I. QUENCHED AND ANNEALED DISORDER

We must first clarify the distinction between annealed and quenched disorder. As a physical example, suppose that we substitute some non-magnetic impurity atoms into a lattice of magnetic ions. For instance, we might mix some non-magnetic impurities and the magnetic ions will remain in thermal equilibrium with each other, and the resulting distribution of impurities will have the proper Boltzmann weight for an energy expression that includes the various interactions

---

1 The significance of these advances was recognized by the 1992 Boltzmann and the 1999 Dirac Medals being awarded to Giorgio Parisi for the mean-field solution of spin-glass models. The discovery that spin and structural glasses are formally related at the mean-field level has recently brought renewed attention to this approach.
between the different kinds of atoms. The resulting distribution of impurities is then dubbed *annealed*. If we study systems over the very long time scales necessary to achieve such equilibrium, then we should compute the partition function by not only sampling the possible orientations of the spins of the magnetic ions, but by also sampling over the positions of the impurities. This treatment would be similar to that of Problem 5.27, in which the different realizations of disorder are equilibrated along with the other contributions to the energy. More specifically, given a generalized realization of disorder for a vector of spins $\sigma$

$$E_{\nu}(J, h) = -\sigma^T J \sigma - h^T \sigma,$$

(1)

where $J$ and $h$ are a matrix and a vector, respectively, with components randomly selected from probability distributions $P_J$ and $P_h$, the partition function of annealed system would be

$$\Omega = \int \text{d}h \text{d}J P_h(h) P_J(J) \sum_{\{\sigma_i = \pm 1\}} e^{-E_{\nu}(J, h)}$$

(2)

where we denote averaging over all the possible realizations of disorder

$$\overline{\Omega} = \int \text{d}h \text{d}J P_h(h) P_J(J) \Omega(J, h).$$

(3)

In practice, however, the mobility of impurities in a solid is so small that the timescales involved for them to reach positional thermal equilibrium are beyond human reach. The more common case thus corresponds to regarding the positions of the impurities as fixed, and sampling over only the magnetic degrees of freedom. This is the *quenched* disorder case that we will consider in this chapter. In principle, different realizations of disorder then correspond to distinct systems. Yet for certain quantities, such as the free energy, taking the large system limit is equivalent to averaging over realizations of disorder, and thus corresponds to averaging over the logarithm of the partition function,

$$\ln \overline{\Omega} = \int_{-\infty}^{\infty} \text{d}h \text{d}J P_h(h) P_J(J) \ln \Omega(J, h).$$

(5)

**II. ANNEALED AVERAGING AND THE REPLICA TRICK**

Computing the average of the logarithm of a partition function is generally no simple task. A common way to do so is to rely on the replica trick. More specifically, we first look at $n$ copies of a system with the same realization of disorder $(J, h)$ to obtain

$$\beta na(J, h) = -\frac{1}{N} \ln \left[ \prod_{\alpha=1}^{n} \sum_{\{\sigma(\alpha) = \pm 1\}} e^{-\beta E_{\nu}(\alpha)(J, h)} \right] = -\frac{1}{N} \ln \left[ \Omega^n(J, h) \right].$$

The replica trick then involves using the following identity

$$\beta \overline{\alpha} = -\frac{1}{N} \ln \overline{\Omega} = -\frac{1}{N} \lim_{n \to 0} \frac{\partial}{\partial n} \overline{\Omega}.$$  

(1)

**Exercise:** Using the series expansion for an exponential and l’Hôpital’s rule, derive the above result.

For a general $n$ this approach is not a simplification of the original problem; however, solving the problem for integer values of $n$ and analytically continuing the result to zero may be. The key difficulty is that this last operation is not mathematically well controlled. Although we use this somewhat uncontrolled trick below, note that rigorous (albeit much more involved) derivations are also available for the systems we consider.

**III. REPLICA CALCULATION EXAMPLE: FULLY-CONNECTED RANDOM-FIELD ISING MODEL (RFIM)**

We first consider a fully-connected system of $N$ ferromagnetically-coupled Ising spins subject to random local fields $h_i$

$$E_{\nu}(h) = -J \sum_{i,j=1}^{N} \sigma_i \sigma_j - \sum_{i} h_i \sigma_i,$$

(1)
where the elements of the vector $\mathbf{h}$ are normally distributed, i.e.,

$$P(h_i)dh_i = \frac{1}{\sqrt{2\pi}\sigma_h^2} e^{-h_i^2/(2\sigma_h^2)} dh_i.$$  \hspace{1cm} (2)

(Recall that for such a system setting $J = 1/N$ guarantees that thermodynamic quantities are extensive, see Appendix.) This mean-field model is known as the random-field Ising model (RFIM), and its average free energy can be obtained by using the replica trick. Following the prescription of the replica trick gives

$$Q^n = \int P(h)Q^n(h) dh$$

$$= \sum_{\{\sigma^{(\alpha)} = \pm 1\}} \exp \left[ \frac{\beta}{N} \sum_{\alpha=1}^n \sum_{i,j=1}^N \sigma_i^{(\alpha)} \sigma_j^{(\alpha)} \right] \int_{-\infty}^{\infty} \frac{dh}{(2\pi\sigma_h^2)^{N/2}} \exp \left[ -\frac{h^2}{2\sigma_h^2} + \frac{\beta}{N} \sum_{\alpha=1}^n \mathbf{h} \cdot \mathbf{\sigma}^{(\alpha)} \right]$$

$$= \sum_{\{\sigma^{(\alpha)} = \pm 1\}} \exp \left[ \frac{\beta}{N} \sum_{\alpha=1}^n \left( \sum_{i=1}^N \sigma_i^{(\alpha)} \right)^2 + \frac{\beta^2 \sigma_h^2}{2} \sum_{i=1}^n \left( \sum_{\alpha=1}^N \sigma_i^{(\alpha)} \right)^2 \right].$$

In order to make progress in simplifying this expression, we use a simple identity for Gaussian integrals known as the Hubbard-Stratonovich transformation

$$\exp \left[ \frac{qy^2}{2} \right] = \frac{1}{(2\pi\epsilon)^{1/2}} \int_{-\infty}^{\infty} \exp \left[ \frac{x^2}{2\epsilon} + xy \right] dx.$$  \hspace{1cm} (3)

**Exercise:** Verify the validity of the Hubbard-Stratonovich transformation.

This transformation allows us to rewrite the first term in the exponential as

$$\prod_{\alpha=1}^n \exp \left[ \frac{1}{2} \times \frac{1}{2N\beta} \left( \sum_{i=1}^N 2\beta \sigma_i^{(\alpha)} \right)^2 \right] =$$

$$\left( \frac{N\beta}{\pi} \right)^{n/2} \int \prod_{\alpha=1}^n dx^{(\alpha)} \exp \left[ -N\beta \sum_{\alpha=1}^n (x^{(\alpha)})^2 + 2\beta \sum_{i=1}^n \sum_{\alpha=1}^N x^{(\alpha)} \sigma_i^{(\alpha)} \right],$$

hence

$$Q^n = \left( \frac{N\beta}{\pi} \right)^{n/2} \sum_{\{\sigma^{(\alpha)} = \pm 1\}} \int \prod_{\alpha=1}^n dx^{(\alpha)}$$

$$\times \exp \left[ -N\beta \sum_{\alpha=1}^n (x^{(\alpha)})^2 + 2\beta \sum_{i=1}^n \sum_{\alpha=1}^N x^{(\alpha)} \sigma_i^{(\alpha)} + \beta^2 \sigma_h^2 \sum_{i=1}^n \left( \sum_{\alpha=1}^N \sigma_i^{(\alpha)} \right)^2 \right]$$

Note that the exponential is now a simple sum over all $N$ sites, with no interaction between them. The partition sum is therefore the $N$th power of the partition sum of a single site, and we can write

$$Q^n = \left( \frac{N\beta}{\pi} \right)^{n/2} \int \prod_{\alpha=1}^n dx^{(\alpha)} \exp \left[ N \left( -\beta \sum_{\alpha=1}^n (x^{(\alpha)})^2 + \ln Q_1(\{x^{(\alpha)\}) \right), \hspace{1cm} (4)\right.$$  

where $Q_1(\{x^{(\alpha)\})$ is the partition sum for all replicas on a single site

$$Q_1(\{x^{(\alpha)\}) = \sum_{\{\sigma^{(\alpha)} = \pm 1\}} \exp \left[ 2\beta \sum_{\alpha=1}^n x^{(\alpha)} \sigma^{(\alpha)} + \beta^2 \sigma_h^2 \left( \sum_{\alpha=1}^n \sigma^{(\alpha)} \right)^2 \right].$$  \hspace{1cm} (5)

Because the argument of the exponential is proportional to $N$, in the thermodynamic limit, $N \to \infty$, the integral over $x$ is dominated by its maximal term. And because the system is perfectly replica symmetric, that is, all the replicas of the system are exactly equivalent, we also get that $x^{\alpha}$ is the same for all copies at the saddle point, $x^{(\alpha)\ast} \equiv m$.

Taking the derivative of the exponent with respect to $x^{(\alpha)}$ gives an extremum (here, the maximum)

$$-2\beta m + \frac{\partial \ln Q_1(\{x^{(\alpha)\})}{\partial x^{(\alpha)}} \bigg|_{x^{(\alpha)}=m} = 0,$$  \hspace{1cm} (6)
and thus
\[ m = \frac{1}{Q_1(m)} \sum_{\{\sigma^{(a)}\} = \pm 1} \sigma^{(a)} e^{\beta a_1[\{\sigma^{(a)}\}, m]}, \] \tag{7}

where
\[ \beta a_1[\{\sigma^{(a)}\}, m] = 2\beta m \sum_{\alpha=1}^{n} \sigma^{(a)} + \frac{\beta^2 \sigma_n^2}{2} \left( \sum_{\alpha=1}^{n} \sigma^{(a)} \right)^2 \] \tag{8}
and \[ Q_1(m) = \sum_{\{\sigma^{(a)}\} = \pm 1} e^{\beta a_1[\{\sigma^{(a)}\}, m]}. \] \tag{9}

gives (after dropping the subexponential terms in \( n \), which disappear upon taking the limit \( n \to 0 \))
\[ \overline{Q^n} \propto \exp \left[ N \left(-n\beta m^2 + \ln Q_1(m)\right) \right]. \] \tag{10}

Note that the expression for \( m \) resembles that for the average magnetization, but for a different Boltzmann weight. Using the Hubbard-Stratonovich transformation once more, we can write
\[ e^{\beta a_1[\{\sigma^{(a)}\}, m]} = \int \frac{dx}{\sqrt{2\pi}} \exp \left[ -\frac{1}{2} x^2 + (2\beta m + \beta \sigma_h x) \sum_{\alpha=1}^{n} \sigma^{(a)} \right], \] \tag{11}

which simplifies the single-site partition function
\[ Q_1(m) = \int \frac{dx}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \prod_{\alpha=1}^{n} \sum_{\sigma^{(a)} = \pm 1} e^{(2\beta m + \beta \sigma_h x)\sigma^{(a)}} \] \tag{12}\[ = \int \frac{dx}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} [2 \cosh(2\beta m + \beta \sigma_h x)]^n \] \tag{13}\[ = \int \frac{dx}{\sqrt{2\pi}} \exp \left\{ -\frac{x^2}{2} + n \ln [2 \cosh(2\beta m + \beta \sigma_h x)] \right\}, \] \tag{14}

and the self-consistent expression for the magnetization (noting again that the average of the copies is the same as the single-copy average, i.e., it is replica symmetric)
\[ m = \frac{1}{Q_1(m)} \sum_{\{\sigma^{(a)}\} = \pm 1} \left( \frac{1}{n} \sum_{\alpha=1}^{n} \sigma^{(a)} \right) e^{\beta a_1[\{\sigma^{(a)}\}, m]} \] \tag{15}\[ = \frac{1}{Q_1(m)} \int \frac{dx}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \frac{1}{n} \frac{\partial}{\partial (2\beta m)} \prod_{\alpha=1}^{n} \sum_{\sigma^{(a)} = \pm 1} e^{(2\beta m + \beta \sigma_h x)\sigma^{(a)}} \] \tag{16}\[ = \frac{1}{Q_1(m)} \int \frac{dx}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \frac{1}{n} \frac{\partial}{\partial (2\beta m)} [2 \cosh(2\beta m + \beta \sigma_h x)]^n \] \tag{17}\[ = \frac{1}{Q_1(m)} \int \frac{dx}{\sqrt{2\pi}} e^{-\frac{x^2}{2} + n \ln [2 \cosh(2\beta m + \beta \sigma_h x)]} \tanh(2\beta m + \beta \sigma_h x). \] \tag{18}

Finally, in order to obtain the disorder-averaged free energy, we use the replica trick as described in Eq. (1),
\[ \beta \tilde{a} = -\frac{1}{N} \frac{\partial \overline{Q^n}}{\partial n} \bigg|_{n=0} = \beta m^2 - \frac{\partial Q_1(m)}{\partial n} \bigg|_{n=0} \] \tag{19}\[ = \beta m^2 - \int \frac{dx}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \ln [2 \cosh(2\beta m + \beta \sigma_h x)], \] \tag{20}

where \( m \) is fixed by evaluating the self-consistent equation (18) at \( n = 0 \),
\[ m = \int \frac{dx}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \tanh(2\beta m + \beta \sigma_h x). \] \tag{21}
After the change of variables $h \equiv \sigma_h x$, the disorder-averaged free energy becomes

$$
\beta \bar{a} = \beta m^2 - \int \frac{dh}{\sqrt{2\pi\sigma_h^2}} e^{-\frac{h^2}{2\sigma_h^2}} \ln \left[ 2 \cosh(2\beta m + \beta h) \right],
$$

(22)

with

$$
m_{SC}(m) = \int \frac{dh}{\sqrt{2\pi\sigma_h^2}} e^{-\frac{h^2}{2\sigma_h^2}} \tanh(2\beta m + \beta h)
$$

(23)

The self-consistent has a trivial solution $m = 0$ for all values of $\sigma_h$ and $\beta$. Similarly to what happened in the mean-field treatment of the standard Ising model in Chapter 5, if $\beta$ is sufficiently large, then there exists a non-trivial solution with $m > 0$, and thus ferromagnetic order emerges. More specifically, the phase boundary between the paramagnetic and the ferromagnetic phase is found at unit slope

$$
2\beta_c \int \frac{dh}{\sqrt{2\pi\sigma_h^2}} \frac{e^{-\frac{h^2}{2\sigma_h^2}}}{(\cosh(\beta_c h))^2} = 1
$$

(24)

The resulting phase diagram is given in Fig. 1. Remarkably, this results indicate the existence of a phase transition driven purely by the fluctuations of the magnetic field, without any thermal noise! The expression for the free energy given above can also be used to deduce more elaborate observables, such as the susceptibility. For example, near the phase transition the magnetization scales as a power-law,

$$
m(T) \sim (T - T_c)^{1/2},
$$

(25)

with the same critical exponent as the mean-field Ising model. This mean-field scaling, however, is here only valid for $d \geq 6$, while the scaling is valid for $d \geq 4$ in the standard Ising model. Similarly, the Imry-Ma argument gives that there is no transition in $d = 2$ for the RFIM, while the standard Ising model still does. Clearly, magnetic field disorder generally weakens the ferromagnetic ordering transition.
IV. PURE STATES

Before considering models with more elaborate types of disorder, we first need to describe pure states. (We encountered these states already in Chapter 5, but did not bother to carefully characterize them at the time.) A pure state is a cluster of microstates that become separated from other clusters upon ergodicity breaking in the system. Once pure states emerge, a given copy of the system ends up only accessing a sub-part of phase space through local rearrangements (a properly defined global rearrangement may obviously bring the system from one state to another). In an Ising ferromagnet, for instance, two pure states with \( \pm \) are separated by an infinite barrier in the large system limit in the absence of magnetic field\(^2\). Going from one state to the other thus requires flipping \( \mathcal{O}(N) \) spins concurrently.

If the system is decomposed in pure states, its partition function can be expressed as

\[
Q = e^{-\beta Na(T)} = \sum_{\nu} e^{-\beta Na_{\nu}(T)} = \sum_{\nu} Q_{\nu},
\]

and in order to obtain the average of an observable \( O \), we have to calculate

\[
\langle O \rangle = \sum_{\nu} w_{\nu} \langle O \rangle_{\nu} \text{ with } w_{\nu} = Q_{\nu}/Q
\]

In the ferromagnetic phase of the Ising model, the average magnetization is indeed

\[
m = \frac{1}{2} m_+ + \frac{1}{2} m_- = 0,
\]

because \( m_+ = -m_- \), by symmetry (see footnote on p. 129). If we have more than a few pure states, however, it is more convenient to rewrite the free energy using a density of pure states

\[
Q(N,T) = e^{-\beta Na(T)} = \sum_{\nu} e^{-\beta Na_{\nu}(T)} = \int da \sum_{\nu} \delta(a - a_{\nu}(T)) e^{-\beta Na}
\]

\[
= \int da \tilde{\Omega}(a) e^{-\beta Na(T)} = \int_{a_{\min}}^{a_{\max}} da e^{N[\Sigma(a) - \beta a]} \sim e^{N[\Sigma(a^*) - \beta a^*]},
\]

where by analogy to Boltzmann’s entropy, we have defined the complexity per spin \( \Sigma(a) \equiv \log \tilde{\Omega}(a)/N \). The complexity measures the “disorder” of phase space, its configurational entropy so to speak. By construction, \( \Sigma \geq 0 \), and \( \Sigma = 0 \) when the number of pure states grows subexponentially with system size; complexity is then thermodynamically negligible. For the Ising model at \( T < T_c \), for instance, \( \Sigma = \ln 2/N \), and thus \( \Sigma = 0 \) in the limit \( N \to \infty \).

The last equality in Eq. (5) is obtained from the maximal term method, which we can use in the thermodynamic limit if the complexity is a monotonically increasing function of \( a \), i.e., if \( \Sigma \) has a single maximum. The partition function can then be approximated by finding the free energy \( a^* \) that maximizes the argument of the exponential, i.e.,

\[
\frac{\partial \Sigma}{\partial a}_{a=a^*} = \beta.
\]

In a glass, unlike in a ferromagnet, states can’t be easily separated by symmetry. To identify the transition, we thus use an altogether different strategy. We look for the temperature at which coupled copies of the system become more similar to each other than copies that are not, in the limit of weak coupling. It is similar to selecting a particular magnetization direction in the ferromagnetic Ising model by applying a very weak field in that direction. We will come back to this point later, but the key aspect for now is that we wish to compute the partition function of \( m \) copies of the system

\[
Q_m = \sum_{\nu} (e^{-\beta Na_{\nu}})^m = \int_{a_{\min}}^{a_{\max}} da e^{N[\Sigma(a) - \beta ma]} \sim e^{N[\Sigma(a^*) - \beta ma^*]},
\]

\(^2\) In finite dimensions, applying \( h \neq 0 \) allows the nucleation of the lower free energy phase by flipping only a finite number, \( \mathcal{O}(N^d) \) of spins, but in infinitely-connected models the barrier is infinite for any \( h \).
where the saddle point is obtained by setting
\[ \frac{\partial \Sigma}{\partial a} \bigg|_{a=a^*} = \beta m. \] (8)

Defining the result of this optimization as
\[ \Phi(m, T) \equiv ma^*(m, T) - T\Sigma[a^*(m, T)], \] (9)

and assuming that an analytical continuation of \( m \) to real values is possible, we obtain
\[ a^*(m, T) = \frac{\partial \Phi(m, T)}{\partial m}, \] (10)
\[ \Sigma[a^*(m, T)] = m\beta a^*(m, T) - \beta \Phi(m, T) = m^2 \frac{\partial m^{-1} \beta \Phi(m, T)}{\partial m}. \] (11)

**Exercise:** Check the above identities.

From the free energy of \( m \) copies of the system in Eq. (9), we can thus also obtain the complexity and the optimal free energy of a single copy of the system.

V. PHYSICAL BEHAVIOR OF THE \( p \)-SPIN GLASS MODEL

As discussed above, adding quenched disorder to the coupling constants results in the interaction energy itself being probabilistic\(^3\). Introducing disorder also weakens phase transitions. It reduces the energy gap at a first-order transition, and can even make a second-order transition disappear altogether, as it does in the RFIM. To increase the probability that a model has a finite-temperature phase transition, we consider the generalized fully-connected \( p \)-spin models\(^4\):

\[ E_\nu = -\sum_{1 \leq i_1 < i_2 < \ldots < i_p \leq N} J_i \sigma_{i_1} \sigma_{i_2} \ldots \sigma_{i_p}, \] (1)

In order to simplify the analysis, we choose \( \sigma_i \in \{-\infty, \infty\} \) to obey the spherical constraint \( N = \sum_i \sigma_i^2 \), and the coupling constants to be normally distributed, i.e.,

\[ P(J_l)dJ_l = \sqrt{\frac{N^{p-1}}{p!\pi J^2}} \exp\left(-\frac{J^2 N^{p-1}}{J^2 p!}\right) dJ_l. \] (2)

Note that the distribution properties have been chosen so that the energy per spin remains finite in the thermodynamic limit \( N \to \infty \).

Before solving this model, let us anticipate the result by first considering the various solution regimes and their microscopic interpretation.

- For (high) \( T > T_d \), the system is in a disordered paramagnetic phase. The free energy of that phase is smaller than any \( a - T\Sigma(a) \) for \( a \in [a_{\text{min}}, a_{\text{max}}] \), so the decomposition in pure states is meaningless.

- For \( T_d \geq T \geq T_K \), the paramagnetic state is obtained from the sum of an exponential number of pure glassy states, where \( a^* \in [a_{\text{min}}, a_{\text{max}}] \). Yet no phase transition occurs at \( T_d \), because the free energy changes smoothly and continuously from that of the paramagnetic phase. Below \( T_d \), however, the system is dynamically confined to a small region of phase space around where it was initialized. It sits within a pure state. Therefore, \( T_d \) is a dynamical transition, because the paramagnetic state is replaced by a glassy state in which phase space is disconnected, i.e., ergodicity is broken.

- For (low) \( T < T_K \), the system is dominated by the single lowest free energy state \( a^* = a_{\text{min}} \) with a complexity \( \Sigma(a^*) = 0 \). At \( T_K \), a phase transition occurs. The free energy and its first derivatives are continuous, but the second derivative jumps, thus giving rise to a second order phase transition, also known as a random first-order phase transition.

\(^3\) Note that this interaction energy contrasts with that of glass formation from simple fluids. In this case disorder arises from the irregular spatial organization of particles, and not from the pair interaction itself. The randomness is then said to be self-generated.

\(^4\) Recall that a fully-connected, \( p \)-spin models undergo a first-order transition for \( p \geq 3 \), while the standard Ising model with \( p = 2 \) has a second-order transition. See Appendix for details.
VI. REPLICA SYMMETRIC CALCULATION OF THE $p$-SPINS GLASS MODEL

In a replica calculation with replica symmetry breaking, we need $mn$ copies of the system: $m$ copies that will have an infinitely small coupling for calculating the free energy as described above, and $n$ copies for using the replica trick. We thus need to calculate the $n$ replicated free energy of $m$ coupled copies of the system, which have

$$\beta a_m(J_1) = -\frac{1}{N} \ln \left[ \int \left( \prod_{\alpha=1}^{m} d\sigma^{(\alpha)} \right) e^{-\beta \sum_{i=1}^{N} E^{(\alpha)}(\sigma)} e^{\beta \varepsilon \sum_{\alpha,\alpha'=1}^{m} \sum_{i=1}^{N} \sigma_i^{(\alpha)} \sigma_{i'}^{(\alpha')}} \right]$$

$$= -\frac{1}{N} \ln (Q_m(J_1)), \quad (1)$$

where the prime denotes the spherical constraint on the $\sigma_i$ values. Because the coupling terms are infinitely small ($\varepsilon \to 0$), we drop them from the calculation for now, but they will back later in the analysis.

In the derivation below we drop all non-exponential prefactors because they only provide a trivial offset to the free energy, and thus do not affect the final phase diagram. The partition function is then

$$Q_m = \int \left( \prod_{\alpha=1}^{mn} d\sigma^{(\alpha)} \right) \prod_{1 \leq i_1 < \ldots < i_p \leq N} \exp \left[ -\beta_0 \frac{N^{p-1}}{p!} + \beta J_1 \sum_{\alpha=1}^{mn} \sigma_{i_1}^{(\alpha)} \ldots \sigma_{i_p}^{(\alpha)} \right] \quad (2)$$

Note that the first integral over Gaussian variables simplifies the expression, but couples the replicas, just as it did for the RFIM. Note also that the last line is obtained by rewriting the product of exponentials into a sum of the exponents.

A. Overlap

Further simplifying the expression requires getting rid of the constrained integral to rewrite the $p$-tuplets in a more convenient form. In order to do so symmetrically, we introduce an overlap function that measures the similarity between two replicas of the system

$$\chi_{\alpha \alpha'} \equiv \frac{1}{N} \sum_{i=1}^{N} \sigma_i^{(\alpha)} \sigma_i^{(\alpha')} \quad (3)$$

For $\alpha = \alpha'$, we obviously get a perfect overlap and recover the spherical constraint, $\chi_{\alpha \alpha'} = 1$. For $\alpha \neq \beta$, the overlap of configurations within a state measures the size of that state in phase space; the larger the overlap, the tighter the cluster, and hence the more similar are the microstates within a pure state. In the low-temperature limit, $T \to 0$, each pure state concentrates to the lowest energy microstate within the cluster, hence $\chi_{\alpha, \alpha'} \to 1$. As $T$ grows, however, more configurations typically contribute to the pure state, and thus $\chi_{\alpha, \alpha'} < 1$. In the (high-temperature) paramagnetic state, the system exhibits no clustering at all, hence all of phase space is accessible and $\chi_{\alpha, \alpha'} = 0$.

Using an integral representation for the $\delta$ function, $\delta(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(itx)dt$, we can rewrite the overlap as

$$1 = \int d\hat{\chi} \prod_{\alpha, \alpha'=1}^{mn} \delta \left( N\chi_{\alpha \alpha'} - \sum_{i=1}^{N} \sigma_i^{(\alpha)} \sigma_i^{(\alpha')} \right) \quad (4)$$

$$= \int d\hat{\chi} \int_{-\infty}^{\infty} d\lambda \prod_{\alpha, \alpha'=1}^{mn} \frac{1}{2\pi i} \exp \left[ N\lambda_{\alpha \alpha'} \chi_{\alpha \alpha'} - \sum_{i=1}^{N} \sigma_i^{(\alpha)} \lambda_{\alpha \alpha'} \sigma_i^{(\alpha')} \right]. \quad (5)$$

Exercise: Verify this last result.
Plugging this result in Eq. (2) gets rid of the integration constraint, substituting them instead by integrals over $\chi_{\alpha\alpha'}$. We therefore get (again dropping subexponential prefactors)

$$Q^n_m = \int \left( \prod_{\alpha=1}^{mn} d\sigma^{(\alpha)} \right) d\hat{\chi} d\hat{\lambda} \times \exp \left[ \frac{\beta^2 N}{4} \sum_{\alpha,\alpha'=1}^{mn} \chi_{\alpha\alpha'}^p + N \sum_{\alpha,\alpha'=1}^{mn} \lambda_{\alpha\alpha'} \chi_{\alpha\alpha'} - \sum_{\alpha,\alpha'=1}^{mn} \sum_i^{N} \sigma_i^{(\alpha)} \lambda_{\alpha\alpha'} \sigma_i^{(\alpha')} \right].$$

Using the spectral theorem (as in Section 4.3) decouples the quadratic terms in the exponent. The integrals over the spins then simplify to a multivariate Gaussian integral, hence we can write

$$Q^n_m = \int d\hat{\chi} d\hat{\lambda} \exp[NX(\hat{\chi}, \hat{\lambda})],$$

where we have implicitly defined for matrices $\hat{\chi}$ and $\hat{\lambda}$

$$X(\hat{\chi}, \hat{\lambda}) = \frac{\beta^2}{4} \sum_{\alpha,\alpha'=1}^{mn} \chi_{\alpha\alpha'}^p + \sum_{\alpha,\alpha'=1}^{mn} \lambda_{\alpha\alpha'} \chi_{\alpha\alpha'} - \frac{1}{2} \log(\det(\hat{\lambda})).$$

**B. Saddle-point evaluation**

Equation (6) lends itself to a saddle point approximation in the thermodynamic limit. If the distribution were bimodal (or worse), this option would not be readily available, but the presence of multiple replicas of the system enables us to capture all minima at once. There are a few additional mathematical difficulties. First, the proper limits we need to take are

$$\beta \Phi(m, \beta) = -\lim_{N \to \infty} \frac{1}{N} \lim_{n \to 0} \frac{\partial}{\partial n} \int d\hat{\chi} d\hat{\lambda} \exp[NX(\hat{\chi}, \hat{\lambda})].$$

As for the RFIM, in order to make a saddle-point approximation we need to invert the two limits. In general, this operation can be mathematically risky, but it is here (provably) correct. Second, we must make sure that the optimum of $X$ is not unstable, i.e., that it is not a (high-order) saddle point. All the eigenvalues of the matrix must thus be negative, which in this case can also be shown to be true. Third, in order to optimize the function $X$ with respect to each element of the matrix $\hat{\lambda}$ we need the linear-algebra identity

$$\frac{\partial}{\partial M_{\alpha\alpha'}} \log(\det M) = (M^{-1})_{\alpha'\alpha}. \quad (9)$$

**Exercise:** Verify the above result.

Setting $\partial X/\partial \lambda_{\alpha\alpha'}|_{\lambda_{\alpha\alpha'}=\lambda_{\alpha\alpha'}} = 0$ in order to identify the extremum of the exponent then gives

$$\chi_{\alpha\alpha'} - (2\hat{\lambda}^{-1})_{\alpha'\alpha} = 0,$$

which upon substitution simplifies the exponent as

$$X(\hat{\chi}) = \frac{\beta^2}{4} \sum_{\alpha,\alpha'=1}^{mn} \chi_{\alpha\alpha'}^p + \frac{1}{2} \log(\det(\hat{\chi})) + O(nm).$$

(11)

The leading order correction is linear in $n$, which provides a constant shift to the free energy upon taking the derivative with respect to $n$. Without loss of generality, we can thus drop it. Higher-order terms in $n$ become exactly zero in the limit $n \to 0$ and can thus also be dropped. The free energy of the replicated system is then

$$\beta \Phi(m, \beta) = -\frac{1}{N} \lim_{n \to 0} \frac{\partial}{\partial m} \exp \left\{ -\frac{\beta^2}{2} \sum_{\alpha,\alpha'=1}^{mn} \chi_{\alpha\alpha'}^p + \log(\det(\hat{\chi})) \right\}. \quad (12)$$
C. The Parisi ansatz (enlightened guess)

We seek a general solution for $\hat{\chi}$, for arbitrary $n$. This seems hard, but cannot be avoided if we are to use the replica trick. As a guess, we postulate that $\hat{\chi}$ takes the form known as the Parisi ansatz, which can be motivated as follows.

At low temperatures, we expect multiple pure states to be possible, but each of the $m$ weakly coupled replicas, as in Eq. (1), should end up in the same state and thus to have an overlap $0 < \chi_{\alpha\alpha'} \equiv q \leq 1$. By contrast, all the other replicas are overwhelmingly likely to end up in different pure states and thus to have a much lower overlap. (It can be shown that for these replicas $\chi_{\alpha\alpha'} = 0$.) If we group replicas by the pure state to which they belong, we obtain a matrix of the form

$$
\hat{\chi} = \begin{pmatrix}
1 & q & q \\
q & 1 & q \\
q & q & 1
\end{pmatrix}
$$

for $n = 2$ sets of $m = 3$ copies. We can then use the relation

$$
\det \begin{pmatrix}
1 & q & q \\
q & 1 & q \\
q & q & 1
\end{pmatrix} = (1 - q)^{m-1}[1 + (m - 1)q]
$$

(13)

(14)

(15)

(16)

and directly compute

$$
\sum_{\alpha\alpha'} \chi_{\alpha\alpha'}^{p} = n[m + m(m - 1)q^{p}].
$$

Implicitly defining

$$
X(\hat{\chi}) \equiv -\beta mn \phi_{1RSB}(m, q, T)
$$

$$
= \frac{mn}{2} \left\{ \frac{\beta^2}{2} [1 + (m - 1)q^{p}] + \frac{m - 1}{m} \log(1 - q) + \frac{1}{m} \log[1 + (m - 1)q] \right\}
$$

(17)

where $q^*$ minimizes the free energy, i.e., its solutions must obey

$$
\left. \frac{\partial \phi(m, q, T)}{\partial q} \right|_{q = q^*} = 0 \Rightarrow (m - 1) \left( \frac{\beta^2}{2} pq^{p-1} - \frac{q}{(1 - q)[(m - 1)q + 1]} \right) = 0.
$$

(18)

Recall that $m$ must be set to unity in the end, because we ultimately care about the free energy of a single replica of the system. These operations must, however, be done with care, because we have analytically continued the free energy expression to unphysical regimes.

---

5 This result is known as the 1RSB solution, because additional subdivision of phase space are possible, which would correspond to a higher order of replica symmetry breaking. The form of the overall matrix is then adjusted so as to capture this richer state structure.
FIG. 2. Phase diagram of the 3-spin spherical model in the \(m-T\) plane. The equilibrium solution is found on the \(m = 1\) line (dashed). The thermodynamic transition takes place at \(T_K = 0.58605\), where \(m_d(T)\) (black line) crosses the \(m = 1\) line and \(q_d = 0.645\). The overlap is therefore quite significant in the ground state. The onset of breaking up of phase space defines the dynamical transition at \(T_d = 0.61237\), where \(m_d(T)\) (blue line) crosses the \(m = 1\) line and \(q_d = 0.5\). The states are then already fairly well-formed. States with \(q \neq 0\) are found in the grey zone, above \(m^*(T)\) (red line). Adapted from F. Zamponi, arXiv:1008.4884.

D. Phase diagram

We have assumed that replicas in different blocks are uncorrelated and have zero overlap, while replicas in the same block are correlated and thus have an overlap \(q\). This correlation is due to the artificial coupling between the replicas, whose role is only to select the solution for which \(q\) has the highest value at any given \(T\). As mentioned above, this approach is similar to using an infinitesimal external field to select for one of the two ferromagnetic states of the Ising model. Without this coupling, the free energy from Eq. (12), which always has a minimum at \(q = 0\), would not detect the \(q^* \neq 0\) stationary points that appear at low \(T\). If we are interested in the partition function of \(m\) replicas in the same state, as we did above to compute the complexity, we should therefore always take the solution with \(q \neq 0\), if it exists. From the final result

\[
\Phi(m, T) = m \phi_{1\text{RSB}}(m, q^*(m, T), T),
\]  

we can draw a phase diagram in the \(m-T\) plane. The phase diagram in Figure 32 for \(p = 3\) results from the following considerations.

- We must identify the region where a solution with \(q \neq 0\) is found, i.e., where pure states exist. This delimits the line \(m^*(T)\). For \(m < m^*(T)\), we have that \(q^*(m, T) = 0\), \(\Phi(m, T)\) is trivial, and the complexity is zero; while for \(m > m^*(T)\), a non-trivial solution \(q^*(m, T)\) is found. In this region we can compute the complexity and the free energy as a function of \(m\) using Eqs. 10 from our discussion of pure states

\[
\Sigma(m, T) = m^2 \frac{\partial}{\partial m} \left[ m^{-1} \beta \Phi(m, T) \right] = m^2 \frac{\partial \beta \phi_{1\text{RSB}}(m, q^*, T)}{\partial m},
\]

\[
a^*(m, T) = \frac{\partial \Phi(m, T)}{\partial m} = \frac{\partial \beta \phi_{1\text{RSB}}(m, q^*, T)}{\partial m}.
\]

Note that these quantities are only defined above the \(m^*(T)\) line.
• Analyzing the behavior of $a^*(m, T)$ and $\Sigma(m, T)$ indicates that starting from $m = \max\{0, m^*(T)\}$ and increasing $m$, both $a^*(m, T)$ and $\Sigma(m, T)$ first increase up to a maximum and then decrease upon further increasing $m$. This non-monotonic behavior results in two branches of the parametric curve $\Sigma(a)$, of which only one is physical: the one that corresponds to both $a$ and $\Sigma$ decreasing with $m$, as can be seen from Eq. (8). We can thus define a line $m = m_d(T)$ where $a^*(m, T)$ is maximum, which corresponds to where the upper threshold values $a_{\text{max}}(T)$, i.e., no states of higher free energy exists at that $T$. The relevant region is thus the one for which $m \geq m_d(T)$.

• We observe that at a fixed $T$, the complexity is finite for $m = m_d(T)$ and decreases upon increasing $m$ above $m_d(T)$, until it vanishes on a second line $m_s(T)$. Above this line, the complexity becomes negative, indicating that pure states do not exist anymore. States are therefore found for $m_s(T) > m > m_d(T)$, with $m_d(T)$ corresponding to $a_{\text{max}}(T)$ and $m_s(T)$ corresponding to $a_{\text{min}}(T)$.

The line $m = 1$ which corresponds to the equilibrium partition function of a single copy of the system. The lines $m_d(T)$ and $m^*(T)$ merge at $m = 1$. When the line $m_d(T)$ crosses $m = 1$, the optimum free energy is $a_{\text{max}}$. At this point the paramagnet breaks into many states, and this crossing defines the temperature $T_d$. Similarly, the point where the line $m_s(T)$ crosses $m = 1$ corresponds to the point where the saddle point is equal to $a_{\text{min}}$, the complexity vanishes and a single state dominates, which corresponds to the phase transition at $T_K$.

For $T < T_K$, the optimal value of the free energy is always $a_{\text{min}}$, as obtained along the $m_s(T)$. Continuing the free energy above $m_s(T)$ is incorrect, because the complexity then becomes negative, which is clearly unphysical. In other words, the system of $m$ replicas undergoes a thermodynamic transition on the line $m_s(T)$, and the value of the free energy $\Phi(m, T)$ on the line $m_s(T)$, which is $\Phi(m, T) = m a_{\text{min}}(T)$, persists up to the $m = 1$ line. Noting that $m < 1$ and recalling that $\Sigma = 0$ along $m_s$, one obtains from Eq. (9) that

$$a_{\text{min}}(T) = \frac{\Phi(m, T)}{m} \bigg|_{m=m_s(T)} = \phi_{\text{1RSB}}(m_s(T), q^*(T), T).$$

This last result is very important. Indeed, for $T > T_K$ the free energy of the system is equal to the free energy of the paramagnet, corresponding simply to $\phi_{\text{1RSB}}(m = 1) = -\beta/4$. Below $T_K$, instead, the free energy of the system is given by extremizing $\phi_{\text{1RSB}}$ with respect to both $m$ and $q^6$.

The above discussion shows that $q$ is the order parameter for the transition. It is zero in the paramagnetic phase, and it jumps to a nonzero value in the glass phase signaling the spontaneous breaking of replica symmetry. This assumption is represented in the structure of the overlap matrix. The nature of the low-temperature phase is to have its symmetry spontaneously broken below a certain temperature, with $q$ going from 0 in the paramagnetic phase, to a finite value at $T_d$. This behavior is similar to the ferromagnetic $p$-spin mean-field model from the first exercise at the end of this chapter, where the magnetization goes from 0 to a finite value, i.e., the state is well formed at the transition for $p \geq 3$. There is no thermodynamic transition at $T_d$ because the sum over all these pure states is equal to the paramagnetic free energy. Yet the system is dynamically trapped within a pure state, which can be shown explicitly by solving the dynamics of this model (see additional references). By contrast, the second-order transition at $T_K$ is thermodynamic, because the free energy then behaves differently from that of the paramagnetic phase.

It can be shown that in general the dynamical transition takes place at $T_d = \sqrt{\frac{2p-2p^3}{2p-2p^3}}$ with $q_d = \frac{\beta}{2}$, while $T_K = \sqrt{\frac{2p}{2}}(1 - y^*)^{1-2/p}$ with $q_K = 1 - y^*$, where $y^*$ is the solution to $\frac{2}{p} = -2y^{1+y\ln y}$. Note that both transitions weaken as $p$ disappear and vanish for $p = 2$. This effect is reminiscent of what happens in systems without disorder.

In summary, the replica method allow us to fully characterize the thermodynamics of the spherical $p$-spin model, by computing the free energy of the paramagnetic phase, the free energy of the glass phase, and the distribution $\Sigma(a)$ of all the metastable states as a function of $a$ and $T$. The main assumption we made in the derivation is that all the states are equivalent, with a self-overlap $q$ and zero mutual overlap, which suggests that they are randomly distributed in phase space. These properties are expressed by the 1RSB structure of the overlap matrix $\lambda_{\alpha\alpha'}$, in which all nonzero entries, correspond to replicas evolving in the same pure state, have an equivalent overlap $q$. This behavior is exact for the spherical $p$-spin glass model, but this is somewhat exceptional. The general description of the transitions is nonetheless quite general.

VII. FINITE $d$ CORRECTIONS

A description of some of the finite-dimensional corrections to the mean-field picture due is contained within the random first-order transition (RFOT) theory of the (spin-)glasses initially formulated by T. Kirkpatrick, D. Thir-
malai, and P. Wolynes in the 1980s. RFOT notably suggests that barriers between states remain finite if the model is considered on a finite-dimensional lattice. Going from one state to another is then possible via the nucleation of a lower free energy state. The barriers then result from the competition between the free energy difference and the interfacial free energy between the various states. Because the scaling of these two quantities converges in high dimension, we recover that the barrier heights diverge in the infinitely-connected limit, so pure states are then well defined. The role of fluctuations in low-dimensions is also non negligible, and adding them to the mean-field description is still an active area of research. In structural glasses, excitations due to local structural defects also affect the system dynamics. This last effect was extensively considered by David Chandler and his collaborators in the 2000s.

ACKNOWLEDGMENTS

I am grateful to the many former students and collaborators who have helped me improve these notes. I am especially indebted to Francesco Zamponi, who has patiently introduced me to these ideas. This work was supported in part by a grant from the Simons Foundation (#454937) and by National Science Foundation grant no. NSF DMR-1055586.

Appendix A: Generalized, fully-connected spin models

In this Appendix, we consider generalized versions of the fully-connected Ising model that are useful for the mean-field study of spin glasses. Consider a system of $N$ spins that are all nearest-neighbors to each other. This arrangement is the fully-connected (or infinite-dimensional) version of the lattice depicted in Fig. 5.1. Note that for such a system the mean-field solution is exact because nearest-neighbor fluctuations have but a negligible impact on the effective field felt by a particle.

We also generalize the Ising model by including interactions between multiple sites. The $p$-spin version of the Ising model has spin variables $\sigma = \pm 1$ coupled in (integer) $p$-plets (pairs for $p = 2$, triplets for $p = 3$, etc.),

$$E_{\nu} = -J \sum_{1 \leq i_1 < i_2 < \ldots < i_p \leq N} \sigma_{i_1} \sigma_{i_2} \ldots \sigma_{i_p}$$  \hspace{1cm} (A1)

where the choice of coupling constant $J \equiv p!/N^{p-1}$ guarantees that the energy per spin remains finite in the thermodynamic limit $N \to \infty$.

Exercise: Use mean-field arguments to show that the free energy of the Ising $p$-spin model in the limit $N \to \infty$ is given by (setting $J = 1$)

$$\beta a \equiv \lim_{N \to \infty} \frac{\beta A}{N} = -\beta m^p + \left[ \left( \frac{1 + m}{2} \right) \log \left( \frac{1 + m}{2} \right) + \left( \frac{1 - m}{2} \right) \log \left( \frac{1 - m}{2} \right) \right],$$  \hspace{1cm} (A2)

where $m = \langle M \rangle / N = \frac{1}{N} \sum_{i=1}^{N} \sigma_i$ is the average magnetization per spin. Hint: Write the expression starting from $A = E - TS$.

Exercise: In order to better grasp the behavior of the above model, find the series expansion of $a$ to order $m^4$ for $p = 2$, and plot the original $a$ as a function of $m$ for $\beta = 5/6, 1, 2, 3$, and 4 for both $p = 2$ and $p = 4$. What are the equilibrium states? For $p = 4$, if the system is prepared in a metastable state, how long will the system remain in that state?

Exercise: By taking the derivative of the Ising $p$-spin model free energy obtained above with respect to the magnetization, show that the self-consistent mean-field relation for the $p = 2$ Ising model is recovered. Find the numerical value of the transition temperature and plot the equilibrium magnetization as a function of temperature for both $p = 2$ and $p = 4$. What is different between the two systems?

Exercise: By looking at the second derivative of the free energy, show that the $p = 2$ behavior is qualitatively different than for $p \geq 3$.

For mathematical convenience we also consider the spherical $p$-spin model, where the spin variable can take any real value, i.e., $\sigma_i \in (-\infty, \infty)$, under the constraint

$$\frac{1}{N} \sum_i \sigma_i^2 = 1.$$  \hspace{1cm} (A3)

Note that this condition is implicitly included in Ising models.
**Exercise:** We will now solve the mean-field problem for the $p$-spin spherical model. Note that in this case, the mean-field energy is also

$$\langle E \rangle = -Nm(\sigma)^p,$$  \hfill (A4)

but the entropy expression cannot be written down as straightforwardly as before. We write the partition function as (setting $k_B = 1$)

$$Q = \int d\sigma \delta \left( \sum_i \sigma_i^2 - N \right) e^{\beta Nm(\sigma)^p}$$

$$= \int dm d\sigma \delta \left( \sum_i \sigma_i^2 - N \right) \delta \left( \sum_i \sigma_i - Nm \right) e^{\beta Nm^p}$$

$$= \int dm e^{Ns(m) + \beta Nm^p},$$

where

$$e^{Ns(m)} = \int d\sigma \delta \left( \sum_i \sigma_i^2 - N \right) \delta \left( \sum_i \sigma_i - Nm \right).$$ \hfill (A5)

Rewrite this expression using the integral representation for the delta function, $\delta(x) = \int_{-i\infty}^{i\infty} \frac{dy}{2\pi} e^{-xy}$, before using a saddle point approximation (for $N \to \infty$) to calculate the resulting integral and obtain

$$\beta a \equiv \lim_{N \to \infty} \frac{\beta A(m)}{N} = -\frac{1}{2} \left( 1 + \ln(2) + \ln(1 - m^2) \right) - \beta m^p.$$ \hfill (A6)

**BIBLIOGRAPHY**

Various additional pedagogical references about the $p$-spin model are available. Part of these notes borrowed heavily from the first one.

- F. Zamponi, “Mean field theory of spin glasses”, arXiv:1008.4844.
- T. Castellani and A. Cavagna, “Spin-glass theory for pedestrians”, *J. Stat. Mech.* P05012 (2005).
- V. Dotsenko, *Introduction to the Replica Theory of Disordered Statistical Systems*, Cambridge University Press, Cambridge (2000).

Solving the dynamical behavior of the spherical $p$-spin glass is beyond the scope of this chapter, but a pedagogical introduction can be found in Castellani and Cavagna’s notes.

Applications of spin glass ideas to protein folding are reviewed in the following.

- J. N. Onuchic, Z. Luthey-Schulten, and P. G. Wolynes, “Theory of protein folding: the energy landscape perspective”, *Annu. Rev. Phys. Chem.* 48, 545 (1997).

The connection of the above ideas with the mean-field description of structural glasses was also recently reviewed.

- P. Charbonneau, J. Kurchan, G. Parisi, P. Urbani, F. Zamponi, “Glass and Jamming Transitions: From Exact Results to Finite-Dimensional Descriptions”, *Annu. Rev. Cond. Matt. Phys.* 8, 265 (2017).

Applications of spin glass ideas to computer science and information theory are presented in the following textbooks.

- M. Mézard and A. Montanari, *Information, Physics, and Computation*, Oxford University Press, Oxford (2009).
- H. Nishimoro, *Statistical Physics of Spin Glasses and Information Processing*, Oxford University Press, Oxford (2001).