, 5, 6]. Physically, dilution is realised by vacancies or impurity atoms in a crystal.

1.5.2 Spin-glass models

Spin-glasses are alloys in which magnetic impurities substitute the original atoms in positions randomly selected during the chemical preparation of the sample [7, 8, 9, 10, 11, 12, 13]. The interactions between the impurities are of RKKY type:

$$V_{\rm rkky} = -J \, \frac{\cos(2k_F r_{ij})}{r_{ij}^3} \, s_i s_j \tag{1.16}$$

with $r_{ij} = |\vec{r}_i - \vec{r}_j|$ the distance between them and s_i a spin variable that represents their magnetic moment. Clearly, the initial location of the impurities varies from sample to sample. The time-scale for diffusion of the magnetic impurities is much longer than the time-scale for spin flips. Thus, for all practical purposes the positions $\vec{r}_i$ can be associated to quenched random variables distributed according to a uniform probability distribution that in turn implies a probability distribution of the exchanges. This is called *quenched disorder.*

In early 70s *Edwards and Anderson* proposed a rather simple model that should capture the main features of spin-glasses [15]. The interactions (1.16) decay with a cubic power of the distance and hence they are relatively short-ranged. This suggests to put the spins on a regular cubic lattice model and to trade the randomness in the positions into random nearest neighbour exchanges, independently and identically distributed according to a Gaussian probability distribution function (pdf):

$$H_J^{\rm EA} = -\sum_{\langle ij \rangle} J_{ij} s_i s_j \qquad \text{with} \qquad P(J_{ij}) = (2\pi\sigma^2)^{-\frac{1}{2}} \, e^{-\frac{J_{ij}^2}{2\sigma^2}} \; . \tag{1.17}$$

The precise form of the probability distribution of the exchanges is supposed not to be important (though some authors claimed that there might be non-universality with respect to it, see however [20] where this is refuted at least in the random field Ising model case).

Another natural choice is to use bimodal exchanges

$$P(J_{ij}) = p\delta(J_{ij} - J_0) + (1-p)\delta(J_{ij} + J_0) \tag{1.18}$$

with the possibility of a bias towards positive or negative interactions depending on the parameter p. A tendency to non-zero average J_{ij} can also be introduced in the Gaussian pdf.

A natural extension of the EA model in which all spins interact has been proposed by Sherrington and Kirkpatrick

$$H_J^{\rm SK} = -\sum_{i \neq j} J_{ij} s_i s_j \tag{1.19}$$

and it is called the *SK model* [16]. The interaction strengths J_{ij} are taken from a Gaussian pdf and they scale with N in such a way that the thermodynamic limit is non-trivial:

$$P(J_{ij}) = (2\pi\sigma_N^2)^{-\frac{1}{2}} \, e^{-\frac{J_{ij}^2}{2\sigma_N^2}} \qquad\qquad \sigma_N^2 = J^2 N \; . \tag{1.20}$$

The first two-moments of the exchange distribution are $[J_{ij}] = 0$ and $[J_{ij}^2] = J^2/N$. This is a case for which a mean-field theory is expected to be exact.

1.5.3 Glass models

A further extension of the EA model is called the *p-spin model*

$$H_J^{\mathrm{p-spin}} = - \sum_{i_1 < \dots < i_p} J_{i_1 \dots i_p} s_{i_1} \dots s_{i_p} \tag{1.21}$$

with $p \geq 3$. The sum can also be written as $\sum_{i_1 < i_2 < \dots < i_p} = 1/p! \sum_{i_1 \neq i_2 \neq \dots \neq i_p}$. The exchanges are now taken from a Gaussian probability distribution

$$P(J_{ij}) = (2\pi\sigma_N^2)^{-\frac{1}{2}}\, e^{-\frac{J_{ij}^2}{2\sigma_N^2}} \qquad\qquad \sigma_N^2 = J^2 p!/(2N^{p-1}) \; . \tag{1.22}$$

with $[J_{i_1 \dots i_p}] = 0$ and $[J_{i_1 \dots i_p}^2] = \frac{J^2 p!}{N^{p-1}}$. Indeed, an extensive free-energy is achieved by scaling $J_{i_1 \dots i_p}$ with $N^{-(p-1)/2}$. This scaling can be justified as follows. Imagine that at low temperatures the spins acquire local equilibrium expectation values that we call m_i. The 'local field' that they induce are $h_i = 1/(p-1)! \sum_{ii_2 \neq i_p} J_{ii_2 \dots i_p} m_{i_2} \dots m_{i_p}$ and they should be of order one. Contrary to ferromagnetic models, the m_i's take plus and minus signs in the disordered case as there is no tendency to align all moments in the same direction. In particular, we estimate the order of magnitude of this term by working at $T = 0$ and taking $m_i = \pm 1$ with probability $\frac{1}{2}$, since there is no external magnetic field nor a non-vanishing mean of the exchanges that could bias the local order in one or the other direction. In order to keep the discussion simple, let us take $p = 2$. In this case, if the strengths J_{ij}, are of order one, h_i is a sum of N *i.i.d.* random variables, with zero mean and unit variance,[1] and h_i is a random variable with zero mean and variance equal to N. Therefore, one can argue that h_i is of order $\sqrt{N}$. To make it finite we then chose J_{ij} to be of order $1/\sqrt{N}$ or, in other words, we impose $[J_{ij}^2] = J^2/N$. The generalization to $p > 2$ is straightforward.

We classify this model in the "glass" class since it has been shown that its behaviour mimics the one of so-called fragile glasses for $p > 2$ [17, 18, 19].

1.5.4 Vector spins

Extensions to *vector spins* with two (XY), three (Heisenberg) or N components also exist. In the former cases can be relevant to describe real samples. One usually keeps the modulus of the spins fixed to be 1 in these cases.

1.5.5 Spherical model

But there is another way to extend the spin variables and it is to use a *spherical*

---

[1] The calculation goes as follow: $\langle F_i \rangle = \sum_j J_{ij} \langle m_j \rangle = 0$ and $\langle F_i^2 \rangle = \sum_{jk} J_{ij} J_{ik} \langle m_j m_k \rangle = \sum_j J_{ij}^2$

constraint [21, 22],

$$-\infty \leq s_i \leq \infty \qquad\qquad \sum_{i=1}^{N} s_i^2 = N \; . \tag{1.23}$$

In this case, the spins s_i are the components of an N-dimensional vector, constrained to be an N-dimensional sphere.

1.5.6 Optimization problems

Cases that find an application in **computer science** [23, 24] are defined on random graphs with fixed or fluctuating finite connectivity. In the latter case one places the spins on the vertices of a graph with links between couples or groups of p spins chosen with a probability c. These are *dilute spin-glasses* on graphs (instead of lattices).

Optimisation problems can usually be stated in a form that requires the minimisation of a cost (energy) function over a large set of variables. Typically these **cost functions** have a very large number of local minima – an exponential function of the number of variables – separated by barriers that scale with N and finding the truly absolute minimum is hardly non-trivial. Many interesting optimisation problems have the great advantage of being defined on random graphs and are then mean-field in nature. The mean-field machinery that we will discuss at length is then applicable to these problems with minor (or not so minor) modifications due to the finite connectivity of the networks.

Let us illustrate this kind of problems with two examples. The **graph partitioning** problem consists in, given a graph $G(N, E)$ with N vertices and E edges, to partition it into smaller components with given properties. In its simplest realisation the uniform graph partitioning problem is how to partition, in the optimal way, a graph with N vertices and E links between them in two (or k) groups of equal size $N/2$ (or N/k) and the minimal number of edges between them. Many other variations are possible. This problem is encountered, for example, in computer design where one wishes to partition the circuits of a computer between two chips. More recent applications include the identification of clustering and detection of cliques in social, pathological and biological networks.

Another example, that we will map to a spin model, is **k-satisfiability (k-SAT)**. The problem is to determine whether the variables of a given Boolean formula can be assigned in such a way to make the formula evaluate to 'TRUE'. Equally important is to determine whether no such assignments exist, which would imply that the function expressed by the formula is identically 'FALSE' for all possible variable assignments. In this latter case, we would say that the function is unsatisfiable; otherwise it is satisfiable.

We illustrate this problem with a concrete example. Let us use the convention x for the requirement $x =$ TRUE and $\overline{x}$ for the requirement $x =$ FALSE. For example, the formula $C_1 : x_1$ OR $\overline{x}_2$ made by a single clause C_1 is satisfiable because one can find the values x_1 = TRUE (and x_2 free) or $x_2 =$ FALSE (and x_1 free), which make $C_1 : x_1$ OR $\overline{x}_2$ TRUE. This formula is so simple that 3 out of 4 possible configurations of the two variables solve it. This example belongs to the $k = 2$ class of satisfiability problems since the clause is made by two literals (involving different variables) only. It has $M = 1$ clauses and $N = 2$ variables.

Harder to decide formulæ are made of M clauses involving k literals required to take the true value (x) or the false value ($\overline{x}$) each, these taken from a pool of N variables. An example in 3-SAT is

$$\mathrm{F} = \begin{cases} C_1 \;:\; x_1 \text{ OR } \overline{x}_2 \text{ OR } x_3 \\ C_2 \;:\; \overline{x}_5 \text{ OR } \overline{x}_7 \text{ OR } x_9 \\ C_3 \;:\; x_1 \text{ OR } \overline{x}_4 \text{ OR } x_7 \\ C_4 \;:\; x_2 \text{ OR } \overline{x}_5 \text{ OR } x_8 \end{cases} \tag{1.24}$$

All clauses have to be satisfied simultaneously so the formula has to be read

$$\mathrm{F}\text{: } C_1 \text{ AND } C_2 \text{ AND } C_3 \text{ AND } C_4 \quad . \tag{1.25}$$

It is not hard to believe that when $\alpha \equiv M/N \gg 1$ the problems typically become unsolvable while many solutions exist for $\alpha \ll 1$. One could expect to find a sharp threshold between a region of parameters $\alpha < \alpha_c$ where the formula is satisfiable and another region of parameters $\alpha \geq \alpha_c$ where it is not.

In **random k-SAT** an instance of the problem, i.e. a formula, is chosen at random with the following procedure: first one takes k variables out of the N available ones. Second one decides to require x_i or $\overline{x}_i$ for each of them with probability one half. Third one creates a clause taking the OR of these k literals. Forth one returns the variables to the pool and the outlined three steps are repeated M times. The M resulting clauses form the final formula.

The Boolean character of the variables in the k-SAT problem suggests to transform them into Ising spins, i.e. x_i evaluated to TRUE (FALSE) will correspond to $s_i = 1$ (-1) . The requirement that a formula be evaluated TRUE by an assignment of variables (i.e. a configuration of spins) will correspond to the ground state of an adequately chosen energy function. In the simplest setting, each clause will contribute zero (when satisfied) or one (when unsatisfied) to this cost function. There are several equivalent ways to reach this goal. The fact that the variables are linked together through the clauses suggests to define k-uplet interactions between them. We then choose the interaction matrix to be

$$J_{ai} = \begin{cases} 0 & \text{if neither } x_i \text{ nor } \overline{x}_i \in C_a \\ 1 & \text{if } x_i \in C_a \\ -1 & \text{if } \overline{x}_i \in C_a \end{cases} \tag{1.26}$$

and the energy function as

$$H_J[\{s_i\}] = \sum_{a=1}^{M} \delta(\sum_{i=1}^{N} J_{ai} s_i, -k) \tag{1.27}$$

where $\delta(x, y)$ is a Kronecker-delta that equals one when the arguments are identical and zero otherwise. This cost function is easy to understand. The Kronecker delta contributes one to the sum over a only if the k non-vanishing terms in the sum $\sum_{i=1}^N J_{ai}s_i$ are equal to -1. This can happen when $J_{ai} = 1$ and $s_i = -1$ or when $J_{ai} = -1$ and $s_i = 1$. In both cases the condition on the variable x_i is not satisfied. Since this is required from all the variables in the clause, the clause itself and hence the formula are not satisfied.

Another way to represent a clause in an energy function is to consider, for instance for C_1 above, the term $(1 - s_1)(1 + s_2)(1 - s_3)/8$. This term vanishes if $s_1 = 1$ or $s_2 = -1$ or $s_3 = 1$ and does not contribute to the total energy, that is written as a sum of terms of this kind. It is then simple to see that the total energy can be rewritten in a way that resembles strongly physical spin models,

$$H_J[\{s_i\}] = \frac{M}{2^K} + \sum_{R=1}^{K} (-1)^R \sum_{i_1 < \cdots < i_R} J_{i_1 \dots i_R} s_{i_1} \dots s_{i_R} \tag{1.28}$$

and

$$J_{i_1 \dots i_R} = \frac{1}{2^K} \sum_{a=1}^{M} J_{a i_1} \dots J_{a i_R} \; . \tag{1.29}$$

These problems are “solved” numerically, with algorithms that do not necessarily respect physical rules. Thus, one can use non-local moves in which several variables are updated at once – as in cluster algorithms of the Swendsen-Wang type used to beat critical slowing down close to phase transitions – or one can introduce a temperature to go beyond cost-function barriers and use dynamic local moves that do not, however, satisfy a detail balance. The problem is that with hard instances of the optimization problem none of these strategies is successful. Indeed, one can expect that glassy aspects, such as the proliferation of metastable states separated by barriers that grow very fast with the number of variables, can hinder the resolutions of these problems in polynomial time, that is to say a time that scales with the system size as N^ζ, for *any* algorithm. These are then **hard combinatorial problems**.

1.5.7 Random bond ferromagnets

Let us now discuss some, *a priori* simpler cases. An example is the *Mattis random magnet* with generic energy (1.21) in which the interaction strengths are given by [25]

$$J_{i_1 \dots i_p} = \xi_{i_1} \dots \xi_{i_p} \qquad \text{with} \qquad \xi_j = \pm 1 \quad \text{with} \quad \text{prob} = 1/2 \tag{1.30}$$

for any p and any kind of graph. In this case a simple *gauge transformation*, $\eta_i \equiv \xi_i s_i$, allows one to transform the disordered model in a ferromagnet, showing that there was no true frustration in the system.

Random bond ferromagnets (RBFMs) are systems in which the strengths of the interactions are not all identical but their sign is always positive. One can imagine such a exchange as the sum of two terms:

$$J_{ij} = J + \delta J_{ij} \; , \qquad \text{with } J > 0 \quad \text{and} \quad \delta J_{ij} \qquad \text{small and random} \; . \tag{1.31}$$

There is no frustration in these systems either. Ising models of this kind have also been used to describe fracture in materials, where the J_{ij} represents the local force needed to break the material and it is assumed the fracture occurs along the surface of minimum total rupture force [26].

As long as all J_{ij} remain positive, this kind of disorder should not change the two bulk phases with a paramagnetic-ferromagnetic second-order phase transition. Moreover the up-down spin symmetry is not broken by the disorder. The disorder just changes the local tendency towards ferromagnetism that can be interpreted as a change in the *local critical temperature*. Consequently, this type of disorder is often called random-T_c disorder, and it admits a Ginzburg-Landau kind of description, with a random distance from criticality, $\delta u(\vec{r})$,

$$F^{\text{random mass}}[m(\vec{r})] = \int d^d r \left\{ -h\, m(\vec{r}) + [r + \delta r(\vec{r})]\, m^2(\vec{r}) + (\nabla\, m(\vec{r}))^2 + u\, m^4(\vec{r}) + \dots \right\} . \tag{1.32}$$

The disorder couples to the m^2 term in the free-energy functional. In quantum field theory, this term is called the mass term and, therefore, random-T_c disorder is also called *random-mass* disorder. (In addition to random exchange couplings, random-mass disorder can also be realized by random dilution of the spins.)

1.5.8 Random field ferromagnets

Link randomness is not the only type of disorder encountered experimentally. Random fields, that couple linearly to the magnetic moments, are also quite common; the classical model is the *ferromagnetic random field Ising model* (RFIM) [20, 27]:

$$H_J^{\text{rfim}} = -J \sum_{\langle ij \rangle} s_i s_j - \sum_i s_i h_i \qquad \text{with} \qquad P(h_i) = (2\pi\sigma^2)^{-\frac{1}{2}}\, e^{-\frac{h_i^2}{2\sigma^2}} . \tag{1.33}$$

The *dilute antiferromagnet in a uniform magnetic field* is believed to behave similarly to the ferromagnetic random field Ising model. Experimental realisations of the former are common and measurements have been performed in samples like $\mathrm{Rb_2Co_{0.7}Mg_{0.3}F_4}$.

Note that the up-down Ising symmetry is not preserved in models in the RFIMm and any spin model such that the disorder couples to the local order parameter.

In the Ginzburg-Landau description this model reads

$$F[m(\vec{r})] = \int d^d r \left\{ -h(\vec{r})\, m(\vec{r}) + r\, m^2(\vec{r}) + (\nabla\, m(\vec{r}))^2 + u\, m^4(\vec{r}) + \dots \right\} \tag{1.34}$$

where $h(\vec{r})$ is the local random variable that breaks the up-down spin symmetry. Whether or not the symmetry is broken globally depends on the probability distribution of the random fields. A particularly interesting situation arises if the distribution is even in h such that the up-down symmetry is globally preserved in the statistical sense.

Random-field disorder is generally stronger than random-mass disorder.

The random fields give rise to many metastable states that modify the equilibrium and non-equilibrium behaviour of the RFIM. In one dimension the RFIM does not order at all, in $d = 2$ there is strong evidence that the model is disordered even at zero temperature, in $d = 3$ it there is a finite temperature transition towards a ferromagnetic state [29]. Whether there is a glassy phase near zero temperature and close to the critical point is still and open problem.

The RFIM at zero temperature has been proposed to yield a generic description of material cracking through a series of avalanches. In this problem one cracking domain triggers others, of which size, depends on the quenched disorder in the samples. In a random magnetic system this phenomenon corresponds to the variation of the magnetisation in discrete steps as the external field is adiabatically increased (the time scale for an avalanche to take place is much shorter than the time-scale to modify the field) and it is accessed using Barkhausen noise experiments [30]. Disorder is responsible for the jerky motion of the domain walls. The distribution of sizes and duration of the avalanches is found to decay with a power law tail and cut-off at a given size. The value of the cut-off size depends on the strength of the random field and it moves to infinity at the critical point.

1.5.9 Random manifolds

Once again, disorder is not only present in magnetic systems. An example that has received much attention is the so-called *random manifold* [31, 32]. This is a d dimensional *directed elastic manifold* moving in an embedding $N + d$ dimensional space under the effect of a quenched random potential. The simplest case with $d = 0$ corresponds to a particle moving in an embedding space with N dimensions. If, for instance $N = 1$, the particle moves on a line, if $N = 2$ it moves on a plane and so on and so forth. If $d = 1$ one has a line that can represent a domain wall, a polymer, a vortex line, *etc.* The fact that the line is directed means it has a preferred direction, in particular, it does not have overhangs. If the line moves in a plane, the embedding space has $(N = 1) + (d = 1)$ dimensions. One usually describes the system with an N-dimensional coordinate, $\vec{\phi}$, that locates in the transverse space each point on the manifold, represented by the internal d-dimensional coordinate $\vec{r}$,

The elastic energy is $H_{\text{elas}} = \gamma \int d^d x \, \sqrt{1 + (\nabla\phi(\vec{r}))^2}$ with γ the deformation cost of a unit surface. Assuming the deformation is small one can linearise this expression and get, upto an additive constant, $H_{\text{elas}} = \frac{\gamma}{2} \int d^d r \, (\nabla\phi(\vec{r}))^2$.

Disorder is introduced in the form of a random potential energy $V(\vec{\phi}(\vec{r}), \vec{r})$ characterised by its pdf.

The random manifold model is then

$$H_V(\vec{\phi}) = \int d^d r \left[\frac{\gamma}{2} (\nabla\phi(\vec{r}))^2 + V(\vec{\phi}(\vec{r}), \vec{r}) \right] . \tag{1.35}$$

If the random potential is the result of a large number of impurities, the central limit theorem implies that its probability density is Gaussian. Just by shifting the energy scale one can set its average to zero, $[\,V\,] = 0$. As for its correlations, one typically assumes, for simplicity, that they exist in the transverse direction only:

$$[\,V(\vec{\phi}(\vec{r}),\vec{r})V(\vec{\phi}'(\vec{r}'),\vec{r}')\,] = \delta^d(\vec{r}-\vec{r}')\mathcal{V}(\vec{\phi},\vec{\phi}') \; . \tag{1.36}$$

If one further assumes that there is a statistical isotropy and translational invariance of the correlations, $\mathcal{V}(\vec{\phi},\vec{\phi}') = W/\Delta^2\,\mathcal{V}(|\vec{\phi}-\vec{\phi}'|/\Delta)$ with Δ a correlation length and $(W\Delta^{d-2})^{1/2}$ the strength of the disorder. The disorder can now be of two types: short-ranged if $\mathcal{V}$ falls to zero at infinity sufficiently rapidly and long-range if it either grows with distance or has a slow decay to zero. An example involving both cases is given by the power law $\mathcal{V}(z) = (\theta+z)^{-\gamma}$ where θ is a short distance cut-off and γ controls the range of the correlations with $\gamma > 1$ being short-ranged and $\gamma < 1$ being long-ranged.

This model also describes *directed domain walls* in random systems. One can derive it in the long length-scales limit by taking the continuum limit of the pure Ising part (that leads to the elastic term) and the random part (that leads to the second disordered potential). In the pure Ising model the second term is a constant that can be set to zero while the first one implies that the ground state is a perfectly flat wall, as expected. In cases with quenched disorder, the long-ranged and short-ranged random potentials mimic cases in which the interfaces are attracted by pinning centres ('random field' type) or the phases are attracted by disorder ('random bond' type), respectively. For instance, random bond disorder is typically described by a Gaussian pdf with zero mean and delta-correlated $[V(\vec{\phi}(\vec{r}),\vec{r}), V(\vec{\phi}'(\vec{r}'),\vec{r}')] = W\Delta^{d-2}\,\delta^d(\vec{r}-\vec{r}')\delta(\vec{\phi}-\vec{\phi}')$.

2 Properties of finite dimensional disordered systems

Once various kinds of quenched disorder introduced, a number of questions on their effect on the equilibrium and dynamic properties arise. Concerning the former:

- Are the equilibrium phases qualitatively changed by the random interactions?
- Is the phase transition still sharp, or is it rendered smoother because different parts of the system undergo the transition independently?
- If there is still a phase transition, does its order (first order *vs.* continuous) change?
- If the phase transition remains continuous, does the critical behavior, *i.e.*, the values of the critical exponents, change?

Now, for the latter:

- Is the dynamic behaviour of the system modified by the quenched randomness?

In the following we explain a series of classical results in this field: the Harris criterium, the proof of non-analyticity of the free-energy of quenched disordered systems close to

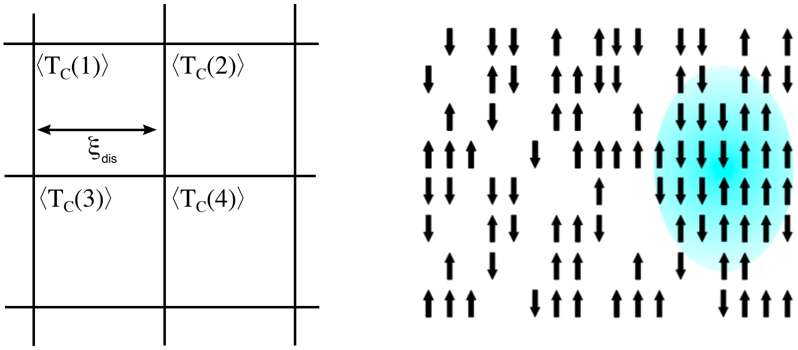

Figure 2.5: Left: scheme of the Harris construction. The disordered system is divided into cells with linear length $\xi_{\rm dis}$, its correlation length. Right: a typical configuration of the dilute Ising ferromagnet. Figures taken from [33].

their critical temperature given by Griffiths, the analysis of droplets and their domain wall stiffness, and the derivation of some exact results using the gauge invariance.

We first focus on impurities or defects that lead to spatial variations with respect to the tendency to order but do not induce new types of order, that is to say, no changes are inflicted on the two phases at the two sides of the transition. Only later we consider the spin-glass case.

2.1 The Harris criterium

The first question to ask is how does the average disorder strength behave under coarse-graining or, equivalently, how is it seen at long distances. This is the question answered by the Harris argument.

The Harris' criterion [34] states that if the specific-heat of a pure system

$$C^{\rm pure}(T) \simeq |T - T_c^{\rm pure}|^{-\alpha_{\rm pure}} \tag{2.37}$$

presents a power-like divergence with

$$\alpha_{\rm pure} > 0 \; , \tag{2.38}$$

the disorder may induce a new universality class. Otherwise, if $\alpha_{\rm pure} < 0$, disorder is irrelevant in a renormalisation group sense and the critical behaviour of the model remains unchanged. The criterium does not decide in the marginal case $\alpha_{\rm pure} = 0$ case. Note that the Harris criterium is a necessary condition for a change in critical behaviour but not a sufficient one.

The hyper-scaling relation $2 - d\nu_{\text{pure}} = \alpha_{\text{pure}}$ allows to rewrite the Harris criterium as

$$\text{critical behaviour} = \begin{cases} \text{unchanged} & \text{if } \nu_{\text{pure}} > 2/d \\ \text{may change} & \text{if } \nu_{\text{pure}} < 2/d \end{cases} \tag{2.39}$$

where ν_{pure} is the correlation length exponent

$$\langle s_0 s_{\vec{r}} \rangle \simeq e^{-r/\xi_{\text{pure}}} \qquad \text{and} \qquad \xi \simeq |T - T_c^{\text{pure}}|^{-\nu_{\text{pure}}} \ , \tag{2.40}$$

of the pure system.

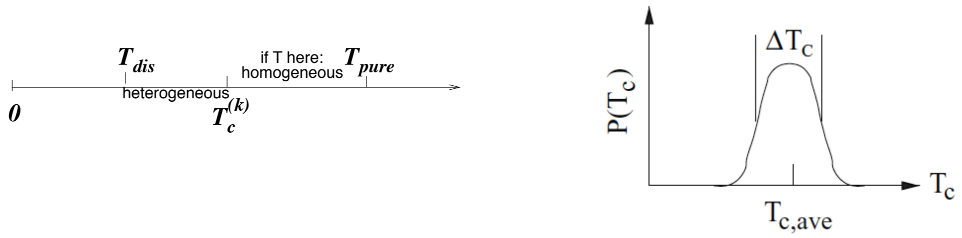

Figure 2.6: The characteristic temperatures. T_{pure} and T_{dis} are the critical temperatures of the pure and disordered systems, respectively. $T_c^{(k)}$ is the critical temperature of the local region with linear size ξ_{dis} labelled k, see the sketch in Fig. 2.5-left. The distance from the disordered critical point is measured by $\Delta T_c^{(k)} = T_c^{(k)} - T_{\text{dis}}$ for the critical temperature of block k and $\Delta T = T - T_{\text{dis}}$ for the working temperature T. Right: the probability distribution function of the local critical temperatures $T_c^{(k)}$. The width depends on ξ_{dis} and clearly decreases with increasing ξ as the local temperatures fluctuate less and less.

The proof of the Harris result is rather simple and illustrates a way of reasoning that is extremely useful [34, 33]. Take the full system with frozen-in disorder at a temperature T slightly above its critical temperature T_c^{dis}. Divide it into equal pieces with linear size ξ_{dis}, the correlation length at the working temperature. By construction, the spins within each of these blocks behave as a *super-spin* since they are effectively parallel. Because of disorder, each block k has its own local critical temperature $T_c^{(k)}$ determined by the interactions (or dilution) within the block. Harris proposes to compare the fluctuations in the local critical temperatures $\Delta T_c^{(k)} \equiv T_c^{(k)} - T_c^{\text{dis}}$ with respect to the global critical one T_c^{dis}, with the distance from the critical point $\Delta T \equiv T - T_c^{\text{dis}} > 0$, taken to be positive:

- If $\Delta T_c^{(k)} < \Delta T$ for all k, all blocks have critical temperature below the working one, $T_c^{(k)} < T$, and the system is 'homogeneous' with respect to the phase transition.
- If $\Delta T_c^{(k)} > \Delta T$ for some k, some blocks are in the disordered (paramagnetic) phase and some are in the ordered (ferromagnetic) phase, making a uniform transition impossible. The inhomogeneity in the system may then be important.

Require now $\Delta T_c^{(k)} < \Delta T$ for all k to have an unmodified critical behaviour. Use also that an unmodified critical behaviour implies $\xi_{\rm dis} = \xi_{\rm pure}$ and, consequently, $\nu_{\rm dis} = \nu_{\rm pure}$.

As this should be the case for all k we call $\Delta T_c^{({\rm loc})}$ the generic one. $\Delta T_c^{({\rm loc})}$ can be estimated using the central limit theorem. Indeed, as each local $T_c^{(k)}$ is determined by an average of a large number of random variables in the block (*e.g.*, the random J_{ij} in the Hamiltonian), its variations decay as the square root of the block volume, $\Delta T_c^{({\rm loc})} \simeq \xi_{\rm pure}^{-d/2}$. On the other hand, $\Delta T \simeq \xi_{\rm pure}^{-1/\nu_{\rm pure}}$. Therefore,

$$\Delta T_c^{({\rm loc})} < \Delta T \qquad \Rightarrow \qquad d\nu_{\rm pure} > 2 \; . \tag{2.41}$$

The interpretation of this inequality is the following. If the Harris criterion $d\nu_{\rm pure} > 2$ is fulfilled, the ratio $\Delta T_c^{({\rm loc})}/\Delta T \simeq \xi_{\rm pure}^{-d/2+1/\nu_{\rm pure}}$ goes to zero as the critical point is approached. The system looks less and less disordered on larger length scales, the effective disorder strength vanishes right at criticality, and the disordered system features the same critical behaviour as the clean one. An example of a transition that fullfills the Harris criterion is the ferromagnetic transition in a three-dimensional classical Heisenberg model. Its clean correlation length exponent is $\nu_{\rm pure} \approx 0.69 > 2/d = 2/3$.

In contrast, if $d\nu_{\rm pure} < 2$, the ratio $\Delta T_c^{({\rm loc})}/\Delta T$ increases upon approaching the phase transition. The blocks differ more and more on larger length scales. Eventually, some blocks are on one side of the transition while other blocks are on the other side. This makes a uniform sharp phase transition impossible. The clean critical behavior is unstable and the phase transition can be erased or it can remain continuous but with different critical behaviour. More precisely, the disordered system can be in a new universality class featuring a correlation length exponent that fullfills the inequality $d\nu_{\rm dis} > 2$. Many phase transitions in classical disordered systems follow this scenario, for example the three-dimensional classical Ising model. Its clean correlation length exponent is $\nu_{\rm pure} \approx 0.63$ which violates the Harris criterion. In the presence of random-mass disorder, the critical behavior changes and $\nu_{\rm dis} \approx 0.68$. (Note, however, that the difference between these exponents is tiny!)

In the marginal case $d\nu_{\rm pure} = 2$, more sophisticated methods are required to decide the stability of the clean critical point.

Chayes et al. [36] turned this argument around to show rigorously that for all the continuous phase transitions in presence of disorder, the correlation-length critical exponent of the disordered system, $\nu_{\rm dis}$ verifies $\nu_{\rm dis} \geq 2/d$, independently of whether or not the critical behaviour is the same as in the uniform system and even when the system does not have a uniform analogue.

Finally, note that the Harris criterion $d\nu_{\rm pure} > 2$ applies to uncorrelated or short-range correlated disorder. If the disorder displays long-range correlations in space, the inequality needs to be modified because the central-limit theorem estimate of $\Delta T_c^{({\rm loc})}$ changes.

Long-range correlated disorder is especially important in quantum phase transitions. The reason is the fact that the statistical properties of quantum systems are studied in an imaginary time formulation that makes a d-dimensional quantum problem equivalent to a $d+1$ dimensional classical one. Along this additional spatial direction, quenched randomness is long-range correlated.

2.2 The Griffiths phase

The critical temperature of a spin system is usually estimated from the high temperature expansion and the evaluation of its radius of convergence (see App. C.1). However, Griffiths showed that the temperature at which the free-energy of models with quenched disorder starts being non-analytical falls above the critical temperature where the order parameter detaches from zero [37]. The argument applies to models with second order phase transitions.

Griffiths explained his argument using the dilute ferromagnetic Ising model. First, he argued that the critical temperature of the disordered model should decrease for increasing p, the probability of *empty* sites. This is 'intuitively obvious' since no spontaneous magnetisation can occur at a finite temperature if the probability of occupied sites is less than the critical percolation probability at which an 'infinite cluster' first appears. See Fig. 2.7 where the phase diagram of the dilute Ising ferromagnet is shown.

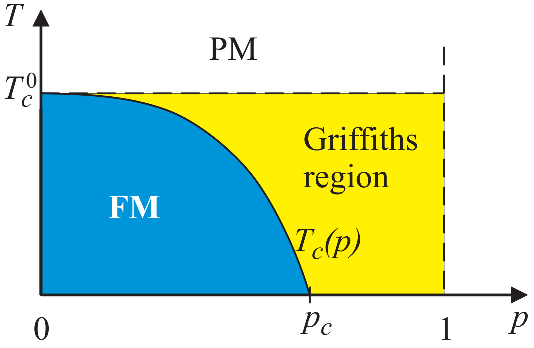

Figure 2.7: The phase diagram of the dilute ferromagnetic Ising model. p is the probability of empty sites in this figure differently from the notation in the main text, figure taken from [33]. With increasing dilution the ordered phase is eventually suppressed.

In the following paragraph we sketch Griffiths' argument and we use his notation in which p is the probability of occupying a site. For any concentration $p < 1$ the magnetisation m is not an analytic function of h at $h = 0$ at any temperature below T_c^{pure}, the critical temperature of the regular Ising model $p = 1$. As he explains, this fact is most easily explained for $p < p_c$. The magnetisation m per lattice site in the thermodynamic limit has the form

$$m = \frac{1}{N}\sum_{i=1}^{N}\langle s_i\rangle = \sum_c P(c)m(c) \tag{2.42}$$

where $P(c)$ is the probability that a particular site on the lattice belongs to a cluster c that is necessarily finite for $p < p_c$, and $m(c)$ is the magnetisation density of the cluster c, that is to say $m(c) = N^{-1}(c) \sum_{i\in c} \langle s_i \rangle$ with $N(c)$ the number of sites in the cluster.

Griffiths uses the Yang-Lee theorem, see App. C, to express $m(c)$ as

$$m(c) = 1 + \frac{2z}{N(c)} \sum_{i\in c} \frac{1}{\xi_i - z} \qquad \text{with} \qquad z = e^{-2\beta h} \tag{2.43}$$

and ξ_i, $i = 1, \ldots, N(c)$, complex numbers with $|\xi_i| = 1$. The total magnetisation density is then of the same form

$$m = 1 + z f(z) \qquad f(z) = \sum_i \eta_i (\xi_i - z)^{-1} \tag{2.44}$$

with $\eta_i = 2P(c)/N(c)$. He then argues that this form is analytic for $z < 1$ but non-analytic at $z = 1$ that corresponds to $h = 0$.

A more intuitive understanding of what is going on in the temperature region above the critical temperature of the disordered model, $T_c^{\rm dis}$, and below the critical temperature the pure one, $T_c^{\rm pure}$, can be reached as follows [33]. The effects of quenched disorder show up already in the paramagnetic phase of finite dimensional systems. Below the critical point of the pure case (no disorder) finite regions of the system can order due to fluctuations in the couplings or, in a dilute ferromagnetic model, they can be regions where all sites are occupied, as shown in Fig. 2.5. As such rare regions are finite-size pieces of the clean system, their spins align parallel to each other below the clean critical temperature $T_c^{\rm pure}$. Because they are of finite size, these regions cannot undergo a true phase transition by themselves, but for temperatures between the actual transition temperature $T_c^{\rm dis}$ and $T_c^{\rm pure}$ they act as large superspins.

Note that using the ideas of percolation theory, one can estimate the scaling of $P(c)$ with its size. Recall the one dimensional case. Take a segment of length $L + 2$ on the lattice. A cluster of size L will occupy the internal sites with empty borders with probability $p^L(1-p)^2$. This is because one needs L contiguous sites to be occupied and its boundary sites be empty. In larger dimensions, this probability will be approximately $p^{L^d}(1-p)^{L^{d-1}}$ with the first factor linked to the filled volume and the second to the empty surface. In the large L limit one can make a harsh approximation and use $\simeq \exp\{\ln[p^{L^d}(1-p)^{L^{d-1}}]\} = \exp[\ln p^{L^d} + \ln(1-p)^{L^{d-1}}] \simeq \exp[-\mathrm{c}(p)L^d]$.

The sum in eq. (2.42) is made of two contributions. On the one hand, there are the large clusters that are basically frozen at the working temperature. On the other, there are the free spins that belong to small clusters and are easy to flip at the working temperature. Let us focus on the former. Their magnetic moment is proportional to their volume $m(c) \simeq \mu L^d$. The energy gain due to their alignment with the field is $\Delta E(c) = -hm(c) = -h\mu L^d$ where h is a small uniform field applied to the system, say to measure its susceptibility.

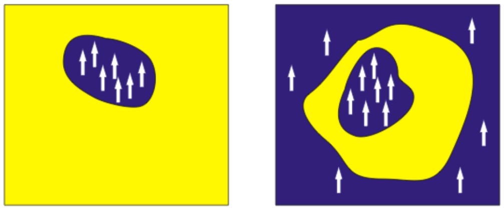

Figure 2.8: Rare regions in a random ferromagnet, figure taken from [33]. On the left, a ferromagnetically ordered region in the paramagnetic bulk ($T > T_c^{\rm dis}$). On the right, a paramagnetic band in a system that is ordered ferromagnetically in a patchwork way ($T < T_c^{\rm dis}$).

The separation of the clusters in the two groups is then controlled by the comparison between $\Delta E(c)$ and the thermal energy: the small clusters with $|\Delta E(c)| < k_B T$ can be flipped by thermal fluctuations, and the large clusters with $|\Delta E(c)| > k_B T$ are frozen.

The effect of the frozen clusters for which $|\Delta E(c)| > k_B T$ is then

$$m_{\rm frozen}(T,h) \approx \sum_{|\Delta E(c)|>k_B T} P(c) m(c) \approx \int_{L_c}^{\infty} dL \; e^{-{\rm c}(p) L^d} \; \mu L^d \tag{2.45}$$

and $L_c^d \approx k_B T/(\mu h)$. This integral can be computed by the saddle-point method and it is dominated by the lower border in the small h limit. The result is

$$m_{\rm frozen}(T,h) \approx e^{-{\rm c}(p) L_c^d} = e^{-{\rm c}(p) k_B T/(\mu h)} \tag{2.46}$$

where we dropped the contribution from μL^d since it is negligible, and this contribution has an essential singularity in the $h \to 0$ limit.

It is important to note that the clusters that contribute to this integral are *rare regions* since they occur with probability $P(c) \simeq e^{-{\rm c}(p) L^d}$ that is exponentially small in their volume. Still they are the cause of the non-analytic behaviour of $m(h)$.

The magnetic susceptibility χ can be analyzed similarly. Each locally ordered rare region makes a Curie contribution $m^2(c)/k_B T$ to χ. The total rare region susceptibility can therefore be estimated as

$$\chi_{\rm rare}(T,h) \approx \int_{L_c}^{\infty} dL \; e^{-{\rm c}(p) L^d} \mu^2 \, L^{2d}/(k_B T) \approx e^{-{\rm c}(p) k_B T/(\mu h)} \; . \tag{2.47}$$

This equation shows that the susceptibility of an individual rare region does not increase fast enough to overcome the exponential decay of the rare region probability with increasing size L. Consequently, large rare regions only make an exponentially small contribution to the susceptibility.

Rare regions also exist on the ordered side of the transition $T < T_c$. One has to consider locally ordered islands inside holes that can fluctuate between up and down because they are only very weakly coupled to the bulk ferromagnet outside the hole, see Fig. 2.8. This conceptual difference entails a different probability for the rare events as one needs to find a large enough vacancy-rich region around a locally ordered island.

There are therefore slight differences in the resulting Griffiths singularities on the two sides of the transition. In the site-diluted Ising model, the ferromagnetic Griffiths phase comprises all of the ferromagnetic phase for $p > 0$. The phase diagram of the dilute ferromagnetic Ising model is sketched in Fig. 2.7 with p denoting the probability of empty sites in the figure.

2.3 Scenario for the phase transitions

The argument put forward by Harris is based on the effect of disorder on average over the local critical temperatures. The intuitive explanation of the Griffiths phase shows the importance of rare regions on the behaviour of global observables such as the magnetisation or the susceptibility. The analysis of the effect of randomness on the phase transitions should then be refined to take into account the effect of *rare regions* (tails in the distributions). Different classes of rare regions can be identified according to their dimension $d_{\rm rr}$. This leaves place for three possibilities for the effect of (still weak in the sense of not having frustration) disorder on the phase transition.

- The rare regions have dimension $d_{\rm rr}$ smaller than the lower critical dimension of the pure problem, $d_{\rm rr} < d_L$; therefore the critical behaviour is not modified with respect to the one of the clean problem.

- When the rare regions have dimension equal to the lower-critical one, $d_{\rm rr} = d_L$, the critical point is still of second order with conventional power law scaling but with different exponents that vary in the Griffiths phase. At the disordered critical point the Harris criterium is satisfied $d\nu_{\rm dis} > 2$.

- Infinite randomness strength, appearing mostly in problems with correlated disorder, lead to a complete change in the critical properties, with unconventional activated scaling. This occurs when $d_{\rm rr} > d_L$.

In the derivation of this scenario the rare regions are supposed to act independently, with no interactions among them. This picture is therefore limited to systems with short-range interactions.

2.4 Domain-wall stiffness and droplets

Let us now just discuss one simple argument that is at the basis of what is needed to derive the results of the *droplet theory* for spin-glasses without entering into the complications of the calculations.

At very high temperature the configurations are disordered and one does not see large patches of ordered spins.

Close but above the critical temperature T_c finite patches of the system are ordered (in all possible low-temperature equilibrium states) but none of these include a finite fraction of the spins in the sample and the magnetization density vanishes. However, these patches are enough to generate non-trivial thermodynamic properties very close to T_c and the richness of critical phenomena.

At criticality one observes ordered domains of the two equilibrium states at all length scales – with *fractal* properties.

Below the critical temperature thermal fluctuations induce the spin reversal with respect to the order selected by the spontaneous symmetry breaking. It is clear that the structure of *droplets*, meaning patches in which the spins point in the opposite direction to the one of the background ordered state, plays an important role in the thermodynamic behaviour at low temperatures.

M. Fisher and others developed a droplet phenomenological theory for critical phenomena in clean systems. Later D. S. Fisher and D. Huse extended these arguments to describe the effects of quenched disorder in spin-glasses and other random systems; this is the so-called *droplet model*.

Domain-wall stiffness

Ordered phases resist spatial variations of their order parameter. This property is called *stiffness* or *rigidity* and it is absent in high-temperature disordered phases.

More precisely, in an ordered phase the *free-energy cost* for changing one part of the system with respect to another part far away is proportional to k_BT and usually diverges as a power law of the system size. In a disordered phase the information about the reversed part propagates only a finite distance (of the order of the correlation length, see below) and the stiffness vanishes.

Concretely, the free-energy cost of installing a *domain-wall* in a system, gives a measure of the stiffness of a phase. The domain wall can be imposed by special boundary conditions. Compare then the free-energy of an Ising model with linear length L, in its ordered phase, with periodic and anti-periodic boundary conditions on one Cartesian direction and periodic boundary conditions on the $d-1$ other directions of a d-dimensional hypercube. The $\pm$ boundary conditions forces an interface between the regions with positive and negative magnetisations. At $T=0$, the minimum energy interface is a $d-1$ flat hyper-plane and the energy cost is

$$\Delta E(L) \simeq \sigma L^{\theta} \qquad \text{with} \qquad \theta = d-1 \tag{2.48}$$

and $\sigma = 2J$ the interfacial energy per unit area or the interfacial tension of the domain wall.

Droplets - generalisation of the Peierls argument

In an ordered system at finite temperature domain walls, surrounding *droplet fluctuations*, or domains with reversed spins with respect to the bulk order, are naturally generated by thermal fluctuations. The study of droplet fluctuations is useful to establish whether an ordered phase can exist at low (but finite) temperatures. One then studies the free-energy cost for creating large droplets with thermal fluctuations that may destabilise the ordered phase, in the way usually done in the simple Ising chain (the Peierls argument).

Indeed, temperature generates fluctuations of different size and the question is whether these are favourable or not. These are the *droplet excitations* made by simply connected regions (domains) with spins reversed with respect to the ordered state. Because of the surface tension, the minimal energy droplets with linear size or radius L will be compact spherical-like objects with volume L^d and surface L^{d-1}. The surface determines their energy and, at finite temperature, an entropic contribution has to be taken into account as well. Simplifying, one argues that the free-energy cost is of the order of L^θ, that is L^{d-1} in the ferromagnetic case but can be different in disordered systems.

Summarising, in system with symmetry breaking the free-energy cost of an excitation of linear size L is expected to scale as

$$\Delta F(L) \simeq \sigma(T) L^\theta \; . \tag{2.49}$$

The sign of θ determines whether thermal fluctuations destroy the ordered phase or not. For $\theta > 0$ large excitations are costly and very unlikely to occur; the order phase is expected to be stable. For $\theta < 0$ instead large scale excitations cost little energy and one can expect that the gain in entropy due to the large choice in the position of these excitations will render the free-energy variation negative. A proliferation of droplets and droplets within droplets is expected and the ordered phase will be destroyed by thermal fluctuations. The case $\theta = 0$ is marginal and its analysis needs the use of other methods.

As the phase transitions is approached from below the surface tension $\sigma(T)$ should vanish. Moreover, one expects that the stiffness should be independent of length close to T_c and therefore, $\theta_c = 0$.

Above the transition the stiffness should decay exponentially

$$\Delta F(L) \simeq e^{-L/\xi} \tag{2.50}$$

with ξ the equilibrium correlation length.

2.5 Stability of ordered phases

A ferromagnet under a magnetic field

Let us study the stability properties of an equilibrium ferromagnetic phase under an applied external field that tends to destabilize it. If we set $T = 0$ the free-energy is just the energy. In the ferromagnetic case the free-energy cost of a spherical droplet of radius

R of the equilibrium phase parallel to the applied field embedded in the dominant one (see Fig. 2.9-left) is

$$\Delta F(R) = -2\Omega_d R^d h m_{\rm eq} + \Omega_{d-1} R^{d-1} \sigma_0 \tag{2.51}$$

where σ_0 is the interfacial free-energy density (the energy cost of the domain wall) and Ω_d is the volume of a d-dimensional unit sphere. We assume here that the droplet has a regular surface and volume such that they are proportional to R^{d-1} and R^d, respectively. The excess free-energy reaches a maximum

$$\Delta F_c = \frac{\Omega_d}{d} \frac{\Omega_{d-1}^d}{\Omega_d^d} \left(\frac{d-1}{2dhm_{\rm eq}} \right)^{d-1} \sigma_0^d \tag{2.52}$$

at the critical radius

$$R_c = \frac{(d-1)\Omega_{d-1}\sigma_0}{2d\Omega_d h m_{\rm eq}} \; , \tag{2.53}$$

see Fig. 2.9-right ($h > 0$ and $m_{\rm eq} > 0$ here, the signs have already been taken into account). The free-energy difference vanishes at

$$\Delta F(R_0) = 0 \qquad \Rightarrow \qquad R_0 = \frac{\Omega_{d-1}\sigma_0}{2\Omega_d h m_{\rm eq}} \; . \tag{2.54}$$

Several features are to be stressed:

- The barrier vanishes in $d = 1$; indeed, the free-energy is a linear function of R in this case.
- Both R_c and R_0 have the same dependence on $hm_{\rm eq}/\sigma_0$: they monotonically decrease with increasing $hm_{\rm eq}\sigma_0$ vanishing for $hm_{\rm eq}/\sigma_0 \to \infty$ and diverging for $hm_{\rm eq}/\sigma_0 \to 0$.
- In dynamic terms, the passage above the barrier is done *via* thermal activation; as soon as the system has reached the height of the barrier it rolls on the right side of the 'potential' ΔF and the favourable phase nucleates.
- As long as the critical size R_c is not reached the droplet is not favorable and the system remains positively magnetised.

The Imry-Ma argument for the random field Ising model at $T = 0$

Take a ferromagnetic Ising model in a random field, defined in eq. (1.33). In zero applied field and low enough temperature, if $d > 1$ there is a phase transition between a ferromagnetic and a paramagnetic phase at a critical value of the variance of the random fields, $\sigma_h^2 = [h_i^2] \propto h^2$, that sets the scale of the values that these random fields can take. Under the effect of a random field with very strong typical strength, the spins align with the local external fields that point in both directions and the system is paramagnetic. It is, however, non-trivial to determine the effect of a relatively weak random field on the

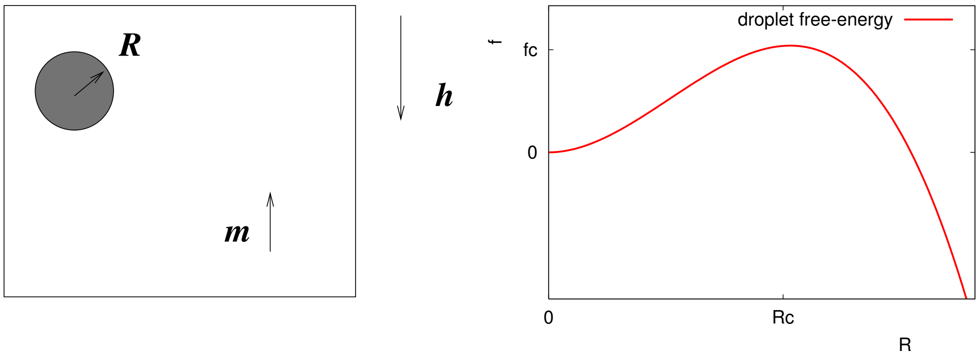

Figure 2.9: Left: the droplet. Right: the free-energy density $f(R)$ of a spherical droplet with radius R.

ferromagnetic phase at sufficiently low temperature. The long-range ferromagnetic order could be preserved or else the field could be enough to break up the system into large but finite domains of the two ferromagnetic phases.

A qualitative argument to decide whether the ferromagnetic phase survives or not in presence of the external random field is due to Imry and Ma [38]. Let us fix $T = 0$ and switch on a random field. If a compact domain $\mathcal{D}$ of the opposite order (say down) is created within the bulk of the ordered state (say up) the system pays an energy due to the unsatisfied links lying on the boundary that is

$$\Delta E_{\text{border}} \sim 2JR^{d-1} \tag{2.55}$$

where R is the radius of the domain and $d-1$ is the dimension of the border of a domain embedded in d a dimensional volume, assuming the interface is not fractal. By creating a domain boundary the system can also gain a magnetic energy in the interior of the domain due to the external field:

$$\Delta E_{\text{random field}} \sim -hR^{d/2} \tag{2.56}$$

since there are $N \propto R^d$ spins inside the domain of linear scale R (assuming now that the bulk of the domain is not fractal) and, using the central limit theorem, $-h\sum_{j\in\mathcal{D}} s_i \sim -h\sqrt{N} \propto -hR^{d/2}$. $h \approx \sigma_h$ is the width of the random field distribution.

One dimensional systems. In $d = 1$ the energy difference is a monotonically decreasing function of R thus suggesting that the creation of droplets is very favourable. There is no barrier to cross to do it. The larger the droplets to form, the better. The system fully disorders.

Dimension lower than two. For any $d < 2$, the random field energy increases faster with R than the domain wall energy. For all non-vanishing random fields, there is a critical R below which forming domains that align with the local random field becomes favourable.

Seen as a function of R, the energy function has a minimum at an R_c that is conveniently written as

$$R_c \sim \left(\frac{4J(d-1)}{hd}\right)^{2/(2-d)} \simeq \left\{ \begin{array}{l} \text{increases with } \; J/h \;\; \uparrow \\ \text{decreases with } \; J/h \;\; \uparrow \end{array} \right. \tag{2.57}$$

and later crosses zero at an R_0 that scales with J/h in the same way. In particular,

$$R_0 \propto R_c \to \infty \qquad \text{for} \qquad h/J \to 0 \; . \tag{2.58}$$

We note that while this energy function is negative, and this is the case for all R in the limit $h \to 0$ in $d < 2$, it is convenient for the system to reverse droplets with radius $R < R_0$ and thus break the ferromagnetic order in pieces eventually disordering the full system. Consequently, *the uniform ferromagnetic state is unstable* against domain formation for arbitrary random field strength. In other words, in dimensions $d < 2$ random-field disorder prevents spontaneous symmetry breaking. In the resulting Imry-Ma state there is only short-range ordering within randomly oriented domains of average size R_0 which depends on the strength of the random field.

Dimension equal to two. This is a marginal case. The function $E(R)$ is linear and the slope depends on the sign of $2J - h$ (but we neglected many numerical factors) and a more refined study is needed to decide what the system does in this case.

Dimension larger than two. The functional form of the total energy variation $\Delta E = \Delta E_{\text{border}} + \Delta E_{\text{random field}}$ as a function of R is characterised by $\Delta E \to 0$ for $R \to 0$ and $\Delta E \to \infty$ for $R \to \infty$. The function has a minimum at

$$R_c \sim \left(\frac{hd}{4J(d-1)}\right)^{2/(d-2)} \simeq \left\{ \begin{array}{l} \text{increases with } \; h/J \;\; \uparrow \\ \text{decreases with } \; h/J \;\; \uparrow \end{array} \right. \tag{2.59}$$

and crosses zero at $R_0 \propto R_c$ to approach ∞ at $R \to \infty$. The comparison between these two energy scales yields

$$2JR_0^{d-1} \sim hR_0^{d/2} \qquad \Rightarrow \qquad R_0 \sim \left(\frac{h}{2J}\right)^{\frac{2}{d-2}} \tag{2.60}$$

In particular,

$$R_0 \propto R_c \to 0 \qquad \text{for} \qquad h/J \to 0 \; . \tag{2.61}$$

Therefore, in $d > 2$ the energy difference also decreases from $\Delta E(R = 0) = 0$ to reach a negative minimum at R_c, and then increases back to pass through zero at R_0 and diverge at infinity. The main difference with the $d < 2$ case is the dependence of R_c and R_0 with h/J, the fact that both increase, and in particular vanish, for $h/J \to 0$ in $d > 2$. In consequence, in $d > 2$, for infinitesimal field, it *is not favourable* to reverse domains and long-range ferromagnetic order can be sustained in the sample.

In the arguments above, it has been very important the change in parameter dependence of the R_0 occurring in $d = 2$,

$$\lim_{h/J \to 0} R_0(h/J) = \begin{cases} 0 & \text{if} \quad d > 2 \ , \\ \infty & \text{if} \quad d < 2 \ . \end{cases} \tag{2.62}$$

With this argument one cannot show the existence of a phase transition at h_c nor the nature of it. The argument is such that it suggests that order can be supported by the system at zero temperature and small fields in $d > 2$.

Again, we stress that these results hold for short-range correlated disorder.

The argument has at least two drawbacks that have been discussed in the literature and shown to be not important for the final conclusions. One is that one should count the number of possible contours with a given length to take into account an entropic contribution to the bubble's free-energy density. Another one is that one should consider the possibility of there being contours within contours. Both problems have been taken care of, see e.g. [29].

There are rigorous proofs that random fields destroy long-range order (and thus prevent spontaneous symmetry breaking) in all dimensions $d \leq 2$ for discrete (Ising) symmetry and in dimensions $d \leq 4$ for continuous (Heisenberg) symmetry [28]. The existence of a phase transition from a FM to a PM state at zero temperature in $3d$ was shown in [29].

An elastic line in a random potential

A similar argument has been put forward by Larkin [31] for the random manifold problem.

The interfacial tension, σ, will tend to make an interface, forced into a system as flat as possible. However, this will be resisted by thermal fluctuations and, in a system with random impurities, by quenched disorder.

Let us take an interface model of the type defined in eq. (1.35) with $N = 1$. If one assumes that the interface makes an excursion of longitudinal length L and transverse length ϕ the elastic energy cost is

$$E_{\rm elast} = \frac{c}{2} \int d^dx \ (\nabla \phi(\vec{x}))^2 \qquad \Rightarrow \qquad \Delta E_{\rm elast} \sim cL^d (L^{-1}\phi)^2 = cL^{d-2}\phi^2 \tag{2.63}$$

Ignore for the moment the random potential. Thermal fluctuations cause fluctuations of the kind shown in Fig. 2.10. The interfaces *roughens*, that is to say, it deviates from being flat. Its mean-square displacement between two point $\vec{x}$ and $\vec{y}$, or its *width* on a scale L satisfies

$$\langle [\phi(\vec{x}) - \phi(\vec{y})]^2 \rangle \simeq T \ |\vec{x} - \vec{y}|^{2\zeta_T} \tag{2.64}$$

with ζ_T the *roughness exponent*.

The elastic energy cost of an excitation of length L is then

$$\Delta E_{\rm elast}(L) \simeq cL^{d-2}\phi^2(L) \simeq cTL^{d-2}L^{2\zeta_T} \tag{2.65}$$

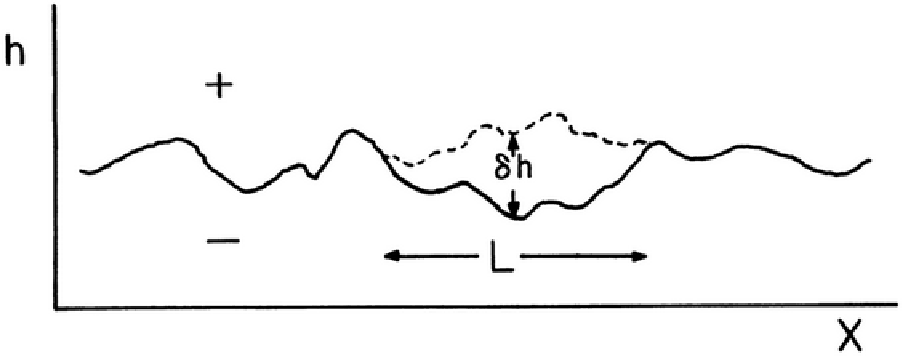

Figure 2.10: Illustration of an interface modeled as a directed manifold. In the example, the domain wall separates a region with positive magnetisation (above) from one with negative magnetisation (below). The line represents a lowest energy configuration that deviates from a flat one due to the quenched randomness. An excitation on a length-scale L is shown with a dashed line. The relative displacement is $\delta h \equiv \delta\phi \simeq L^{\alpha}$ and the excitation energy $\Delta E(L) \simeq L^{\theta}$. Figure taken from [39].

and this is of order one if

$$\zeta_T = \frac{2-d}{2} \; . \tag{2.66}$$

In the presence of quenched randomness, the deformation energy cost competes with gains in energy obtained from finding more optimal regions of the random potential. Naively, the energy gain due to the randomness is

$$\int d^dx \; V \simeq [W^2 L^d]^{1/2} \simeq W L^{d/2} \tag{2.67}$$

and the balance with the elastic cost, assumed to be the same as with no disorder, yields

$$cTL^{d-2}L^{2\zeta_D} \simeq WL^{d/2} \qquad \Rightarrow \qquad \zeta_D = \frac{4-d}{2} \tag{2.68}$$

This result turns out to be an upper bound of the exponent value [39]. It is called the *Flory exponent* for the roughness of the surface. One then concludes that for $d > 4$ disorder is irrelevant and the interface is flat ($\phi \to 0$ when $L \to \infty$). Since the linearization of the elastic energy [see the discussion leading to eq. (1.35)] holds only if $\phi/L \ll 1$, the result (2.68) may hold only for $d > 1$ where $\alpha < 1$.

Destruction of first order phase transitions under randomness

A first order phase transition is characterized by macroscopic phase coexistence at the transition point. For example, at the liquid-gas phase transition of a fluid, a macroscopic

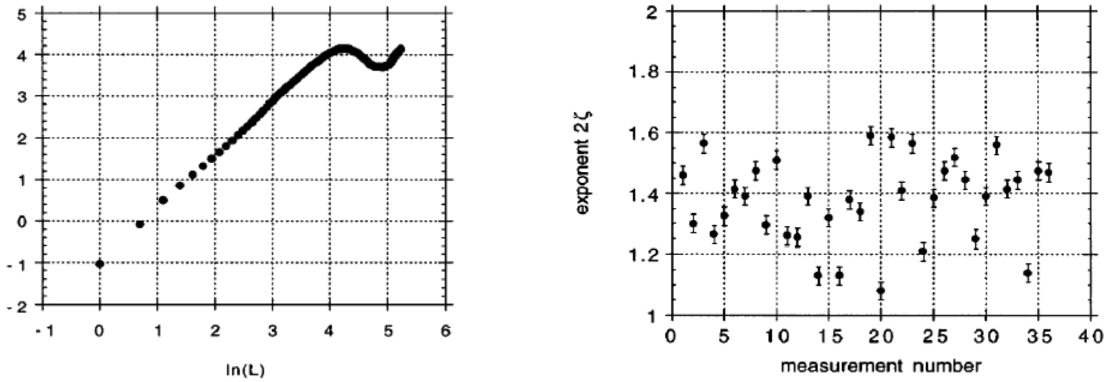

Figure 2.11: The interface width and the roughness exponent in a magnetic domain wall in a thin film. The value measured $\zeta_D \simeq 0.6$ is compatible with the Flory value 2/3 expected for a one dimensional domain wall in a two dimensional space ($N = 1$ and $d = 1$ in the calculations discussed in the text.) [32].

liquid phase coexists with a macroscopic vapour phase. Random-mass disorder locally favours one phase over the other. The question is whether the macroscopic phases survive in the presence of disorder or the system forms domains (droplets) that follow the local value of the random-mass.

Consider a single domain or droplet (of linear size L) of one phase embedded in the other phase. The free energy cost due to forming the surface is

$$\Delta F_{\rm surf} \sim \sigma L^{d-1} \tag{2.69}$$

where σ is the surface energy between the two phases. The energy gain from the random-mass disorder can be estimated via the central limit theorem, resulting in a typical magnitude of

$$|\Delta F_{\rm dis}| \sim W^{1/2} L^{d/2} \tag{2.70}$$

where W is the variance of the random-mass disorder.

The macroscopic phases are stable if $|\Delta F_{\rm dis}| < \Delta F_{\rm surf}$, but this is impossible in dimensions $d \leq 2$ no matter how weak the disorder is. In dimensions $d > 2$, phase coexistence is possible for weak disorder but will be destabilized for sufficiently strong disorder.

We thus conclude that random-mass disorder destroys first-order phase transitions in dimensions $d \leq 2$. In many examples, the first-order transition is replaced by ('rounded to') a continuous one, but more complicated scenarios cannot be excluded.

The 3d Edwards-Anderson model in a uniform magnetic field

A very similar reasoning is used to argue that there cannot be spin-glass order in an Edwards-Anderson model in an external field [40, 41]. The only difference is that the domain wall energy is here assumed to be proportional to L^y with an *a priori* unknown d-dependent exponent y that is related to the geometry of the domains.

Comments

These arguments are easy to implement when one knows the equilibrium states (or one assumes what they are). They cannot be used in models in which the energy is not a slowly varying function of the domain wall position.

2.6 Consequences of the gauge invariace

H. Nishimori used the gauge transformation to derive a series of exact results for averaged observables of finite dimensional disordered systems [11].

The idea follows the steps by which one easily proves, for example, that the averaged local magnetisation of a ferromagnetic Ising model vanishes, that is to say, one applies a transformation of variables within the partition sum and evaluates the consequences over the averaged observables. For example,

$$\langle s_i \rangle = \sum_{\{s_j=\pm 1\}} s_i \; e^{\beta J \sum_{ij} s_i s_j} = \sum_{\{s_j=\pm 1\}} (-s_i) \; e^{\beta J \sum_{ij} s_i s_j} = -\langle s_i \rangle \; . \tag{2.71}$$

This immediately implies $\langle s_i \rangle = 0$ and, more generally, the fact that the average of any odd function under $\{s_i\} \to \{-s_i\}$ vanishes exactly.

In the case of disordered systems, one is interested in observables that are averaged over the random variables weighted with their probability distribution. The gauge transformation that leaves the Hamiltonian unchanged involves a change of spins accompanied by a transformation of the exchanges:

$$\overline{s}_i = \eta_i s_i \qquad \overline{J}_{ij} = \eta_i \eta_j J_{ij} \tag{2.72}$$

with $\eta_i = \pm 1$. The latter affects the couplings probability distribution as this one, in general, is not gauge invariant. For instance, the bimodal pdf $P(J_{ij}) = p\delta(J_{ij} - J) + (1 - p)\delta(J_{ij} + J)$ can be rewritten as

$$P(J_{ij}) = \frac{e^{K_p J_{ij}/J}}{2\cosh K_p} \qquad \text{with} \qquad e^{2K_p} = \frac{p}{1-p} \; , \tag{2.73}$$

as one can simply check. $\tau_{ij} \equiv J_{ij}/J$ are just the signs of the J_{ij}. Under the gauge transformation $P(J_{ij})$ transforms as

$$\overline{P}(\overline{J}_{ij}) d\overline{J}_{ij} = P(J_{ij}) dJ_{ij} \qquad \Rightarrow \qquad \overline{P}(\overline{J}_{ij}) = P(J_{ij}(\overline{J}_{ij})) \, \frac{dJ_{ij}}{d\overline{J}_{ij}} \tag{2.74}$$

that implies

$$\overline{P}(\overline{J}_{ij}) = \frac{e^{K_p \overline{J}_{ij}/(\eta_i\eta_j J)}}{2\cosh K_p} \frac{1}{\eta_i\eta_j} \qquad \Rightarrow \qquad \overline{P}(\overline{J}_{ij}) = \eta_i\eta_j \, \frac{e^{K_p \overline{J}_{ij}\eta_i\eta_j/J}}{2\cosh K_p} \tag{2.75}$$

For instance, applying the gauge transformation to the internal energy of an Ising spin-glass model with bimodal disorder, after a series of straightforward transformations one finds

$$[\langle H_J\rangle]_J = -N_B J \tanh K_p \tag{2.76}$$

with N_B the number of bonds in the lattice, under the condition $\beta J = K_p$. This relation holds for any lattice. The constraint $\beta J = K_p$ relates the inverse temperature $J/(k_B T)$ and the probability $p = (\tanh K_p + 1)/2$. The curve $\beta J = K_p$ connects the points $(p = 1, T = 0)$ and $(p = 1/2, T \to \infty)$ in the (p, T) phase diagram and it is called the *Nishimori line*.

The proof of the relation above goes as follows. The full pdf of the interactions is

$$P(\{J_{ij}\}) = \prod_{\langle ij\rangle} P(J_{ij}) \tag{2.77}$$

and the average of any disorder dependent quantity is expressed as

$$[A_J] = \sum_{\{J_{ij}=\pm J\}} \prod_{\langle ij\rangle} P(J_{ij}) A_J \tag{2.78}$$

The disorder average Hamiltonian reads

$$[\langle H_J\rangle]_J = \sum_{\{J_{ij}\}} \frac{e^{K_p \sum_{\langle ij\rangle} J_{ij}/J}}{(2\cosh K_p)^{N_B}} \; \frac{\sum_{\{s_i\}}(-\sum_{ij} J_{ij} s_i s_j)\; e^{\beta \sum_{\langle ij\rangle} J_{ij} s_i s_j}}{\sum_{\{s_i\}} e^{\beta \sum_{\langle ij\rangle} J_{ij} s_i s_j}} \tag{2.79}$$

with N_B the number of bonds in the graph or lattice. Performing the gauge transformation

$$[\langle H_J\rangle]_J = \sum_{\{J_{ij}\}} \frac{e^{K_p \sum_{\langle ij\rangle} J_{ij}\eta_i\eta_j/J}}{(2\cosh K_p)^{N_B}} \; \frac{\sum_{\{s_i\}}(-\sum_{ij} J_{ij} s_i s_j)\; e^{\beta \sum_{\langle ij\rangle} J_{ij} s_i s_j}}{\sum_{\{s_i\}} e^{\beta \sum_{\langle ij\rangle} J_{ij} s_i s_j}} \tag{2.80}$$

where gauge invariance of the Hamiltonian has been used and the spins and interactions have been renamed J_{ij} and s_i. As this is independent of the choice of the parameters $\{\eta_i\}$ used in the transformation, one can sum over all possible 2^N choices and divide by this number keeping the result unchanged:

$$[\langle H_J\rangle]_J = \frac{1}{2^N} \sum_{\{J_{ij}\}} \frac{\sum_{\{\eta_i\}} e^{K_p \sum_{\langle ij\rangle} J_{ij}\eta_i\eta_j/J}}{(2\cosh K_p)^{N_B}} \; \frac{\sum_{\{s_i\}}(-\sum_{ij} J_{ij} s_i s_j)\; e^{\beta \sum_{\langle ij\rangle} J_{ij} s_i s_j}}{\sum_{\{s_i\}} e^{\beta \sum_{\langle ij\rangle} J_{ij} s_i s_j}} \tag{2.81}$$

If β is chosen to be $\beta = K_p/J$ the sum over the spins in the denominator (the partition sum in the normalisation) cancels out the sum over the parameters η_i introduced via the gauge transformation. The sum over J_{ij} and the remaining sum over the spin configurations can be rewritten

$$[\langle H_J\rangle]_J \;=\; \frac{1}{2^N} \frac{1}{(2\cosh K_p)^{N_B}} \left(-\frac{\partial}{\partial \beta}\right) \sum_{\{s_i\}} \prod_{\langle ij\rangle} \sum_{\{J_{ij}=\pm J\}} e^{\beta J_{ij} s_i s_j} \; . \tag{2.82}$$

Changing now variables in the sum over $J_{ij} = \pm J$ to $\tau_{ij} = J_{ij} s_i s_j = \pm J$,

$$\begin{aligned} [\langle H_J \rangle]_J &= \frac{1}{2^N} \frac{1}{(2\cosh K_p)^{N_B}} \left(-\frac{\partial}{\partial \beta} \right) \sum_{\{s_i\}} \prod_{\langle ij \rangle} \sum_{\tau_{ij}=\pm J} e^{\beta \tau_{ij}} \\ &= \frac{1}{2^N} \frac{1}{(2\cosh K_p)^{N_B}} \left(-\frac{\partial}{\partial \beta} \right) 2^N (2\cosh K_p)^{N_B} \ , \end{aligned} \tag{2.83}$$

where the sum over the spin configurations yields the 2^N factor and the sum over the independent τ_{ij} configurations yields the last factor. Finally, taking the derivative with respect to β:

$$[\langle H_J \rangle]_J = -N_B J \tanh K_p \tag{2.84}$$

with $K_p = \beta J$, defining the Nishimori line in the phase diagram.

For Gaussian distributed quenched randomness there also exists a Nishimori line and the averaged internal energy can also be computed exactly on this line.

Many other relations of this kind exist and are explained in [11]. A timely application appeared recently [46] where the gauge transformation was used to put bounds on the Jarzynski relation [47] for the work done in a non-equilibrium transformation of a spin-glass on the Nishikori line.

3 Mean-field models

In the previous section we analysed finite dimensional models, with approximation methods and a few exact ones. Here, we focus on what are called mean-field models, or models defined on the complete graph in such a way that the mean-field treatment becomes exact. We present to two techniques: the Thouless-Anderson-Palmer (TAP) approach and the replica trick.

3.1 The TAP approach

Disordered models have quenched random interactions. Due to the fluctuating values of the exchanges, one expects that the equilibrium configurations be such that *in each equilibrium state* the spins freeze in different directions. The local averaged magnetizations need not be identical, on the contrary one expects $\langle s_i \rangle = m_i$ and, if many states exist, each of them can be identified by the vector $(m_1, \dots, m_N)$.

Let us focus on the Sherrington-Kirkpatrick model, defined by

$$H_J^{\rm SK} = -\frac{1}{2} \sum_{i \neq j} J_{ij} s_i s_j - \sum_i h_i^{\rm ext} s_i \tag{3.85}$$

with interaction strengths J_{ij} taken from a Gaussian pdf and scaled with N in such a way that the thermodynamic limit is non-trivial:

$$P(J_{ij}) = (2\pi\sigma_N^2)^{-\frac{1}{2}} \, e^{-\frac{J_{ij}^2}{2\sigma_N^2}} \tag{3.86}$$

and external applied field $h_i^{\rm ext}$. The first two-moments of the exchange distribution are $[J_{ij}]=0$ and $[J_{ij}^2]=\sigma_N^2=J^2/(2N)$.

One may try to use the naive mean-field equations generalised to local variational parameters m_i, to characterise the low temperature properties of these models at *fixed quenched disorder*:

$$m_i=\tanh\left(\beta h_i^{\rm loc}\right)=\tanh\left(\sum_{j(\neq i)}\beta J_{ij}m_j+\beta h_i^{\rm ext}\right)\;.$$

and determine then the different $\{m_i^\alpha\}=(m_1^\alpha,\dots,m_N^\alpha)$ values from them, with the label α indicating the possibility of there being many solutions to these equations. It is important to reckon that, in this discussion, the $m_i=\langle s_i\rangle$ are assumed to be average in each thermodynamic state (with no mixture between them).

It has been shown by Thouless-Anderson-Palmer (TAP) [42] that these equations are not completely correct even in the fully-connected disordered case: a term which is called the *Onsager reaction term* is missing. This term represents the reaction of the spin i: the magnetisation of the spin i produces a field $h'_{j(i)}=J_{ji}m_i=J_{ij}m_i$ on spin j; this field induces a magnetisation $m'_{j(i)}=\chi_{jj}h'_{j(i)}=\chi_{jj}J_{ij}m_i$ on the spin j. This magnetisation, in turn, produces a field $h'_{i(j)}=J_{ij}m'_{j(i)}=J_{ij}\chi_{jj}J_{ij}m_i=\chi_{jj}J_{ij}^2m_i$ on the site i. The equilibrium fluctuation-dissipation relation between susceptibilities and connected correlations implies $\chi_{jj}=\beta\,\langle\,(s_j-\langle\,s_j\,\rangle)^2\,\rangle=\beta(1-m_j^2)$ and one then has $h'_{i(j)}=\beta(1-m_j^2)J_{ij}^2m_i$. The idea of Onsager – or *cavity method* – is that one has to study the ordering of the spin i in the absence of its own effect on the rest of the system. Thus, the total field produced by the sum of $h'_{i(j)}=\beta(1-m_j^2)J_{ij}^2m_i$ over all the spins j with which it can connect, has to be subtracted from the mean-field created by the other spins in the sample, *i.e.*

$$h_i^{\rm loc}=\sum_j J_{ij}m_j+h_i^{\rm ext}-\beta m_i\sum_j J_{ij}^2(1-m_j^2) \tag{3.87}$$

where $h_i^{\rm ext}$ is the external field. The equations then read

$$m_i=\tanh\left(\sum_{j(\neq i)}\beta J_{ij}m_j-\beta^2m_iJ_{ij}^2(1-m_j^2)+\beta h_i^{\rm ext}\right)\;.$$

The reason why the reaction term does not appear in the mean-field equations for ferromagnets is that it is sub-leading with respect to the first one. We now discuss why it is not in the disordered case. Let us study the order of magnitude, as powers of N of each term in the r.h.s. In the first term

$$\sum_{j(\neq i)}J_{ij}m_j\simeq\sum\frac{1}{\sqrt{N}}m_j\simeq 1 \tag{3.88}$$

because of the central limit theorem. In the second term

$$\sum_{j(\neq i)} J_{ij}^2(1-m_j^2) \simeq \sum \frac{1}{N}(1-m_j^2) \simeq 1 \tag{3.89}$$

because all terms in the sum are positive definite. Thus, in disordered systems the reaction term is of the same order of the usual mean-field; a correct mean-field description has to include it. In the ferromagnetic case this term can be neglected since it is sub-leading in N, since $J_{ij}^2 = J^2/N^2$ in this case.

The argument leading to the Onsager reaction term can be generalised to include the combined effect of the magnetisation of spin i on a sequence of spins (or $p-1$ spins) in the sample, *i.e.* the effect on i on j and then on k that comes back to i in the SK model. These higher order terms are indeed negligible only if the series of all higher order effects does not diverge. The ensuing condition is $1 > \beta^2\,(1-2q_{\rm EA} + N^{-1}\sum_i m_i^4)$ for the SK model.

Using the fact that there is a sum over a very large number of elements, J_{ij}^2 can be replaced by its site-independent variance $[J_{ij}^2] = p!J^2/(2N)$ in the last term in (3.88). Introducing the Edwards-Anderson parameter $q_{\rm EA} = N^{-1}\sum_{i=1} m_i^2$ the TAP equations follow:

$$m_i = \tanh\left(\beta\sum_{j\neq i} J_{ij}m_j - \beta^2J^2m_i(1-q_{\rm EA}) + \beta h_i\right)\,. \tag{3.90}$$

The generalisation of the argument leading to the reaction term to p spin interactions

$$H_J[\{s_i\}] = -\sum_{i_1<\dots<i_p} J_{i_1\dots i_p}s_{i_1}\dots s_{i_p} \qquad [\,J^2_{i_1\dots i_p}\,] = \frac{J^2p!}{2N^{p-1}} \tag{3.91}$$

is not so straightforward. An alternative derivation has been given by Biroli [43]. The TAP equations for p-spin fully connected models read

$$m_i = \tanh\left[\sum_{i_2\neq\dots\neq i_p}\left(\frac{\beta}{(p-1)!}J_{ii_2\dots i_p}m_{i_2}\dots m_{i_p} - \beta^2 m_i J^2_{ii_2\dots i_p}(1-m_{i_2}^2)\dots(1-m_{i_p}^2)\right)\right]$$

where we set $h_i^{\rm ext} = 0$. The first contribution to the internal field is proportional to $J_{i1_2\dots i_p} \sim N^{-(p-1)/2}$ and once the $p-1$ sums performed it is of order one. The reaction term instead is proportional to $J^2_{ii_2\dots i_p}$ and, again, a simple power counting shows that it is $O(1)$. Using the fact that there is a sum over a very large number of elements, $J^2_{i_1\dots i_p}$ can be replaced by its site-independent variance $[J^2_{i_1\dots i_p}] = p!J^2/(2N^{p-1})$ in the last term in (3.92). Introducing the Edwards-Anderson parameter $q_{\rm EA} = N^{-1}\sum_{i=1} m_i^2$ the TAP equations follow:

$$m_i = \tanh\left(\frac{\beta}{(p-1)!}\sum_{i_2\neq\dots\neq i_p} J_{ii_2\dots i_p}m_{i_2}\dots m_{i_p} - \frac{\beta^2J^2p}{2}m_i(1-q_{\rm EA})^{p-1} + \beta h_i\right)\,.$$

The importance of the reaction term becomes clear from the analysis of the linearised equations, expected to describe the second order critical behaviour for the SK model ($p = 2$) in the absence of an applied field. The TAP equations become

$$m_i \sim \beta \sum_j J_{ij} m_j - \beta^2 J^2 m_i + \beta h_i \; . \tag{3.92}$$

A change of basis to the one in which the J_{ij} matrix is diagonal leads to $m_\lambda \sim \beta(\lambda - \beta J^2) m_\lambda + \beta h_\lambda$. The notation we use is such that J_λ is an eigenvalue of the J_{ij} matrix associated to the eigenvector $\vec{v}_\lambda$. m_λ represents the projection of $\vec{m}$ on the eigenvector m_λ, $m_\lambda = \vec{v}_\lambda \cdot \vec{m}$, with $\vec{m}$ the N-vector with components m_i. The *staggered susceptibility* then reads

$$\chi_\lambda \equiv \left. \frac{\partial m_\lambda}{\partial h_\lambda} \right|_{h=0} = \beta \, \left(1 - \beta J_\lambda + (\beta J)^2\right)^{-1} \; . \tag{3.93}$$

Random matrix theory tells us that the eigenvalues of the random matrix J_{ij} are distributed with the semi-circle law. For the normalisation of the J_{ij}'s that we used, the largest eigenvalue is $J_\lambda^{\max} = 2J$ [44]. The staggered susceptibility for the largest eigenvalue diverges at $\beta_c J = 1$. Note that without the reaction term the divergence appears at the inexact value $T^* = 2T_c$ (see Sect. 3.3 for the replica solution of the SK model).

The TAP equations are the extremisation conditions on the TAP free-energy density:

$$\begin{aligned} f_J^{\rm tap}(\{m_i\}) &= -\frac{1}{p!} \sum_{i_1 \neq \cdots \neq i_p} J_{i_1 \ldots i_p} m_{i_1} \ldots m_{i_p} - \frac{\beta}{4p} \sum_{i_1 \neq \cdots \neq i_p} J^2_{i_1 \ldots i_p} (1 - m^2_{i_1}) \ldots (1 - m^2_{i_p}) \\ & -\sum_i h_i m_i + T \sum_{i=1}^N \left[\frac{1+m_i}{2} \ln \frac{1+m_i}{2} + \frac{1-m_i}{2} \ln \frac{1-m_i}{2} \right] \end{aligned} \tag{3.94}$$

where presented for the generalised p-spin model. The free-energy density as a function of the local magnetizations m_i defines what is usually called the *free-energy landscape*. Note that this function depends on $N \gg 1$ variables, m_i, and these are not necessarily identical in the disordered case in which the interactions between different groups of spins are different. The stability properties of each extreme $\{m_l^\alpha\}$ are given by the eigenvalues of the Hessian matrix

$$\mathcal{H}^J_{ij} \equiv \left. \frac{\partial f_J^{\rm tap}(\{m_k\})}{\partial m_i \partial m_j} \right|_{\{m_l^\alpha\}} \; . \tag{3.95}$$

The number of positive, negative and vanishing eigenvalues determine then the number of directions in which the extreme is a minimum, a maximum or marginal. The sets $\{m_l^\alpha\}$ for which $f_J^{\rm tap}(\{m_l^\alpha\})$ is the absolute minima yield a first definition of equilibrium or pure states.

The TAP equations apply to $\{m_i\}$ and not to the configurations $\{s_i\}$. The values of the $\{m_i^\alpha\}$ are determined as extrema of the TAP free-energy density, $f_J^{\rm tap}(\{m_i\})$, and they not need to be the same as those of the energy, $H_J(\{s_i\})$, a confusion sometimes encountered in the glassy literature. The coincidence of the two can only occur at $T \to 0$.

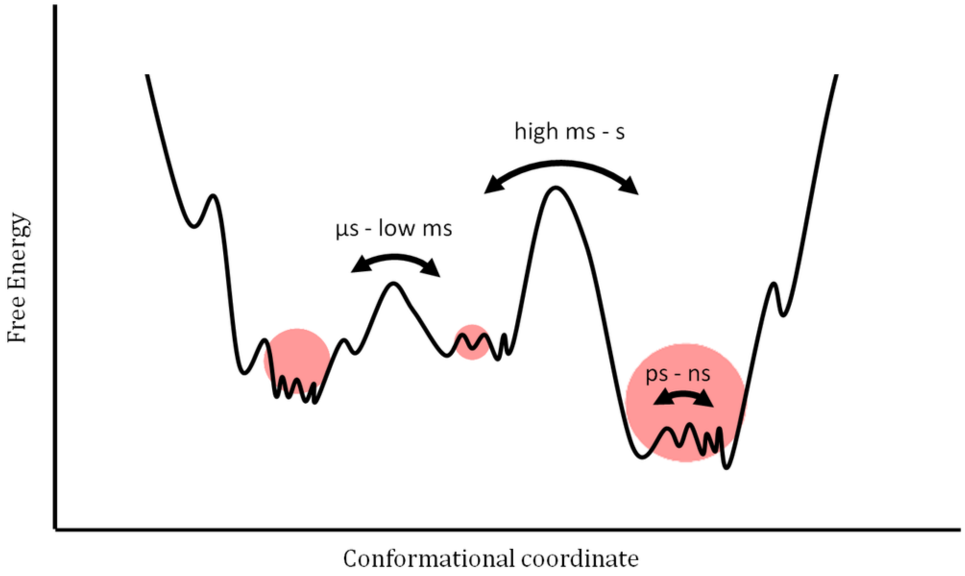

Figure 3.12: Schematic representation of a rugged free-energy landscape. Application to protein folding [45].

3.1.1 The complexity or configurational entropy

There are a number of interesting questions about the extrema of the TAP free-energy landscape, or even its simpler version in which the Onsager term is neglected, that help us understanding the static behaviour of disordered systems:

- For a given temperature, T, how many solutions to the mean-field equations exist? The *total number of solutions* can be calculated using

  $$\mathcal{N}_J(T) = \prod_i \int_{-1}^{1} dm_i \, \delta(m_i - m_i^\alpha) = \prod_i \int_{-1}^{1} dm_i \, \delta(\text{eq}_i^J) \left| \det \frac{\partial \text{eq}_i^J}{\partial m_j} \right| . \qquad (3.96)$$

  $\{m_i^\alpha\}$ are the solutions to the TAP equations that we write as $\{\text{eq}_i^J = 0\}$. The last factor is the normalization of the delta function after the change of variables, it ensures that we count one each time the integration variables touch a solution to the TAP equations independently of their stability properties. We made explicit the fact the this quantity depends on temperature.

  We define the *complexity* or *configurational entropy* as the logarithm of the number of solutions at temperature T divided by N:

  $$\Sigma_J(T) \equiv N^{-1} \ln \mathcal{N}_J(T) \; . \qquad (3.97)$$

  The normalization with N suggests that the number of solutions is actually an exponential of N. We shall come back to this very important point below.

- Does $\mathcal{N}_J(T)$ depend on T and does it change abruptly at particular values of T that may or may not coincide with static and dynamic phase transitions?
- One can define a free-energy level dependent complexity

  $$\Sigma_J(f,T) \equiv N^{-1} \ln \mathcal{N}_J(f,T) \tag{3.98}$$

  where $\mathcal{N}_J(f,T)$ is the number solutions in the interval $[f, f+df]$ at temperature T.
- From these solutions, one can identify the minima as well as all saddles of different type, *i.e.* with different indices K. These are different kinds of metastable states. Geometry constrains the number of metastable states to satisfy Morse theorem that states $\sum_{l=1}^{\mathcal{N}_J}(-1)^{\kappa_l} = 1$, where κ_l is the number of negative eigenvalues of the Hessian evaluated at the solution l, for any continuous and well-behaved function diverging at infinity in all directions. (For example, in a one-dimensional double-well function, one has $(-1)^0 + (-1)^1 + (-1)^0 = 1$.)

  One can then count the number of solutions to the TAP equations of each index, $\mathcal{N}_J(K,T)$, and define the corresponding complexity

  $$\Sigma_J(K,T) \equiv N^{-1} \ln \mathcal{N}_J(K,T) \; , \tag{3.99}$$

  or even work at fixed free-energy density

  $$\Sigma_J(K,f,T) \equiv N^{-1} \ln \mathcal{N}_J(K,f,T) \; . \tag{3.100}$$

  Even more interestingly, one can analyse how are the free-energy densities of different saddles organized. For instance, one can check whether all maxima are much higher in free-energy density than saddles of a given type, *etc.*
- What is the barrier, $\Delta f = f_1 - f_0$, between ground states and first excited states? How does this barrier scale with the actual free-energy difference, Δf between these states? To answer this question one has to estimate the nucleation radius for the reversal of a droplet under an applied field, for instance.

The definitions of complexity given above are disorder-dependent. One might then expect that the complexity will show sample-to-sample fluctuations and be characterized by a probability distribution. The *quenched complexity*, Σ^{quenched}, is then the most likely value of Σ_J, and it is defined through

$$\max P(\Sigma_J) = P(\Sigma^{\text{quenched}}) \; . \tag{3.101}$$

In practice, this is very difficult to compute. Most analytic results concern the *annealed complexity*

$$\Sigma_{\text{ann}} \equiv N^{-1} \ln [\, \mathcal{N}_J \,] = N^{-1} \ln [\, e^{N\Sigma_J} \,] \; . \tag{3.102}$$

One can show that the annealed complexity is smaller or equal than the quenched one.

3.1.2 Weighted averages

Having identified many solutions to the TAP equations one needs to determine now how to compute statistical averages. A natural proposal is to give a probability weight to each solution, w_α, and to use it to average the value of the observable of interest:

$$\langle\, O\,\rangle = \sum_\alpha w_\alpha^J \, O_\alpha \qquad \text{with} \qquad O_\alpha = O(\{m_i^\alpha\}) \tag{3.103}$$

where α labels the TAP solutions, O_α is the value that the observable O takes in the TAP solution α, and w_α^J are their statistical weights, satisfying the normalization condition $\sum_\alpha w_\alpha^J = 1$. Two examples can illustrate the meaning of this average. In a spin-glass problem, if $O = s_i$, then $O_\alpha = m_i^\alpha$. In an Ising model in its ferromagnetic phase, if $O = s_i$, then $O_\alpha = m_i^\alpha = \pm m$ and $w_\alpha = 1/2$. Within the TAP approach one proposes

$$w_\alpha^J = \frac{e^{-\beta F_\alpha^J}}{\sum_\gamma e^{-\beta F_\gamma^J}} \tag{3.104}$$

with F_α^J the total free-energy of the α-solution to the TAP equations. The discrete sum can be transformed into an integral over free-energy densities, introducing the degeneracy of solutions quantified by the free-energy density dependent complexity:

$$\langle\, O\,\rangle = \frac{1}{Z_J} \int df \; e^{-N\beta f} \; \mathcal{N}_J(f,T)\; O(f) = \frac{1}{Z_J} \int df \; e^{-N(\beta f - \Sigma_J(f,T))} \; O(f)\;. \tag{3.105}$$

The normalization is the ‘partition function’

$$Z_J = \int df \; e^{-N\beta f} \; \mathcal{N}_J(f,T) = \int df \; e^{-N(\beta f - \Sigma_J(f,T))} \;. \tag{3.106}$$

We assumed that the labelling by α can be traded by a labelling by f that implies that at the same free-energy density level f the observable O takes the same value. In the $N \to \infty$ limit these integrals can be evaluated by saddle-point, provided the parenthesis is positive. In order to simplify the calculations, the disorder-dependent complexity is generally approximated with the annealed value introduced in eq. (3.102).

The equilibrium free-energy

The total equilibrium free-energy density, using the saddle-point method to evaluate the partition function Z_J in eq. (3.106), reads

$$-\beta f_{\rm eq}^J = N^{-1} \ln Z_J = \min_f \left[f - k_B T \Sigma_J(f,T)\right] \equiv \min_f \Phi_J(f,T)\;. \tag{3.107}$$

It is clear that $\Phi_J(f,T)$ is the *Landau free-energy density* of the problem with f playing the rôle of the energy and Σ_J of the entropy. If we use $f = (E - k_B T S)/N = e - Ts$ with E the actual energy and S the microscopic entropy one has

$$\Phi_J(f,T) = e - k_B T\left(s + \Sigma_J(f,T)\right)\;. \tag{3.108}$$

Thus, Σ_J is an extra contribution to the total entropy that is due to the exponentially large number of metastable states. Note that we do not distinguish here their stability.

Note that Σ_J is subtracted from TAP free-energy level f. Thus, it is possible that in some cases states lying at a *higher free-energy density* f but being *very numerous* have a *lower total Landau free-energy density* Φ than lower lying states that are less numerous. Collectively, higher states dominate the equilibrium measure in these cases. This phenomenon actually occurs in p-spin models, as explained below.

The order parameter

Now that we know that there can be a large number of states (defined as extrema of the TAP free-energy) we have to be careful about the definition of the spin-glass order parameter.

The *Edwards-Anderson parameter* is understood as a *property of a single state*. Within the TAP formalism one then has

$$q_{J\mathrm{EA}}^{\alpha} = N^{-1} \sum_i (m_i^\alpha)^2 \qquad \text{with} \qquad m_i^\alpha = \langle s_i \rangle_\alpha \tag{3.109}$$

being restricted to spin configurations in state α. An average of this quantity over all extrema of the free-energy density yields $\sum_\alpha w_\alpha^J \, q_{J\mathrm{EA}}^\alpha = \sum_\alpha w_\alpha^J \, N^{-1} \sum_i (m_i^\alpha)^2$.

Instead, the statistical *equilibrium magnetisation*, $m_i = \langle s_i \rangle = \sum_\alpha w_\alpha^J m_i^\alpha$, squared is

$$q_J \equiv \langle \, s_i \, \rangle^2 = m_i^2 = \left(\sum_\alpha w_\alpha^J m_i^\alpha \right)^2 = \sum_{\alpha\beta} w_\alpha^J w_\beta^J \, m_i^\alpha m_i^\beta \; . \tag{3.110}$$

If there are multiple phases, the latter sum has crossed contributions from terms with $\alpha \neq \beta$. These sums, as in a usual paramagnetic-ferromagnetic transition have to be taken over half space-space, otherwise global up-down reversal would imply the cancellation of all cross-terms.

Clearly

$$q_{J\mathrm{EA}}^\alpha \neq q_J \qquad \text{and} \qquad \sum_\alpha w_\alpha^J \, q_{J\mathrm{EA}}^\alpha \neq q_J \; . \tag{3.111}$$

3.2 Metastable states in two families of models

3.2.1 High temperatures

For all models, at high temperatures $f(m_i)$ is characterized by a single stable absolute minimum in which all local magnetizations vanish, as expected; this is the paramagnetic state. The $m_i = 0$ for all i minimum continues to exist at all temperatures. However, even if it is still the global absolute minimum of the TAP free-energy density, f_JTAP, at low temperatures it becomes unstable thermodynamically, and it is substituted as the equilibrium state, by other non-trivial configurations with $m_i \neq 0$ that are the absolute minima of Φ. Note the difference with the ferromagnetic problem for which the paramagnetic solution is no longer a minimum below T_c.

3.2.2 Low temperatures

At low temperature many equilibrium states appear (and not just two as in an Ising ferromagnetic model) and they are not related by symmetry (as spin reversal in the Ising ferromagnet or a rotational symmetry in the Heisenberg ferromagnet). These are characterized by non-zero values of the local magnetizations m_i that are different in different states.

At low-temperatures both the naive mean-field equations and the TAP equations have an *exponential in N number of solutions* and still an exponential in N number of them correspond to absolute minima of the m_i-dependent free-energy density. This means that $\Sigma_J(T)$ and even $\Sigma_J(0, f_0, T)$ are quantities $O(1)$. These minima can be identified as *different* states that could be accessed by applying the corresponding site-dependent pinning fields.

The derivation and understanding of the structure of the TAP free-energy landscape is quite subtle and goes beyond the scope of these Lectures. Still, we shall briefly present their structure for the SK and p-spin models to give a flavor of their complexity.

The SK model

The first calculation of the complexity in the SK model appeared in 1980 [48, 49]. After 35 years of research the structure of the free-energy landscape in this system is still a matter of discussion. At present, the picture that emerges is the following. The temperature-dependent annealed complexity is a decreasing function of temperature that vanishes only at T_c but takes a very small value already at $\sim 0.6\,T_c$. Surprisingly enough, at finite but large N the TAP solutions come in pairs of minima and saddles of type one, that is to say, extrema with only one unstable direction. These states are connected by a mode that is softer the larger the number of spins: they coalesce and become marginally stable in the limit $N \to \infty$. Numerical simulations show that starting from the saddle-point and following the 'left' direction along the soft mode one falls into the minimum; instead, following the 'right' direction along the same mode one falls into the paramagnetic solution. See Fig. 3.13 for a sketch of these results. The free-energy difference between the minimum and saddle decreases for increasing N and one finds, numerically, an averaged $\Delta f \sim N^{-1.4}$. The extensive complexity of minima and type-one saddles is identical in the large N limit, $\Sigma_J(0, T) = \Sigma_J(1, T) + O(N^{-1})$ [50] in such a way that the Morse theorem is respected. The free-energy dependent annealed complexity is a smooth function of f with support on a finite interval $[f_0, f_1]$ and maximum at $f_{\max}$. The Bray and Moore annealed calculation (with supersymmetry breaking) yields $f_{\max} = -0.654$, $\Sigma_J^{\max} = 0.052$, $\Sigma''(f_{\max}) = 8.9$. The probability of finding a solution with free-energy density f can be

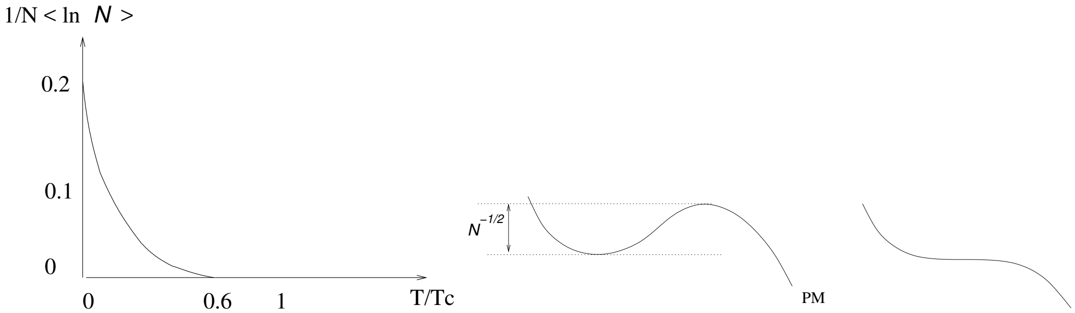

Figure 3.13: Left: sketch of the temperature dependent complexity, $\Sigma_J(T)$, of the SK. It actually vanishes only at T_c but it takes a very small value already at $\sim 0.6\ T_c$. Right: pairs of extrema in the SK model with N large and $N \to \infty$ limit.

expressed as

$$p_J(f,T) = \frac{\mathcal{N}_J(f,T)}{\mathcal{N}_J(T)} = \frac{e^{N\Sigma_J(f,T)}}{\mathcal{N}_J(T)} \sim \sqrt{\frac{N\Sigma''_J(f_{\max})}{2\pi}}\ e^{-\frac{N}{2}|\Sigma''_J(f_{\max})|(f-f_{\max})^2}\ , \tag{3.112}$$

where we evaluated the total number of solutions, $\mathcal{N}_J(T) = \int df\ e^{N\Sigma_J(f,T)}$, by steepest descent. The complexity, approximated by its annealed value, vanishes linearly close to f_0: $\Sigma_J(f,T) \sim \lambda(f - f_0)$ with $\lambda < \beta$.

Only the lowest lying TAP solutions contribute to the statistical weight. The complexity does not contribute to Φ in the large N limit:

$$\begin{aligned} \Phi &= \beta f - \Sigma_{\rm ann}(f,T) \simeq \beta f - (f - f_0)\lambda \\ \frac{\partial \Phi}{\partial f} &\simeq \beta - \lambda > 0 \qquad \text{iff} \qquad \beta > \lambda \end{aligned} \tag{3.113}$$

and $\Phi_{\min} \simeq \beta f_{\min} = \beta f_0$. See Fig. 3.14. The 'total' free-energy density in the exponential is just the free-energy density of each low-lying solution.

3.3 The replica method

A picture that is consistent with the one arising from the naive mean-field approximation but contradicts the initial assumption of the droplet model arises from the *exact* solution of fully-connected spin-glass models. These results are obtained using a method called the *replica trick* that we will briefly present below.

In Sect. 1.2.4 we argued that the typical properties of a disordered system can be computed from the disorder averaged free-energy [14]

$$[\,F_J\,] \equiv \int dJ\ P(J) F_J\ . \tag{3.114}$$

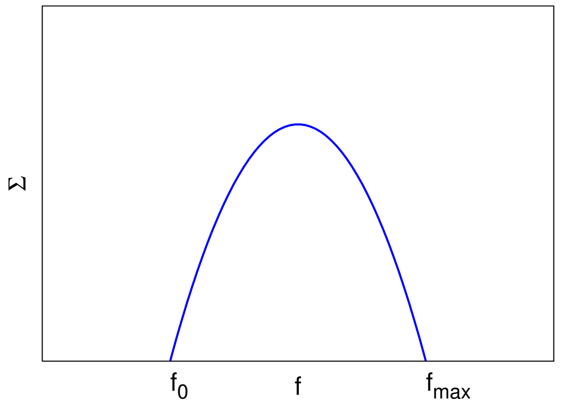

Figure 3.14: The complexity as a function of f for the SK model.

One then needs to average the logarithm of the partition funtion. In the *annealed approximation* one exchanges the ln with the average over disorder and, basically, considers the interactions equilibrated at the same temperature T as the spins:

$$[\,\ln Z_J\,] \sim \ln[\,Z_J\,] \; . \tag{3.115}$$

This approximation turns out to be correct at high temperatures but incorrect at low ones.

The evaluation of the disordered average free-energy density is difficult for at least two reasons: firstly, virtually all configurations are not translationally invariant; secondly, no factorisation helps one easily reduce the partition sum over one acting on independent variables.

The replica method allows one to compute $[\,F_J\,]$ for fully-connected models. It is based on the smart use of the identity

$$\ln Z_J = \lim_{n\to 0} \frac{Z_J^n - 1}{n} \; . \tag{3.116}$$

The idea is to compute the right-hand-side for finite and integer $n = 1, 2, \ldots$ and then perform the analytic continuation to $n \to 0$. Care should be taken in this step: for some models the analytic continuation may not be unique. (Recall the calculation done using the Potts model with $q \to 1$ that allows one to recover results for the percolation problem.) It turns out that this is indeed the case for the emblematic Sherrington-Kirkpatrick model, as discussed by van Hemmen and Palmer [55] though it has also been recently shown that the free-energy $f(T)$ obtained by Parisi [56] with the replica trick is exact! [57, 58]

The disorder averaged free-energy is given by

$$-\beta[\,F_J\,] = -\int dJ\, P(J) \ln Z_J = -\lim_{n\to 0} \frac{1}{n} \left(\int dJ\, P(J) Z_J^n - 1 \right) \; , \tag{3.117}$$

where we have exchanged the limit $n \to 0$ with the integration over the exchanges. For *integer* n the *replicated partition function*, Z_J^n, reads

$$Z_J^n = \sum_{\{s_i^a\}} e^{-\beta[H_J(\{s_i^1\})+\dots+H_J(\{s_i^n\}]} \; . \tag{3.118}$$

Here $\sum_{\{s_i^a\}} \equiv \sum_{\{s_i^1=\pm 1\}} \cdots \sum_{\{s_i^n=\pm 1\}}$. Z_J^n corresponds to n identical copies of the original system, that is to say, all of them with the same realisation of disorder. Each copy is characterised by an ensemble of N spins, $\{s_i^a\}$. We label the copies with a replica index $a = 1, \dots, n$. For p-spin disordered spin models Z_J^n takes the form

$$Z_J^n = \sum_{\{s_i^a\}} e^{\frac{\beta}{p!} \sum_{a=1}^n \left[\sum_{i_1 \neq \dots \neq i_p} J_{i_1 \dots i_p} s_{i_1}^a \dots s_{i_p}^a + \sum_i h_i s_i^a \right]} \; . \tag{3.119}$$

The average over disorder amounts to computing a Gaussian integral for each set of spin indices $i_1 \neq \dots \neq i_p$, with $[J_{i_1\dots i_p}^2] = J^2 p!/(2N^{p-1})^2$. One finds

$$[\, Z_J^n \,] = \sum_{\{s_i^a\}} e^{\frac{\beta^2 J^2}{2N^{p-1}p!} \sum_{i_1 \neq \dots \neq i_p} (\sum_a s_{i_1}^a \dots s_{i_p}^a)^2 + \beta \sum_a \sum_i h_i s_i^a} \equiv \sum_{\{s_i^a\}} e^{-\beta F(\{s_i^a\})} \; . \tag{3.120}$$

The function $\beta F(\{s_i^a\})$ is not random. It depends on the spin variables only but it includes terms that couple different replica indices. Indeed,

$$\sum_{i_1 \neq \dots \neq i_p} \sum_a s_{i_1}^a \dots s_{i_p}^a \sum_b s_{i_1}^b \dots s_{i_p}^b = \sum_{ab} \sum_{i_1 \neq \dots \neq i_p} (s_{i_1}^a s_{i_1}^b) \dots (s_{i_p}^a s_{i_p}^b) \tag{3.121}$$

We first note that all terms are identical to one for $a = b$ since $s_i^2 = 1$. The sum over the spin indices (ignoring the constraint $i_1 \neq \dots \neq i_p$ that will, in any case give a subdominant contribution in the $N \to \infty$ limit) and the replica indices of such terms yields $N^p n$. Focusing then on the cases $a \neq b$, Ignoring the constraint $i_1 \neq \dots \neq i_p$, there are N factors here

$$\beta F(\{s_i^a\}) \approx -\frac{N\beta^2 J^2}{2p!} \left[\sum_{a \neq b} \left(\frac{1}{N} \sum_i s_i^a s_i^b \right)^p + n \right] - \beta \sum_a \sum_i h_i s_i^a \; . \tag{3.122}$$

In writing the last expression we have dropped terms that are sub-leading in N. (In complete analogy with what is done for the pure p spin ferromagnet. They correspond to adding terms with self-interactions in the Hamiltonian.) The constant term $-Nn\beta^2 J^2/2$ originates in the terms with $a = b$, for which $(s_i^a)^2 = 1$.

To summarise, we started with an interacting spin model. Next, we enlarged the number of variables from N spins to $N \times n$ replicated spins by introducing n non-interacting

---

[2] We use $\int dJ_\alpha e^{-J_\alpha^2/(2\sigma^2) - J_\alpha x} \propto e^{x^2\sigma^2/2}$

copies of the system. By integrating out the disorder we *decoupled the sites* but we payed the price of *coupling the replicas*. Hitherto the replica indices act as a formal tool introduced to compute the average over the bond distribution. Nothing distinguishes one replica from another and, in consequence, the ‘free-energy’ $F(\{s_i^a\})$ is invariant under permutations of the replica indices.

The next step to follow is to identify the order parameters and transform the free-energy into an order-parameter dependent expression to be rendered extremal at their equilibrium values. In a spin-glass problem we already know that the order parameter is not the global magnetisation as in a pure magnetic system but the parameter q – or more generally the overlap between states. Within the replica calculation an *overlap between replicas*

$$q_{ab} \equiv N^{-1} \sum_i s_i^a s_i^b \tag{3.123}$$

naturally appeared in eq. (3.122). The idea is to write the free-energy density as a function of the order parameter q_{ab} and look for their extreme in complete analogy with what is done for the fully-connected ferromagnet This is, of course, a tricky business, since the order parameter is here a matrix with number of elements n going to zero! A recipe for identifying the form of the order parameter (or the correct saddle-point solution) has been proposed by G. Parisi in the late 70s and early 80s [56]. This solution has been recently proven to be exact for mean-field models by two mathematical physics, F. Guerra [57] and M. Talagrand [58]. Whether the very rich physical structure that derives from this rather formal solution survives in finite dimensional systems remains a subject of debate.

Introducing the Gaussian integral (Hubbard-Stratonovich transformation)

$$\int dq_{ab}\; e^{\beta J q_{ab} \sum_i s_i^a s_i^b - \frac{N}{2} q_{ab}^2} = e^{\frac{N}{2}\left(\frac{1}{N}\beta J \sum_i s_i^a s_i^b\right)^2} \tag{3.124}$$

for each pair of replica indices $a \neq b$, one decouples the site indices, i, and the averaged replicated partition function can be rewritten as

$$[Z_J^n] = \int \prod_{a\neq b} dq_{ab}\; e^{-\beta F(q_{ab})} \tag{3.125}$$

and

$$\begin{aligned}
\beta F(q_{ab}) &= -\frac{N\beta^2 J^2}{2}\left[-\sum_{a\neq b} q_{ab}^p + n\right] - N \ln \zeta(q_{ab})\;, && (3.126)\\
\zeta(q_{ab}) &= \sum_{s_a} e^{-\beta H(q_{ab}, s_a)}\;, && (3.127)\\
H(q_{ab}, s_a) &= -J \sum_{ab} q_{ab} s_a s_b - h \sum_a s_a\;, && (3.128)
\end{aligned}$$

where for simplicity we set $h_i = h$. The factor N in front of $\ln\zeta$ comes from the decoupling of the site indices. Note that the transformation (3.124) serves to uncouple the sites and to obtain the very useful factor N in front of the exponential. The partition function

$$Z(q_{ab}) = \sum_{\{s_a\}} e^{-\beta H(q_{ab},s_a)} \tag{3.129}$$

is the one of a fully-connected Ising model with interaction matrix q_{ab}. As it is posed, this problem remains unsolvable. However, important steps forward will be possible taking advantage of the $n \to 0$ limit.

Saddle-point evaluation

Having extracted a factor N in the exponential suggests to evaluate the integral over q_{ab} with the saddle-point method. This, of course, involves the *a priori* dangerous exchange of limits $N \to \infty$ and $n \to 0$. The replica theory relies on this assumption. One then writes

$$\lim_{N\to\infty} -[\, f_J \,] \to -\lim_{n\to 0} \frac{1}{n} f(q_{ab}^{\rm sp}) \tag{3.130}$$

and searches for the solutions to the $n(n-1)/2$ extremization equations

$$\left.\frac{\delta f(q_{ab})}{\delta q_{cd}}\right|_{q_{ef}^{\rm sp}} = 0 \; . \tag{3.131}$$

In usual saddle-point evaluations the saddle-point one should use is (are) the one(s) that correspond to absolute minima of the free-energy density. In the replica calculation the number of variables is $n(n-1)/2$ that becomes negative! when $n < 1$ and makes the saddle-point evaluation tricky. In order to avoid unphysical complex results one needs to focus on the saddle-points with positive (or at least semi-positive) definite Hessian

$$\mathcal{H} \equiv \left.\frac{\partial f(q_{ab})}{\partial q_{cd}\partial q_{ef}}\right|_{q_{gh}^{\rm sp}} \; , \tag{3.132}$$

and these sometimes corresponds to *maxima* (instead of minima) of the free-energy density.

The saddle-point equations are also self-consistency equations

$$q_{ab}^{\rm sp} = \langle s_a s_b \rangle_{H(q_{ab},\{s_a\})} = [\, \langle s_a s_b \rangle \,] \tag{3.133}$$

where the second member means that the average is performed with the *single site many-replica Hamiltonian* $H(q_{ab},\{s_a\})$ and the third member is just one of the averages we would like to compute.

The partition function in eq. (3.128) cannot be computed for generic q_{ab} since there is no large n limit to exploit. On the contrary, $n \to 0$. Thus, one usually looks for solutions to eqs. (3.131) within a certain family of matrices q_{ab}. We discuss below the relevant parametrizations.

Replica symmetry (RS)

In principle, nothing distinguishes one replica from another one. This is the reason why Sherrington and Kirkpatrick looked for solutions that preserve replica symmetry:

$$q_{ab} = q \ , \qquad \text{for all } a \neq b \ . \tag{3.134}$$

Inserting this *Ansatz* in (3.126) and (3.128) and taking $n \to 0$ one finds

$$q = \int_{-\infty}^{\infty} \frac{dz}{\sqrt{2\pi}} \, e^{-z^2/2} \, \tanh^2\left(\beta\sqrt{\frac{pq^{p-1}}{2}}z + \beta h\right) \ . \tag{3.135}$$

This equation resembles strongly the one for the magnetisation density of the p-spin ferromagnet, eq. (**??**). The free-energy density for $p = 2$ is

$$f = -\frac{\beta}{4}(1-q)^2 - \int_{-\infty}^{\infty} \frac{dz}{\sqrt{2\pi}} \, e^{-z^2/2} \, \ln[2\cosh(\beta q^{1/2} z + \beta h)] \ . \tag{3.136}$$

Let us first discuss the case $p = 2$, *i.e.* the SK model. In the absence of a magnetic field, one finds a second order phase transition at $T_s = J$ from a paramagnetic ($q = 0$) to a spin-glass phase with $q \neq 0$. In the presence of a field there is no phase transition. SK soon realized though that there is something wrong with this solution: the entropy at zero temperature is negative, $S(0) = -1/(2\pi)$, and this is impossible for a model with discrete spins, for which S is strictly positive. de Almeida and Thouless later showed that the reason for this failure is that the replica symmetric saddle-point is not stable, since the Hessian (3.132) is not positive definite and has negative eigenvalues [59]. The eigenvalue responsible for the instability of the replica symmetric solution is called the *replicon*.

Comparison with the TAP equations shows that the RS *Ansatz* corresponds to the assumption that the *local fields* $h_i = \sum_{i_1 \dots i_p} J_{i_1 \dots i_p} m_{i_1} \dots m_{i_p} + h$ are independent and have a *Gaussian* distribution with average h and variance $\sigma^2 = J^2 q^{p-1}$. Numerical simulations clearly show that this assumption is invalid.

Interestingly enough, the numerical values for several physical quantities obtained with the replica symmetric solution do not disagree much with numerical results. For instance, the ground state zero-temperature energy density is $E_{\rm gs}^{\rm rs} = -0.798$ while with numerical simulations one finds $E_{\rm gs} \sim -0.76$.

For the $p > 2$ model one finds that the replica symmetric solution is stable at all temperatures. However, the problem of the negative entropy remains and should be solved by another solution. The transition must then have aspects of a first-order one, with another solution appearing at low temperatures and becoming the most convenient one at the transition.

One step replica symmetry breaking

The next challenge is to device a replica symmetry breaking *Ansatz*, in the form of a matrix q_{ab} that is not invariant under permutations of rows or columns. There is no

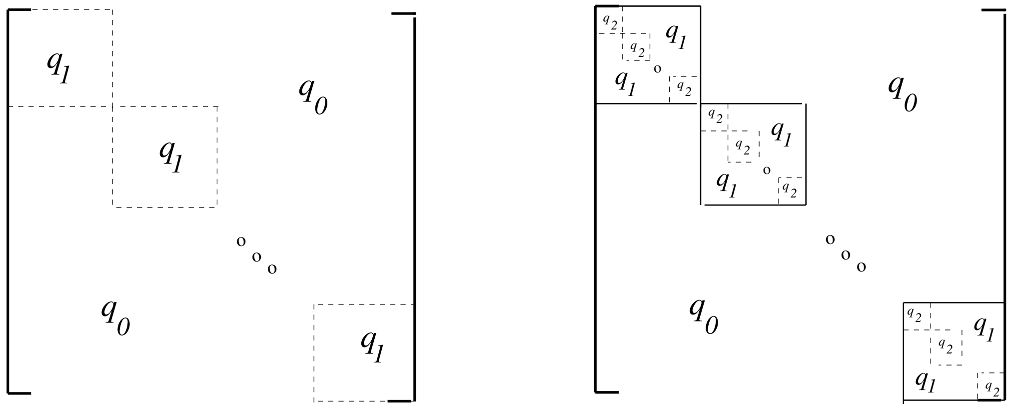

Figure 3.15: Left: a one-step replica symmetry breaking (1RSB) *Ansatz*. Right: a two-step replica symmetry breaking *Ansatz*. The elements on the main diagonal vanish identically. In the 1RSB case the diagonal blocks have size $m \times m$. In the 2RSB the proceudre is repeated and one has blocks of size $m_1 \times m_1$ with smaller diagonal blocks of size $m_2 \times m_2$.

first principles way of doing this, instead, the structure of the *Ansatz* is the result of trial and error. Indeed, a kind of minimal way to break the replica symmetry is to propose a structure in blocks as the one shown in Fig. 3.15-left. The diagonal elements are set to zero as in the RS case. Square blocks of linear size m close to the main diagonal are filled with a paramater q_1. The elements in the rest of the matrix take a different value q_0 and one takes $0 \leq q_0 \leq q_1$. The matrix q_{ab} depends on three parameters q_0, q_1, m and one has to find the values such that the free-energy density is *maximized*! The conditions for a extreme are

$$\frac{\partial f(q_0, q_1, m)}{\partial q_0} = \frac{\partial f(q_0, q_1, m)}{\partial q_1} = \frac{\partial f(q_0, q_1, m)}{\partial m} = 0 \; . \tag{3.137}$$

In the SK model ($p = 2$) the 1RSB *Ansatz* yields a second order phase transition ($q_0 = q_1 = 0$ and $m = 1$ at criticality) at a critical temperature $T_s = J$, that remains unchanged with respect to the one predicted by the RS *Ansatz*. The 1RSB solution is still unstable below T_s and in all the low temperature phase. One notices, however, that the zero temperature entropy, even if still negative and incorrect, takes a value that is closer to zero, $S(T = 0) \approx -0.01$, the ground state energy is closer to the value obtained numerically, and the replicon eigenvalue even if still negative has an absolute value that is closer to zero. All this suggest that the 1RSB *Ansatz* is closer to the exact solution.

Instead, in all cases with $p \geq 3$ the 1RSB *Ansatz* is stable below the static critical temperature T_s and all the way up to a new characteristic temperature $0 < T_f < T_s$. Moreover, one can prove that in this range of temperatures the model is solved exactly by this *Ansatz*. The critical behaviour is quite peculiar: while the order parameters q_0 and q_1 jump at the transition from a vanishing value in the paramagnetic phase to a non-zero value right below T_s, all thermodynamic quantities are continuous since $m = 1$ at T_s and all q_0 and q_1 dependent terms appear multiplied by $1 - m$. This is a mixed type of transition that has been baptised *random first-order*. Note that disorder weakens the critical behaviour in the $p \geq 3$-spin models. In the limit $p \to \infty$ the solutions become $m = T/T_c$, $q_0 = 0$ and $q = 1$.

k-step replica symmetry breaking

The natural way to generalize the 1RSB *Ansatz* is to propose a k-step one. In each step the off-diagonal blocks are left unchanged while the diagonal ones of size m_k are broken as in the first step thus generating smaller square blocks of size m_{k+1}, close to the diagonal. At a generic k-step RSB scheme one has

$$0 \leq q_0 \leq q_1 \leq \cdots \leq q_{k-1} \leq q_k \leq 1 \; , \tag{3.138}$$

$$n = m_0 \geq m_1 \geq \cdots \geq m_k \geq m_{k+1} \; , \tag{3.139}$$

parameters. In the $n \to 0$ limit the ordering of the parameters m is reversed

$$0 = m_0 \leq m_1 \leq \cdots \leq m_k \leq m_{k+1} \; . \tag{3.140}$$

In the SK model one finds that any finite k-step RSB *Ansatz* remains unstable. However, increasing the number of breaking levels the results continue to improve with, in particular, the zero temperature entropy getting closer to zero. In the $p \geq 3$ case instead one finds that the 2RSB *Ansatz* has, as unique solution to the saddle-point equations, one that boils down to the 1RSB case. This suggests that the 1RSB *Ansatz* is stable as can also be checked with the analysis of the Hessian eigenvalues: the replicon is stricly positive for all $p \geq 3$.

Full replica symmetry breaking

In order to construct the full RSB solution the breaking procedure is iterated an infinite number of times. The full RSB *Ansatz* thus obtained generalizes the block structure to an infinite sequence by introducing a function

$$q(x) = q_i \; , \qquad m_{i+1} < x < m_i \tag{3.141}$$

with $0 \leq x \leq 1$. Introducing $q(x)$ sums over replicas are traded by integrals over x; for instance

$$\frac{1}{n} \sum_{a \neq b} q_{ab}^l = \int_0^1 dx \; q^l(x) \; . \tag{3.142}$$

The free-energy density becomes a functional of the function $q(x)$. The extremization condition is then a hard functional equation. A Landau expansion – expected to be valid close to the assumed second order phase transition – simplifies the task of solving it. For the SK model one finds

$$q(x) = \begin{cases} \frac{x}{2} \; , & 0 \leq x \leq x_1 = 2q(1) \; , \\ q_{\rm EA} \equiv q_{\max} = q(1) \; , & x_1 = 2q(1) \leq x \leq 1 \; , \end{cases} \tag{3.143}$$

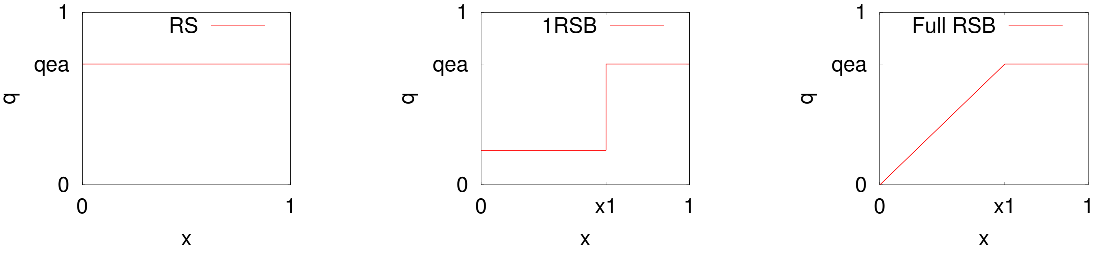

Figure 3.16: The function $q(x)$ for a replica symmetric (left), one step replica symmetry breaking (center) and full replica symmetry breaking *Ansätze*.

at first order in $|T - T_c|$, with $q(1) = |T - T_c|/T_c$ and $x_1 = 2q(1)$. The stability analysis yields a vanishing replicon eigenvalue signalling that the full RSB solution is *marginally stable*.

One can also recover the particular case of the 1RSB using a $q(x)$ with two plateaux, at q_0 and q_1 and the breaking point at $x = m$.

Marginality condition

In the discussion above we chose the extreme that *maximize* the free-energy density since we were interested in studying equilibrium properties. We could, instead, use a different prescription, though *a priori* not justified, and select other solutions. For example, we can impose that the solution is *marginally stable* by requiring that the replicon eigenvalue vanishes. In the $p = 2$ this leads to identical results to the ones obtained with the usual prescription since the full-RSB *Ansatz* is in any case marginally stable. In the p-spin models with $p \geq 3$ instead it turns out that the averaged properties obtained in this way correspond to the asymptotic values derived with the stochastic dynamics starting from random initial conditions. This is quite a remarkable result.

3.3.1 Interpretation of replica results

Let us now discuss the implications of the solution to fully-connected disordered models obtained with the, for the moment, rather abstract replica formalism.

The interpretation uses heavily the identification of *pure states*. Their definition is a tricky matter that we shall not discuss in detail here. We shall just assume it can be done and use the analogy with the ferromagnetic system – and its two pure states – and the TAP results at fixed disorder. As we already know, which are the pure states, its properties, number, *etc.* can depend on the quenched disorder realization and fluctuate from sample to sample. We shall keep this in mind in the rest of our discussion.

Let us then distinguish the averages computed within a pure state and over all configuration space. In a ferromagnet with no applied magnetic field this is simple to grasp: at high temperatures there is just one state, the paramagnet, while at low temperatures there are two, the states with positive and negative magnetization. If one computes the averaged magnetization restricted to the state of positive (negative) magnetization one finds $m_{\rm eq} > 0$ ($m_{\rm eq} < 0$); instead, summing over all configurations $m_{\rm eq} = 0$ even at low temperatures. Now, if one considers systems with more than just two pure states, and one labels them with Greeks indices, averages within such states are denoted $\langle O\rangle_\alpha$ while averages taken with the full Gibbs measure are expressed as

$$\langle\, O\,\rangle = \sum_\alpha w_\alpha^J \,\langle\, O\,\rangle_\alpha \; . \tag{3.144}$$

w_α^J is the probability of the α state given by

$$w_\alpha^J = \frac{e^{-\beta F_\alpha^J}}{Z_J} \; , \qquad \text{with} \qquad Z_J = \sum_\alpha e^{-\beta F_\alpha^J} \tag{3.145}$$

and thus satisfying the normalization condition $\sum_\alpha w_\alpha^J = 1$. F_α^J can be interpreted as the total free-energy of the state α. These probabilities, as well as the state dependent averages, will show sample-to-sample fluctuations.

One can then define an *overlap between states*:

$$q_{J\alpha\beta} \equiv N^{-1} \sum_i \langle s_i\rangle_\alpha \langle s_i\rangle_\beta = N^{-1}\sum_i m_i^\alpha m_i^\beta \tag{3.146}$$

and rename the *self-overlap* the Edwards-Anderson parameter

$$q_{J\alpha\alpha} \equiv N^{-1}\sum_i \langle s_i\rangle_\alpha \langle s_i\rangle_\alpha \equiv q_{J{\rm EA}} \tag{3.147}$$

(assuming the result is independent of α). The statistics of possible overlaps is then characterized by a probability function

$$P_J(q) \equiv \sum_{\alpha\beta} w_\alpha^J w_\beta^J \; \delta(q - q_{\alpha\beta}^J) \; , \tag{3.148}$$

where we included a subindex J to stress the fact that this is a strongly sample-dependent quantity. Again, a ferromagnetic model serves to illustrate the meaning of $P_J(q)$. First, there is no disorder in this case so the J label is irrelevant. Second, the high-T equilibrium phase is paramagnetic, with $q = 0$. $P(q)$ is then a delta function with weight 1 (see the left panel in Fig. 3.17). In the low-T phase there are only two pure states with identical statistical properties and $q_{\rm EA} = m^2$. Thus, $P(q)$ is just the sum of two delta functions with weight 1/2 (central panel in Fig. 3.17).

Next, one can consider averages over quenched disorder and study

$$[\, P_J(q)\,] \equiv \int dJ\, P(J) \, \sum_{\alpha\beta} w_\alpha^J w_\beta^J \; \delta(q - q_{\alpha\beta}^J) \; . \tag{3.149}$$

How can one access $P_J(q)$ or $[\,P_J(q)\,]$? It is natural to reckon that

$$P_J(q) = Z^{-2} \sum_{\sigma s} e^{-\beta H_J(\{\sigma_i\})} e^{-\beta H_J(\{s_i\})}\; \delta\left(N^{-1} \sum_i \sigma_i s_i - q\right) \tag{3.150}$$

that is to say, $P_J(q)$ is the probability of finding an overlap q between two *real replicas* of the system with identical disordered interactions in equilibrium at temperature T. This identitiy gives a way to compute $P_J(q)$ and its average in a numerical simulation: one just has to simulate two independent systems with identical disorder in equilibrium and calculate the overlap.

But there is also, as suggested by the notation, a way to relate the pure state structure to the replica matrix q_{ab}. Let us consider the simple case

$$\begin{aligned}[\,m_i\,] \;&=\; \left[Z_J^{-1} \sum_{\{s_i\}} s_i \; e^{-\beta H_J(\{s_i\})} \right] = \left[\frac{Z_J^{n-1}}{Z_J^n} \sum_{\{s_i^1\}} s_i^1 \; e^{-\beta H_J(\{s_i^1\})} \right] \\ &=\; \left[\frac{1}{Z_J^n} \sum_{\{s_i^a\}} s_i^1 e^{-\beta \sum_{a=1}^n H_J(\{s_i^a\})} \right]\end{aligned} \tag{3.151}$$

where we singled out the replica index of the spin to average. This relation is valid for all n, in particular for $n \to 0$. In this limit the denominator approaches one and the average over disorder can be simply evaluated

$$[\,m_i\,] \;=\; \sum_{\{s_i^a\}} s_i^1 e^{-\beta H_{\rm eff}(\{s_i^a\})} \tag{3.152}$$

and introducing back the normalization factor $Z^n = 1 = \sum_{\{s_i^a\}} e^{-\beta \sum_{a=1}^n H_J(\{s_i^a\})}$ that becomes $Z^n = [\,\sum_{\{s_i^a\}} e^{-\beta \sum_{a=1}^n H_J(\{s_i^a\})}\,] = e^{-\beta H_{\rm eff}(\{s_i^a\})}$ we have

$$[\,m_i\,] = \langle\, s_i^a \,\rangle_{H_{\rm eff}} \tag{3.153}$$

with a any replica index. The average is taken over the Gibbs measure of a system with effective Hamiltonian $H_{\rm eff}$. In a replica symmetric problem in which all replicas are identical this result should be independent of the label a. Instead, in a problem with replica symmetry breaking the averages on the right-hand-side need not be identical for all a. This could occur in a normal vectorial theory with dimension n in which not all components take the same expected value. It is reasonable to assume that the full thermodynamic average is achieved by the sum over all these cases,

$$[\,m_i\,] = \lim_{n\to 0} \frac{1}{n} \sum_{a=1}^n \langle\, s_i^a \,\rangle_{H_{\rm eff}} \;. \tag{3.154}$$

Let us now take a less trivial observable and study the spin-glass order parameter q

$$\begin{aligned} q &\equiv [\,\langle\, s_i \,\rangle^2\,] = \left[Z_J^{-1} \sum_{\{s_i\}} s_i \; e^{-\beta H_J(\{s_i\})} \; Z_J^{-1} \sum_{\{\sigma_i\}} \sigma_i \; e^{-\beta H_J(\{\sigma_i\})} \right] \\ &= \left[\frac{Z^{n-2}}{Z^n} \sum_{\{s_i\},\{\sigma_i\}} s_i \sigma_i \; e^{-\beta H_J(\{s_i\}) - \beta H_J(\{\sigma_i\})} \right] \\ &= \left[\frac{1}{Z_J^n} \sum_{\{s_i^a\}} s_i^1 s_i^2 \; e^{-\beta \sum_{a=1}^n H_J(\{s_i^a\})} \right] \end{aligned} \tag{3.155}$$

In the $n \to 0$ limit the denominator is equal to one and one can then perform the average over disorder. Introducing back the normalization one then has

$$q = \langle\, s_i^a s_i^b \,\rangle_{E_{\rm eff}(\{s_i^a\})} \tag{3.156}$$

for any arbitrary pair of replicas $a \neq b$ (since $\langle\, s_i^a s_i^a \,\rangle = 1$ for Ising spins). The average is done with an effective theory of n interacting replicas characterized by $E_{\rm eff}(\{s_i^a\})$. Again, if there is replica symmetry breaking the actual thermal average is the sum over all possible pairs of replicas:

$$q = \lim_{n \to 0} \frac{1}{n(n-1)} \sum_{a \neq b} q^{ab} \; . \tag{3.157}$$

A similar argument allows one to write

$$q^{(k)} = [\,\langle\, s_{i_1} \dots s_{i_k} \,\rangle^2\,] = \lim_{n \to 0} \frac{1}{n(n-1)} \sum_{a \neq b} q_{ab}^k \; . \tag{3.158}$$

One can also generalize this argument to obtain

$$P(q) = [\, P_J(q) \,] = \lim_{n \to 0} \frac{1}{n(n-1)} \sum_{a \neq b} \delta(q - q^{ab}) \tag{3.159}$$

Thus, the replica matrix q_{ab} can be ascribed to the overlap between pure states.

Note that a small applied field, though uncorrelated with a particular pure state, is necessary to have non-zero local magnetizations and then non-zero q values.

The function $P(q)$ then extends the concept of order parameter to a function. In zero field the symmetry with respect to simultaneous reversal of all spins translates into the fact that $P_J(q)$ must be symmetric with respect to $q = 0$. $[\, P_J(q) \,]$ can be used to distinguish between the droplet picture prediction for finite dimensional spin-glasses – two pure states – that simply corresponds to

$$[\, P_J(q) \,] = \frac{1}{2}\delta(q - q_{\rm EA}) + \frac{1}{2}\delta(q + q_{\rm EA}) \tag{3.160}$$

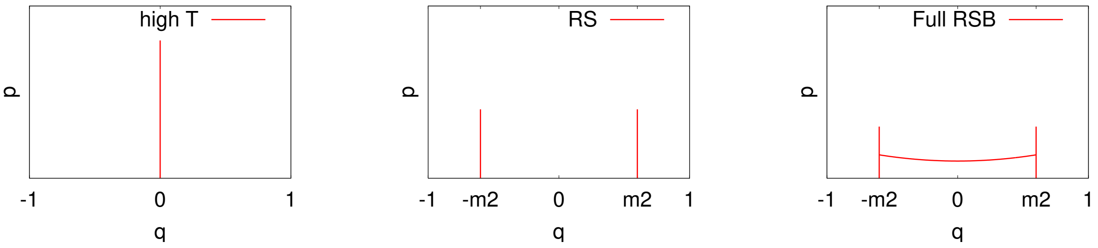

Figure 3.17: $[\,P_J(q)\,]$ in a paramagnet (left), in a ferromagnet or a replica symmetric system (centre) and for system with full RSB (right).

(see the central panel in Fig. 3.17) and a more complicated situation in which $[\,P_J(q)\,]$ has the two delta functions at $\pm q_{\rm EA}$ plus non-zero values on a finite support (right panel in Fig. 3.17) as found in mean-field spin-glass models.

The linear susceptibility

Taking into account the multiplicity of pure states, the magnetic susceptibility, eq. (**??**), and using (3.144) becomes

$$T\chi = T[\,\chi_J\,] = 1 - \frac{1}{N}\sum_i [\,\langle\, s_i\,\rangle^2\,] = 1 - \sum_{\alpha\beta} [\,w_\alpha^J w_\beta^J\,]\, q_{\alpha\beta} = \int dq\,(1-q)\,P(q)\;. \qquad (3.161)$$

There are then several possible results for the susceptibility depending on the level of replica symmetry breaking in the system:

- In a replica symmetric problem or, equivalently, in the droplet model,

  $$\chi = \beta(1-q_{\rm EA})\;. \qquad (3.162)$$

  This is also the susceptibility within a pure state of a system with a higher level of RSB.

- At the one step RSB level, this becomes

  $$\chi = \beta\,[1-(1-m)q_{\rm EA}]\;. \qquad (3.163)$$

- For systems with full RSB one needs to know the complete $P(q)$ to compute χ, as in (3.161).

Note that in systems with RSB (one step or full) the susceptibility is larger than $\beta(1-q_{\rm EA})$.

A system with $q_{\rm EA} = 1$ in the full low-temperature phase (as the REM model or $p \to \infty$ limit of the p spin model, see below) has just one configuration in each state. Systems with $q_{\rm EA} < 1$ below T_c have states formed by a number of different configurations that is exponentially large in N. (Note that $q_{\rm EA} < 1$ means that the two configurations differ in a number of spins that is proportional to N.) The logarithm of this number is usually called the intra-state entropy.

Even if the number of pure states can be very large (exponential in N) only a fraction of them can have a non-negligible weight. This is the case if one finds, for example, $\sum_\alpha w_\alpha^2 < +\infty$

Symmetry and ergodicity breaking

In all $p \geq 2$ spin models there is a phase transition at a finite T_s at which the rather abstract *replica symmetry* is broken. This symmetry breaking is accompanied by ergodicity breaking as in the usual case. Many pure states appear at low temperatures, each one has its reversed $s_i \to -s_i$ counterpart, but not all of them are related by real-space symmetry properties.

The one-step RSB scenario

In this case the transition has first-order and second-order aspects. The order parameters q_0 and q_1 jump at the critical point as in a first-order transition but the thermodynamic quantities are continuous.

The full RSB scenario

Right below T_c an exponential in N number of equilibrium states appear. The transition is continuous, the order parameter approaches zero right below T_c. Lowering further the temperature each ergodic component breaks in many other ones. In this sense, the full spin-glass phase, $T < T_c$, is 'critical' and not only the single point T_c.

3.3.2 The pinning field

We can nevertheless choose a possible direction, given by another field $\sigma(x)$, and compute the free–energy of our system when it is weakly pinned by this external quenched field

$$F_\phi\left[\sigma, g, \beta\right] = -\frac{1}{\beta} \log \int d\phi(x)\ e^{-\beta H[\phi] - \frac{g}{2}\int dx (\sigma(x)-\phi(x))^2} \qquad (3.164)$$

where $g > 0$ denotes the strength of the coupling. This free-energy (3.164) will be small when the external perturbing field $\sigma(x)$ lies in a direction corresponding to the bottom of a well of the unperturbed free-energy. Therefore, we should be able to obtain useful information about the free-energy landscape by scanning the entire space of the configurations $\sigma(x)$ to locate all the states in which the system can freeze after spontaneous ergodicity breaking ($g \to 0$). According to this intuitive idea, we now consider the field $\sigma(x)$ as a thermalized variable with the "Hamiltonian" $F_\phi\left[\sigma, g, \beta\right]$. The free-energy of the

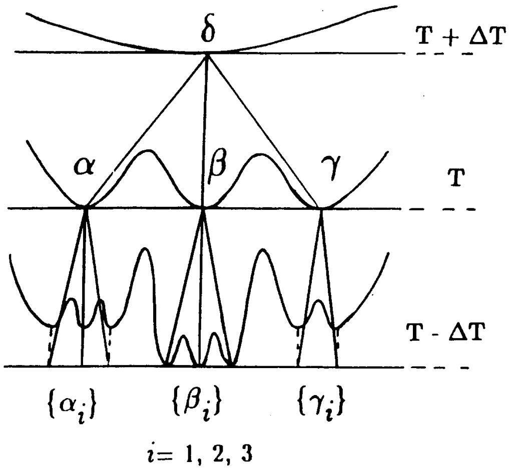

Figure 3.18: The susequent phase transitions in the SK model.

field σ at inverse temperature βm where m is a positive free parameter therefore reads

$$F_\sigma(m,\beta) = \lim_{g\to 0^+} -\frac{1}{\beta m}\log\int d\sigma(x)\ e^{-\beta m F_\phi[\sigma,g,\beta]} \tag{3.165}$$

When the ratio m between the two temperatures is an integer, one can easily integrate $\sigma(x)$ in Eq.(3.165) after having introduced m copies $\phi^\rho(x)$ ($\rho = 1...m$) of the original field to obtain the relation

$$F_\sigma(m,\beta) = \lim_{g\to 0^+} -\frac{1}{\beta m}\log\int\prod_{\rho=1}^{m} d\phi^\rho(x)\ e^{-\beta\sum_\rho H[\phi^\rho]+\frac{1}{2}\sum_{\rho,\lambda} g^{\rho\lambda}\int dx\phi^\rho(x)\phi^\lambda(x)} \tag{3.166}$$

where $g^{\rho\lambda} = g(\frac{1}{m} - \delta^{\rho\lambda})$. Let us define two more quantities related to the field σ : its internal energy $W(m,\beta) = \frac{\partial(mF_\sigma)}{\partial m}$ and its entropy $S(m,\beta) = \beta m^2\frac{\partial F_\sigma}{\partial m}$. Since the case $m = 1$ will be of particular interest, we shall use hereafter $F_{hs}(\beta) \equiv W(m = 1,\beta)$ and $\mathcal{S}_{hs}(\beta) \equiv S(m = 1,\beta)$ where hs stands for "hidden states". We stress that $S(m,\beta)$ and $\beta^2\frac{\partial F_\phi}{\partial\beta}$ which are respectively the entropies of the fields σ and ϕ are two distinct quantities with different physical meanings.

When the pinning field $\sigma(x)$ is thermalized at the same temperature as $\phi(x)$, that is when $m = 1$, one sees from Eq.(3.166) that $F_\phi(\beta) = F_\sigma(m = 1,\beta)$. The basic idea of this letter is to decompose F_σ into its energetic and entropic contributions to obtain

$$\mathcal{S}_{hs}(\beta) = \beta\Big[F_{hs}(\beta) - F_\phi(\beta)\Big] \tag{3.167}$$

To get some insights on the significance of the above relation, we shall now turn to the particular case of disordered mean-field systems. We shall see how it rigorously gives back some analytical results derived within the mean-field TAP and dynamical approaches. We shall then discuss the physical meaning of identity (3.167) for the general case of glassy systems.

3.3.3 Coupling replicas and the effective potential

Let us take a spin-configuration, $\{s\}$, in equilibrium at temperature T', that is to say, drawn from the canonical probability distribution $P[\{s\}] = \exp(-\beta' H[\{s\}])/Z(T')$. One computes the free-energy cost to keep the system at a fixed overlap $\tilde{p} = q_{s,\sigma}$ with $\{s\}$ at a temperature T (in general different from T'):

$$\begin{aligned} V_J(\beta, \tilde{p}, \{s\}) &= -\frac{T}{N} \ln Z_J(\beta, \tilde{p}, \{s\}) - f_J(T); && (3.168)\\ Z_J(\beta, \tilde{p}, \{s\}) &\equiv \sum_{\{\sigma\}} e^{-\beta H_J[\{\sigma\}]}\, \delta\left(\tilde{p} - q_{s,\sigma}\right) && (3.169)\\ \beta N f_J(T) &= \ln Z_J(\beta) = \ln \sum_{\{s\}} e^{-\beta H_J[\{s\}]}\,. && (3.170)\end{aligned}$$

($f_J(T)$ is the disorder-dependent free-energy density without constraint.) In this problem the spins s_i are quenched variables on the same footing as the random interactions in the Hamiltonian. One then assumes that V is self-averaging with respect to the quenched disorder and the probability distribution of the reference configuration $\{s\}$. One then computes the two averages:

$$NV(\beta, \beta', \tilde{p}) \equiv N[V_J(\beta, \tilde{p}, \{s\})]_{J,\{s\}} = \left[\sum_{\{s\}} \frac{e^{-\beta' H_J[\{s\}]}}{Z_J(\beta')} \;\left(-T \ln Z_J(\beta, \tilde{p}, \{s\}) - f_J(T)\right)\right]_J\,. \tag{3.171}$$

This average can be done using the replica method:

$$NV(\beta, \beta', \tilde{p}) = -T \lim_{n\to 0} \lim_{m\to 0} \left[\sum_{\{s\}} e^{-\beta' H_J[\{s\}]} Z_J(\beta')^{n-1} \left(\frac{Z_J(\{s\}; \tilde{p}, \{s\})^m - 1}{m}\right)\right]_J\,. \tag{3.172}$$

The analytic continuation is performed from integer n and m. One then has

$$Z^{(n,m)} = \left[\sum_{\{s^a\}} \sum_{\{\sigma^\alpha\}} \exp\left[\beta' \sum_{a=1}^{n} H[\{s^a\}] + \beta \sum_{\alpha=1}^{m} H[\{\sigma^\alpha\}]\right] \prod_{\alpha=1}^{m} \delta\left(\sum_i s_i^1 \sigma_i^\alpha - N\tilde{p}\right)\right]_J\,. \tag{3.173}$$

After averaging over the disorder strength distribution one introduces the order parameters:

$$Q_{ab} = \frac{1}{N}\sum_i s_i^a s_i^b \;, \qquad R_{\alpha\gamma} = \frac{1}{N}\sum_i \sigma_i^\alpha \sigma_i^\gamma \qquad P_{a\alpha} = \frac{1}{N}\sum_i s_i^a \sigma_i^\alpha \;, \tag{3.174}$$

with $a, b = 1, ..., n$ and $\alpha, \gamma = 1, ..., m$. Combining the order parameters in the single $(n+m) \times (n+m)$ matrix

$$Q = \begin{pmatrix} Q & P \\ P^T & R \end{pmatrix} \tag{3.175}$$

one finds

$$\frac{1}{N}\log Z^{(n,m)} = \frac{1}{2}\left[\beta'^2 \sum_{a=1,b=1}^{n} Q_{ab}^p + \beta^2 \sum_{\alpha=1\gamma=1}^{m} R_{\alpha,\gamma}^p) + 2\beta\beta' \sum_{a=1}^{n}\sum_{\alpha=1}^{m} P_{a\alpha}^p)\right] + \frac{1}{2}\mathrm{Tr}\ln Q. \tag{3.176}$$

We shall not present the details of the RSB *Ansatz* here.

One studies different ranges of β and β' and analyses the minima of V with respect to $\tilde{p}$.

The effective potential for four different temperatures, $T = T'$ for $p = 3$ is shown in [52] From top to bottom, the curves represent the potential at temperature higher then T_d, equal to T_d between T_d and T_S, and right at T_S. We can see from the figure that for $T > T_d$ the potential is monotonically increasing, and the only extremum of the potential is the minimum at $\tilde{p} = 0$. At the temperature T_d where the dynamical transition happens, the potential develops for the first time a minimum with $\tilde{p} \equiv r \neq 0$. It is interesting to observe that the energy in this flex point is equal to the asymptotic value of the energy in the out-of-equilibrium dynamics. The same is true for the parameter r which turns out to be equal to the dynamical Edward-Anderson parameter.

The condition for the potential of having a flex coincides with the marginality condition. Indeed the flex implies a zero eigenvalue in the longitudinal sector and at $x = 1$ the replicon and the longitudinal eigenvalues are degenerate. The marginality condition is well known to give exact results for the transition temperatures in p-spin spherical models.

We have observed that in general more then one minimum can be present in the potential. In the p-spin model it happens that two minima develop at the same temperature T_d. The rightmost one, that we will call primary is the one with $\tilde{p} = r$, while the other, secondary, has $\tilde{p} < r$. For temperatures smaller than T_d the minima have a finite depth, i.e. are separated by extensive barriers from the absolute minimum.

The primary minimum is easily interpreted. There the system denoted by s is in the same pure state as the system σ. In the region $T_S < T < T_D$ the number of pure states is exponentially large in N: $\mathcal{N} = e^{N\Sigma(T)}$. Consequently the probability of finding two system in the same state is exponentially small and proportional to $e^{-N\Sigma(T)}$. The free energy cost to constrain two systems to be in the same state is then proportional to the logarithm of this probability, namely we have

$$V_{primary} = T\Sigma(T). \tag{3.177}$$

Coherently at the statical transition temperature $T = T_S$ one finds $V_{primary} = 0$. The quantity Σ has been computed for the p-spin model as the number of solution of the TAP equation with given free energy and coincides with our calculation. The secondary minima, could aso be associated to metastable states, but at present we do not have an interpretation for them. This conclusion on the equivalence of the potential with the number of solution of the TAP equation hold also in the ROM.

The study of the potential for temperatures smaller than T_S would require to take into account RSB effects, which would complicate a bit the analysis. However it is physically clear that the shape of the potential in that region it is not different qualitatively from the one at $T = T_S$. It has a minimum where $r = \tilde{p}$ are equal to the Edwards Anderson parameter and the value of potential is zero.

The study of the effective potential at different gives information about the chaotic properties of the models. We shall not develop it here.

A Jenssen's inequality

Jensen's inequality relates the value of a convex function of an integral to the integral of the convex function. In its simplest form the inequality states that the convex transformation of a mean is less than or equal to the mean applied after convex transformation; it is a simple corollary that the opposite is true of concave transformations.

In probability theory, the Jenssen's inequality implies that, for x a random variable and ϕ a convex function, then

$$\phi(\mathbb{E}(x)) \leq \mathbb{E}[\phi(x)] \tag{A.1}$$

where $\mathbb{E}[\dots]$ is the expectation value of ...

B The central limit theorem

In probability theory, the central limit theorem (CLT) establishes that, in most situations, when independent random variables are added, their properly normalized sum tends toward a normal (Gaussian) distribution (informally a "bell curve") even if the original variables themselves are not normally distributed. More precisely, for x_i *i.i.d.* with average μ and variance σ^2,

$$X = \frac{1}{N}\sum_i x_i \tag{B.1}$$

is a Gaussian distributed with average $[X] = \mu$ and variance $[(X - [X])^2] = \sigma^2/N$.

They all express the fact that a sum of many independent and identically distributed (*i.i.d.*) random variables, or alternatively, random variables with specific types of dependence, will tend to be distributed according to one of a small set of attractor distributions. When the variance of the i.i.d. variables is finite, the attractor distribution is the normal distribution. In contrast, the sum of a number of $i.i.d.$ random variables with power law tail distributions decreasing as $|x|^{-\alpha-1}$ where $0 < \alpha < 2$ (and therefore having infinite variance) will tend to an alpha-stable distribution with stability parameter (or index of stability) of α as the number of variables grows.

C Classical results in statistical physics

We recall here some classical results in statistical physics.

C.1 High temperature expansion

The partition function of the Ising ferromagnet reads

$$Z = \sum_{s_i=\pm 1} e^{\beta J \sum_{\langle ij \rangle} s_i s_j} = \sum_{s_i=\pm 1} \prod_{\langle ij \rangle} e^{\beta J s_i s_j} \tag{C.2}$$

Using the identity $e^{\beta J s_i s_j} = a(1+bs_i s_j)$ with $a = \cosh(\beta J)$ and $b = \tanh(\beta J)$ and the fact that b is order β, an expansion if powers of b can be established. The average of products of the spins s's that remains can be non-zero only if each spin appears an even number of times s. The expansion can then be represented as graphs on the lattice, a representation that makes the enumeration of terms easier.

C.2 Lee-Yang theorem

The LeeÐYang theorem states that if partition functions of models with ferromagnetic interactions are considered as functions of an external field, then all zeros are purely imaginary (or on the unit circle after a change of variable) [61].

C.3 Critical behaviour

Second order phase transitions are characterised by the diverge of the correlation length. In normal conditions, far from the critical point, the correlation function of the fluctuations of an observable decay as an exponential of the distance between the measuring points:

$$C(\vec{r}) \equiv \langle [O(\vec{r}+\vec{r}') - \langle (O(\vec{r}+\vec{r}') \rangle][O(\vec{r}') - \langle (O(\vec{r}') \rangle \rangle] \rangle \simeq e^{-r/\xi} \; . \tag{C.3}$$

ξ is the correlation length that diverges at the critical point as

$$\xi \simeq |T - T_c|^{-\nu} \tag{C.4}$$

with ν a critical exponent. A power-law singularities in the length scales leads to power-law singularities in observable quantities. We summarise in Table 1 all the critical exponents associated to various quantities in a second order phase transition. The values of the critical exponents generally do not depend on the microscopic details but only on the space dimensionality and the symmetries of the system under consideration.

The collection of all these power laws characterizes the critical point and is usually called the critical behavior.

Whether fluctuations influence the critical behavior depends on the space dimensionality d. In general, fluctuations become less important with increasing dimensionality.

In sufficiently low dimensions, *i.e.* below the lower critical dimension d_l, fluctuations are so strong that they completely destroy the ordered phase at all (nonzero) temperatures and there is no phase transition. Between d_l and the upper critical dimension d_u, order at low temperatures is possible, there is a phase transition, and the critical exponents are influenced by fluctuations (and depend on d). Finally, for $d > d_u$, fluctuations are unimportant for the critical behavior, and this is well described by mean-field theory. The exponents become independent of d and take their mean-field values. For example, for Ising ferromagnets, $d_l = 1$ and $d_u = 4$, for Heisenberg ferromagnets $d_l = 2$ and $d_u = 4$.

	exponent	definition	conditions
Specific heat	α	$c \propto \|u\|^{-\alpha}$	$u \to 0,\ h = 0$
Order parameter	β	$m \propto (-u)^{\beta}$	$u \to 0-,\ h = 0$
Susceptibility	γ	$\chi \propto \|u\|^{-\gamma}$	$u \to 0,\ h = 0$
Critical isotherm	δ	$h \propto \|m\|^{\delta}\mathrm{sign}(m)$	$h \to 0,\ u = 0$
Correlation length	ν	$\xi \propto \|r\|^{-\nu}$	$r \to 0,\ h = 0$
Correlation function	η	$G(\vec{r}) \propto \|\vec{r}\|^{-d+2-\eta}$	$r = 0,\ h = 0$

Table 1: Definitions of the commonly used critical exponents. m is the order parameter, *e.g.* the magnetization, h is an external conjugate field, *e.g.* a magnetic field, u denotes the distance from the critical point, *e.g.* $|T - T_c|$, and d is the space dimensionality.

References

[1] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, *Complex networks: Structure and dynamics*, Phys. Rep. **424**, 175 (2006).

[2] A. Barrat, M. Barthelemy, and A. Vespignani, *Dynamical processes on complex networks* (Cambridge University Press, Cambridge, 2008).

[3] A. L. Efros, *Physics and geometry of disorder: percolation theory*, (Mir Publishers, Mockba, 1986)

[4] D. Stauffer and A. Aharony, *Introduction To Percolation Theory*, (Taylor and Francis, London,1994)

[5] K. Christensen *Percolation theory*, unpublished (2002).

[6] A. A. Saberi, *Recent advances in percolation theory and its applications*, Phys. Rep. **578**, 1 (2015).

[7] J. A. Mydosh, *Spin Glasses* (Taylor & Francis, 1985).

[8] M. Mézard, G. Parisi, and M. A. Virasoro, *Spin glasses and beyond; an intriduction to the replica method and its applications*, (World Scientific, Singapore, 1986).

[9] K. Binder and A. P. Young, *Spin glasses: Experimental facts, theoretical concepts, and open question*, Rev. Mod. Phys. **58** 801 (1986).

[10] K. H. Fischer and J. Hertz, *Spin glasses* (Cambridge Univ. Press, 1991).

[11] H. Nishimori, *Statistical phyiscs of spin glasses and information processing*, (Oxford Univ. Press, 2001).

[12] P. Nordblad, *Spin Glasses: Model systems for non-equilibrium dynamics*, arXiv:cond-mat/0309602.

[13] E. Vincent, J. Hammann, M. Ocio, J.-P. Bouchaud and L. F. Cugliandolo, *Slow dynamics and aging in spin glasses*, Complex Behaviour of Glassy Systems, 184-219 (1997).

[14] R. Brout, Phys Rev. **115**, 824 (1959).

[15] S. F. Edwards and P. W. Anderson, *Theory of spin glasses*, J. Phys. F: Metal Phys. **5**, 965 (1975).

[16] D. S. Sherrington and S. Kirkpatrick, *Solvable model of a spin-glass*, Phys. Rev. Lett. **35**, 1792 (1975).

[17] J.-P. Bouchaud, L. F. Cugliandolo, J. Kurchan, and M. Mezard, *Out of equilibrium dynamics in spin-glasses and other glassy systems*, in “Spin-glasses and random fields”, A. P. Young Ed. (World Scientific), arXiv:cond-mat/9702070.

[18] A. Cavagna, *Supercooled Liquids for Pedestrians*, Phys. Rep. **476**, 51 (2009).

[19] L. Berthier and G. Biroli *Theoretical perspective on the glass transition and amorphous materials*, Rev. Mod. Phys. **83**, 587 (2011).

[20] N. G. Fytas, V. Martín-Mayor, M. Picco, and N. Sourlas, *Review of recent developments in the random-field Ising model*, arXiv:1711:09597

[21] T. H. Berlin and M. Kac, *The spherical model of a ferromagnet*, Phys. Rev. **86**, 821 (1952).

[22] J. M. Kosterlitz, D. J. Thouless, and R. C. Jones, *Spherical Model of a Spin-Glass*, Phys. Rev. Lett. **36**, 1217 (1976).

[23] P. Clifford and A. Sudbury, *Model for spatial conflict*, Biometrika 60, 581 (1973).

[24] M. Mézard and A. Montanari, *Information, Physics, and Computation*, (Oxford University Press, 2009).

[25] D. C. Mattis, *Solvable Spin Systems with Random Interactions*, Phys. Lett. A **56**, 421 (1976).

[26] See, for example, M. Kardar in Disorder and Fracture, ed. J. C. Charmet, et al. (Plenum, New York, 1990).

[27] T. Nattermann, *Theory of the Random Field Ising Model*, in "Spin Glasses and Random Fields", ed. P. Young (World Scientific), arXiv:cond-mat/9705295.

[28] M. Aizenman and Wehr, Phys. Rev. Lett. **62**, 2503 (1989).

[29] J. Bricmont and A. Kupiainen, *Phase transition in the 3d random field Ising model* Comm. Math. Phys. **116**, 539 (1988).

[30] J. P. Sethna, O. Perkovic, and K.A. Dahmen, *Hysteresis, Avalanches, and Barkhausen Noise*, in "Scale Invariance and Beyond", B. Dubrulle, F. Graner, and D. Sornette eds., Les Houches Workshop, March 10?14, 1997.

[31] A. I. Larkin, Sov. Phys. JETP **31**, 784 (1970).

[32] S. Lemerle, J. Ferré, C. Chappert, V. Mathet, T. Giamarchi, and P. Le Doussal, *Domain Wall Creep in an Ising Ultrathin Magnetic Film*, Phys. Rev. Lett. **80**, 849 (1998).

[33] T. Vojta, *Phases and phase transitions in disordered quantum systems*, arXiv:1301.7746

[34] A. B. Harris, *Effects of random defects on critical behaviour of Ising models*, J. Phys. C **7**, 1671 (1974).

[35] A. B. Harris, *The Harris criterion lives on* J. Phys.: Condens. Matter **28**, 421006 (2016).

[36] J. T. Chayes, L. Chayes, D. S. Fisher and T. Spencer, *Correlation length bounds for disordered Ising ferromagnets*, Comm. Math. Phys. **120**, 501 (1989).

[37] R. B. Griffiths, *Non-analytic behavior above the critical point in a random Ising ferromagnet*, Phys. Rev. Lett. **23**, 17 (1969).

[38] Y. Imry and S-k Ma, *Random-Field Instability of the Ordered State of Continuous Symmetry*, Phys. Rev. Lett. **35**, 1399 (1975).

[39] D. S. Fisher, *Low temperature phases, ordering and dynamics in random media*, in *Phase transitions and relaxation in system with competing energy scales*, T. Riste and D. S. Sherrington eds. (Kruwer Academic Series, 1993).

[40] A. J. Bray and M. A. Moore, *Metastable states in spin glasses*, J. Phys. C: Solid State Phys. **13**, L469 (1980), *Chaotic nature of the spin-glass phase*, Phys. Rev. Lett. **58**, 57 (1987).

[41] D. S. Fisher and D. A. Huse, *Equilibrium behavior of the spin-glass ordered phase*, Phys. Rev. B **38**, 386 (1988).

[42] D. J. Thouless, P. W. Anderson and R. G. Palmer, *Solution of 'Solvable model of a spin glass'* Phil. Mag. **35**, 593 (1977).

[43] G. Biroli, *Dynamical TAP approach to mean field glassy systems*, J. Phys. A: Math. Gen. **32**, 8365 (1999).

[44] M. L. Mehta, *Random matrices* (Academic Press, 2004).

[45] C. Göbl and N. Tjandra, *Application of Solution NMR Spectroscopy to Study Protein Dynamics*, Entropy **14**, 581 (2012).

[46] C. Yamaguchi, *Exact Results for the Jarzynski Equality in Ising Spin Glass Models Derived by Using a Gauge Symmetry*, arXiv:1203.3332.

[47] C. Jarzynski, Phys. Rev. Lett. **78**, 2690 (1997). Phys. Rev. E **56**, 5018 (1997).

[48] F. Tanaka and S. F. Edwards, J. Phys. F **10**, 2769 (1980).

[49] A. J. Bray and M. A. Moore, J. Phys. C **13**, L469 (1980).

[50] T. Aspelmeier, R. A. Blythe, A. J. Bray, and M. A. Moore, *Free energy landscapes, dynamics and the edge of chaos in mean-field models of spin glasses*, Phys. Rev. B **74**, 184411 (2006).

[51] R. Monasson, *Structural glass transition and the entropy of metastable states*, Phys. Rev. Lett. **75**, 2847 (1995).

[52] S. Franz and G. Parisi, *Recipes for metastable states in spin-glasses*, J. Phys. I **5**, 1401 (1995).

[53] L. F. Cugliandolo and J. Kurchan, *Analytical solution of the off-equilibrium dynamics of a long-range spin-glass model*, Phys. Rev. Lett. **71**, 173 (1993).

[54] D. Mehta, D. A. Stariolo, and M. Kastner, *Energy landscape of the finite-size mean-field 3-spin spherical model*, Phys. Rev. E **87**, 052143 (2013).

[55] J. L. van Hemmen and R. G. Palmer, *The Replica Method and a Solvable Spin Glass Model*, J. Phys. A **12**, 563 (1979).

[56] G. Parisi, *The order parameter for spin glasses: a function on the interval 0-1*, J. Phys. A: Math. Gen. 13: 1101Ð1112 (1980).

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

[58] M. Talagrand, *Replica symmetry breaking and exponential inequalities for the SherringtonÐKirkpatrick model*, Annals of Probability **28**, 1018 (2000).

[59] J. R. L. de Almeida and D. J. Thouless, *Stability of the Sherrington-Kirkpatrick solution of a spin-glass model*, J. Phys. A. **11**, 983 (1978).

[60] J.-P. Bouchaud, L. F. Cugliandolo, J. Kurchan and M. Mézard, *Out of equilibrium dynamics in spin-glasses and other glassy systems*, in *Spin glasses and random fields*, A. P. Young ed. (World Scientific, 1997).

[61] T. D. Lee and C. N. Yang, *Statistical Theory of Equations of State and Phase Transitions. II. Lattice Gas and Ising Model*, Phys. Rev. **87**, 410 (1952).