UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Fisica e Astronomia "Galileo Galilei"

Models Of Theoretical Physics

Gradenigo's Lectures on:

Ergodicity Breaking Transitions in Disordered Systems

Written by: **Zanola Andrea, Rosset Lorenzo**

A.A: **2020/2021**

Contents

1 Introduction

Here we are going to introduce a topic which arises from the study of condensed-matter, spin glasses and super-fluids, but it turns out to be a very important theoretical framework also for understanding the problem of Inference and Learning Algorithms.

We will start by recalling the concept of ergodic system and what we mean when we talk about ergodicity breaking. Then, we're going to introduce the p-spin model, which is the test bench for our discussion. After computing the free energy for this model, we will introduce the Franz-Parisi potential, a very useful concept for exploring the thermodynamic landscape of our system. Finally, we are going to show how the concepts we developed can be also useful in order to understand some concepts related to information theory and machine learning.

1.1 Ergodic Systems and Ergodicity Breaking

We say that a system is **ergodic** if the dynamics, for almost all initial configurations, visits all regions of the available phase space with uniform probability. If the ergodicity hypothesis holds, we can associate to each micro-state a probability to occur, which from a canonical perspective is given by the Gibbs weight

$$P(\mathscr{C}_0) = \frac{e^{-\beta H[\mathscr{C}_0]}}{\int \mathscr{D}\mathscr{C}\, e^{-\beta H[\mathscr{C}]}}$$

Given any macroscopic observable $\mathcal{A}$, we have that the average value computed over the dynamics of the system corresponds to the weighted average computed over all the possible configurations, namely the ensemble average

$$\mathbb{E}[\mathcal{A}] \equiv \lim_{t\to\infty} \frac{1}{t} \int_0^t d\tau \mathcal{A}(\mathscr{C}(\tau)|\mathscr{C}_0) = \int \mathscr{D}\mathscr{C}\, \mathcal{A}(\mathscr{C}) P(\mathscr{C}) \equiv \langle \mathcal{A} \rangle \tag{1}$$

Of particular interest are the systems for which, under certain conditions, the equivalence between thermodynamic and dynamic averages written in Eq. (1) does not hold.

A simple example is given by magnetic systems: above the critical temperature T_c all the configurations of spins are allowed, and on average the magnetization of the system is zero. However, once the temperature goes below T_c, the system choices one of the new equilibrium states with non-zero magnetization. The **ergodicity is broken**, because now the available states are only those that agree with the magnetization of the system. Theoretically, due to the thermal fluctuations, it would be always possible that the system switches from one equilibrium state to the other, but in practice the time required for this to happen is much larger than the observation time.

Other systems, such as the spin glasses, display a much more complicated behaviour, and the ergodicity breaking comes out in very fancy fashions.

2 P-Spins Model

2.1 Introduction

The variables of our model are continuous spins $\sigma_i \in [-\infty, +\infty]$, $i = 1...N$.

The Hamiltonian for a generic p-spin model is:

$$H_{J,p} = -\sum_{i_1<i_2<...<i_p} J_{i_1,i_2,...,i_p}\sigma_{i_1}\sigma_{i_2}\cdots\sigma_{i_p} \tag{2}$$

In our discussion we're going to consider the case $p = 3$, with Hamiltonian

$$H_J = -\sum_{i<j<k} J_{ijk}\sigma_i\sigma_j\sigma_k \tag{3}$$

where the coefficients $\{J_{ijk}\}_{i<j<k} \equiv \vec{J}$ represents what we call the *disorder* of the system. We define the configuration vector of the spins, namely the micro-state of the system, as $\vec{\sigma} = \{\sigma_1, ..., \sigma_N\}$. The probability for a given configuration to occur is then

$$P_J[\vec{\sigma}] = \frac{e^{-\beta H_J[\vec{\sigma}]}}{Z_J[\beta]}, \tag{4}$$

where $\beta = \frac{1}{T}$.

The partition function is given by

$$Z_J[\beta] = \int_{-\infty}^{+\infty} \prod_{i=1}^{N} d\sigma_i \, e^{-\beta H_J[\vec{\sigma}]} \delta\left(\sum_{i=1}^{N} \sigma_i^2 - N\right) = \int_{\mathbb{R}} \mathscr{D}\sigma \, e^{-\beta H_J[\vec{\sigma}]} \tag{5}$$

where we defined the measure

$$\mathscr{D}\sigma = \prod_{i=1}^{N} d\sigma_i \; \delta\left(\sum_{i=1}^{N} \sigma_i^2 - N\right). \tag{6}$$

The spherical constraint $\sum_{i=1}^{N} \sigma_i^2 = N$ implemented in Eq. (6) is needed for technical reasons, i.e., in order to have a bounded energy.

Given an observable $\mathcal{A}[\vec{\sigma}] : \mathbb{R}^N \to \mathbb{R}$, we can compute its average value as:

$$\langle \mathcal{A} \rangle = \int_{\mathbb{R}} \mathscr{D}\sigma \, \mathcal{A}[\vec{\sigma}] P_J[\vec{\sigma}] \tag{7}$$

Once we manage to compute the partition function, the free energy of the system for a given disorder realization is straightforwardly given by

$$F_J = -\frac{1}{\beta} \log\big(Z_J[\beta]\big) \tag{8}$$

and this is the most important object that we want to study.
The quenched disorder J_{ijk} are random variables that follow a gaussian distribution

$$P(J_{ijk}) = \frac{1}{\sqrt{2\pi}\sigma_N} e^{-\frac{1}{2}\left(\frac{J_{ijk}}{\sigma_N}\right)^2} \tag{9}$$

where σ_N^2 is the variance (the letter N remarks the explicit dependence on the number of spins), defined as

$$\sigma_N^2 = \frac{p!}{2N^{p-1}} \tag{10}$$

This definition is important in order to guarantee the extensivity of F_J, that is $F_J \sim N$. Note also that $P(J_{ijk}) \xrightarrow{N\gg 1} \delta(J_{ijk})$.

2.2 Quenched Disorder

When we talk about quenched disorder (in italian "disordine fissato"), we refer to the idea of *thermalize the system by keeping the disorder configuration fixed.*
The physical observable (thermodynamic potential) is the total free energy, which from the thermodynamic we know to be given by $F = U - TS$.
Giving the fact that we can have many possible disorder configurations, we want to define the **quenched free energy** as

$$F_J = -\frac{1}{\beta}\log\big(Z_J[\beta]\big) \to \overline{F_J} = -\frac{1}{\beta}\overline{\log\big(Z_J[\beta]\big)} \tag{11}$$

where the over-bar means "averaged over disorder". More explicitly, this can be written as

$$\overline{F_J} = -\frac{1}{\beta}\int \mathscr{D}J P\big[\vec{J}\big]\log\left(\int_{\mathbb{R}} \mathscr{D}\sigma\, e^{-\beta H_J[\vec{\sigma}]}\right) \tag{12}$$

where we denoted:

$$\mathscr{D}J = \prod_{i<j<k} dJ_{ijk} \tag{13}$$

Another way to do define the average free energy is what is called **annealed free energy**:

$$\overline{F_J} = -\frac{1}{\beta}\log\left[\int \mathscr{D}J\mathscr{D}\sigma P\big[\vec{J}\big]\, e^{-\beta H_J}\right] \tag{14}$$

The difference between the two definitions is that the quenched free energy is obtained by thermalizing the system with a fixed disorder and then taking the average over all possible disorder configurations, while the annealed free energy is obtained by making the disorder evolve with the system. In general the annealed free energy is easier to calculate, because the logarithm is outside the integral, but gives certainly correct results when ergodicity is not broken (high temperature regime). If for example we have a problem in which there is disorder but it's a dynamic variable, and it evolves with the same time scale of the other variables of the system, then the disorder is just only another variable, so we can average it as usual.

For the p-spin model the free energy has the important property to be **self-averaging**. Indeed, given that each configuration $\vec{J}$ is a random variable, we have that the free energy for a given quenched disorder configuration is also a random variable whose mean value is given by

$$\overline{F_J} = \int \mathscr{D}JP\big[\vec{J}\big]F_J \tag{15}$$

and the variance is:

$$\mathrm{Var}(F) = \overline{F_J^2} - \overline{F}_J^2 \tag{16}$$

The self-averaging property reads

$$\lim_{N\to+\infty} \frac{\sqrt{\mathrm{Var}(F)}}{\overline{F_J}} \sim \frac{1}{\sqrt{N}} \Rightarrow P(F) \xrightarrow{N\to+\infty} \delta(F - \overline{F_J}) \tag{17}$$

This is important, because it allows to compute a well defined free energy per spin in the large system size limit:

$$\boxed{\lim_{N\to+\infty} \frac{F_J}{N} = \lim_{N\to+\infty} \frac{\overline{F_J}}{N}} \tag{18}$$

To compute $\overline{F_J}$ we will need to use the (in-)famous **replica trick**:

$$f = \lim_{N\to+\infty} \frac{F_J}{N} = \lim_{N\to+\infty} \frac{\overline{F_J}}{N} = \lim_{N\to+\infty} -\frac{1}{\beta N}\overline{\log\big(Z_J\big)} = \lim_{\substack{N\to+\infty \\ n\to 0}} -\frac{1}{\beta N n}\big(\overline{Z_J^n} - 1\big) \tag{19}$$

which is nothing more than a Taylor expansion of an exponential:

$$\lim_{n\to 0} \frac{x^n - 1}{n} = \lim_{n\to 0} \frac{e^{n\log(x)} - 1}{n} = \lim_{n\to 0} \frac{1 + n\log(x) + \mathcal{O}(n^2) - 1}{n} = \log(x) + \mathcal{O}(n) \tag{20}$$

2.3 Phase space and dynamics to explore it

The phase space of the system is the $N-1$ hyper-surface $\sum_{i=1}^N \sigma_i^2 = N$. We say that ergodicity breaking transition occurs when the phase space breaks down into disjoint ergodic components (in condensed matter we talk about *glass transition*).
In the p-spin model we have two kinds of ergodicity breaking transitions:

- Ergodicity broken dynamically (DB) at temperature T_D
- Ergodicity broken thermodynamically (TB) at the Kauzmann temperature T_K

It is always true that $T_K < T_D$. We have a dynamic breaking when, if we start from a configuration $\vec{\sigma_0}$, we cannot explore the whole phase space, but we are constrained inside an egodic connected component. For $T_K < T < T_D$ we have DB without TB. In this region the probability for a configuration to belong to a given ergodic component, say A, is zero: $P(\vec{\sigma} \in A) = 0$. In other words, if we sample the system according to $P \sim \exp(-\beta H[\vec{\sigma}])$, the phase space volume associated to A is negligible.

Instead, for $T < T_K$ the thermodynamic breaking occurs, and we have disjoint ergodic components with a non-zero phase space volume. Of course, the thermodynamic breaking always implies dynamical breaking, while the vice-versa is not true.

2.3.1 Dynamic ergodicity breaking

For our system we can consider a Langevin dynamics

$$\dot{\sigma}_i(t) = -\frac{\partial H_J}{\partial \sigma_i} + \eta_i(t)$$

where $\eta_i(t)$ is the white noise associated to thermal fluctuations

$$\langle \eta_i(t) \rangle = 0 \qquad \langle \eta_i(t)\, \eta_j(t') \rangle = 2T\delta_{ij}\delta(t-t')$$

Formally, if $\phi_s(\vec{\sigma_0})$ is the flux of the Langevin dynamics, when $T < T_D$ the average over the dynamics is no more equal to the ensemble average:

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t ds\, \mathcal{A}[\phi_s(\vec{\sigma_0})] \neq \langle \mathcal{A} \rangle$$

In practice, what happens for $T < T_D$ is that the number of disjoint ergodic components $\mathcal{N}$ grows as $\mathcal{N} \sim e^{N\Sigma}$, where Σ is the *configurational entropy*, hence it is exponentially unlikely to fall in one of them. By looking at the energy landscape, we have that the number of minima grows exponentially, and becomes more and more unlikely to end up in one of them.

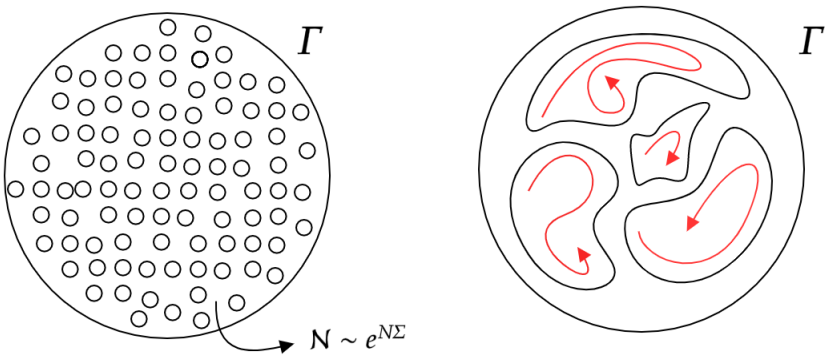


Figure 1: Phase space landscape for the dynamic breaking (left) and the thermodynamic breaking (right). Don't be deceived by the dimension of the areas: the number of microstates within each component doesn't change; the point is that only a subset of the energy minima survives.

Now we are going to show that for $T_K < T < T_D$ we do have states (disjoint ergodic components), but they are statistically irrelevant. Let us compute the partition function for a given disorder configuration:

$$\begin{aligned} Z_J(\beta) &= \int \mathscr{D}\sigma\; e^{-\beta H_J[\vec{\sigma}]} \overset{(\star)}{=} \sum_{\alpha=1}^{\mathcal{N}} \int_{\vec{\sigma}\in\mathbb{B}_\alpha} \mathscr{D}\sigma\; e^{-\beta H_J[\vec{\sigma}]} \\ &= \sum_{\alpha=1}^{\mathcal{N}} Z_\alpha = \sum_{\alpha=1}^{\mathcal{N}} e^{-\beta N f_\alpha}, \end{aligned} \tag{21}$$

where in $(\star)$ we took into account the fact that we are studying a system composed by $\mathcal{N}$ disjoint components $\{\mathbb{B}_\alpha\}_{\alpha=1,\dots,\mathcal{N}}$. Since each configuration $\vec{\sigma}$ takes values only inside his connected component $\mathbb{B}_\alpha$ the integration can be split as shown in the first line of Eq. (21).

Now we introduce a function *density of states* that counts the number of states with a given energy f

$$\mathcal{N}(f) = \sum_{\alpha=1}^{\mathcal{N}} \delta\big(f - f_\alpha\big). \tag{22}$$

Since the number of "states" is exponentially large in the systems size, we can define the so-called **configurational entropy** as:

$$\Sigma(f) = \frac{1}{N}\ \log \mathcal{N}(f), \tag{23}$$

where N is the number of degrees of freedom (spins) in the systems. The partition function can be written as:

$$\begin{aligned} Z_J(\beta) &= \int_0^\infty df\ e^{-\beta N f} \mathcal{N}(f) = \int_0^\infty df\ e^{-\beta N f + N\Sigma(f)} = \\ &= \int_0^\infty df\ e^{-\beta N \Phi(\beta, f)} \approx e^{-\beta N \Phi(\beta, f^*)} \end{aligned} \tag{24}$$

where we defined a new thermodynamic potential

$$\Phi(\beta, f) \equiv f - T\Sigma(f) = U - TS - T\Sigma(f), \tag{25}$$

and where f^* is defined as the solution of the saddle-point equation

$$\frac{\partial \Phi(f,\beta)}{\partial f} = 0 \qquad \Longrightarrow \qquad \frac{1}{T} = \left.\frac{\partial \Sigma}{\partial f}\right|_{f=f^*}.$$

We can get rid of the integration over free energy in Eq. (24) by means of a saddle-point approximation thanks to the large N limit we are interested in.

The probability associated to a given state α is given by

$$P_\alpha = \frac{Z_\alpha}{Z} = \frac{Z_\alpha}{\sum_{\gamma=1}^{\mathcal{N}} Z_\gamma} = \frac{e^{-\beta f_\alpha N}}{\sum_{\gamma=1}^{\mathcal{N}} e^{-\beta f_\gamma N}} \overset{(\star)}{=} \frac{e^{-\beta f^* N}}{e^{-\beta f^* N + N\Sigma(f^*)}} = e^{-N\Sigma(f^*)} \tag{26}$$

And hence we have that the probability is exponentially small. In $(\star)$ we used the saddle point approximation at the denominator, and at the numerator we assumed that, in the large-N limit, all typical configurations at a given temperature T have the same free-energy, i.e.

$$\lim_{N\to\infty} \frac{F_J(\beta)}{N} = f^*(\beta) + \mathcal{O}\left(\frac{1}{\sqrt{N}}\right) \tag{27}$$

apart from subleading corrections. This is equivalent to say that, in the large-N limit, by fixing the temperature T we are uniquely fixing even the free energy. This assumption is not generally true for all models, but it is true for the p-spin model considered here. Models where this hypothesis does not hold are said to have *"chaos in temperature"*.

2.4 Free Energy Calculation

In order to compute the free energy for the p-spin model we can exploit equation (19) and focus on the *replicated partition function* $\overline{Z_J^n}$. The non-averaged version is given by

$$Z_J^n = \int_{-\infty}^{+\infty} D\sigma \ \exp\left\{\beta \sum_{i<j<k} J_{ijk} \sum_{\alpha=1}^{n} \sigma_i^\alpha \sigma_j^\alpha \sigma_k^\alpha\right\} \ \prod_{\alpha=1}^{n} \delta\left(\sum_{i=1}^{N} (\sigma_i^\alpha)^2 - N\right) \tag{28}$$

where we denoted

$$D\sigma \equiv \prod_{i=1}^{N} \prod_{\alpha=1}^{n} d\sigma_i^\alpha \tag{29}$$

The average over the disorder reads as

$$\overline{Z_J^n} = \int_{\mathbb{R}} \mathscr{D}J \ P\big[\vec{J}\,\big] \ Z_J^n \tag{30}$$

but since $P(J_{ijk})$ is a gaussian distribution, in practice we have to compute the product of

$$N(N-1)...(N-p+1) = \binom{N}{p} \tag{31}$$

identical integrals as

$$\begin{aligned} I &= \int_{\mathbb{R}} dJ_{ijk} \frac{1}{\sqrt{2\pi}\sigma_N} \exp\left\{-J_{ijk}^2 \frac{N^{p-1}}{p!} + \beta J_{ijk} \sum_{\alpha=1}^{n} \sigma_i^\alpha \sigma_j^\alpha \sigma_k^\alpha\right\} \\ &= \sqrt{\frac{\pi p!}{N^{p-1}}} \frac{1}{\sqrt{2\pi}\sigma_N} \exp\left\{\frac{\beta^2 p!}{4N^{p-1}} \left(\sum_{\alpha=1}^{n} \sigma_i^\alpha \sigma_j^\alpha \sigma_k^\alpha\right)^2\right\} \\ &= \exp\left\{\frac{\beta^2 p!}{4N^{p-1}} \sum_{\alpha\beta=1}^{n} \sigma_i^\alpha \sigma_i^\beta \ \sigma_j^\alpha \sigma_j^\beta \ \sigma_k^\alpha \sigma_k^\beta\right\} \end{aligned}$$

where recall that $\frac{N^{p-1}}{p!}$ is the variance of P_{ijk}, so here we can appreciate the fact that this choice of P guarantees the extensivity of F. Indeed, otherwise there would be no way to apply the thermodynamic limit at the end, because of the presence of the pre-factor coming out from the gaussian integral. The corresponding term in equation (30) becomes

$$\begin{aligned} &= \exp\left\{\frac{\beta^2 p!}{4N^{p-1}} \sum_{i<j<k} \sum_{\alpha\beta=1}^{n} \sigma_i^\alpha \sigma_i^\beta \ \sigma_j^\alpha \sigma_j^\beta \ \sigma_k^\alpha \sigma_k^\beta\right\} \\ &= \exp\left\{\frac{\beta^2}{4N^{p-1}} \sum_{ijk} \sum_{\alpha\beta=1}^{n} \sigma_i^\alpha \sigma_i^\beta \ \sigma_j^\alpha \sigma_j^\beta \ \sigma_k^\alpha \sigma_k^\beta\right\} \\ &= \exp\left\{\frac{\beta^2}{4N^{p-1}} \sum_{\alpha\beta=1}^{n} \left(\sum_{i=1}^{N} \sigma_i^\alpha \sigma_i^\beta\right)^p\right\} \\ &= \exp\left\{\frac{\beta^2 N}{4} \sum_{\alpha\beta=1}^{n} \left(\frac{1}{N} \sum_{i=1}^{N} \sigma_i^\alpha \sigma_i^\beta\right)^p\right\} \end{aligned}$$

The averaged partition function can now be written as

$$\overline{Z_J^n} = \int D\sigma \ \exp\left\{\frac{\beta^2 N}{4}\sum_{\alpha\beta=1}^{n}\left(\frac{1}{N}\sum_{i=1}^{N}\sigma_i^\alpha\sigma_i^\beta\right)^p\right\}\ \prod_{\alpha=1}^{n}\delta\left(\sum_{i=1}^{N}(\sigma_i^\alpha)^2 - N\right) \tag{32}$$

where we can recognize the **overlap** between the configurations α and β:

$$Q_{\alpha\beta} = \frac{1}{N}\sum_{i=1}^{N}\sigma_i^\alpha\sigma_i^\beta \tag{33}$$

Remark 1: since $Q_{\alpha\beta}$ is an intensive quantity, the argument of the exponential in equation (32) is an extensive variable that can be interpreted as a free energy.

Remark 2: we started from an expression in which we had *independent replicas and coupled spins*:

$$-\beta\sum_{i<j<k}J_{ijk}\sum_{\alpha=1}^{n}\sigma_i^\alpha\sigma_j^\alpha\sigma_k^\alpha \tag{34}$$

and, after these calculations, we got a form in which *spins are decoupled but replicas are coupled*:

$$\frac{\beta^2 N}{4}\sum_{\alpha\beta=1}^{n}\left(\frac{1}{N}\sum_{i=1}^{N}\sigma_i^\alpha\sigma_i^\beta\right)^p \tag{35}$$

and note also that the disorder J_{ijk} has disappeared in (35), as we represented in Figure 2.

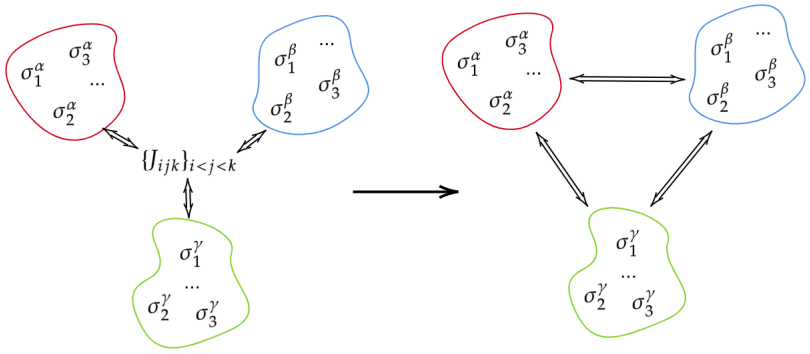


Figure 2: Pictorial way to show that, by using the replica trick, we decouple spins and couple replicas.

Now we want to change variables, switching to a description in terms of our order parameter Q. The trick is the same used to quantize gauge theories, and it consists on writing the identity $\mathbb{1}$ in a clever way as a function $\mathbb{1}[Q,\Lambda]$ of the overlap parameter matrix $Q = \{Q_{\alpha\beta}\}_{\alpha,\beta=1,\dots,n}$ and

of the related Lagrange multipliers matrix $\Lambda = \{\lambda_{\alpha\beta}\}_{\alpha,\beta=1,\dots,n}$:

$$\begin{aligned}
\mathbb{1}[Q,\Lambda] &\overset{(a)}{=} \int dQ_{\alpha\beta}\ \delta\Big(NQ_{\alpha\beta} - \sum_{i=1}^{N}\sigma_i^\alpha\sigma_i^\beta\Big) \\
&\overset{(b)}{=} \int \prod_{\alpha<\beta} dQ_{\alpha\beta}\ \prod_{\alpha<\beta}\delta\Big(NQ_{\alpha\beta} - \sum_{i=1}^{N}\sigma_i^\alpha\sigma_i^\beta\Big) \\
&\overset{(c)}{=} \int \prod_{\alpha<\beta} dQ_{\alpha\beta}\ \int_{\lambda^0_{\alpha\beta}-i\infty}^{\lambda^0_{\alpha\beta}+i\infty} d\lambda_{\alpha\beta}\ \exp\Big\{N\sum_{\alpha<\beta}\lambda_{\alpha\beta}\,Q_{\alpha\beta}\Big\}\exp\Big\{-\sum_{\alpha<\beta}\lambda_{\alpha\beta}\sum_{i=1}^{N}\sigma_i^\alpha\sigma_i^\beta\Big\}
\end{aligned}$$

where in (a) we used the overlap definition, in (b) we've done nothing but multiplying many times $\mathbb{1}$ and in (c) we opened the Dirac delta integrating through a straight vertical line in the complex plane, where $\lambda^0_{\alpha\beta}$ is the intersection between the integrating path and the real axis; the idea is to remove the hard constraint imposed by the delta by introducing an integral.
By multiplying the last identity to (32) we can insert the definition of $Q_{\alpha\beta}$:

$$\overline{Z_J^n} = \int D\sigma\ \exp\Big\{\frac{\beta^2 N}{4}\sum_{\alpha\beta=1}^{n}\big(Q_{\alpha\beta}\big)^p\Big\}\ \prod_{\alpha=1}^{n}\delta\left(Q_{\alpha\alpha}-1\right)\ \cdot\ \mathbb{1}[Q,\Lambda] \tag{36}$$

In the expression of the partition function in Eq. (36) we have exploited the constraint imposed within $\mathbb{1}[Q,\Lambda]$ to write

$$\prod_{\alpha=1}^{n}\delta\Big(\sum_{i=1}^{N}\big(\sigma_i^\alpha\big)^2 - N\Big)\ =\ \prod_{\alpha=1}^{n}\delta\Big(Q_{\alpha\alpha}-1\Big), \tag{37}$$

where we have neglected irrelevant N factors coming out from the Dirac deltas. We can thus define

$$\mathscr{D}\Lambda = \prod_{\alpha<\beta}^{n} d\lambda_{\alpha\beta} \tag{38}$$

$$\mathscr{D}Q = \prod_{\alpha<\beta}^{n} dQ_{\alpha\beta}\prod_{\alpha=1}^{n}\delta\Big(Q_{\alpha\alpha}-1\Big) \tag{39}$$

In conclusion, we can rewrite $\overline{Z_J^n}$ as:

$$\overline{Z_J^n} = \int \mathscr{D}Q\ \mathscr{D}\Lambda\ D\sigma \exp\Big\{\frac{\beta^2 N}{4}\sum_{\alpha\beta=1}^{n}\big(Q_{\alpha\beta}\big)^p\ +\ N\sum_{\alpha<\beta}\lambda_{\alpha\beta}\,Q_{\alpha\beta}\ -\ \sum_{\alpha<\beta}\lambda_{\alpha\beta}\sum_{i=1}^{N}\sigma_i^\alpha\sigma_i^\beta\Big\} \tag{40}$$

First, we want to integrate over the spins, and in particular we have N identical gaussian

integrals:

$$\begin{aligned}
I &= \int_{\mathbb{R}} D\sigma \exp\Big\{ - \sum_{\alpha<\beta} \lambda_{\alpha\beta} \sum_{i=1}^{N} \sigma_i^\alpha \sigma_i^\beta \Big\} \\
&= \int_{\mathbb{R}} \prod_{\alpha=1}^{n} \prod_{i=1}^{N} d\sigma_i^\alpha \exp\Big\{ - \sum_{\alpha<\beta} \lambda_{\alpha\beta} \sum_{i=1}^{N} \sigma_i^\alpha \sigma_i^\beta \Big\} \\
&= \Big[\int_{\mathbb{R}} \prod_{\alpha=1}^{n} d\sigma^\alpha \exp\Big\{ -\frac{1}{2} \sum_{\alpha\beta} \lambda_{\alpha\beta}\ \sigma^\alpha \sigma^\beta \Big\} \Big]^N \\
&= \Big[\int_{\mathbb{R}^n} d\vec{\sigma} \exp\Big\{ -\frac{1}{2} \vec{\sigma}^T \Lambda \vec{\sigma} \Big\} \Big]^N = \Big[\frac{(2\pi)^{n/2}}{\sqrt{\det \Lambda}} \Big]^N \\
&= \exp\big\{ -\frac{N}{2} \log(\det \Lambda) + \frac{Nn}{2} \log(2\pi)] \big\} \\
&= \exp\big\{ -\frac{N}{2} \operatorname{Tr}\big(\log(\Lambda)\big) \big\}
\end{aligned}$$

where in the last passage we neglected the constant term and we used the following property:

Prop: $\log(\det B) = \operatorname{Tr}\big(\log(B)\big)$

Proof. We just have to prove that $\det\big(\exp A\big) = \exp\big(\operatorname{Tr} A\big)$, then if we choose $A = \log(B)$ and we take the log at both sides, we are done.
Recall that for every complex squared matrix A there exists a matrix S such that $A = S^{-1}JS$, where J is the Jordan form. Then:

$$\begin{aligned}
\det\big(\exp A\big) &= \det\big(\exp\big(S^{-1}JS\big)\big) = \det\big(S^{-1}\exp\big(J\big)S\big) \\
&= \det S^{-1} \det\big(\exp J\big) \det S = \det S^{-1} \det S \det\big(\exp J\big) \\
&= \det\big(\exp J\big) \overset{(1)}{=} \prod_i e^{J_{ii}} = e^{\sum_i J_{ii}} \\
&= \exp\big(\operatorname{Tr}(J)\big) = \exp\big(\operatorname{Tr}(A)\big)
\end{aligned}$$

where in (1) we used the property that for a 4-blocks matrix

$$J = \begin{pmatrix} D_1 & E \\ C & D \end{pmatrix}$$

the determinant is $\det J = \det D_1 \det\big(D - CD_1^{-1}E\big)$. In a Jordan matrix, E and C are non-squared zero matrices, while in the diagonal we have the Jordan matrices D_j. Hence the previous formula can be applied iteratively and becomes $\det J = \prod_j \det D_j$, where j goes through the number of diagonal blocks. But the determinant of the j-esim block is just $\prod_{i|j} J_{ii|j}$ because of the form of D_j

$$D_j = \begin{pmatrix} \lambda_j & 1 & 0 & \dots & 0 \\ 0 & \lambda_j & 1 & \ddots & \vdots \\ 0 & 0 & \lambda_j & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 1 \\ 0 & 0 & 0 & \dots & \lambda_j \end{pmatrix}$$

□

So far we have obtained:

$$\overline{Z_J^n} = \int \mathscr{D}Q\mathscr{D}\Lambda \exp\Big\{ \frac{\beta^2 N}{4} \sum_{\alpha\beta=1}^{n} \big(Q_{\alpha\beta}\big)^p \; + \; \frac{N}{2} \operatorname{Tr}\big(\Lambda Q\big) \; - \; \frac{N}{2} \operatorname{Tr}\big(\log(\Lambda)\big) \Big\} \tag{41}$$

Now we can introduce the **Effective Action $\hat{S}$** such that

$$\overline{Z_J^n} = \int \mathscr{D}Q\mathscr{D}\Lambda \exp\big\{ N\hat{S}(\Lambda, Q)\big\} \tag{42}$$

and since we are interested in the large N limit, we can perform as usual a saddle point approximation

$$\overline{Z_J^n} = \int \mathscr{D}Q \;\; \exp\big\{ N\hat{S}(\Lambda^*(Q), Q)\big\} \tag{43}$$

where Λ^* is the solution of the saddle-point equations

$$\begin{aligned} \frac{\partial \hat{S}}{\partial \Lambda_{\alpha\beta}} = 0 \quad &\Longrightarrow \quad \frac{\partial}{\partial \Lambda}\left[\frac{1}{2} \operatorname{Tr}\big(\Lambda Q\big) - \frac{1}{2} \operatorname{Tr}\big(\log(\Lambda)\big) \right] = \mathbf{0} \\ &\Longrightarrow \quad \left[\operatorname{Tr}\big(Q\big) - \operatorname{Tr}\left(\frac{1}{\Lambda}\right) \right] = \mathbf{0} \\ &\Longrightarrow \quad \Lambda^* = Q^{-1}, \end{aligned}$$

where, having assumed that $\lambda_{\alpha\beta}$ are real and symmetric we can take derivatives and ignore trace operators as if Λ was a number. Once we put the saddle-point solution $\Lambda^* = Q^{-1}$ inside (43) we get

$$\begin{aligned} \overline{Z_J^n} &= \int \mathscr{D}Q \exp\Big\{ \frac{\beta^2 N}{4} \sum_{\alpha\beta=1}^{n} \big(Q_{\alpha\beta}\big)^p + \frac{N}{2} \operatorname{Tr}\big(\mathbb{1}_{n\times n}\big) + \frac{N}{2} \operatorname{Tr}\big(\log(Q)\big) \Big\} \\ &= \int \mathscr{D}Q \exp\Big\{ \frac{\beta^2 N}{4} \sum_{\alpha\beta=1}^{n} \big(Q_{\alpha\beta}\big)^p + \frac{N}{2} \log(\det(Q)) + \frac{Nn}{2} \Big\}. \end{aligned} \tag{44}$$

Finally, we can define an **Effective Free Energy** as:

$$A[Q] = -\frac{1}{n}\left[\frac{\beta^2}{4} \sum_{\alpha\beta=1}^{n} \big(Q_{\alpha\beta}\big)^p + \frac{1}{2} \log(\det(Q)) \right] \tag{45}$$

and we can ignore as usual the irrelevant constant terms. We have introduced a prefactor $1/n$ because we expect that in the limit $n \to 0$ the argument of the exponential in Eq. (44) will be proportional to n. Indeed, what we want to calculate is:

$$\lim_{N\to\infty} \frac{\overline{F_J}}{N} = \lim_{N\to\infty} \lim_{n\to 0} -\frac{1}{nN\beta}(\overline{Z_J^n} - 1) = \lim_{N\to\infty} \lim_{n\to 0} -\frac{1}{nN\beta} \left(\int \mathscr{D}Q \, e^{-nNA[Q]} - 1 \right)$$

Exchanging the order of the two limits (we are physicists, hence we can) allows us to use the saddle point approximation

$$\lim_{N\to\infty}\frac{\overline{F_J}}{N}\simeq\lim_{n\to 0}\lim_{N\to\infty}-\frac{1}{nN\beta}\left(e^{-nNA[Q^*]}-1\right)$$

so that, finally:

$$\lim_{N\to\infty}\frac{\overline{F_J}}{N}\simeq\lim_{n\to 0}\lim_{N\to\infty}-\frac{1}{nN\beta}\left(e^{-nNA[Q^*]}-1\right)\simeq\lim_{n\to 0}\lim_{N\to\infty}-\frac{1}{nN\beta}\left(1-nNA[Q^*]-1\right)$$
$$\Rightarrow\boxed{\lim_{N\to\infty}\frac{\overline{F_J}}{N}=\frac{1}{\beta}A\big[Q^*\big]}$$

Notice that in both the saddle point we've ignored the prefactor because at the end these give a contribute $\sim O\big(\frac{\log(N)}{N}\big)$.
To proceed further we have to make an ansatz over the structure of Q.

2.4.1 Case $T > T_K$: metastable phase and ergodic phase

It turns out that for $T > T_K$ the right hypothesis for the structure of Q is the **Replica symmetric ansatz** (RS):

$$Q^{RS}=\begin{pmatrix}1 & q_0 & \dots & q_0\\ q_0 & 1 & \dots & q_0\\ \vdots & \vdots & \ddots & \vdots\\ q_0 & \dots & q_0 & 1\end{pmatrix}\tag{46}$$

Where we can appreciate that the constraint $\delta\left(\sum_{i=1}^N\sigma_i^2-N\right)$ ensures that we have 1 in the Q-diagonal. Without this constraint, we would have had ∞ in the Q-diagonal, because we're computing the self-overlap of a configuration in a continuous and unlimited domain. Also, since $A[Q]$ depends on the elements of Q, the effective free energy wouldn't have been bounded, and we couldn't have taken the saddle point approximation of a quantity that doesn't have any stationary point.

For such a matrix we have:

$$\det(Q)=\det\left[\underbrace{\begin{pmatrix}1-q_0 & 0 & \dots & 0\\ 0 & 1-q_0 & \dots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & \dots & 0 & 1-q_0\end{pmatrix}}_{A}+\underbrace{\begin{pmatrix}q_0 & q_0 & \dots & q_0\\ q_0 & q_0 & \dots & q_0\\ \vdots & \vdots & \ddots & \vdots\\ q_0 & \dots & q_0 & q_0\end{pmatrix}}_{uv^T}\right]$$
$$=(1-q_0)^n+nq_0(1-q_0)^{n-1}=(1-q_0)^n\left(1+\frac{nq_0}{1-q_0}\right)$$

where here we've used the *Matrix Determinant Lemma*:
Prop: Suppose A to be an invertible square matrix, then: $\det(A+uv^T)=\det(A)\big(1+v^TA^{-1}u\big)$

where in our case $\vec{v} = \begin{bmatrix} 1 & \dots & 1 \end{bmatrix}^T$ and $\vec{u} = q_0 \begin{bmatrix} 1 & \dots & 1 \end{bmatrix}^T$

$$\Rightarrow \log(\det(Q)) = n\log(1-q_0) + \log\left(1 + \frac{nq_0}{1-q_0}\right) \overset{n\to 0}{=} n\log(1-q_0) + n\frac{q_0}{1-q_0} + \mathcal{O}(n^2)$$

$$\sum_{\alpha,\beta} Q^p_{\alpha\beta} = n + n(n-1)q_0^p = n(1-q_0^p)$$

$$\begin{aligned} \Rightarrow A\big[Q_{RS}\big] &= -\frac{1}{n}\left[\frac{\beta^2}{4}n(1-q_0^p) + \frac{n}{2}\log(1-q_0) + \frac{n}{2}\frac{q_0}{1-q_0}\right] \\ &= -\left[\frac{\beta^2}{4}(1-q_0^p) + \frac{1}{2}\log(1-q_0) + \frac{1}{2}\frac{q_0}{1-q_0}\right] \end{aligned}$$

$$f_{RS}(\beta) = \lim_{N\to\infty}\frac{\overline{F_J}}{N} = \max_{q_0}\left\{-\frac{1}{2\beta}\left[\frac{\beta^2}{2}(1-q_0^p) + \log(1-q_0) + \frac{q_0}{1-q_0}\right]\right\} \tag{47}$$

2.4.2 Case $T < T_K$: glass phase

For $T < T_K$ the right choice for the overlap matrix is instead the **one step replica symmetry breaking ansatz** (1RSB), in which we assume

$$Q^{1RSB} = \underbrace{\left(\begin{array}{ccc|ccc|c} 1 & & q_1 & & & & \\ & \ddots & & & q_0 & & q_0 \\ q_1 & & 1 & & & & \\ \hline & & & 1 & & q_1 & \\ & q_0 & & & \ddots & & q_0 \\ & & & q_1 & & 1 & \\ \hline & q_0 & & & q_0 & & \ddots \\ & & & & & & \end{array}\right)}_{n\times n} \tag{48}$$

Each block is of size $m \times m$, with $m < n$. After some calculations one gets

$$f_{1RSB}(\beta) = \lim_{N\to\infty}\frac{\overline{F_J}}{N} = -\frac{1}{2\beta}\max_{q_0,q_1,m}\left\{\frac{\beta^2}{4}\left[1 - mq_0^p - (1-m)q_1^p\right] + \right. \tag{49}$$

$$+ \frac{m-1}{m}\log(1-q_0) + \frac{1}{m}\log\left[1 - mq_0 - (1-m)q_1\right] + \tag{50}$$

$$\left. + \frac{q_o}{1 - mq_0 - (1-m)q_1}\right\} \tag{51}$$

In the limit $n \to 0$ it turns out that $m \in [0,1]$. Thus, for 1RSB we have three variational parameters: q_0, q_1, m. Why is that?
For $T > T_K$ we have just one ergodic component. If we try to measure the overlap between any two states, say $\vec{\sigma}$ and $\vec{\tau}$, we obtain $q_0 = \frac{1}{N}\vec{\sigma}\cdot\vec{\tau}$. Instead, in the glass phase it is different

to compute the overlap between two states that are inside the same connected component or between two states belonging to different components. For this reason we have two overlap values, q_0 and q_1. Moreover, we find that $m \in [0,1]$ becomes the probability to pick up two states at random (according to the Boltzmann weight) that belong to different clusters.

Remark: notice that both in equations (47) end (49) we seek the maximum of the free energy with respect to the variational parameters. This can look strange, because we often want to minimize the free energy in order to get the equilibrium configuration. We will go deeper in this aspect further on.

The conditions we require to take the maximum of the free energy over the variational parameters are:

A) RS: $\frac{\partial f_{RS}}{\partial q_0} = 0 \rightarrow$ 1 equation

B) 1RSB: $\frac{\partial f_{1RSB}}{\partial q_0} = 0$; $\frac{\partial f_{1RSB}}{\partial q_1} = 0$; $\frac{\partial f_{1RSB}}{\partial m} = 0$; $\rightarrow$ 3 equations

- For $T > T_D$ B) has no solutions, so the only possible free energy is f_{RS}
- For $T_K < T < T_D$ both A) and B) have solutions, but $f_{1RSB} > f_{RS}$. Hence f_{1RSB} is a metastable phase
- For $T < T_K$ f_{1RSB} becomes the stable phase

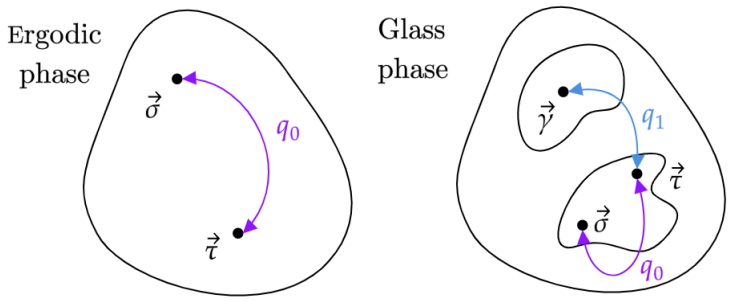


Figure 3: On the left the ergodic phase with just one variational parameter; on the left the glass phase.

2.5 Machine Learning

In the machine learning framework, the target is to learn a rule by doing inference on something. The ML scheme can be summarized as follows:

1. An *Architect* generates data according to some rule.
2. The rule used for generating data is thrown away, but you are given the data.
3. You try to find out the rule according to some protocol (deterministic or probabilistic algorithm).

Question: do we succeed or not? **Answer**: There are phase transitions. Namely, depending on the amount of data or on some parameters of the rule used to generate them we may be able or not to find out the rule.

2.6 Phase Transitions in Inference Algorithms

Before proceeding further, it is useful to introduce what a *Stochastic Monte Carlo Dynamics* (for a disordered system) is.
Suppose that we start with a configuration vector at discrete time t, $\vec{\sigma}(t) = \{\sigma_1(t), ..., \sigma_N(t)\}$, and we want define an update rule $\sigma_i(t+1) = Rule[\sigma_i(t)]$. The Hamiltonian associated to $\vec{\sigma(t)}$ is

$$H_J[\vec{\sigma}(t)] = -\sum_{i<j<k} J_{ijk}\sigma_i(t)\sigma_j(t)\sigma_k(t) \quad \text{with} \quad \sum_{i=1}^{N}\sigma_i^2(t) = N$$

The **Metropolis update Rule** consists on proposing a σ_i^{new}, such that the Hamiltonian of the new configuration is

$$H_J^{new}[\vec{\sigma}(t+1)] = H_J[\sigma_1(t), ..., \sigma_i^{new}, ..., \sigma_N(t)] \tag{52}$$

where $\sigma_i(t+1) = \sigma_i^{new}$. Next, we compute the "energy cost" for the new configuration as

$$\Delta E = H_J^{new} - H_J \tag{53}$$

and we decide whether to keep or to reject the new configuration on the base of the following condition:

$$\begin{aligned}
&\text{if } \Delta E < 0 \longrightarrow \sigma_i(t+1) = \sigma_i^{new} \\
&\text{if } \Delta E > 0 \longrightarrow \text{pick } r \sim U(0,1) \\
&\qquad\qquad \text{if } e^{-\beta\Delta E} > r \longrightarrow \sigma_i(t+1) = \sigma_i^{new} \\
&\qquad\qquad \text{else } \longrightarrow \sigma_i(t+1) = \sigma_i(t)
\end{aligned}$$

In this way we can simulate a stochastic Monte Carlo dynamics.

2.7 Example of an inference problem

An example of an inference problem is given by the following *game*:

1. Choose a *root* spin configuration $\vec{\sigma}^R = \{\sigma_1^R, ..., \sigma_N^R\}$
2. Sort out a set of couplings $\vec{J} = \{J_{ijk}\}$ *correlated to* $\vec{\sigma}^R$, that is, such that $\vec{\sigma}^R$ is an equilibrium state according to a Boltzmann weight

$$P(\vec{J}, \vec{\sigma}^R) = \exp\left(-\beta \sum_{ijk} J_{ijk}\sigma_i^R\sigma_j^R, \sigma_k^R\right) \tag{54}$$

3. Now you throw $\vec{\sigma}^R$ away and retain $\{J_{ijk}\}_{i<j<k}$

Question: Can I retrieve $\vec{\sigma}^R$ by means of a probabilistic Monte Carlo dynamics, that is, sampling configurations with weight $P(\vec{J}, \vec{\sigma})$?
Answer: not always, it depends on the value of β that we used in the conditional probability $P(\vec{J}, \vec{\sigma}^R)$

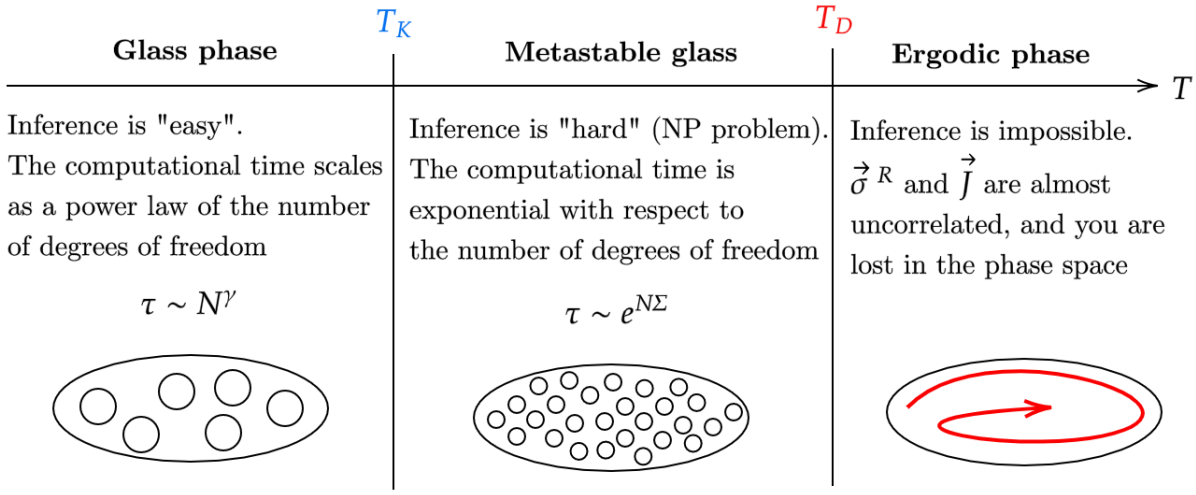


Figure 4: Pictorial representation of the ergodicity breaking in inference problems.

3 Franz-Parisi Potential

The Franz-Parisi (FP) potential $V(q)$, where q is the overlap between two replicas, is a function of overlap which depends parametrically on temperature and which allows to characterize the thermodynamics of the system. The characteristic behaviour of $V(q)$ at different temperatures is shown in Fig. 5.

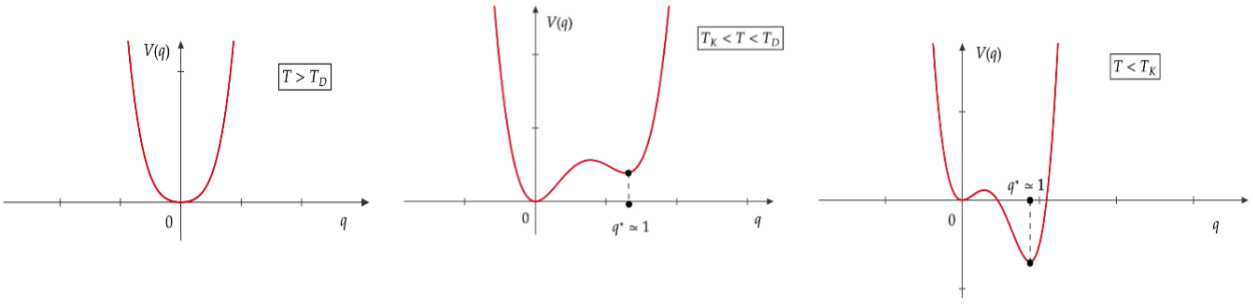


Figure 5: Behaviour of the Franz-Parisi potential. From left to right: ergodic phase, metastable phase, glass phase.

We can see that in the ergodic phase ($T > T_D$), on average the states are totally uncorrelated (since the minimum of V is for $q = 0$). In the metastable phase ($T_K < T < T_D$) for certain initial conditions it is possible to get stuck in the metastable minimum $q^* \simeq 1$, whereas in the glass phase q^* becomes the stable value, and the states are maximally correlated.

The glass transition at T_K is also known as *Random First-Order Transition*, due to its mixed-order character. In fact, it displays both features of a 1^{st} order transition, namely the appearance of a second minimum at a finite value of the order parameter, $q^* > 0$) and of a 2^{nd} order one, the absence of latent heat at the transition point (the first derivative with respect to temperature of the free energy is continuous at the transition).

For the p-spin model there is no *simple* field along which the system aligns at small temperatures $T < T_K$ (i.e. that plays the role of the magnetization for the Ising model). The idea of F-P is to pick up a state $\vec{S}$ with probability $\sim \exp(-\beta H_J[\vec{S}])$ that will be the state along which the system *freezes* for $T < T_K$.

First, we choose a given configuration $\vec{S}$ and define the un-normalized probability

$$P_{J,S}(q) \propto \int_{-\infty}^{+\infty} \left\{ \prod_{i=1}^{N} d\sigma_i \right\} \exp\Big(-\beta H_J[\vec{\sigma}] \Big) \delta\Big(qN - \sum_{i=1}^{N} \sigma_i S_i \Big) \tag{55}$$

then we choose the normalization such that $P_{J,S}(q=0) \overset{T\to\infty}{\longrightarrow} 1$, because we know that in the ergodic phase the states are uncorrelated. Hence we require:

$$P_{J,S}(q=0) = \int \mathscr{D}\sigma \; e^{-\beta H_J[\vec{\sigma}]} \; \delta\big(\vec{\sigma}\cdot\vec{S}\big) \overset{T\to\infty}{\longrightarrow} \int \mathscr{D}\sigma \; e^{-\beta H_J[\vec{\sigma}]} = Z_J(\beta) \tag{56}$$

because if the states are uncorrelated $\vec{\sigma}\cdot\vec{S} = 0 \; \forall \, \vec{S}$, and the Dirac delta is always active.

$$\Rightarrow P_{J,S}(q) = \frac{Z_{J,S}(q)}{Z_J(\beta)} = \frac{1}{Z_J(\beta)} \int \mathscr{D}\sigma \; e^{-\beta H_J[\vec{\sigma}]} \delta\big(qN - \vec{\sigma}\cdot\vec{S}\big) = \exp\left[-\beta\left(F_{J,S}(q) - F_J\right)\right] \tag{57}$$

Since we are interested in the large size limit $N \to \infty$, we can exploit the self-averaging property of the free energy:

$$P_{J,S}(q) \overset{N\to\infty}{\longrightarrow} \exp\left[-\beta N\left(\overline{\langle f_{J,S}(q)\rangle} - \overline{f_J}\right)\right] \tag{58}$$

where, as usual, $\langle\cdot\rangle$ indicates the ensemble average, while $\bar{\cdot}$ is the average over the disorder. We can now define the **Franz-Parisi potential** as

$$\boxed{V(q) = \overline{\langle f_{J,S}(q)\rangle} - \overline{f_J}} \tag{59}$$

that can be interpreted as the free energy cost to keep an equilibrium configuration at a given overlap q with another, previously chosen, configuration $\vec{S}$.

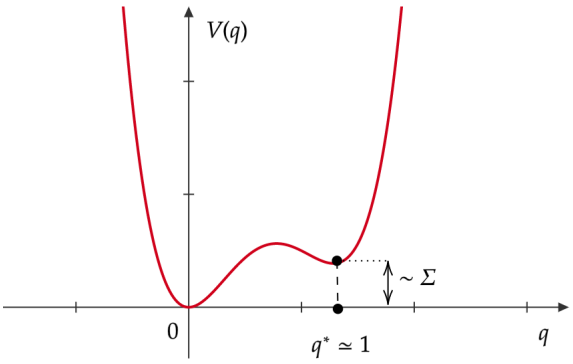
Let us focus now in the case $T_K < T < T_D$. We already know that the number of states scales as $\mathcal{N} \sim e^{\Sigma N}$, and thus the probability to choose a configuration in basin of α is $P_\alpha \sim e^{-\Sigma N}$. But this is equivalent to say that P_α is the probability to choose a configuration with high overlap with a pre-assigned configuration $\vec{S}^\alpha$ belonging to basin of α, that is $q = \frac{1}{N}\vec{S}^\alpha \cdot \vec{\sigma} \simeq 1$.
The probability to have a generic overlap q is $P_{J,S}(q) = e^{-N\beta V(q)}$, hence we have:

$$P_\alpha = P_{J,S}(q = q^* \simeq 1) = e^{-N\beta V(q^*)} = e^{-N\Sigma} \Rightarrow \Sigma = \beta V(q = q^*)$$

We therefore find the important result

$$\boxed{V(q = q^*) = T\Sigma} \tag{60}$$

This means that if we are able to compute $V(q^*)$ we can also quantify the configurational entropy Σ, namely we can "count" the number of metastables glassy states.



3.1 Computation of the Franz-Parisi potential

We want to compute

$$V(q) = \overline{\langle f_{J,S}(q)\rangle} - \overline{f_J} \tag{61}$$

for $T > T_K$. From the previous section we already know that

$$\overline{f_J} = -\frac{1}{\beta}\left[\frac{\beta^2}{4}(1-(q_0^*)^p) + \frac{1}{2}\log(1-q_0^*) + \frac{1}{2}\frac{q_0^*}{1-q_0^*}\right] \quad \text{with} \quad \left.\frac{\partial f_J}{\partial q}\right|_{q=q^*} = 0$$

We can therefore focus on the first term

$$f_{J,S}(q) = -\frac{1}{N\beta}\log\big[Z_{J,S}(q)\big] = -\frac{1}{N\beta}\log\left[\int \mathscr{D}\sigma e^{-\beta H_J[\vec{\sigma}]}\delta\big(qN - \vec{\sigma}\cdot\vec{S}\big)\right]$$

$$\langle f_{J,S}(q)\rangle = \frac{1}{Z_J(\beta)}\int \mathscr{D}S e^{-\beta H_J[\vec{S}]} f_{J,S}(q) = \underbrace{\frac{1}{Z_J(\beta)}}_{\text{(a)}}\int \mathscr{D}S e^{-\beta H_J[\vec{S}]}\left[-\frac{1}{N\beta}\underbrace{\log\big(Z_{J,S}(q)\big)}_{\text{(b)}}\right] \tag{62}$$

$$\overline{\langle f_{J,S}(q)\rangle} = \int \prod_{i<j<k} dJ_{ijk} P(J_{ijk})\langle f_{J,S}(q)\rangle$$

Now, when we try to take the average over the disorder, problems arise both on (a) and (b) terms of equation (62). We can tackle the computation introducing replicas in two points of the expression:

$$\text{(a)} \Rightarrow \frac{1}{Z_J(\beta)} = \lim_{n\to 0} Z_J^{n-1}(\beta) \qquad \text{(b)} \Rightarrow \log\big[Z_{J,S}(q)\big] = \lim_{m\to 0}\frac{\overline{Z_{J,S}^m} - 1}{m} \tag{63}$$

Overall, we have:

$$\overline{\langle f_{J,S}(q)\rangle} = \lim_{N\to\infty}\lim_{n,m\to\infty} -\frac{1}{\beta N}\overline{\left[\int \mathscr{D}S\, e^{-\beta H_J[\vec{S}]} Z_J^{n-1}(\beta)\left(\frac{Z_{J,S}^m(q)-1}{m}\right)\right]} \tag{64}$$

$$= \lim_{N\to\infty}\lim_{n,m\to\infty}\left\{-\frac{1}{\beta m N}\underbrace{\overline{\left[\int \mathscr{D}S\, e^{-\beta H_J[\vec{S}]} Z_J^{n-1}(\beta) Z_{J,S}^m(q)\right]}}_{\overline{[Z^n]}} + \frac{Z_J^{n-1}(\beta)}{m\beta N}\underbrace{\int \mathscr{D}S\, e^{-\beta H_J[\vec{S}]}}_{Z_J(\beta)}\right\} \tag{65}$$

In the second term we can already take the limit $n \to 0$, and the two partition functions cancel out. What remains is

$$\overline{\langle f_{J,S}(q)\rangle} = \lim_{N\to\infty}\lim_{n,m\to\infty} -\frac{1}{\beta N m}\left(\overline{[Z^n]} - 1\right) \tag{66}$$

$$\overline{[Z^n]} = \overline{\left\{\int \mathscr{D}S^1\ e^{-\beta H_J[\vec{S}^1]}\cdot\left[\int \mathscr{D}S\ \exp\left(-\beta\underbrace{\sum_{a=2}^{n}}_{n-1} H_J[\vec{S}^a]\right)\right]\cdot\int \mathscr{D}\sigma\ \exp\left(-\beta\underbrace{\sum_{\alpha=1}^{m}}_{m} H_J[\vec{\sigma}^\alpha]\right)\cdot \prod_{\alpha=1}^{m}\delta\left(qN - \vec{\sigma^\alpha}\cdot\vec{S^1}\right)\right\}}$$

where we denoted by $\vec{S}^1$ the equilibrium configuration which plays the role of a random pinning field for the replica $\vec{\sigma}$

$$\overline{\left[Z^n\right]} = \overline{\left\{\int \mathscr{D}S \exp\left(-\beta \sum_{a=1}^{n} H_J[\vec{S}^a]\right) \cdot \int \mathscr{D}\sigma \exp\left(-\beta \sum_{\alpha=1}^{m} H_J[\vec{\sigma}^\alpha]\right) \cdot \prod_{\alpha=1}^{m} \delta\left(qN - \vec{\sigma^\alpha} \cdot \vec{S^1}\right)\right\}}$$

and recall that the measure is

$$\mathscr{D}\sigma = \prod_{\alpha=1}^{m} \prod_{i=1}^{N} d\sigma_i^\alpha \prod_{\alpha=1}^{m} \delta\left(\sum_{i=1}^{N} (\sigma_i^\alpha)^2 - N\right)$$

Let us unfold the integrand:

$$\exp\left\{-\beta\left[\sum_{a=1}^{n} H_J[\vec{S}^a] + \sum_{\alpha=1}^{m} H_J[\vec{\sigma}^\alpha]\right]\right\} = \exp\left\{\beta\left[\sum_{i<j<k} J_{ijk}\left(\sum_{a=1}^{n} S_i^a S_j^a S_k^a + \sum_{\alpha=1}^{m} \sigma_i^\alpha \sigma_j^\alpha \sigma_k^\alpha\right)\right]\right\}$$

Computing the average over the disorder involves $\binom{N}{p}$ gaussian integrals of the kind:

$$\int_{-\infty}^{+\infty} dJ_{ijk} P(J_{ijk}) \exp \beta\left[\sum_{i<j<k} J_{ijk}\left(\sum_{a=1}^{n} \underbrace{S_i^a S_j^a S_k^a}_{\text{master}} + \sum_{\alpha=1}^{m} \underbrace{\sigma_i^\alpha \sigma_j^\alpha \sigma_k^\alpha}_{\text{slave}}\right)\right] =$$

$$= \exp\left[\frac{p!\beta^2}{4N^{p-1}}\left(\sum_{a,b=1}^{n} S_i^a S_i^b S_j^a S_j^b S_k^a S_k^b + \sum_{\alpha,\beta=1}^{m} \sigma_i^\alpha \sigma_i^\beta \sigma_j^\alpha \sigma_j^\beta \sigma_k^\alpha \sigma_k^\beta + 2 \sum_{a,\alpha=1}^{n,m} S_i^a \sigma_i^\alpha S_j^a \sigma_j^\alpha S_k^a \sigma_k^\alpha\right)\right]$$

Now we proceed along the line of what we've done for the free energy computation and eventually we get

$$\overline{\left[Z^n\right]} = \int \mathscr{D}\sigma \mathscr{D}S \exp\left\{\frac{\beta^2 N}{4} D[\mathbf{S}, \boldsymbol{\sigma}]\right\} \tag{67}$$

where $D[\mathbf{S}, \boldsymbol{\sigma}]$ is defined as:

$$D[\mathbf{S}, \boldsymbol{\sigma}] = \sum_{a,b}^{n}\left(\frac{1}{N}\sum_{i=1}^{N} S_i^a\, S_i^b\right)^p + \sum_{\alpha,\,\beta}^{m}\left(\frac{1}{N}\sum_{i=1}^{N} \sigma_i^a\, \sigma_i^b\right)^p + \sum_{a,\,\alpha}^{n,m}\left(\frac{1}{N}\sum_{i=1}^{N} S_i^a\, \sigma_i^\alpha\right)^p + \sum_{a,\,\alpha}^{n,m}\left(\frac{1}{N}\sum_{i=1}^{N} \sigma_i^a\, S_i^\alpha\right)^p \tag{68}$$

In the expression above $\vec{S}^a$ are the Master Replica, which acts as a random pinning field chosen according to the equilibrium distribution $P(\vec{S}) \sim \frac{1}{Z} e^{-\beta H_J}$. Instead, $\vec{\sigma}^\alpha$ are the Slave Replica, equilibrated under the constraint of a given overlap with $\vec{S}$. Thus we can introduce three order parameters:

- Master-Master overlap

$$Q_{a,\,b} = \frac{1}{N}\sum_{i=1}^{N} S_i^a\, S_i^b \;\in \mathcal{M}_{n\times\, n} \tag{69}$$

- Slave-Slave overlap

$$R_{\alpha,\,\beta} = \frac{1}{N}\sum_{i=1}^{N} \sigma_i^\alpha\, \sigma_i^\beta \;\in \mathcal{M}_{m\times\, m} \tag{70}$$

- Master-Slave overlap

$$P_{a,\beta} = \frac{1}{N}\sum_{i=1}^{N} S_i^a\,\sigma_i^\alpha \;\in \mathcal{M}_{n\times m} \tag{71}$$

Now we can insert these definitions in equation (67), and in order to do that we can use the usual trick already seen for the free energy calculations:

$$\mathbb{1} = \int \prod_{a<b} dQ_{a\,b} \prod_{a<b} \int_{\widehat{Q}^0_{ab}-i\infty}^{\widehat{Q}^0_{ab}+i\infty} d\widehat{Q}_{a\,b}\, \exp\Big\{N\sum_{a<b}\widehat{Q}_{ab}\,Q_{ab}\Big\}\exp\Big\{-\sum_{a<b}\widehat{Q}_{ab}\sum_{i=1}^{N} S_i^a\,S_i^b\Big\}$$

by doing the same also for R and P, we end up with:

$$\overline{\big[Z^n\big]} = \int \mathscr{D}Q\;\mathscr{D}R\;\mathscr{D}P\;\mathscr{D}\widehat{Q}\;\mathscr{D}\widehat{R}\;\mathscr{D}\widehat{P}\prod_{i=1}^{N}\mathscr{D}\vec{x}_i\; e^{\frac{\beta^2 N}{4}\mathcal{K}+\frac{N}{2}\mathcal{T}-\frac{1}{2}\sum_{i=1}^{N}\vec{x}_i^{\,T}\,\widehat{\mathbb{Q}}\,\vec{x}_i}\;\prod_{\alpha=1}^{m}\delta\left(P_{1\alpha}-q\right) \tag{72}$$

where $\mathcal{K}$ and $\mathcal{T}$ are defined as:

$$\mathcal{K} = \sum_{a,b}^{n}\big(Q_{ab}\big)^p + \sum_{\alpha,\beta}^{m}\big(R_{\alpha\beta}\big)^p + 2\sum_{a,\alpha}^{n,m}\big(P_{a\alpha}\big)^p \tag{73}$$

$$\mathcal{T} = \mathrm{Tr}\big[\widehat{Q}Q\big] + \mathrm{Tr}\big[\widehat{R}R\big] + \mathrm{Tr}\big[\widehat{P}^T P\big] + \mathrm{Tr}\big[\widehat{P}P^T\big] \tag{74}$$

whereas $\vec{x}_i$ is a vector in the replica space (both master and slave)

$$\vec{x}_i = \big(S_i^1,...,S_i^n,\sigma_i^1,...,\sigma_i^m\big)\;\in\mathcal{V}_{n+m} \tag{75}$$

and $\widehat{\mathbb{Q}}$ and $\mathbb{Q}$ are two block matrices that represent all the order parameters:

$$\widehat{\mathbb{Q}} = \begin{pmatrix}\widehat{Q} & \widehat{P}\\ \widehat{P}^T & \widehat{R}\end{pmatrix} \in \mathcal{M}_{n+m\,\times\,n+m} \tag{76}$$

$$\mathbb{Q} = \begin{pmatrix}Q & P\\ P^T & R\end{pmatrix} \in \mathcal{M}_{n+m\,\times\,n+m} \tag{77}$$

The integrals over the spin variables $\vec{x}_i$ are gaussian:

$$I_{\mathbf{x}_i} = \int\prod_{i=1}^{N}\mathscr{D}\vec{x}_i\exp\Big\{-\frac{1}{2}\sum_{i=1}^{N}\vec{x}_i^{\,T}\,\widehat{\mathbb{Q}}\,\vec{x}_i\Big\} = e^{-\frac{N}{2}\log\left(\det\widehat{\mathbb{Q}}\right)} = e^{-\frac{N}{2}\,\mathrm{Tr}\left(\log\widehat{\mathbb{Q}}\right)} \tag{78}$$

Thanks to the block-structure of $\widehat{\mathbb{Q}}$, equation (74) is nothing but $\mathrm{Tr}\big(\widehat{\mathbb{Q}}\mathbb{Q}\big)$. If we define

$$\hat{S}(\mathbb{Q},\widehat{\mathbb{Q}}) \equiv \mathrm{Tr}\big(\widehat{\mathbb{Q}}\mathbb{Q}\big) - \mathrm{Tr}\big[\log\big(\widehat{\mathbb{Q}}\big)\big] \tag{79}$$

the integration over $\widehat{\mathbb{Q}}$ reads

$$I_{\widehat{\mathbb{Q}}} = \int\mathscr{D}_{\widehat{\mathbb{Q}}}\exp\Big\{\frac{N}{2}\hat{S}(\mathbb{Q},\widehat{\mathbb{Q}})\Big\} \tag{80}$$

In the large size limit we can perform a saddle point approximation. The maximum is given by:

$$0 = \frac{\partial\widehat{S}}{\partial\widehat{\mathbb{Q}}} = \frac{\partial}{\partial\widehat{\mathbb{Q}}}\Big[\mathrm{Tr}\big(\widehat{\mathbb{Q}}\mathbb{Q}\big) - \mathrm{Tr}\big(\log\widehat{\mathbb{Q}}\big)\Big]\;\Rightarrow\;\widehat{\mathbb{Q}} = \mathbb{Q}^{-1} \tag{81}$$

Now we can use the following identity:
Prop. Given the structure of $\mathbb{Q}$ in equation (77), then:

$$\operatorname{Tr}\log\big(\mathbb{Q}\big) = \operatorname{Tr}\log\big(Q\big) + \operatorname{Tr}\log\big(R - P^T Q^{-1} P\big) \tag{82}$$

Proof. We have

$$\begin{aligned}\det\mathbb{Q} &= \det\big(QR - PP^T\big)\\ &= \det\big[Q\big(R - Q^{-1}PP^T\big)\big]\\ &= \det\big(Q\big)\det\big(R - P^TQ^{-1}P\big)\end{aligned}$$

and using the result we already proved we can conclude □

Recollecting all the pieces, we get:

$$\overline{\big[Z^n\big]} = \int \mathscr{D}Q\ \mathscr{D}R\ \mathscr{D}P\ \exp\{NS(Q,R,P)\}\ \prod_{\alpha=1}^{m}\delta\left(P_{a\alpha} - q\right) \tag{83}$$

where

$$S(Q,R,P) = \frac{\beta^2}{4}\mathcal{K} + \operatorname{Tr}\log\big(Q\big) + \operatorname{Tr}\log\big(R - P^TQ^{-1}P\big) \tag{84}$$

As before we can introduce an effective free energy as:

$$A[Q,R,P] = -\frac{1}{m}\,S[Q,R,P] \tag{85}$$

and we are ready to come back to equation (66) and to perform a saddle point approximation

$$\overline{\langle f_{J,S}(q)\rangle} = \lim_{\substack{N\to+\infty\\ n\to0\\ m\to0}} -\frac{1}{mN\beta}\big(\overline{\big[Z^n\big]} - 1\big) \simeq \lim_{\substack{N\to+\infty\\ n\to0\\ m\to0}} -\frac{1}{mN\beta}\Big[\exp\big\{-nNA\big[Q^*,R^*,q\big]\big\} - 1\Big], \tag{86}$$

where the only elements of $P_{a\alpha}$ different from zero are those fixed from the constraint to be equal to q (see the explicit form of the Ansatz below), so that there is no saddle point approximation to do with respect to P, just to explicitate its dependence on q. We end up with the Franz-Parisi potential definition:

$$\boxed{\beta V(q) = \overline{\langle f_{J,S}(q)\rangle} - \overline{f_J} = \frac{1}{\beta}\max_{Q_{ab}\,R_{\alpha\beta}} A\big[Q,R,q\big] - \overline{f_J}} \tag{87}$$

Where does q arise from the previous expression? Recalling that we're considering the system

at temperature $T > T_K$, the saddle point solutions of equation (86) are:

$$Q^*_{ab} = \begin{pmatrix} 1 & & q_0 \\ & \ddots & \\ q_0 & & 1 \end{pmatrix} \tag{88}$$

$$R^*_{\alpha\beta} = \begin{pmatrix} 1 & & r \\ & \ddots & \\ r & & 1 \end{pmatrix} \tag{89}$$

$$P^*_{a\alpha} = \underbrace{\begin{pmatrix} q & \dots & q \\ 0 & \dots & 0 \\ 0 & \dots & 0 \end{pmatrix}}_{m \times n} \tag{90}$$

where q is the element that appears in the constraint:

$$\delta\big(qN - \sum_{i=1}^{N} \sigma_i^\alpha \, S_i^1\big) \tag{91}$$

A couple of words are now in order to explain some subtelties related to the strategy to find solutions for the saddle-point equations Eq. (87). By considering the two limits $n \to 0$ and $m \to 0$ which allow to pass from the expression in Eq. (86) to the one in Eq. (87) we need to clarify one important point: though considering that $n/m \to 1$ when both $n, m \to 0$, still in the expression of the effective action we regard terms of order $\mathcal{O}(m)$ subleading with respect to terms of order $\mathcal{O}(n)$. This allows us for instance to approximate the saddle-point equation for Q_{ab} in the following manner

$$\begin{aligned} &\frac{\partial}{\partial Q_{ab}} \left[\frac{\beta^2}{4} \sum_{ab} Q^p_{ab} + \log\det Q + \log\det(R - P^T Q^{-1} P) \right] = 0 \approx \\ &\approx \frac{\partial}{\partial Q_{ab}} \left[\frac{\beta^2}{4} \sum_{ab} Q^p_{ab} + \log\det Q \right] = 0, \end{aligned} \tag{92}$$

for the reason that Q is an $n \times n$ matrix, so that both $\sum_{ab} Q^p_{ab}$ and $\log\det Q$, once the replica symmetric ansatz of Eq. (88) is plugged into the expression, turns out to be of order $\mathcal{O}(n)$ while the term $\log\det(R - P^T Q^{-1} P)$ is of order $\mathcal{O}(m)$, because R and $P^T Q^{-1} P$ are $m \times m$ matrices. The "physical" argument to say that terms $\mathcal{O}(m)$ are sub-leading with respect to term $\mathcal{O}(n)$ then goes back to the meaning of Eq. (65); while the limit $n \to 0$ is related to the equilibrium average over the *master replicas* S_i's, the limit $m \to 0$ is related the equilibrium average of *slave replicas* σ_i's, while keeping fixed S_i. This means that the average over slave replicas, physically, comes *first* with respect to the average over master replicas. The σ_i's are thermalized under the effect of a quenched random pinning field represented by one of the S_i's. Correspondingly, this means that the limit $m \to 0$ must be taken first and, consequently terms of order $\mathcal{O}(m)$ are sub-leading with respect to terms of order $\mathcal{O}(n)$. That is why the saddle point equation for Q_{ab} decouples from the one for $R_{\alpha\beta}$. This notwithstanding, at the end both limits $n \to 0$ and

$m \to 0$ must be considered, so that the n at the numerator cancels out with the $1/m$ prefactor. We are thus left with

$$\begin{aligned}\beta V(q) &= -\frac{1}{2\beta}\left[\frac{\beta^2}{2}(1-(q_0^*)^p) + \log(1-q_0^*) + \frac{q_0^*}{1-q_0^*}\right] \\ &+ \frac{1}{\beta}\max_{R_{\alpha\beta}}\left\{-\frac{1}{m}\left[\frac{\beta^2}{4}\sum_{\alpha\beta}R_{\alpha\beta}^p + \log\det\left(R - P^T Q^{-1} P\right)\right]\right\} - \overline{f_J} \\ &= \frac{1}{\beta}\max_{R_{\alpha\beta}}\left\{-\frac{1}{m}\left[\frac{\beta^2}{4}\sum_{\alpha\beta}R_{\alpha\beta}^p + \log\det\left(R - P^T Q^{-1} P\right)\right]\right\}. \end{aligned} \tag{93}$$

since we known from the p-spin model calculation that we have precisely

$$\overline{f_J} = -\frac{1}{2\beta}\left[\frac{\beta^2}{2}(1-(q_0^*)^p) + \log(1-q_0^*) + \frac{q_0^*}{1-q_0^*}\right]. \tag{94}$$

We have therefore shown how the calculation of $\overline{\langle f_{J,S}(q)\rangle}$ yields a contribution which perfectly cancels out with $\overline{f_J}$. In the limit of $m \to 0$ the *energetic* term with $R_{\alpha\beta}$ becomes

$$\frac{\beta^2}{4m}\sum_{\alpha\beta}R_{\alpha\beta}^p = \frac{\beta^2}{4m}\left[m + m(m-1)r^p\right] \overset{m\to 0}{=} \frac{\beta^2}{4}(1-r^p). \tag{95}$$

For the entropic terms involving R we have first to find out who is the matrix $R - P^T Q^{-1} P$. Since the solution of the saddle point equation for Q_{ab} for temperatures $T > T_K$ is $Q_{ab} = \mathbf{1}_{n\times n}$ we have, taking into account the Ansatz in Eq. (90) for P,

$$\left(P^T Q^{-1} P\right)_{\alpha\beta} = \sum_{ab} P_{\alpha a}(Q^{-1})_{ab} P_{b\beta} = \sum_{ab} P_{\alpha a}\delta_{ab} P_{b\beta} = \sum_{a=1}^{n} P_{\alpha a} P_{a\beta} = q^2, \tag{96}$$

so that

$$\begin{aligned} R - P^T Q^{-1} P &= \begin{pmatrix} 1-q^2 & r-q^2 & \dots & r-q^2 \\ r-q^2 & 1-q^2 & \dots & r-q^2 \\ \vdots & \vdots & \ddots & \vdots \\ r-q^2 & \dots & r-q^2 & 1-q^2 \end{pmatrix} \\ &= \begin{pmatrix} 1-r & 0 & \dots & 0 \\ 0 & 1-r & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 1-r \end{pmatrix} + \begin{pmatrix} r-q^2 & r-q^2 & \dots & r-q^2 \\ r-q^2 & r-q^2 & \dots & r-q^2 \\ \vdots & \vdots & \ddots & \vdots \\ r-q^2 & \dots & r-q^2 & r-q^2. \end{pmatrix} \end{aligned} \tag{97}$$

We exploit once again the formula of the determinant of a $m \times m$ matrix $C = A + B$ where A is diagonal with all elements equal to a, $A_{m\times m} = \mathrm{diag}(a, \dots, a)$, and B is completely filled with elements all identical to b, $B = b\ \mathbf{1}_m \otimes \mathbf{1}_m$:

$$\det(C) = \det(A+B) = a^m + m\ b\ a^{m-1}. \tag{98}$$

The last formula, applied to $R - P^T Q^{-1} P$ yields:

$$\begin{aligned}\det(R - P^T Q^{-1} P) &= (1-r)^m + m\ (r-q^2)\ (1-r)^{m-1} \\ &= (1-r)^m\ \left[1 + m\ \frac{r-q^2}{1-r}\right],\end{aligned} \tag{99}$$

from wich we have

$$\begin{aligned}\lim_{m\to 0}\ \frac{1}{m}\log\det(R - P^T Q^{-1} P) &= \log(1-r) + \lim_{m\to 0}\ \frac{1}{m}\log\left(1 + m\ \frac{r-q^2}{1-r}\right) \\ &= \log(1-r) + \frac{r-q^2}{1-r}.\end{aligned} \tag{100}$$

At the end, collecting all pieces, what we get is:

$$\beta V(q) = \frac{1}{\beta}\max_r\left\{-\frac{\beta^2}{4}(1-r^p) - \frac{\beta^2 q^p}{2} - \log(1-r) - \frac{r-q^2}{1-r}\right\} \tag{101}$$

and the maximization has to be performed numerically. Again, notice that we are looking for the maximum of the free energy instead of the minimum. The reason comes from the the replica limit $m \to 0$.

3.2 Maximization vs Minimization Argument over Q_{ab}

Consider the free energy of a magnetic system (for example the Curie-Weiss model):

$$F(M) = U(M) - TS(M) \tag{102}$$

Usually in statistical mechanics you look to configurations which minimize the free energy, and we want to be consistent with that; note also that F and U agree in sign. Now we have to recognize what is the internal energy of our p-spin model. The expression we found for the internal energy is

$$f(\beta) = \frac{1}{\beta}A[Q^{SP}] = -\frac{1}{n}\left[\overbrace{\frac{\beta}{4}\sum_{\alpha,\beta=1}^{n}(Q^{SP})^p_{\alpha\beta}}^{\text{Energetic term}} + \overbrace{\frac{1}{2\beta}\log\Big(\det(Q^{SP})\Big)}^{\text{Entropic term}}\right]$$

And we therefore make the assignment

$$U(Q_{\alpha\beta}) = -\frac{\beta}{4}\sum_{\alpha\beta}^{n}\big(Q_{\alpha\beta}\big)^p \tag{103}$$

When $n > 1$ in principle you should minimize the free energy f with respect to $Q_{\alpha\beta}$; the point is that when you take the limit $n \to 0$ the internal energy changes sign. Indeed, if we consider the simplest possible ansatz, the replica symmetric one, it turns out that:

$$U(Q_{\alpha\beta}) = -\frac{\beta}{4}\sum_{\alpha\beta}^{n}\big(Q_{\alpha\beta}\big)^p = \frac{\beta n}{4}\Big[(1-n)q_0^p - 1\Big] \tag{104}$$

and we can see that the value of q_0 that minimize the function for $n > 1$ is the one that maximize the function for $n < 1$, as shown in figure (6).
The change in the sign essentially comes out when we try to count the off-diagonal terms, which are $n(n-1)$, and we realize that this quantity becomes negative for $n \to 0$.
Maximization has been proved rigorously by Guerra Francesco (Rome) in the paper: *Comm. Math. Phys. (2003).*

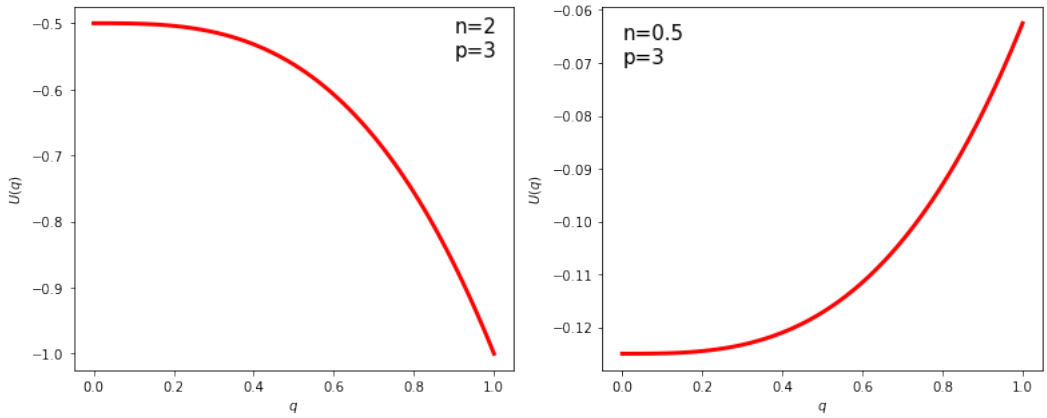


Figure 6: Behaviour of the internal energy depending on the parameter n. As soon as $n < 1$ the function changes its concavity

4 Statistical mechanics and inference problems

The main reference for this chapter is *Statistical Physics and Inference Problems. Notes: M.C. Angelini, F. Caltagirone, F. Krzakala.*
Let us consider as a building block for what follows the discrete p-spin model with nearest neighbours interaction

$$H_J(\vec{J},\vec{\sigma}) = -\sum_{<ijk>} J_{ijk}\sigma_i\sigma_j\sigma_k \tag{105}$$

where J_{ijk} and σ_i have become binary variables of values in $\{-1,+1\}$. If we denote by $\square$ the couplings and by $\bigcirc$ the spin variables, the geometry of the system can be described by the **regular random graph** where each spin is attached to three couplings, and each coupling is attached to three spins.

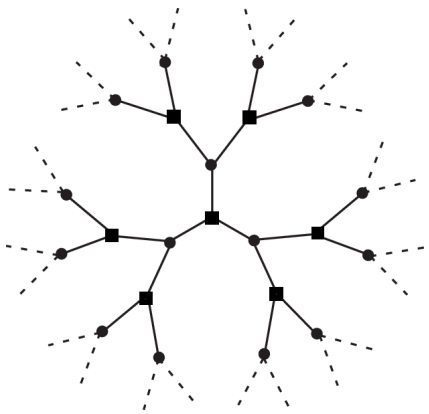


Figure 7: Bethe Lattice.

It can be proved that this graph is the only type of geometry with finite connectivity that with our choice of H (105) displays a glass transition (for example, in a square lattice this wouldn't happen). In particular, this geometry with discrete variables has the same kind of physics of the p-spin model with continuous variables, with the three different phases and the two temperatures T_K and T_D.
This pattern is called *Locally Tree-Like Random Graph* and it's well described in *Information Physics Computation. M. Mézard, A. Montanari (2009).*

4.1 Bayesian Inference

Now we want play the "inference game" already mentioned in section 2.5.

Recall that:

1. INPUT: random sequence of digits $\vec{\sigma^0} = \{\sigma_1^0, ..., \sigma_N^0\}$, $\sigma_i^0 = \pm 1$ (the signal).
2. OUTPUT: coupling coefficients $\vec{J^0} = \{J_{ijk}\}_{<ijk>}$
3. Probabilistic RULE to get $\vec{J^0}$ from $\vec{\sigma^0}$: $\vec{J^0}$ are such that the spins $\vec{\sigma^0}$ are an equilibrium configuration at temperature $T = \beta^{-1}$

$$P(J_{ijk}^0) = \frac{1}{2\cosh\beta}\exp\{\beta J_{ijk}^0\sigma_i^0\sigma_j^0\sigma_k^0\} \tag{106}$$

We thus have a rule depending on a parameter β. By tuning the value of β, the output will be more or less correlated to the input.

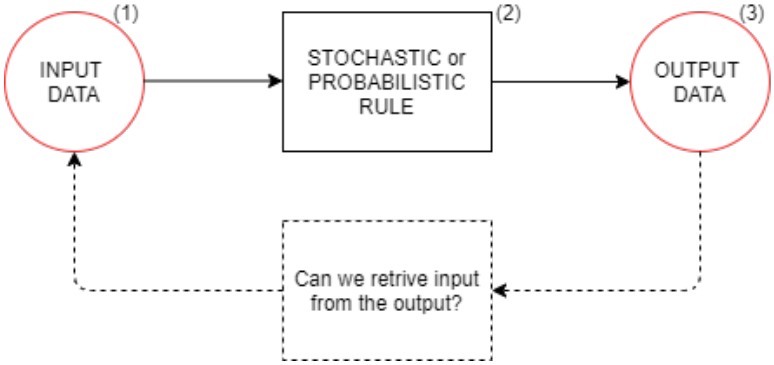


Figure 8: Schematic diagram of inference game.

The conditional probability of the output $\vec{J^0}$, given the input $\vec{\sigma^0}$, is the *Likelihood* function:

$$P\left(\vec{J^0}|\vec{\sigma^0}\right) = e^{-\beta H\left[\vec{J^0},\vec{\sigma^0}\right]} \tag{107}$$

while $P(\vec{\sigma^0})$ is the *Prior Probability*. What we are going to do in what follows is called **bayesian inference**.
The bayesian inference consists on throwing away $\vec{\sigma^0}$ and trying to infer it from the knowledge of $\vec{J^0}$. Bayesian inference exploits the Bayes formula for conditional probabilites:

$$\underbrace{P\left[\vec{\sigma}|\vec{J}\right]}_{Posterior} P\left[\vec{J}\right] = \underbrace{P\left[\vec{J}|\vec{\sigma}\right]}_{Likelihood} \underbrace{P\left[\vec{\sigma}\right]}_{Prior} \tag{108}$$

and we want to find the Posterior Probability, namely:

$$P\left[\vec{\sigma}|\vec{J}\right] = \frac{P\left[\vec{J}|\vec{\sigma}\right]P\left[\vec{\sigma}\right]}{P\left[\vec{J}\right]} = \frac{P\left[\vec{J}|\vec{\sigma}\right]}{P\left[\vec{J}\right]} \tag{109}$$

where the last step is justified by the choice of a uniform prior.
The $P\left[\vec{J}\right]$ term can be obtained by a simple normalization argument:

$$1 = \sum_{\{\vec{\sigma}\}} P\left[\vec{\sigma}|\vec{J}\right] \Rightarrow P\left[\vec{J}\right] = \sum_{\{\vec{\sigma}\}} P\left[\vec{J}\,|\vec{\sigma}\right] = Z\left(\vec{J}\right) \tag{110}$$

and therefore we know the posterior:

$$P\left[\vec{\sigma}|\vec{J^0}\right] = \frac{1}{Z\left(\vec{J^0}\right)} e^{-\beta H\left[\vec{J^0},\vec{\sigma}\right]} \tag{111}$$

This is called *Planted Ensemble*, because the couplings we sampled were correlated to $\vec{\sigma^0}$ through the relation (106).

Now we look at the problem from a machine learning point of view, that is, we still know the "right answer" $\vec{\sigma^0}$ and we want to train our algorithm in order to minimize the number of errors, i.e. the loss function.

- We sample the configurations $\vec{\sigma}$ according to the probability $P\big[\vec{\sigma}|\vec{J^0}\big]$
- The loss function ("energy") is

$$E = \frac{1}{N}\sum_{i=1}^{N}\big(1-\delta_{\sigma_i,\sigma_i^0}\big) \tag{112}$$

- To minimize E we use the Argmax algorithm

$$\widehat{\sigma_i} = \text{argmax}_{\sigma_i}\ \ \mu\big(\sigma_i|\vec{J^0}\big) \tag{113}$$

  where μ is the marginal probability distribution:

$$\mu\big(\sigma_i|\vec{J^0}\big) = \sum_{\{\sigma_j|j\neq i\}}\frac{1}{Z\big(\vec{J^0}\big)}e^{-\beta H[\,\vec{\sigma}|\vec{J^0}\,]} = \sum_{\{\sigma_j|j\neq i\}} P\big[\vec{\sigma}|\vec{J^0}\big] \tag{114}$$

The configuration, iteratively selected by the Argmax algorithm is, by definition, the one that minimizes the free energy:

$$f = -\frac{1}{N\beta}\log\Bigg(\sum_{\{\vec{\sigma}\}}\exp\big\{-\beta H[\,\vec{\sigma}|\vec{J^0}\,]\big\}\Bigg) \tag{115}$$

In order to estimate how the chances to perform a successful inference depend on the parameter β, it is useful to study how the free energy depends on the **similarity (overlap)** between the answer of our inference and with the signal to be retrieved ($\vec{\sigma^0}$). We can define the overlap as follows:

$$q = \frac{1}{N}\sum_{i=1}^{N}\big(1-\delta_{\sigma_i\,\sigma_i^0}\big) - \frac{1}{2} = E\big[\vec{\sigma},\vec{\sigma^0}\big] - \frac{1}{2} \tag{116}$$

and the associated free energy is:

$$f_{\vec{J^0},\vec{\sigma^0}}(q) = \log\Bigg(\sum_{\{\vec{\sigma}\}}\exp\big\{-\beta H[\,\vec{\sigma}|\vec{J^0}\,]\big\}\delta\big(qN-\vec{\sigma}\cdot\vec{\sigma^0}\big)\Bigg) \tag{117}$$

But, on average, we know that $\overline{\big\langle f_{\vec{J^0},\vec{\sigma^0}(q)}\big\rangle} \sim V(q)$, so our inference problem is strictly connected with the Franz-Parisi potential. In particular, we know that $P(q) \sim e^{-\beta N V(q)}$, and therefore we can state that inference is successful when for $N\to\infty$ we have $p(q\simeq 1) > 0$, i.e. when we have a finite probability to retrieve the correct $\vec{\sigma^0}$.

The shape of the potential for various temperatures is shown in Figure 5). We can conclude that if $T > T_D$ then inference is impossible, if $T < T_K$ inference is possible with finite probability (easy) and in between $T_D < T < T_K$ inference is a hard problem, exponentially unlikely, because we already now that in this case $\tau_{alg} \sim \exp\{N\Sigma\}$.

Remark 1: As already said, the above argument works because we are supposed to know what $\vec{\sigma^0}$ is, otherwise we cannot even compute the error function E. Here we have assumed that first we train our algorithm (supervised learning) and we check if inference is possible and, if so, we

eventually can forget about our input and make predictions over data belonging to the same distribution.

Remark 2: Glass transitions in algorithms are quite common, and they are not restricted to the p-spin model we used for our discussion. In supervised learning you are given the training data and the labels and you want to learn an associative rule. In statistical mechanics this corresponds to an inverse problem, where you know the configurations and you want to learn the coupling coefficients, so you are moving in this phase space of coefficients.

In ML the phase space is the Hypothesis class $\mathscr{H}$ of our predictor rules, and the size of training sample m_s plays the role of the temperature. If we increase the size of the training set our learning improves, and this is due to the fact that we add more constraints to the phase space. Since $\mathscr{H}$ has an entropy very similar to the configurational entropy, the transition beyond which you can learn the problem with good probability occours when the entropy of this phase-space goes to zero.

For example, in the Perceptron algorithm or in Neural Networks this idea works.

5 References

Books and papers suggested by professor Giacomo Gradenigo about this topic:

Recipes for metastable states in spin glasses
S Franz, G Parisi, Journal de Physique I 5 (11), 1401-1415 (1995)
https://arxiv.org/pdf/cond-mat/9503167.pdf

Spin-glass theory for pedestrians
T Castellani, A Cavagna Journal of Statistical Mechanics: Theory and Experiment 2005 (05), P05012
https://iopscience.iop.org/article/10.1088/1742-5468/2005/05/P05012/pdf

Statistical physics of spin-glasses and information processing
Hidetoshi Nishimori (2001)
http://inis.jinr.ru/sl/P_Physics/PT_Thermodynamics,%20statistical%20physics/PTin_Information%20theory/Nishimori%20Statistical.pdf

Spin glass theory and beyond
Marc Mezard, Giorgio Parisi, Miguel Angel Virasoro, World Scientific (1987)

Information, Physics and Computation
Marc Mezard, Andrea Montanari (2009)

Random Fields and spin glasses: a Field-Theory approach
Irene Giardina, Cirano De Dominicis.