5

THE RANDOM ENERGY MODEL

{ch:rem}

The random energy model (REM) is probably the simplest statistical physics model of a disordered system which exhibits a phase transition. It is not supposed to give a realistic description of any physical system, but it provides a workable example on which various concepts and methods can be studied in full details. Moreover, due the its simplicity, the same mathematical structure appears in a large number of contexts. This is witnessed by the examples from information theory and combinatorial optimization presented in the next two chapters. The model is defined in Sec. 5.1 and its thermodynamic properties are studied in Sec. 5.2. The simple approach developed in these section turns out to be useful in a large varety of problems. A more detailed (and also more involved) study of the low temperature phase is given in Sec. 5.3. Section 5.4 provides an introduction to the so-called annealed approximation, which will be useful in more complicated models.

5.1 Definition of the model

{se:rem_def}

A statistical mechanics model is defined by a set of configurations and an energy function defined on this space. In the REM there are $M = 2^N$ configurations (like in a system of N Ising spins) to be denoted by indices $i, j, \dots \in \{1, \dots, 2^N\}$. The REM is a **disordered model**: the energy is not a deterministic function but rather a stochastic process. A particular realization of such a process is usually called a **sample** (or **instance**). In the REM, one makes the simplest possible choice for this process: the energies $\{E_i\}$ are i.i.d. random variables (the energy of a configuration is also called an **energy level**). For definiteness we shall keep here to the case where they have Gaussian distribution with zero mean and variance $N/2$, but other distributions could be studied as well[9]. The pdf for the energy E_i of the state i is thus given by

$$P(E) = \frac{1}{\sqrt{\pi N}}\, e^{-E^2/N} \ , \tag{5.1}$$

{eq:pde_rem}

Given an instance of the REM, which consists of the 2^N real numbers E_j drawn from the pdf (5.1), one assigns to each configuration i a Boltzmann probability p_i in the usual way:

$$p_j = \frac{1}{Z} \exp\left(-\beta E_j\right) \tag{5.2}$$

{eq:bolt_rem}

[9] The scaling with N of the distribution should be chosen in such a way that thermodynamic potentials are extensive

where $\beta = 1/T$ is the inverse of the temperature, and the normalization factor Z (the partition function) equals:

$$Z = \sum_{j=1}^{2^N} \exp\left(-\beta E_j\right) \, . \tag{5.3}$$

{eq:rem_zdef}

Notice that Z depends upon the temperature β, the ‘sample size’ N, and the particular realization of the energy levels $E_1 \dots E_M$. We dropped all these dependecies in the above formula.

It is important not to be confused by the existence of two levels of probabilities in the REM, as in all disordered systems. We are interested in the properties of a probability distribution, the Boltzmann distribution (5.2), which is itself a random object because the energy levels are random variables.

Physically, a particular realization of the energy function corresponds to a given sample of some substance whose microscopic features cannot be controlled experimentally. This is what happens, for instance, in a metallic alloy: only the proportions of the various components can be controlled. The precise positions of the atoms of each species are described as random variables. The expectation value with respect to the sample realization will be denoted in the following by $\mathbb{E}(\cdot)$. For a given sample, Boltzmann’s law (5.2) gives the probability of occupying the various possible configurations, according to their energies. The average with respect to Boltzmann distribution will be denoted by $\langle \cdot \rangle$. In experiments one deals with a single (or a few) sample(s) of a given disordered material. One could therefore be interested in computing the various thermodynamic potential (free energy F_N, internal energy U_N , or entropy S_N) for *this given* sample. This is an extremely difficult task. However, we shall see that, as $N \to \infty$, the probability distributions of intensive thermodynamic potentials concentrate around their expected values:

$$\lim_{N\to\infty} \mathbb{P}\left[\left| \frac{X_N}{N} - \mathbb{E}\left(\frac{X_N}{N}\right) \right| \geq \theta \right] = 0 \tag{5.4}$$

for any potential X ($X = F, S, U, \dots$) and any tolerance $\theta > 0$. The quantity X is then said to be **self-averaging**. This essential property can be summarized plainly by saying that almost all large samples “behave” in the same way [10]. Often the convergence is exponentially fast in N (this happens for instance in the REM): this means that the expected value $\mathbb{E}\, X_N$ provide a good description of the system already at moderate sizes.

{se:thermo_rem}

5.2 Thermodynamics of the REM

In this Section we compute the thermodynamic potentials of the REM in the thermodynamic limit $N \to \infty$. Our strategy consists first in estimating the

[10] This is the reason why different samples of alloys with the same chemical composition have the same thermodynamic properties

microcanonical entropy density, which has been introduced in Sec. 2.4. This knowledge is then used for computing the partition function Z to exponential accuracy at large N.

5.2.1 *Direct evaluation of the entropy*

{se:MicroREM}

Let us consider an interval of energies $\mathcal{I} = [N\varepsilon, N(\varepsilon+\delta)]$, and call $\mathcal{N}(\varepsilon, \varepsilon+\delta)$ the number of configurations i such that $E_i \in \mathcal{I}$. Each energy 'level' E_i belongs to $\mathcal{I}$ independently with probability:

$$P_{\mathcal{I}} = \sqrt{\frac{N}{\pi}} \int_{\varepsilon}^{\varepsilon+\delta} e^{-Nx^2/2}\, dx\,. \tag{5.5}$$

Therefore $\mathcal{N}(\varepsilon, \varepsilon+\delta)$ is a binomial random variable, and its expectation and variance are given by:

$$\mathbb{E}\,\mathcal{N}(\varepsilon,\varepsilon+\delta) = 2^N\, P_{\mathcal{I}}, \qquad \mathrm{Var}\,\mathcal{N}(\varepsilon,\varepsilon+\delta) = 2^N\, P_{\mathcal{I}}[1-P_{\mathcal{I}}]\,, \tag{5.6}$$

Because of the appropriate scaling with N of the interval $\mathcal{I}$, the probability $P_{\mathcal{I}}$ depends exponentially upon N. To exponential accuracy we thus have

$$\mathbb{E}\,\mathcal{N}(\varepsilon,\varepsilon+\delta) \doteq \exp\left\{N \max_{x\in[\varepsilon,\varepsilon+\delta]} s_{\mathrm{a}}(x)\right\}\,, \tag{5.7}$$

$$\frac{\mathrm{Var}\mathcal{N}(\varepsilon,\varepsilon+\delta)}{[\mathbb{E}\,\mathcal{N}(\varepsilon,\varepsilon+\delta)]^2} \doteq \exp\left\{-N \max_{x\in[\varepsilon,\varepsilon+\delta]} s_{\mathrm{a}}(x)\right\} \tag{5.8}$$

where $s_{\mathrm{a}}(x) \equiv \log 2 - x^2$. Notice that $s_{\mathrm{a}}(x) \geq 0$ if and only if $x \in [-\varepsilon_*, \varepsilon_*]$, with $\varepsilon_* = \sqrt{\log 2}$.

The intuitive content of these equalities is the following: When ε is outside the interval $[-\varepsilon_*, \varepsilon_*]$, the typical density of energy levels is exponentially small in N: for a generic sample there is no configuration at energy $E_i \approx N\varepsilon$. On the contrary, when $\varepsilon \in]-\varepsilon_*, \varepsilon_*[$, there is an exponentially large density of levels, and the fluctuations of this density are very small. This result is illustrated by a small numerical experiment in Fig. 5.1. We now give a more formal version of this statement.

Proposition 5.1 *Define the entropy function*

{propo:REMdos}

$$s(\varepsilon) = \begin{cases} s_{\mathrm{a}}(\varepsilon) = \log 2 - \varepsilon^2 & \textit{if } |\varepsilon| \leq \varepsilon_*, \\ -\infty & \textit{if } |\varepsilon| > \varepsilon_*. \end{cases} \tag{5.9}$$

Then, for any couple ε and δ, with probability one:

$$\lim_{N\to\infty} \frac{1}{N} \log \mathcal{N}(\varepsilon,\varepsilon+\delta) = \sup_{x\in[\varepsilon,\varepsilon+\delta]} s(x)\ \ . \tag{5.10}$$

Proof: The proof makes a simple use of the two moments of the number of energy levels in $\mathcal{I}$, found in (5.7,5.8).

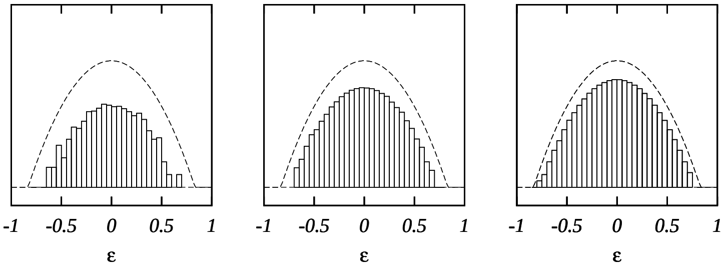


FIG. 5.1. Histogram of the energy levels for three samples of the random energy model with increasing sizes: from left to right $N = 10, 15$ and 20. Here we plot $N^{-1} \log \mathcal{N}(\varepsilon, \varepsilon + \delta)$ versus ε, with $\delta = 0.05$. The dashed curve gives the $N \to \infty$ analytical prediction (5.9).

{fig:remexp}

Let us first assume that the interval $[\varepsilon, \varepsilon + \delta]$ is disjoint from $[-\varepsilon_*, \varepsilon_*]$. Then $\mathbb{E}\mathcal{N}(\varepsilon, \varepsilon + \delta) \doteq e^{-AN}$, with $A = -\sup_{x \in [\varepsilon, \varepsilon+\delta]} s_{\rm a}(x) > 0$. As $\mathcal{N}(\varepsilon, \varepsilon + \delta)$ is an integer, we have the simple inequality

$$\mathbb{P}[\mathcal{N}(\varepsilon, \varepsilon + \delta) > 0] \le \mathbb{E}\mathcal{N}(\varepsilon, \varepsilon + \delta) \doteq e^{-AN} \,. \tag{5.11}$$

In words, the probability of having an energy level in any fixed interval outside $[-\varepsilon_*, \varepsilon_*]$ is exponentially small in N. The inequality of the form (5.11) goes under the name of **Markov inequality**, and the general strategy is sometimes called the **first moment method**. A general introduction to this approach is provided in App. **???**.

Assume now that the intersection between $[\varepsilon, \varepsilon + \delta]$ and $[-\varepsilon_*, \varepsilon_*]$ is a finite length interval. In this case $\mathcal{N}(\varepsilon, \varepsilon + \delta)$ is tightly concentrated around its expectation $\mathbb{E}\mathcal{N}(\varepsilon, \varepsilon + \delta)$ as can be shown using Chebyshev inequality. For any fixed $C > 0$ one has

$$\mathbb{P}\left\{\left|\frac{\mathcal{N}(\varepsilon, \varepsilon + \delta)}{\mathbb{E}\mathcal{N}(\varepsilon, \varepsilon + \delta)} - 1\right| > C\right\} \le \frac{\mathrm{Var}\,\mathcal{N}(\varepsilon, \varepsilon + \delta)^2}{C^2[\mathbb{E}\mathcal{N}(\varepsilon, \varepsilon + \delta)]^2} \doteq e^{-BN} \,, \tag{5.12}$$

with $B = \sup_{x \in [\varepsilon, \varepsilon+\delta]} s_{\rm a}(x) > 0$. A slight variation of the above reasoning is often referred to as the **second moment method**, and will be further discussed in App. **????**.

Finally, the statement (5.10) follows from the previous estimates through a straightfoward application of Borel-Cantelli Lemma. □

Exercise 5.1 Large deviations: let $\mathcal{N}_{\rm out}(\delta)$ be the total number of configurations j such that $|E_j| > N(\varepsilon_* + \delta)$, with $\delta > 0$. Use Markov inequality to show that the fraction of samples in which there exist such configurations is exponentially small.

Besides being an interesting mathematical statement, Proposition 5.1 provides a good quantitative estimate. As shown in Fig. 5.1, already at $N = 20$, the outcome of a numerical experiment is quite close to the asymptotic prediction. Notice that, for energies in the interval $]-\varepsilon_*, \varepsilon_*[$, most of the discrepancy is due to the fact that we dropped subexponential factors in $\mathbb{E}\mathcal{N}(\varepsilon, \varepsilon+\delta)$. It is easy $\star$ to show that this produces corrections of order $\Theta(\log N/N)$ to the asymptotic behavior (5.10). The contribution due to fluctuations of $\mathcal{N}(\varepsilon, \varepsilon+\delta)$ around its average is instead exponentially small in N.

5.2.2 *Thermodynamics and phase transition*

From the previous result on the microcanonical entropy density, we now compute the partition function $Z_N(\beta) = \sum_{i=1}^{2^N} \exp(-\beta E_i)$. In particular, we are interested in intensive thermodynamic potentials like the free entropy density $\phi(\beta) = \lim_{N\to\infty}[\log Z_N(\beta)]/N$. We start with a fast (and loose) argument, using the general approach outlined in Sec. 2.4. It amounts to discretizing the energy axis using some step δ, and counting the energy levels in each interval with (5.10). Taking in the end the limit $\delta \to 0$ (after the limit $N \to \infty$), one expects to get, to leading exponential order:

$$Z_N(\beta) \doteq \int_{-\varepsilon_*}^{\varepsilon_*} d\varepsilon \;\exp\left[N\left(s_{\rm a}(\varepsilon) - \beta\varepsilon\right)\right] \ . \tag{5.13}$$

{eq:rem_zcanon}

The rigorous formulation of the result can be obtained in analogy[11] with the general equivalence relation stated in Proposition 2.6. We find the free entropy density:

$$\phi(\beta) = \max_{\varepsilon\in[-\varepsilon_*,\varepsilon_*]} \left[s_{\rm a}(\varepsilon) - \beta\varepsilon\right], \tag{5.14}$$

Notice that although every sample of the REM is a new statistical physics system, which might have its own thermodynamic potentials, we have found that almost all samples have the same free entropy density (5.14), and thus the same energy ,entropy, and free energy densities. More precisely, for any fixed tolerance $\theta > 0$, we have $|(1/N)\log Z_N(\beta) - \phi(\beta)| < \theta$ with probability approaching one as $N \to \infty$.

Let us now discuss the physical content of the result (5.14). The optimization problem on the right-hand side can be solved through the geometrical construction illustrated in Fig. 5.2. One has to find a tangent to the curve $s_{\rm a}(\varepsilon) = \log 2 - \varepsilon^2$ with slope $\beta \geq 0$. Call $\varepsilon_{\rm a}(\beta) = -\beta/2$ the abscissa of the tangent point. If $\varepsilon_{\rm a}(\beta) \in [-\varepsilon_*, \varepsilon_*]$, then the max in Eq. (5.14) is realized in $\varepsilon_{\rm a}(\beta)$. In the other case $\varepsilon_{\rm a}(\beta) < -\varepsilon_*$ (because $\beta \geq 0$) and the max is realized in $-\varepsilon_*$. Therefore:

Proposition 5.2 *The free energy of the REM, $f(\beta) = -\phi(\beta)/\beta$, is equal to:*

$$f(\beta) = \begin{cases} -\frac{1}{4}\beta - \log 2/\beta & \text{if } \beta \leq \beta_{\rm c}\,, \\ -\sqrt{\log 2} & \text{if } \beta > \beta_{\rm c}\,, \end{cases} \qquad \text{where} \qquad \beta_{\rm c} = 2\sqrt{\log 2}\ . \tag{5.15}$$

[11] The task is however more difficult here, because the density of energy levels $\mathcal{N}(\varepsilon, \varepsilon+\delta)$ is a random function whose fluctuations must be controlled.

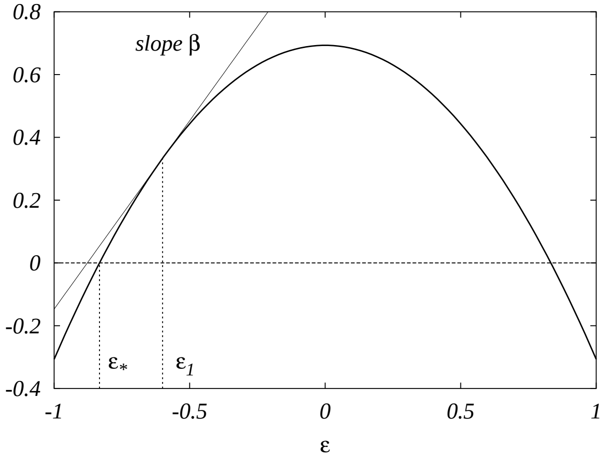


FIG. 5.2. The ‘annealed’ entropy density $s_{\rm a}(\varepsilon)$ of the REM as a function of the energy density ε, see Eq. (5.14). The canonical entropy density $s(\beta)$ is the ordinate of the point with slope $ds_{\rm a}/d\varepsilon = \beta$ when this point lies within the interval $[-\varepsilon_*, \varepsilon_*]$ (this is for instance the case at $\varepsilon = \varepsilon_1$ in the plot), and $s(\beta) = 0$ otherwise. This gives rise to a phase transition at $\beta_{\rm c} = 2\sqrt{\log 2}$. In the ‘annealed’ approximation, the phase transition is not seen, and the $s_{\rm a}(\varepsilon) < 0$ part of the curve is explored, due to the contribution of rare samples to the partition function, see Sec. 5.4.

{fig:rem_sde}

This shows that a phase transition (i.e. a non-analyticity of the free energy density) takes place at the inverse critical temperature $\beta_{\rm c} = 1/T_{\rm c} = 2\sqrt{\log 2}$. It is a second order phase transition in the sense that the derivative of $f(\beta)$ is continuous, but because of the condensation phenomenon which we will discuss in Sec. 5.3 it is often called a ‘random first order’ transition. The other thermodynamic potentials are obtained through the usual formulas, cf. Sec. 2.2. They are plotted in Fig. 5.3.

The two temperature regimes -or ‘phases’- , $\beta \leq$ or $> \beta_{\rm c}$, have distinct qualitative properties which are most easily characterized through the thermodynamic potentials.

- In the high temperature phase $T \geq T_{\rm c}$ (or, equivalently, $\beta \leq \beta_{\rm c}$), the energy and entropy densities are given by: $u(\beta) = -\beta/2$ and $s(\beta) = \log 2 - \beta^2/4$. the configurations which are relevant in Boltzmann’s measure are those with energy $E_i \approx -N\beta/2$. There is an exponentially large number of configurations having such an energy density (the microcanonical entropy density $s(\varepsilon)$ is strictly positive at $\varepsilon = -\beta/2$), and the Boltzmann measure is roughly equidistributed among such configurations.

  In the high temperature limit $T \to \infty$ ($\beta \to 0$) Boltzmann’s measure becomes uniform, and one finds as expected $u(\beta) \to 0$ (because nearly all

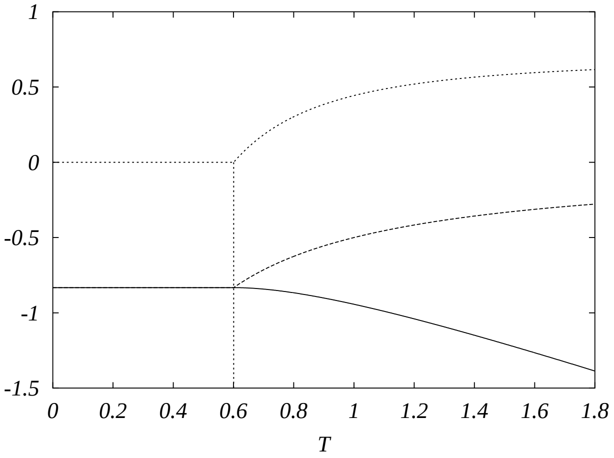


FIG. 5.3. Thermodynamics of the REM: the free energy density (full line), the energy density (dashed line) and the entropy density (dotted line) are plotted versus temperature $T = 1/\beta$. The phase transition takes place at $T_{\rm c} = 1/(2\sqrt{\log 2}) \approx 0.6005612$.

{fig:rem_thermo}

configurations have an energy E_i/N close to 0) and $s \to \log 2$.

- In the low temperature phase $T < T_{\rm c}$ ($\beta > \beta_{\rm c}$), the thermodynamic potentials are constant: $u(\beta) = -\varepsilon_*$ and $s(\beta) = 0$. The relevant configurations are the ones with the lowest energy density, namely with $E_i/N \approx -\varepsilon_*$. The thermodynamics becomes dominated by a relatively small set of configurations, which is not exponentially large in N (the entropy density vanishes).

Exercise 5.2 From the original motivation of the REM as a simple version of a spin glass, one can define a generalization of the REM in the presence of a magnetic field B. The 2^N configurations are divided in $N+1$ groups. Each group is labelled by its 'magnetization' $M \in \{-N, -N+2, \dots, N-2, N\}$, and includes $\binom{N}{(N+M)/2}$ configurations. Their energies $\{E_j\}$ are indipendent Gaussian variables with variance $\sqrt{N/2}$ as in (5.1), and mean $\mathbb{E}\, E_j = -MB$ which depends upon the group j belongs to. Show that there exists a phase transition line $\beta_c(B)$ in the plane β, B such that:

$$\frac{1}{N}\mathbb{E}\, M = \begin{cases} \tanh\left[\beta B\right] & \text{when} \quad \beta \le \beta_c(B)\,, \\ \tanh\left[\beta_c(B) B\right] & \text{when} \quad \beta > \beta_c(B)\,, \end{cases} \tag{5.16}$$

and plot the magnetic susceptibility $\left.\frac{dM}{dB}\right|_B = 0$ versus $T = 1/\beta$.

Exercise 5.3 Consider a generalization of the REM where the pdf of energies, instead of being Gaussian, is $P(E) \propto \exp\left[-C|E|^{\delta}\right]$, where $\delta > 0$. Show that, in order to have extensive thermodynamic potentials, one should scale C as $C = N^{1-\delta}\widehat{C}$ (i.e. the thermodynamic limit $N \to \infty$ should be taken at fixed $\widehat{C}$). Compute the critical temperature and the ground state energy density. What is the qualitative difference between the cases $\delta > 1$ and $\delta < 1$?

{se:rem_cond}

5.3 The condensation phenomenon

In the low temperature phase a smaller-than-exponential set of configurations dominates Boltzmann's measure: we say that the measure **condensates** onto these configurations. This is a scenario that we will encounter again in some other glass phases [12], and it usually leads to many difficulties in finding the relevant configurations. In order to quantify the condensation, one can compute a **participation ratio** $Y_N(\beta)$ defined from Boltzmann's weights (5.2) as:

{eq:rem_Ydef}

$$Y_N(\beta) \equiv \sum_{j=1}^{2^N} p_j^2 \; = \; \left[\sum_j e^{-2\beta E_j}\right]\left[\sum_j e^{-\beta E_j}\right]^{-2} . \tag{5.17}$$

One can think of $1/Y_N(\beta)$ as giving some estimate of the 'effective' number of configurations which contribute to the measure. If the measure were equidistributed on r levels, one would have $Y_N(\beta) = 1/r$.

The participation ratio can be expressed as $Y_N(\beta) = Z_N(2\beta)/Z_N(\beta)^2$, where $Z_N(\beta)$ is the partition function at inverse temperature β. The analysis in the previous Section showed that $Z_N(\beta) \doteq \exp[N(\log 2 + \beta^2/4)]$ with very small fluctuations (see discussion at the end of Sec. 5.2.1) when $\beta < \beta_{\rm c}$, while $Z_N(\beta) \doteq \exp[N\beta\sqrt{\log 2}]$ when $\beta > \beta_{\rm c}$. This indicates that $Y_N(\beta)$ is exponentially small in N for almost all samples in the high temperature phase $\beta < \beta_c$, in agreement with the fact that the measure is not condensed at high temperatures. In the low temperature phase, on the contrary, we shall see that $Y_N(\beta)$ is finite and fluctuates from sample to sample.

The computation of $\mathbb{E}\,Y$ (we drop hereafter its arguments N and β) in the low temperature phase is slightly involved. It requires having a fine control of the energy levels E_i with $E_i/N \approx -\varepsilon_*$. We sketch here the main lines of computation, and leave the details to the reader as an exercise. ⋆ Using the integral representation $1/Z^2 = \int_0^\infty dt\, t \exp(-tZ)$, one gets (with $M = 2^N$):

$$\mathbb{E}\,Y = M\,\mathbb{E}\int_0^\infty dt \;\; t\exp\left[-2\beta E_1\right]\; \exp\left[-t\sum_{i=1}^{M} e^{-\beta E_i}\right] = \tag{5.18}$$

$$= M\int_0^\infty dt \;\; t\; a(t)\;\left[1-b(t)\right]^{M-1} , \tag{5.19}$$

[12] We also call the low temperature phase of the REM a glass phase, by analogy with similar situations that we will encounter later on

where

$$a(t) \equiv \int dP(E) \;\; \exp\left[-2\beta E - te^{-\beta E}\right] \tag{5.20}$$

$$b(t) \equiv \int dP(E) \;\; [1 - \exp(-te^{-\beta E})] \, , \tag{5.21}$$

and $P(E)$ is the Gaussian distribution (5.1). For large N the leading contributions to $\mathbb{E}\, Y$ come from the regions $E = -N\varepsilon_0 + u$ and $t = \theta \exp(-N\beta\varepsilon_0)$, where u and θ are finite as $N \to \infty$, and we defined

$$\varepsilon_0 = \varepsilon_* - \frac{1}{2\varepsilon_*} \log \sqrt{\pi N} \, . \tag{5.22}$$

Notice that ε_0 has been fixed by the condition $2^N P(-N\varepsilon_0) = 1$ and can be thought as a refined estimate for the energy density of the lowest energy configuration. In the region $E = -N\varepsilon_0 + u$, the function $P(E)$ can be substituted by $2^{-N} e^{\beta_c u}$. One gets:

$$\begin{aligned}
a(t) &\approx \frac{1}{M}\, e^{2N\beta\varepsilon_0} \int_{-\infty}^{+\infty} du \;\; e^{\beta_c u - 2\beta u - z e^{-\beta u}} = \frac{e^{2N\beta\varepsilon_0}}{M\beta} \;\; z^{\beta_c/\beta - 2}\, \Gamma(2 - \beta_c/\beta) \, , \\
b(t) &\approx \frac{1}{M} \int_{-\infty}^{+\infty} du \;\; e^{\beta_c u} \, [1 - \exp(-z e^{-\beta u})] = -\frac{1}{M\beta} \;\; z^{\beta_c/\beta}\, \Gamma(-\beta_c/\beta) \, ,
\end{aligned} \tag{5.23}$$

where $\Gamma(x)$ is Euler's Gamma function. Notice that the substitution of $2^{-N} e^{\beta_c u}$ to $P(E)$ is harmless because the resulting integrals (5.23) and (5.23) converge at large u.

At large N, the expression $[1 - b(t)]^{M-1}$ in (5.19) can be approximated by $e^{-Mb(t)}$, and one finally obtains:

$$\begin{aligned}
\mathbb{E}\, Y &= M \int_0^\infty dt \;\; t\, a(t)\, e^{-Mb(t)} \;\; = \\
&= \frac{1}{\beta} \Gamma\left(2 - \frac{\beta_c}{\beta}\right) \int_0^\infty dz \quad z^{\beta_c/\beta - 1} \;\; \exp\left[\frac{1}{\beta}\Gamma\left(-\frac{\beta_c}{\beta}\right) z^{\beta_c/\beta}\right] = \; 1 - \beta_c/\beta \, ,
\end{aligned} \tag{5.24}$$

where we used the approximate expressions (5.23), (5.23) and equalities are understood to hold up to corrections which vanish as $N \to \infty$.

We obtain therefore the following:

{prop:condensation_rem}

Proposition 5.3 *In the REM, the expectation value of the participation ratio is:*

$$\mathbb{E}\, Y = \begin{cases} 0 & \textit{when } T > T_c \, , \\ 1 - T/T_c & \textit{when } T \le T_c \, . \end{cases} \tag{5.25}$$

This gives a quantitative measure of the degree of condensation of Boltzmann's measure: when T decreases, the condensation starts at the phase transition T_c

temperature. At lower temperatures the participation ratio Y increases, meaning that the measure concentrates onto fewer and fewer configurations, until at $T = 0$ only one configuration contributes and $Y = 1$.

With the participation ratio we have a first qualitative and quantitative characterization of the low temperature phase. Actually the energies of the relevant configurations in this phase have many interesting probabilistic properties, to which we shall return in Chapter **??**.

{se:rem_ann}

5.4 A comment on quenched and annealed averages

In the previous section we have found that the self-averaging property holds in the REM, which allowed us to discuss the thermodynamics of a generic sample.

Self-averaging of the thermodynamic potentials is a very frequent property, but in more complicated systems it is often difficult to compute them exactly. We discuss here an approximation which is frequently used in such cases, the so-called **annealed average**. When the free energy density is self averaging, the value of f_N is roughly the same for almost all samples and can be computed as its expectation, called the **quenched average** $f_{N,\mathrm{q}}$:

$$f_{N,\mathrm{q}} = \mathbb{E}\, f_N = -\frac{T}{N}\mathbb{E}\log Z_N \tag{5.26}$$

Since f_N is proportional to the logarithm of the partition function, this average is in general hard to compute and a much easier task is to compute the **annealed average**:

$$f_{N,\mathrm{a}} = -\frac{T}{N}\log(\mathbb{E}\, Z) \tag{5.27}$$

Let us compute it for the REM. Starting from the partition function (8.1), we find:

$$\mathbb{E}\, Z_N = \mathbb{E}\sum_{i=1}^{2^N} e^{-\beta E_i} = 2^N \mathbb{E}\, e^{-\beta E} = 2^N e^{N\beta^2/4}\ , \tag{5.28}$$

yielding $f_{N,\mathrm{a}}(\beta) = -\beta/4 - \log 2/\beta$.

Let us compare this with the correct free energy density found in (5.15). The annealed free energy density $f_{\mathrm{a}}(\beta)$ is always smaller than the correct one, as it should because of Jensen inequality (remember that the logarithm is a concave function). In the REM, and a few other particularly simple problems, it gives the correct result in the high temperature phase $T > T_{\mathrm{c}}$, but fails to identify the phase transition, and predicts wrongly a free energy density in the low temperature phase which is the analytic prolongation of the one at $T > T_{\mathrm{c}}$. In particular, it finds a *negative entropy density* $s_{\mathrm{a}}(\beta) = \log 2 - \beta^2/4$ for $T < T_{\mathrm{c}}$ (see Fig. 5.2).

A negative entropy is impossible in a system with finite configuration space, as can be seen from the definition of entropy. It thus signals a failure, and the reason is easily understood. For a given sample with free energy density f, the partition function behaves as $Z_N = \exp(-\beta N f_N)$. Self-averaging means that f_N

has small sample to sample fluctuations. However these fluctuations exist and are amplified in the partition function because of the factor N in the exponent. This implies that the annealed average of the partition function can be dominated by some very rare samples (those with an anomalously low value of f_N). Consider for instance the low temperature limit. We already know that in almost all samples the configuration with the lowest energy density is found at $E_i \approx -N\varepsilon_*$. However, there exist exceptional samples with one configuration with a smaller minimum $E_i = -N\varepsilon$, $\varepsilon > \varepsilon_*$. These samples are exponentially rare (they occur with a probability $\doteq 2^N e^{-N\varepsilon^2}$), they are irrelevant as far as the quenched average is concerned, but they dominate the annealed average.

Let us add a short semantic note. The terms 'quenched' and 'annealed' originate in the thermal processing of materials used for instance in metallurgy of alloys: a quench corresponds to preparing a sample by bringing it suddenly from high to low temperatures. Then the position of the atoms do not move: a given sample is built from atoms at some random positions (apart from some small vibrations). On the contrary in an annealing process one gradually cools down the alloy, and the various atoms will find favorable positions. In the REM, the energy levels E_i are quenched: for each given sample, they take certain fixed values (like the positions of atoms in a quenched alloy). In the annealed approximation, one treats the configurations i and the energies E_i on the same footing: they adopt a joint probability distribution which is given by Boltzmann's distribution. One says that the E_i variables are thermalized (like the positions of atoms in an annealed alloy).

In general, the annealed average can be used to find a lower bound on the free energy in any system with finite configuration space. Useful results can be obtained for instance using the two simple relations, valid for all temperatures $T = 1/\beta$ and sizes N:

$$f_{N,\mathrm{q}}(T) \geq f_{N,\mathrm{a}}(T) \quad ; \quad \frac{df_{N,\mathrm{q}}(T)}{dT} \leq 0\,. \tag{5.29}$$ {eq:IneqAnnealed}

The first one follows from Jensen as mentioned above, while the second can be obtained from the positivity of canonical entropy, cf. Eq. (2.22), after averaging over the quenched disorder.

In particular, if one is interested in optimization problems (i.e. in the limit of vanishing temperature), the annealed average provides the general bound: {propo:annealed_bound}

Proposition 5.4 *The ground state energy density*

$$u_N(T=0) \equiv \frac{1}{N}\mathbb{E}\left[\min_{\underline{x}\in\mathcal{X}^N} E(\underline{x})\right]\,. \tag{5.30}$$

satisfies the bound $u_N(0) \geq \max_{T\in[0,\infty]} f_{N,\mathrm{a}}(T)$

Proof: Consider the annealed free energy density $f_{N,\mathrm{a}}(T)$ as a function of the temperature $T = 1/\beta$. For any given sample, the free energy is a concave function of T because of the general relation (2.23). It is easy to show that the same property holds for the annealed average. Let T_* be the temperature at which $f_{N,\mathrm{a}}(T)$ achieves its maximum, and $f^*_{N,\mathrm{a}}$ be its maximum value. If $T_* = 0$, then $u_N(0) = f_{N,\mathrm{q}}(0) \geq f^*_{N,\mathrm{a}}$. It $T_* > 0$, then

$$u_N(0) = f_{N,\mathrm{q}}(0) \geq f_{N,\mathrm{q}}(T_*) \geq f_{\mathrm{a}}(T_*) \tag{5.31}$$

where we used the two inequalities (5.29). □

In the REM, this result immediately implies that $u(0) \geq \max_\beta[-\beta/4 - \log 2/\beta] = -\sqrt{\log 2}$, which is actually a tight bound.

5.5 Notes

Notes

The REM was invented by Derrida in 1980 (Derrida, 1980), as an extreme case of some spin glass system. Here we have followed his original solution which makes use of the microcanonical entropy. Many more detailed computations can be found in (Derrida, 1981), including the solution to Exercise 2.

The condensation formula (5.3) appears first in (Gross and Mézard, 1984) as an application of replica computations which we shall discuss in Chapter **??**. The direct estimate of the participation ratio presented here and its fluctuations were developed in (Mézard, Parisi and Virasoro, 1985) and (Derrida and Toulouse, 1985). We shall return to some fascinating (and more detailed) properties of the condensed phase in Chapter **??**.

Exercise 3 shows a phase transition which goes from second order for $\delta > 1$ to first order when $\delta < 1$. Its solution can be found in (Bouchaud and Mézard, 1997).

As a final remark, let us notice that in most of the physics litterature, people don't explicitely write down all the rigorous mathematical steps leading for instance to Eq. (5.13), preferring a smoother presentation which focuses on the basic ideas. In much more complicated models it may be very difficult to fill the corresponding mathematical gaps. The recent book by Talagrand (Talagrand, 2003) adopts a fully rigorous point of view, and it starts with a presentation of the REM which nicely complements the one given here and in Chapter **??**.

6

RANDOM CODE ENSEMBLE

{ch:RandomCodeEnsemble}

As already explained in Sec. 1.6, one of the basic problem of information theory consists in communicating reliably through an unreliable communication channel. Error correcting codes achieve this task by systematically introducing some form of redundancy in the message to be transmitted. One of the major breakthrough accomplished by Claude Shannon was to understand the importance of codes *ensembles*. He realized that it is much easier to construct ensembles of codes which have good properties with high probability, rather than exhibit explicit examples achieving the same performances. In a nutshell: 'stochastic' design is much easier than 'deterministic' design.

At the same time he defined and analyzed the simplest of such ensembles, which has been named thereafter the random code ensemble (or, sometimes, Shannon ensemble). Despite its great simplicity, the random code ensemble has very interesting properties, and in particular it achieves optimal error correcting performances. It provides therefore a prove of the 'direct' part of the channel coding theorem: it is possible to communicate with vanishing error probability as long as the communication rate is smaller than the channel capacity. Furthermore, it is the prototype of a code based on a random construction. In the following Chapters we shall explore several examples of this approach, and the random code ensemble will serve as a reference.

We introduce the idea of code ensembles and define the random code ensemble in 6.1. Some properties of this ensemble are described in Sec. 6.2, while its performances over the BSC are worked out in Sec. 6.3. We generalize these results to a general discrete memoryless channel in Sec. 6.4. Finally, in Sec. 6.5 we show that the random code ensemble is optimal by a simple sphere-packing argument.

6.1 Code ensembles

{se:CodeEnsembles}

An error correcting code is defined as a couple of encoding and decoding maps. The encoding map is applied to the information sequence to get an encoded message which is transmitted through the channel. The decoding map is applied to the (noisy) channel output. For the sake of simplicity, we shall assume throughout this Chapter that the message to be encoded is given as a sequence of M bits and that encoding produces a redundant sequence $N > M$ of bits. The possible codewords (i.e. the 2^M points in the space $\{\mathtt{0},\mathtt{1}\}^N$ which are all the possible outputs of the encoding map) form the **codebook** $\mathfrak{C}_N$. On the other hand, we denote by $\mathcal{Y}$ the output alphabet of the communication channel. We use the notations

$$\underline{x} : \{\mathtt{0},\mathtt{1}\}^M \to \{\mathtt{0},\mathtt{1}\}^N \quad \text{encoding map}\,, \tag{6.1}$$

$$\underline{x}^{\rm d} : \quad \mathcal{Y}^N \to \{\mathtt{0},\mathtt{1}\}^N \quad \text{decoding map}\,. \tag{6.2}$$

Notice that the definition of the decoding map is slightly different from the one given in Sec. 1.6. Here we consider only the difficult part of the decoding procedure, namely how to reconstruct from the received message the codeword which was sent. To complete the decoding as defined in Sec. 1.6, one should get back the original message knowing the codeword, but this is supposed to be an easy task (encoding is assumed to be injective).

The customary recipe for designing a **code ensemble** is the following: (*i*) Define a subset of the space of encoding maps (6.1); (*ii*) Endow this set with a probability distribution; (*iii*) Finally, for each encoding map in the ensemble, define the associated decoding map. In practice, this last step is accomplished by declaring that one among a few general 'decoding strategies' is adopted. We shall introduce a couple of such strategies below.

Our first example is the **random code ensemble (RCE)**. Notice that there exist 2^{N2^M} possible encoding maps of the type (6.1): one must specify N bits for each of the 2^M codewords. In the RCE, any of these encoding maps is picked with uniform probability. The code is therefore constructed as follows. For each of the possible information messages $m \in \{\mathtt{0},\mathtt{1}\}^M$, we obtain the corresponding codeword $\underline{x}^{(m)} = (x_1^{(m)}, x_2^{(m)}, \dots, x_N^{(m)})$ by throwing N times an unbiased coin: the i-th outcome is assigned to the i-th coordinate $x_i^{(m)}$.

Exercise 6.1 Notice that, with this definition the code is not necessarily injective: there could be two information messages $m_1 \neq m_2$ with the same codeword: $\underline{x}^{(m_1)} = \underline{x}^{(m_2)}$. This is an annoying property for an error correcting code: each time that we send either of the messages m_1 or m_2, the receiver will not be able to distinguish between them, even in the absence of noise. Happily enough these unfortunate coincidences occur rarely, i.e. their number is much smaller than the total number of codewords 2^M. What is the expected number of couples m_1, m_2 such that $\underline{x}^{(m_1)} = \underline{x}^{(m_2)}$? What is the probability that all the codewords are distinct?

Let us now turn to the definition of the decoding map. We shall introduce here two among the most important decoding schemes: word MAP (MAP stands here for maximum *a posteriori* probability) and symbol MAP decoding, which can be applied to most codes. In both cases it is useful to introduce the probability distribution $P(\underline{x}|\underline{y})$ for $\underline{x}$ to be the channel input conditional to the received message $\underline{y}$. For a memoryless channel with transition probability $Q(y|x)$, this probability has an explicit expression as a consequence of Bayes rule:

$$P(\underline{x}|\underline{y}) = \frac{1}{Z(\underline{y})} \prod_{i=1}^{N} Q(y_i|x_i)\, P_0(\underline{x})\,. \tag{6.3}$$

Here $Z(\underline{y})$ is fixed by the normalization condition $\sum_{\underline{x}} P(\underline{x}|\underline{y}) = 1$, and $P_0(\underline{x})$ is the *a priori* probability for $\underline{x}$ to be the transmitted message. Throughout this book, we shall assume that the sender choses the codeword to be transmitted with uniform probability. Therefore $P_0(\underline{x}) = 1/2^M$ if $\underline{x} \in \mathfrak{C}_N$ and $P_0(\underline{x}) = 0$ otherwise. In formulas

$$P_0(\underline{x}) = \frac{1}{|\mathfrak{C}_N|}\, \mathbb{I}(\underline{x} \in \mathfrak{C}_N)\,. \tag{6.4}$$

It is also useful to define the marginal distribution $P^{(i)}(x_i|\underline{y})$ of the i-th bit of the transmitted message conditional to the output message. This is obtained from the distribution (6.3) by marginalizing over all the bits x_j with $j \neq i$:

$$P^{(i)}(x_i|\underline{y}) = \sum_{\underline{x}_{\setminus i}} P(\underline{x}|\underline{y})\,, \tag{6.5}$$

where we introduced the shorthand $\underline{x}_{\setminus i} \equiv \{x_j : j \neq i\}$. **Word MAP** decoding outputs the most probable transmitted codeword, i.e. it maximizes[13] the distribution (6.3)

$$\underline{x}^{\rm w}(\underline{y}) = \arg\max_{\underline{x}} P(\underline{x}|\underline{y})\,. \tag{6.6}$$

A strongly related decoding strategy is **maximum-likelihood** decoding. In this case one maximize $Q(\underline{y}|\underline{x})$ over $\underline{x} \in \mathfrak{C}_N$. This coincide with word MAP decoding whenever the *a priori* distribution over the transmitted codeword $P_0(\underline{x})$ is taken to be uniform as in Eq. (6.4).

Symbol (or bit) MAP decoding outputs the sequence of most probable transmitted bits, i.e. it maximizes the marginal distribution (6.5):

$$\underline{x}^{\rm b}(\underline{y}) = \left(\arg\max_{x_1} P^{(1)}(x_1|\underline{y}), \dots, \arg\max_{x_N} P^{(N)}(x_N|\underline{y})\right)\,. \tag{6.7}$$

Exercise 6.2 Consider a code of block-length $N = 3$, and codebook size $|\mathfrak{C}| = 4$, with codewords $\underline{x}^{(1)} = \texttt{001}$, $\underline{x}^{(1)} = \texttt{101}$, $\underline{x}^{(1)} = \texttt{110}$, $\underline{x}^{(1)} = \texttt{111}$. What is the code rate? This code is used to communicate over a binary symmetric channel (BSC) with flip probability $p < 0.5$. Suppose that the channel output is $\underline{y} = \texttt{000}$. Show that the word MAP decoding finds the codeword 001. Now apply symbol MAP decoding to decode the first bit x_1: Show that the result coincides with the one of word MAP decoding only when p is small enough.

It is important to notice that each of the above decoding schemes is optimal with respect a different criterion. Word MAP decoding minimizes the average

[13] We do not specify what to do in case of ties (i.e. if the maximum is degenerate), since this is irrelevant for all the coding problems that we shall consider. The scrupulous reader can chose his own convention in such cases.

block error probability $\mathrm{P_B}$ already defined in Sec. 1.6.2. This is the probability, with respect to the channel distribution $Q(\underline{y}|\underline{x})$, that the decoded codeword $\underline{x}^{\rm d}(\underline{y})$ is different from the transmitted one, averaged over the transmitted codeword:

$$\mathrm{P_B} \equiv \frac{1}{|\mathfrak{C}|}\sum_{\underline{x}\in\mathfrak{C}} \mathbb{P}[\underline{x}^{\rm d}(\underline{y}) \neq \underline{x}]\,. \tag{6.8}$$

Bit MAP decoding minimizes the **bit error probability**, or **bit error rate** (BER) $\mathrm{P_b}$. This is the fraction of incorrect bits, averaged over the transmitted codeword:

$$\mathrm{P_b} \equiv \frac{1}{|\mathfrak{C}|}\sum_{\underline{x}\in\mathfrak{C}} \frac{1}{N}\sum_{i=1}^{N} \mathbb{P}[x_i^{\rm d}(\underline{y}) \neq x_i]\,. \tag{6.9}$$

⋆ We leave to the reader the easy exercise to show that word MAP and symbol MAP decoding are indeed optimal with respect to the above criteria.

{se:GeometryRCE}

6.2 Geometry of the Random Code Ensemble

We begin our study of the random code ensemble by first working out some of its geometrical properties. A code from this ensemble is defined by the codebook, a set $\mathfrak{C}_N$ of 2^M points (all the codewords) in the **Hamming space** $\{\mathtt{0},\mathtt{1}\}^N$. Each one of these points is drawn with uniform probability over the Hamming space. The simplest question one may ask on $\mathfrak{C}_N$ is the following. Suppose you sit on one of the codewords and look around you. How many other codewords are there at a given Hamming distance[14]?

This question is addressed through the **distance enumerator** $\mathcal{N}_{\underline{x}_0}(d)$ with respect to a codeword $\underline{x}_0 \in \mathfrak{C}_N$, defined as the number of codewords in $\underline{x} \in \mathfrak{C}_N$ whose Hamming distance from $\underline{x}_0$ is equal to d: $d(\underline{x},\underline{x}_0) = d$.

We shall now compute the typical properties of the weight enumerator for a random code. The simplest quantity to look at is the average distance enumerator $\mathbb{E}\mathcal{N}_{\underline{x}_0}(d)$, the average being taken over the code ensemble. In general one should further specify *which one* of the codewords is $\underline{x}_0$. Since in the RCE all codewords are drawn independently, and each one with uniform probability over the Hamming space, such a specification is irrelevant and we can in fact fix $\underline{x}_0$ to be the **all zeros codeword**, $\underline{x}_0 = \mathtt{000}\cdots\mathtt{00}$. Therefore we are asking the following question: take 2^M-1 point at random with uniform probability in the Hamming space $\{\mathtt{0},\mathtt{1}\}^N$; what is the average number of points at distance d form the $\mathtt{00}\cdots\mathtt{0}$ corner? This is simply the number of points (2^M-1), times the fraction of the Hamming space 'volume' at a distance d from $\mathtt{000}\cdots\mathtt{0}$ $(2^{-N}\binom{N}{d})$:

$$\mathbb{E}\,\mathcal{N}_{\underline{x}_0}(d) = (2^M-1)\,2^{-N}\binom{N}{d} \quad \doteq \quad 2^{N[R-1+\mathcal{H}_2(\delta)]}\,. \tag{6.10}$$

[14] The **Hamming distance** of two points $\underline{x},\underline{y} \in \{\mathtt{0},\mathtt{1}\}^N$ is the number of coordinates in which they differ.

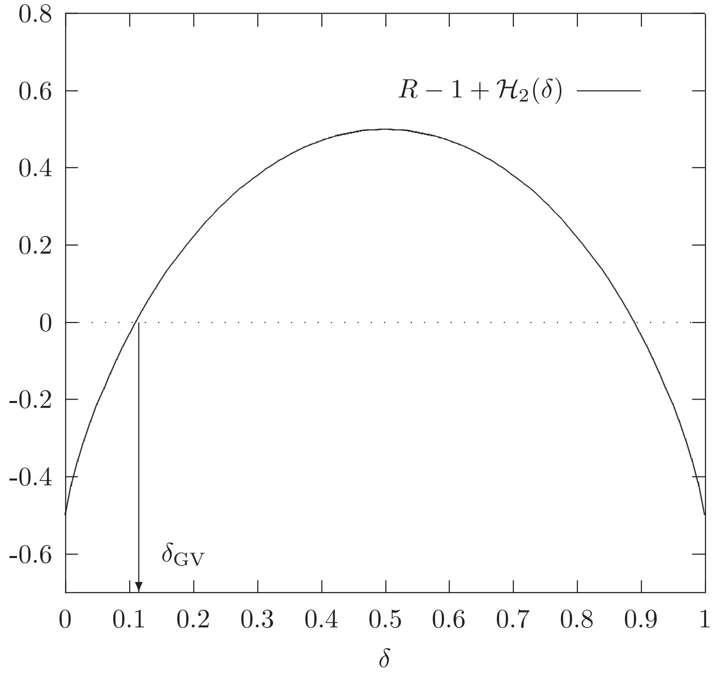
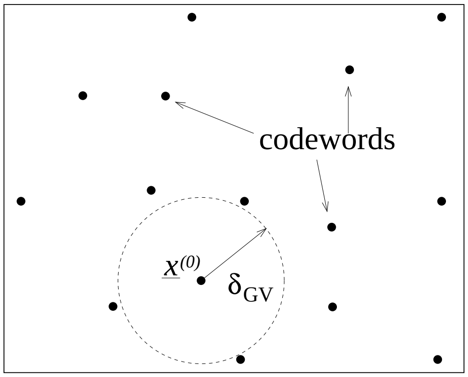


FIG. 6.1. Growth rate of the distance enumerator for the random code ensemble with rate $R = 1/2$ as a function of the Hamming distance $d = N\delta$.

fig:RCEWeightEnumerator}

In the second expression we introduced the fractional distance $\delta \equiv d/N$ and the rate $R \equiv M/N$, and considered the $N \to \infty$ asymptotics with these two quantities kept fixed. In Figure 6.1 we plot the function $R - 1 + \mathcal{H}_2(\delta)$ (which is sometimes called the **growth rate** of the distance enumerator). For δ small enough, $\delta < \delta_{GV}$, the growth rate is negative: the average number of codewords at small distance from $\underline{x}_0$ vanishes exponentially with N. By Markov inequality, the probability of having any codeword at all at such a short distance vanishes as $N \to \infty$. The distance $\delta_{\rm GV}(R)$, called the **Gilbert Varshamov distance**, is the smallest root of $R-1+\mathcal{H}_2(\delta) = 0$. For instance we have $\delta_{\rm GV}(1/2) \approx 0.110278644$.

Above the Gilbert Varshamov distance, $\delta > \delta_{\rm GV}$, the average number of codewords is exponentially large, with the maximum occurring at $\delta = 1/2$: $\mathbb{E}\mathcal{N}_{\underline{x}_0}(N/2) \doteq 2^{NR} = 2^M$. It is easy to show that the weight enumerator $\mathcal{N}_{\underline{x}_0}(d)$ is sharply concentrated around its average in this whole regime $\delta_{\rm GV} < \delta < 1 - \delta_{\rm GV}$, using arguments similar to those developed in Sec.5.2 for the random energy model (REM configurations become codewords in the present context and the role of energy is played by Hamming distance; finally, the Gaussian distribution of the energy levels is replaced here by the binomial distribution). A pictorial interpretation of the above result is shown in Fig. 6.2 (notice that it is often misleading to interpret phenomena occurring in spaces with a large num- ⋆



{fig:RCEHammingSpace} FIG. 6.2. A pictorial view of a typical code from the random code ensemble. The codewords are random points in the Hamming space. If we pick a codeword at random from the code and consider a ball of radius $N\delta$ around it, the ball will not contain any other codeword as long as $\delta < \delta_{\rm GV}(R)$, it will contain exponentially many codewords when $\delta > \delta_{\rm GV}(R)$

ber of dimensions using finite dimensional images: such images must be handled with care!).

Exercise 6.3 The random code ensemble can be easily generalized to other (non binary) alphabets. Consider for instance a q-ary alphabet, i.e. an alphabet with letters $\{\mathtt{0}, \mathtt{1}, \mathtt{2}, \dots, \mathtt{q}-\mathtt{1}\} \equiv \mathcal{A}$. A code $\mathfrak{C}_N$ is constructed by taking 2^M codewords with uniform probability in $\mathcal{A}^N$. We can define the distance between any two codewords $d_q(\underline{x}, \underline{y})$ to be the number of positions in which the sequence $\underline{x}$, $\underline{y}$ differ. The reader will easily show that the average distance enumerator is now

$$\mathbb{E}\,\mathcal{N}_{\underline{x}_0}(d) \doteq 2^{N[R-\log_2 q+\delta\log_2(q-1)+\mathcal{H}_2(\delta)]}\,, \tag{6.11}$$

with $\delta \equiv d/N$ and $R \equiv M/N$. The maximum of the above function is no longer at $\delta = 1/2$. How can we explain this phenomenon in simple terms?

6.3 Communicating over the Binary Symmetric Channel

{se:RCEBSC}

We shall now analyze the performances of the RCE when used for communicating over the binary symmetric channel (BSC) already defined in Fig. 1.4. We start by considering a word MAP (or, equivalently, maximum likelihood) decoder, and we analyze the slightly more complicated symbol MAP decoder afterwards. Finally, we introduce another generalized decoding strategy inspired by the statistical physics analogy.

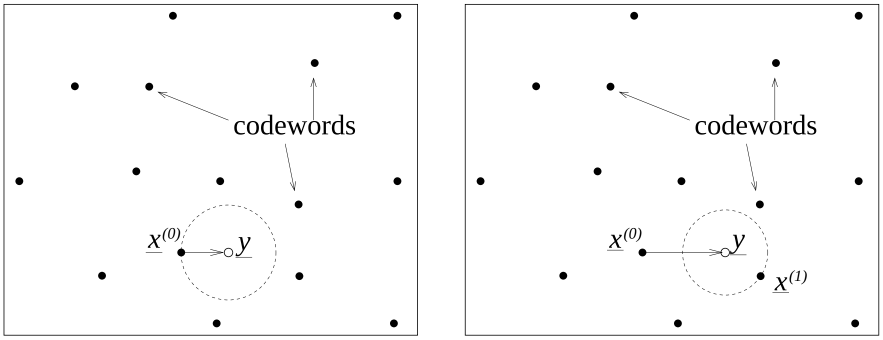


FIG. 6.3. A pictorial view of word MAP decoding for the BSC. A codeword $\underline{x}_0$ is chosen and transmitted through a noisy channel. The channel output is $\underline{y}$. If the distance between $\underline{x}_0$ and $\underline{y}$ is small enough (left frame), the transmitted message can be safely reconstructed by looking for the closest codeword to $\underline{y}$. In the opposite case (right frame), the closest codeword $\underline{x}_1$ does not coincide with the transmitted one. {fig:RCEMaxLikelihood}

6.3.1 *Word MAP decoding*

For a BSC, both the channel input $\underline{x}$ and output $\underline{y}$ are sequences of bits of length N. The probability for the codeword $\underline{x}$ to be the channel input conditional to the output $\underline{y}$, defined in Eqs. (6.3) and (6.4), depends uniquely on the Hamming distance $d(\underline{x},\underline{y})$ between these two vectors. Denoting by p the channel flip probability, we have

$$P(\underline{x}|\underline{y}) = \frac{1}{C}\, p^{d(\underline{x},\underline{y})}(1-p)^{N-d(\underline{x},\underline{y})}\, \mathbb{I}(\underline{x}\in\mathfrak{C}_N)\,, \tag{6.12}$$

C being a normalization constant which depends uniquely upon $\underline{y}$. Without loss of generality, we can assume $p < 1/2$. Therefore word MAP decoding, which prescribes to maximize $P(\underline{x}|\underline{y})$ with respect to $\underline{x}$, outputs the codeword which is the closest to the channel output.

We have obtained a purely geometrical formulation of the original communication problem. A random set of points $\mathfrak{C}_N$ is drawn in the Hamming space $\{0,1\}^N$ and one of them (let us call it $\underline{x}_0$) is chosen for communicating. The noise perturbs this vector yielding a new point $\underline{y}$. Decoding consists in finding the closest to $\underline{y}$ among all the points in $\mathfrak{C}_N$ and fails every time this is not $\underline{x}_0$. The block error probability is simply the probability for such an event to occur. This formulation is illustrated in Fig. 6.3.

This description should make immediately clear that the block error probability vanishes (in the $N\to\infty$ limit) as soon as p is below some finite threshold. In the previous Section we saw that, with high probability, the closest codeword $\underline{x}'\in\mathfrak{C}_N\backslash\underline{x}_0$ to $\underline{x}_0$ lies at a distance $d(\underline{x}',\underline{x}_0)\simeq N\delta_{\rm GV}(R)$. On the other hand $\underline{y}$ is obtained from $\underline{x}_0$ by flipping each bit independently with probability p, therefore $d(\underline{y},\underline{x}_0)\simeq Np$ with high probability. By the triangle inequality $\underline{x}_0$

is surely the closest codeword to $\underline{y}$ (and therefore word MAP decoding is successful) if $d(\underline{x}_0,\underline{y}) < d(\underline{x}_0,\underline{x}')/2$. If $p < \delta_{\rm GV}(R)/2$, this happens with probability approaching one as $N\to\infty$, and therefore the block error probability vanishes.

However the above argument overestimates the effect of noise. Although about $N\delta_{\rm GV}(R)/2$ incorrect bits may cause an unsuccessful decoding, they must occur in the appropriate positions for $\underline{y}$ to be closer to $\underline{x}'$ than to $\underline{x}_0$. If they occur at uniformly random positions (as it happens in the BSC) they will be probably harmless. The difference between the two situations is most significant in large-dimensional spaces, as shown by the analysis provided below.

The distance between $\underline{x}^{(0)}$ and $\underline{y}$ is the sum of N i.i.d. Bernoulli variables of parameter p (each bit gets flipped with probability p). By the central limit theorem, $N(p-\varepsilon) < d(\underline{x}^{(0)},\underline{y}) < N(p+\varepsilon)$ with probability approaching one in the $N\to\infty$ limit, for any $\varepsilon>0$. As for the remaining 2^M-1 codewords, they are completely uncorrelated with $\underline{x}^{(0)}$ and, therefore, with $\underline{y}$: $\{\underline{y},\underline{x}^{(1)},\cdots,\underline{x}^{(2^M-1)}\}$ are 2^M iid random points drawn from the uniform distribution over $\{\mathtt{0},\mathtt{1}\}^N$. The analysis of the previous section shows that with probability approaching one as $N\to\infty$, none of the codewords $\{\underline{x}^{(1)},\cdots,\underline{x}^{(2^M-1)}\}$ lies within a ball of radius $N\delta$ centered on $\underline{y}$, when $\delta<\delta_{\rm GV}(R)$. In the opposite case, if $\delta>\delta_{\rm GV}(R)$, there is an exponential (in N) number of these codewords within a ball of radius $N\delta$.

The performance of the RCE is easily deduced (see Fig. 6.4) : If $p<\delta_{\rm GV}(R)$, the transmitted codeword $\underline{x}^{(0)}$ lies at a shorter distance than all the other ones from the received message $\underline{y}$: decoding is successful. At a larger noise level, $p>\delta_{\rm GV}(R)$ there is an exponential number of codewords closer to $\underline{y}$ than the transmitted one: decoding is unsuccessful. Note that the condition $p<\delta_{\rm GV}(R)$ can be rewritten as $R<C_{\rm BSC}(p)$, where $C_{\rm BSC}(p)=1-\mathcal{H}_2(p)$ is the capacity of a BSC with flip probability p.

6.3.2 *Symbol MAP decoding*

In symbol MAP decoding, the i-th bit is decoded by first computing the marginal $P^{(i)}(x_i|\underline{y})$ and then maximizing it with respect to x_i. Using Eq. (6.12) we get

$$P^{(i)}(x_i|\underline{y}) = \sum_{\underline{x}_{\setminus i}} P(\underline{x}|\underline{y}) = \frac{1}{Z}\sum_{\underline{x}_{\setminus i}} \exp\{-2B\, d(\underline{x},\underline{y})\}\,, \tag{6.13}$$

where we introduced the parameter

$$B \equiv \frac{1}{2}\,\log\left(\frac{1-p}{p}\right)\,, \tag{6.14}$$

and the normalization constant

$$Z \equiv \sum_{\underline{x}\in\mathfrak{C}_N} \exp\{-2B\, d(\underline{x},\underline{y})\}\,. \tag{6.15}$$

Equation (6.13) shows that the marginal distribution $P(x_i|\underline{y})$ gets contributions from all the codewords, not only from the one closest to $\underline{y}$. This makes the

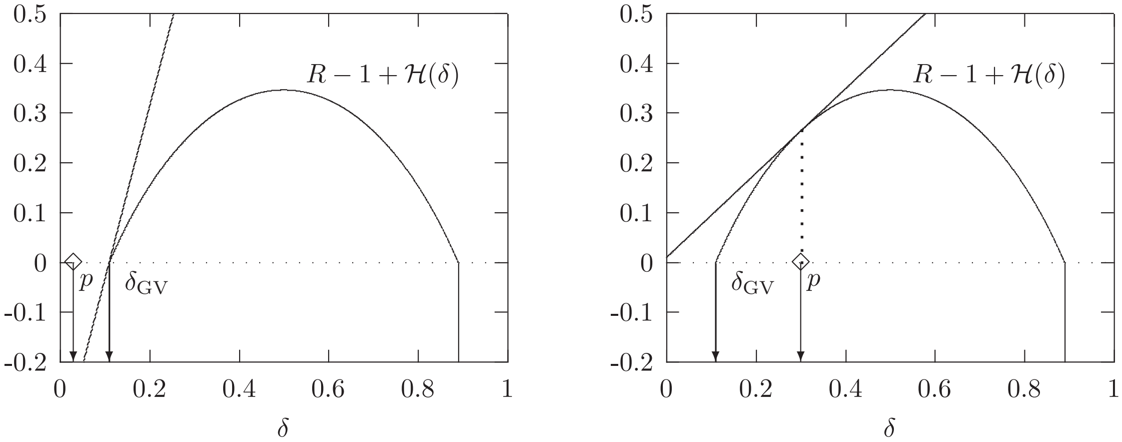


FIG. 6.4. Logarithm of the distance enumerator $\widehat{\mathcal{N}}_{\underline{y}}(d)$ (counting the number of codewords at a distance $d = N\delta$ from the received message) divided by the block--length N. Here the rate is $R = 1/2$. We also show the distance of the transmitted codeword for two different noise levels: $p = 0.03 < \delta_{\rm GV}(1/2) \approx 0.110278644$ (left) and $p = 0.3 > \delta_{\rm GV}(R)$ (right). The tangent lines with slope $2B = \log[(1-p)/p]$ determine which codewords dominate the symbol MAP decoder.

{fig:RCEMicroCanonical}

analysis of symbol MAP decoding slightly more involved than the word MAP decoding case.

Let us start by estimating the normalization constant Z. It is convenient to separate the contribution coming from the transmitted codeword $\underline{x}^{(0)}$ from the one of the *incorrect* codewords $\underline{x}^{(1)}, \dots, \underline{x}^{(2^M-1)}$:

$$Z = e^{-2Bd(\underline{x}^{(0)},\underline{y})} + \sum_{d=0}^{N} \widehat{\mathcal{N}}_{\underline{y}}(d)\, e^{-2Bd} \equiv Z_{\rm corr} + Z_{\rm err}\,, \qquad (6.16)$$

where we denoted by $\widehat{\mathcal{N}}_{\underline{y}}(d)$ the number of incorrect codewords at a distance d from the vector $\underline{y}$. The contribution of $\underline{x}^{(0)}$ in the above expression is easily estimated. By the central limit theorem $d(\underline{x}^{(0)},\underline{y}) \simeq Np$ and therefore $Z_{\rm corr}$ is close to e^{-2NBp} with high probability. More precisely, for any $\varepsilon > 0$, $e^{-N(2Bp+\varepsilon)} \le Z_{\rm corr} \le e^{-N(2Bp-\varepsilon)}$ with probability approaching one in the $N \to \infty$ limit.

As for $Z_{\rm err}$, one proceeds in two steps: first compute the distance enumerator $\widehat{\mathcal{N}}_{\underline{y}}(d)$, and then sum over d. The distance enumerator was already computed in Sec. 6.2. As in the word MAP decoding analysis, the fact that the distances are measured with respect to the channel output $\underline{y}$ and not with respect to a codeword does not change the result, because $\underline{y}$ is independent from the incorrect codewords $\underline{x}^{(1)} \cdots \underline{x}^{(2^M-1)}$. Therefore $\widehat{\mathcal{N}}_{\underline{y}}(d)$ is exponentially large in the interval $\delta_{\rm GV}(R) < \delta \equiv d/N < 1-\delta_{\rm GV}(R)$, while it vanishes with high probability outside the same interval. Moreover, if $\delta_{\rm GV}(R) < \delta < 1-\delta_{\rm GV}(R)$, $\widehat{\mathcal{N}}_{\underline{y}}(d)$ is tightly concentrated around its mean given by Eq. (6.10). The summation over d in Eq. (6.16) can then be evaluated by the saddle point method. This calculation is very similar to the estimation of the free energy of the random energy model, cf. Sec. 5.2. Roughly speaking, we have

$$Z_{\rm err} = \sum_{d=0}^{N} \widehat{\mathcal{N}}_{\underline{y}}(d)\, e^{-2Bd} \simeq N \int_{\delta_{\rm GV}}^{1-\delta_{\rm GV}} e^{N[(R-1)\log 2 + \mathcal{H}(\delta) 2B\delta]}\, d\delta \doteq e^{N\phi_{\rm err}}\, , \tag{6.17}$$

where

$$\phi_{\rm err} \equiv \max_{\delta\in[\delta_{\rm GV}, 1-\delta_{\rm GV}]} [\,(R-1)\log 2 + \mathcal{H}(\delta) - 2B\delta\,]\, . \tag{6.18}$$

⋆ The reader will easily complete the mathematical details of the above derivation along the lines of Sec. 5.2. The bottom-line is that $Z_{\rm err}$ is close to $e^{N\phi_{\rm err}}$ with high probability as $N\to\infty$.

Let us examine the resulting expression (6.18) (see Fig. 6.4). If the maximum is achieved on the interior of $[\delta_{\rm GV}, 1-\delta_{\rm GV}]$, its location δ_* is determined by the stationarity condition $\mathcal{H}'(\delta_*) = 2B$, which implies $\delta_* = p$. In the opposite case, it must be realized at $\delta_* = \delta_{\rm GV}$ (remember that $B > 0$). Evaluating the right hand side of Eq. (6.18) in these two cases, we get

$$\phi_{\rm err} = \begin{cases} -\delta_{\rm GV}(R)\log\left(\frac{1-p}{p}\right) & \text{if } p < \delta_{\rm GV},\\ (R-1)\log 2 - \log(1-p) & \text{otherwise.}\end{cases} \tag{6.19}$$

We can now compare $Z_{\rm corr}$ and $Z_{\rm err}$. At low noise level (small p), the transmitted codeword $\underline{x}^{(0)}$ is close enough to the received one $\underline{y}$ to dominate the sum in Eq. (6.16). At higher noise level, the exponentially more numerous incorrect codewords overcome the term due to $\underline{x}^{(0)}$. More precisely, with high probability we have

$$Z = \begin{cases} Z_{\rm corr}[1+e^{-\Theta(N)}] & \text{if } p < \delta_{\rm GV},\\ Z_{\rm err}[1+e^{-\Theta(N)}] & \text{otherwise,}\end{cases} \tag{6.20}$$

where the $\Theta(N)$ exponents are understood to be positive.

We consider now Eq. (6.13), and once again separate the contribution of the transmitted codeword:

$$P^{(i)}(x_i|\underline{y}) = \frac{1}{Z}\,[Z_{\rm corr}\,\mathbb{I}(x_i = x_i^{(0)}) + Z_{{\rm err},x_i}]\, , \tag{6.21}$$

where we have introduced the quantity

$$Z_{{\rm err},x_i} = \sum_{\underline{z}\in\mathfrak{C}_N\backslash\underline{x}^{(0)}} e^{-2Bd(\underline{z},\underline{y})}\,\mathbb{I}(z_i = x_i)\ . \tag{6.22}$$

Notice that $Z_{{\rm err},x_i} \le Z_{\rm err}$. Together with Eq. (6.20), this implies, if $p < \delta_{\rm GV}(R)$: $P^{(i)}(x_i = x_i^{(0)}|\underline{y}) = 1 - e^{-\Theta(N)}$ and $P^{(i)}(x_i \neq x_i^{(0)}|\underline{y}) = e^{-\Theta(N)}$. In this low p

situation the symbol MAP decoder correctly outputs the transmitted bit $x_i^{(0)}$. It is important to stress that this result holds with probability approaching one as $N \to \infty$. Concretely, there exists bad choices of the code $\mathfrak{C}_N$ and particularly unfavorable channel realizations $\underline{y}$ such that $P^{(i)}(x_i = x_i^{(0)}|\underline{y}) < 1/2$ and the decoder fails. However the probability of such an event (i.e. the bit-error rate $\mathrm{P_b}$) vanishes as $N \to \infty$.

What happens for $p > \delta_{\rm GV}(R)$? Arguing as for the normalization constant Z, it is easy to show that the contribution of incorrect codewords dominates the marginal distribution (6.21). Intuitively, this suggests that the decoder fails. A more detailed computation, sketched below, shows that the bit error rate in the $N \to \infty$ limit is:

$$\mathrm{P_b} = \begin{cases} 0 & \text{if } p < \delta_{\rm GV}(R), \\ p & \text{if } \delta_{\rm GV}(R) < p < 1/2. \end{cases} \tag{6.23}$$

Notice that, above the threshold $\delta_{\rm GV}(R)$, the bit error rate is the same as if the information message were transmitted without coding through the BSC: the code is useless.

A complete calculation of the bit error rate $\mathrm{P_b}$ in the regime $p > \delta_{\rm GV}(R)$ is rather lengthy (at least using the approach developed in this Chapter). We shall provide here an heuristic, albeit essentially correct, justification, and leave the rigorous proof as the exercise below. As already stressed, the contribution $Z_{\rm corr}$ of the transmitted codeword can be safely neglected in Eq. (6.21). Assume, without loss of generality, that $x_i^{(0)} = \mathtt{0}$. The decoder will be successful if $Z_{\rm err,\mathtt{0}} > Z_{\rm err,\mathtt{1}}$ and fail in the opposite case. Two cases must be considered: either $y_i = \mathtt{0}$ (this happens with probability $1-p$), or $y_i = \mathtt{1}$ (probability p). In the first case we have

$$\begin{aligned} Z_{\rm err,\mathtt{0}} &= \sum_{\underline{z} \in \mathfrak{C}_N \backslash \underline{x}^{(0)}} \mathbb{I}(z_i = \mathtt{0})\, e^{-2B d_i(\underline{y},\underline{z})} \\ Z_{\rm err,\mathtt{1}} &= e^{-2B} \sum_{\underline{z} \in \mathfrak{C}_N \backslash \underline{x}^{(0)}} \mathbb{I}(z_i = \mathtt{1})\, e^{-2B d_i(\underline{y},\underline{z})}\,, \end{aligned} \tag{6.24}$$

where we denoted by $d_i(\underline{x},\underline{y})$ the number of of positions j, distinct form i, such that $x_j \neq y_j$. The sums in the above expressions are independent identically distributed random variables. Moreover they are tightly concentrated around their mean. Since $B > 0$, this implies $Z_{\rm err,\mathtt{0}} > Z_{\rm err,\mathtt{1}}$ with high probability. Therefore the decoder is successful in the case $y_i = \mathtt{0}$. Analogously, the decoder fails with high probability if $y_i = \mathtt{1}$, and hence the bit error rate converges to $\mathrm{P_b} = p$ for $p > \delta_{\rm GV}(R)$.

Exercise 6.4 From a rigorous point of view, the weak point of the above argument is the lack of any estimate of the fluctuations of $Z_{\mathrm{err},\mathtt{0}/\mathtt{1}}$. The reader may complete the derivation along the following lines:

- Define $X_{\mathtt{0}} \equiv Z_{\mathrm{err},\mathtt{0}}$ and $X_{\mathtt{1}} \equiv e^{2B}\, Z_{\mathrm{err},\mathtt{1}}$. Prove that $X_{\mathtt{0}}$ and $X_{\mathtt{1}}$ are independent and identically distributed.
- Define the correct distance enumerators $\mathcal{N}_{\mathtt{0}/\mathtt{1}}(d)$ such that a representation of the form $X_{\mathtt{0}/\mathtt{1}} = \sum_d \mathcal{N}_{\mathtt{0}/\mathtt{1}}(d)\, \exp(-2Bd)$ holds.
- Show that a significant fluctuation of $\mathcal{N}_{\mathtt{0}/\mathtt{1}}(d)$ from its average is highly (more than exponentially) improbable (within an appropriate range of d).
- Deduce that a significant fluctuation of $X_{\mathtt{0}/\mathtt{1}}$ is highly improbable (the last two points can be treated along the lines already discussed for the random energy model in Chap. 5).

initeTemperatureDecoder}

6.3.3 *Finite-temperature decoding*

The expression (6.13) for the marginal $P(x_i|\underline{y})$ is strongly reminiscent of a Boltzmann average. This analogy suggests a generalization which interpolates between the two 'classical' MAP decoding strategies discussed so far: **finite-temperature decoding**. We first define this new decoding strategy in the context of the BSC context. Let β be a non-negative number playing the role of an inverse temperature, and $\underline{y} \in \{\mathtt{0},\mathtt{1}\}^N$ the channel output. Define the probability distribution $P_\beta(\underline{x})$ to be given by

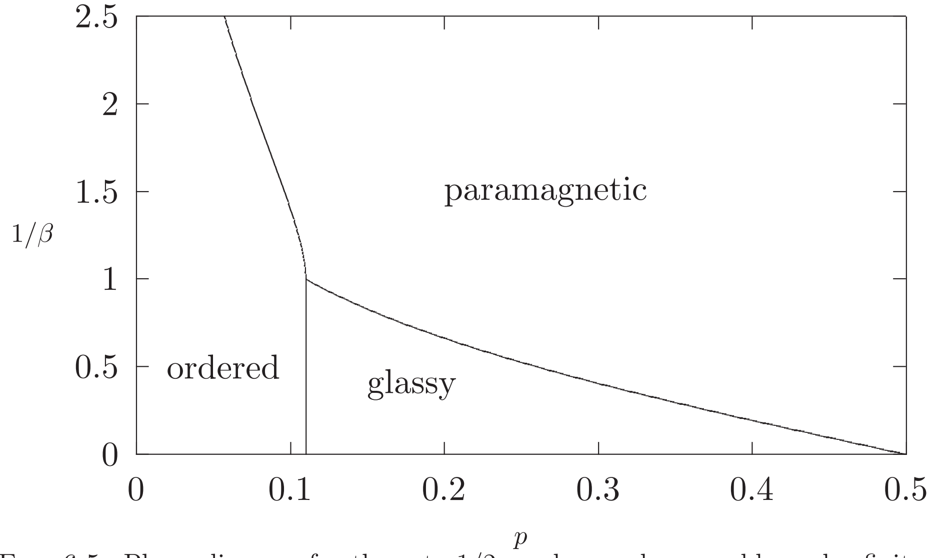
$$P_\beta(\underline{x}) = \frac{1}{Z(\beta)}\, e^{-2\beta B d(\underline{y},\underline{x})}\, \mathbb{I}(x \in \mathfrak{C}_N)\,, \qquad Z(\beta) \equiv \sum_{\underline{x}\in\mathfrak{C}_N} e^{-2\beta B d(\underline{x},\underline{y})}\,, \tag{6.25}$$

where B is always related to the noise level p through Eq. (6.14). This distribution depends upon the channel output $\underline{y}$: for each received message $\underline{y}$, the finite-temperature decoder constructs the appropriate distribution $P_\beta(\underline{x})$. For the sake of simplicity we don't write this dependence explicitly. Let $P^{(i)}_\beta(x_i)$ be the marginal distribution of x_i when $\underline{x}$ is distributed according to $P_\beta(\underline{x})$. The new decoder outputs

$$\underline{x}^\beta = \left(\arg\max_{x_1} P^{(1)}_\beta(x_1)\,, \dots, \arg\max_{x_N} P^{(N)}_\beta(x_N)\right)\,. \tag{6.26}$$

As in the previous Sections, the reader is free to choose her favorite convention in the case of ties (i.e. for those i's such that $P^{(i)}_\beta(\mathtt{0}) = P^{(i)}_\beta(\mathtt{1})$).

Two values of β are particularly interesting: $\beta = 1$ and $\beta = \infty$. If $\beta = 1$ the distribution $P_\beta(\underline{x})$ coincides with the distribution $P(\underline{x}|\underline{y})$ of the channel input conditional to the output, see Eq. (6.12). Therefore, for any $\underline{y}$, symbol MAP decoding coincides with finite-temperature decoding at $\beta = 1$: $\underline{x}_i^{\beta=1} = \underline{x}^{\mathrm{b}}$.



{fig:BSCPhases} FIG. 6.5. Phase diagram for the rate 1/2 random code ensemble under finite temperature decoding. Word MAP and bit MAP decoding correspond (respectively) to $1/\beta = 0$ and $1/\beta = 1$. Notice that the phase boundary of the error-free (ordered) phase is vertical in this interval of temperatures.

If $\beta = \infty$, the distribution (6.25) concentrates over those codewords which are the closest to $\underline{y}$. In particular, if there is a unique closest codeword to $\underline{y}$, finite-temperature decoding at $\beta = \infty$ coincides with word MAP decoding: $\underline{x}^{\beta=\infty} = \underline{x}^{\rm w}$.

The performances of finite-temperature decoding for the RCE at any β, in the large N limit, can be analyzed using the approach developed in the previous Section . The results are summarized in Fig. 6.5 which give the finite-temperature decoding phase diagram. There exist three regimes which are distinct phases with very different behaviors. ⋆

1. A 'completely ordered' phase at low noise ($p < \delta_{\rm GV}(R)$) and low temperature (large enough β). In this regime the decoder works: the probability distribution $P_\beta(\underline{x})$ is dominated by the transmitted codeword $\underline{x}^{(0)}$. More precisely $P_\beta(\underline{x}^{(0)}) = 1 - \exp\{-\Theta(N)\}$. The bit and block error rates vanish as $N \to \infty$.
2. A 'glassy' phase at higher noise ($p > \delta_{\rm GV}(R)$) and low temperature (large enough β). The transmitted codeword has a negligible weight $P_\beta(\underline{x}^{(0)}) = \exp\{-\Theta(N)\}$. The bit error rate is bounded away from 0, and the block error rate converges to 1 as $N \to \infty$. The measure $P_\beta(\underline{x})$ is dominated by the closest codewords to the received message $\underline{y}$ (which are disctinct from the correct one). Its Shannon entropy $H(P_\beta)$ is sub-linear in N. This situation is closely related to the 'measure condensation' phenomenon occurring in

the low-temperature phase of the random energy model.

3. An 'entropy dominated' (paramagnetic) phase at high temperature (small enough β). The bit and block error rates behave as in the glassy phase, and $P_\beta(\underline{x}^{(0)}) = \exp\{-\Theta(N)\}$. However the measure $P_\beta(\underline{x})$ is now dominated by codewords whose distance $d \simeq N\delta_*$ from the received message is larger than the minimal one: $\delta_* = p^\beta/[p^\beta + (1-p)^\beta]$. In particular $\delta_* = p$ if $\beta = 1$, and $\delta_* = 1/2$ if $\beta = 0$. In the first case we recover the result already obtained for symbol MAP decoding. In the second one, $P_{\beta=0}(\underline{x})$ is the uniform distribution over the codewords and the distance from the received message under this distribution is, with high probability, close to $N/2$. In this regime, the Shannon entropy $H(P_\beta)$ is linear in N.

The definition of **finite-temperature decoding** is easily generalized to other channel models. Let $P(\underline{x}|\underline{y})$ be the distribution of the transmitted message conditional to the channel output, given explicitly in Eq. (6.3). For $\beta > 0$, we define the distribution[15]

$$P_\beta(\underline{x}) = \frac{1}{Z(\beta)}\, P(\underline{x}|\underline{y})^\beta\,, \qquad Z(\beta) \equiv \sum_{\underline{x}} P(\underline{x}|\underline{y})^\beta\,. \tag{6.27}$$

Once more, the decoder decision for the i-th bit is taken according to the rule (6.26). The distribution $P_\beta(\underline{x})$ is a 'deformation' of the conditional distribution $P(\underline{x}|\underline{y})$. At large β, more weight is given to highly probable transmitted messages. At small β the most numerous codewords dominate the sum. A little thought $\star$ shows that, as for the BSC, the cases $\beta = 1$ and $\beta = \infty$ correspond, respectively, to symbol MAP and word MAP decoding. The qualitative features of the finite-temperature decoding phase diagram are easily generalized to any memoryless channel. In particular, the three phases described above can be found in such a general context. Decoding is successful in low noise-level, large β phase.

{se:RCEGeneral}

6.4 Error-free communication with random codes

As we have seen, the block error rate $\mathrm{P_B}$ for communicating over a BSC with a random code and word MAP decoding vanishes in the large blocklength limit as long as $R < C_{\mathrm{BSC}}(p)$, with $C_{\mathrm{BSC}}(p) = 1 - \mathcal{H}_2(p)$ the channel capacity. This establishes the 'direct' part of Shannon's channel coding theorem for the BSC case: error-free communication is possible at rates below the channel capacity. This result is in fact much more general. We describe here a proof for general memoryless channels, always based on random codes.

For the sake of simplicity we shall restrict ourselves to memoryless channels with binary input and discrete output. These are defined by a transition probability $Q(y|x)$, $x \in \{\mathtt{0},\mathtt{1}\}$ and $y \in \mathcal{Y}$ with $\mathcal{Y}$ a finite alphabet. In order to handle this case, we must generalize the RCE: each codeword $\underline{x}^{(m)} \in \{\mathtt{0},\mathtt{1}\}^N$,

[15] Notice that the partition function $Z(\beta)$ defined here differs by a multiplicative constant from the one defined in Eq. (6.25) for the BSC.

$m = 0, \dots, 2^M - 1$, is again constructed independently as a sequence of N i.i.d. bits $x_1^{(m)} \cdots x_N^{(m)}$. But $x_i^{(m)}$ is now drawn from an arbitrary distribution $P(x)$, $x \in \{\mathtt{0}, \mathtt{1}\}$ instead of being uniformly distributed. It is important to distinguish $P(x)$ (which is an arbitrary single bit distribution defining the code ensemble and will be chosen at our convenience for optimizing it) and the *a priori* source distribution $P_0(\underline{x})$, cf. Eq. (6.3) (which is a distribution over the codewords and models the information source behavior). As in the previous Sections, we shall assume the source distribution to be uniform over the codewords, cf. Eq. (6.4). On the other hand, the codewords themselves have been constructed using the single-bit distribution $P(x)$.

We shall first analyze the RCE for a generic distribution $P(x)$, under word MAP decoding. The main result is:

{thm:GeneralDirectShannon_lem

Theorem 6.1 *Consider communication over a binary input discrete memoryless channel with transition probability $Q(y|x)$, using a code from the RCE with input bit distribution $P(x)$ and word MAP decoding. If the code rate is smaller than the mutual information $I_{X,Y}$ between two random variables X, Y with joint distribution $P(x)Q(y|x)$, then the block error rate vanishes in the large block-length limit.*

Using this result, one can optimize the ensemble performances over the choice of the distribution $P(\cdot)$. More precisely, we maximixe the maximum achievable rate for error-free communication: $I_{X,Y}$. The corresponding optimal distribution $P^*(\cdot)$ depends upon the channel and can be thought as **adapted** to the channel. Since the channel capacity is in fact defined as the maximum mutual information between channel input and channel output, cf. Eq. (1.37), the RCE with input bit distribution $P^*(\cdot)$ allows to communicate error-free up to channel capacity. The above Theorem implies therefore the 'direct part' of Shannon's theorem **??**.

Proof: Assume that the codeword $\underline{x}^{(0)}$ is transmitted through the channel and the message $\underline{y} \in \mathcal{Y}^N$ is received. The decoder constructs the probability for $\underline{x}$ to be the channel input, conditional to the output $\underline{y}$, see Eq. (6.3). Word MAP decoding consists in minimizing the cost function

$$E(\underline{x}) = -\sum_{i=1}^{N} \log_2 Q(y_i|x_i) \tag{6.28}$$

over the codewords $\underline{x} \in \mathfrak{C}_N$ (note that we use here natural logarithms). Decoding will be successful if and only if the minimum of $E(\underline{x})$ is realized over the transmitted codeword $\underline{x}^{(0)}$. The problem consists therefore in understanding the behavior of the 2^M random variables $E(\underline{x}^{(0)}), \dots, E(\underline{x}^{(2^M-1)})$.

Once more, it is necessary to single out $E(\underline{x}^{(0)})$. This is the sum of N iid random variables $-\log Q(y_i|x_i^{(0)})$, and it is therefore well approximated by its mean

$$\mathbb{E}\, E(\underline{x}^{(0)}) = -N \sum_{x,y} P(x) Q(y|x) \log_2 Q(y|x) = N H_{Y|X}\,. \tag{6.29}$$

In particular $(1-\varepsilon)NH_{Y|X} < E(\underline{x}^{(0)}) < (1+\varepsilon)NH_{Y|X}$ with probability approaching one as $N \to \infty$.

As for the $2^M - 1$ incorrect codewords, the corresponding log-likelihoods $E(\underline{x}^{(1)}), \dots, E(\underline{x}^{(2^M-1)})$ are iid random variables. We can therefore estimate the smallest among them by following the approach developed for the REM and already applied to the RCE on the BSC. In Appendix 6.7, we prove the following large deviation result on the distribution of these variables:

{lem:SH_rce}

Lemma 6.2 *Let $\varepsilon_i = E(\underline{x}^{(i)})/N$. Then $\varepsilon_1, \dots, \varepsilon_{2^M-1}$ are iid random variables and their distribution satisfy a large deviation principle of the form $P(\varepsilon) \doteq 2^{-N\psi(\varepsilon)}$. The rate function is given by:*

$$\psi(\varepsilon) \equiv \min_{\{p_y(\cdot)\} \in \mathfrak{P}_\varepsilon} \left[\sum_y Q(y) D(p_y||P) \right] , \tag{6.30}$$

where the minimum is taken over the set of probability distributions $\{p_y(\cdot),\ y \in \mathcal{Y}\}$ in the subspace $\mathfrak{P}_\varepsilon$ defined by the constraint:

{eq:GeneralChannelRate}

$$\varepsilon = -\sum_{xy} Q(y) p_y(x) \log_2 Q(y|x) \,, \tag{6.31}$$

and we defined $Q(y) \equiv \sum_x Q(y|x)P(x)$.

The solution of the minimization problem formulated in this lemma is obtained through a standard Lagrange multiplier technique:

$$p_y(x) = \frac{1}{z(y)} P(x) Q(y|x)^\gamma \,, \tag{6.32}$$

where the (ε dependent) constants $z(y)$ and γ are chosen in order to verify the normalizations $\forall y :\ \sum_x p_y(x) = 1$, and the constraint (6.31).

The rate function $\psi(\varepsilon)$ is convex with a global minimum (corresponding to $\gamma = 0$) at $\varepsilon_* = -\sum_{x,y} P(x)Q(y)\log_2 Q(y|x)$ where its value is $\psi(\varepsilon_*) = 0$. This implies that, with high probability all incorrect codewords will have costs $E(\underline{x}^{(i)}) = N\varepsilon$ in the range $\varepsilon_{\rm min} \le \varepsilon \le \varepsilon_{\rm max}$, $\varepsilon_{\rm min}$ and $\varepsilon_{\rm max}$ being the two solutions of $\psi(\varepsilon) = R$. Moreover, for any ε inside the interval, the number of codewords with $E(\underline{x}^{(i)}) \simeq N\varepsilon$ is exponentially large (and indeed close to $2^{NR-N\psi(\varepsilon)}$). The incorrect codeword with minimum cost has a cost close to $N\varepsilon_{\rm min}$ (with high probability). Since the correct codeword has cost close to $NH_{Y|X}$, maximum likelihood decoding will find it with high probability if and only if $H_{Y|X} < \varepsilon_{\rm min}$.

The condition $H_{Y|X} < \varepsilon_{\rm min}$ is in fact equivalent to $R < I_{X,Y}$, as it can be shown as follows. A simple calculation shows that the value $\varepsilon = H_{Y|X}$ is obtained using $\gamma = 1$ in Eq. (6.32) and therefore $p_y(x) = P(x)Q(y|x)/Q(y)$. The corresponding value of the rate function is $\psi(\varepsilon = H_{Y|X}) = [H_Y - H_{Y|X}] = I_{Y|X}$. The condition for error free communication, $H_{Y|X} < \varepsilon_{\rm min}$, can thus be rewritten as $R < \psi(H_{Y|X})$, or $R < I_{X,Y}$. □

Example 6.3 Reconsider the BSC with flip probability p. We have

$$E(\underline{x}) = -(N - d(\underline{x},\underline{y}))\,\log(1-p) - d(\underline{x},\underline{y})\,\log p\,. \tag{6.33}$$

Up to a rescaling the cost coincides with the Hamming distance from the received message. If we take $P(\mathtt{0}) = P(\mathtt{1}) = 1/2$, the optimal types are, cf. Eq. (6.32),

$$p_{\mathtt{0}}(\mathtt{1}) = 1 - p_{\mathtt{0}}(\mathtt{0}) = \frac{p^{\gamma}}{(1-p)^{\gamma} + p^{\gamma}}\,, \tag{6.34}$$

and analogously for $p_{\mathtt{1}}(x)$. The corresponding cost is

$$\varepsilon = -(1-\delta)\log(1-p) - \delta\log p\,, \tag{6.35}$$

where we defined $\delta = p^{\gamma}/[(1-p)^{\gamma} + p^{\gamma}]$. The large deviations rate function is given, parametrically, by $\psi(\varepsilon) = \log 2 - \mathcal{H}(\delta)$. The reader will easily recognize the results already obtained in the previous Section.

Exercise 6.5 Consider communication over a discrete memoryless channel with finite input output alphabets $\mathcal{X}$, and $\mathcal{Y}$, and transition probability $Q(y|x)$, $x \in \mathcal{X}$, $y \in \mathcal{Y}$. Check that the above proof remains valid in this context.

6.5 Geometry again: sphere packing

{se:Packing}

Coding has a lot to do with the optimal packing of spheres, which is a general problem of considerable interest in various branches of science. Consider for instance the communication over a BSC with flip probability p. A code of rate R and blocklength N consists of 2^{NR} points $\{\underline{x}^{(1)}\cdots\underline{x}^{(2^{NR})}\}$ in the hypercube $\{\mathtt{0},\mathtt{1}\}^N$. To each possible channel output $\underline{y} \in \{\mathtt{0},\mathtt{1}\}^N$, the decoder associates one of the codewords $\underline{x}^{(i)}$. Therefore we can think of the decoder as realizing a partition of the Hamming space in 2^{NR} decision regions $\mathfrak{D}^{(i)}$, $i \in \{1\dots 2^{NR}\}$, each one associated to a distinct codeword. If we require each decision region $\{\mathfrak{D}^{(i)}\}$ to contain a sphere of radius ρ, the resulting code is *guaranteed* to correct *any* error pattern such that less than ρ bits are flipped. One often defines the **minimum distance** of a code as the smallest distance between any two codewords[16]. If a code has minimal distance d, the Hamming spheres of radius $\rho = \lfloor (d-1)/2 \rfloor$ don't overlap and the code can correct ρ errors, whatever are their positions.

We are thus led to consider the general problem of sphere packing on the hypercube $\{\mathtt{0},\mathtt{1}\}^N$. A (Hamming) sphere of center $\underline{x}_0$ and radius r is defined as the set of points $\underline{x} \in \{\mathtt{0},\mathtt{1}\}^N$, such that $d(\underline{x},\underline{x}_0) \le r$. A packing of spheres

[16] This should not be confused with the minimal distance from one given codewords to all the other ones

of radius r and cardinality $\mathcal{N}_{\mathcal{S}}$ is specified by a set of centers $\underline{x}_1, \dots, \underline{x}_{\mathcal{N}_{\mathcal{S}}} \in \{\mathtt{0},\mathtt{1}\}^N$, such that the spheres of radius r centered in these points are disjoint. Let $\mathcal{N}_N^{\max}(\delta)$ be the maximum cardinality of a packing of spheres of radius $N\delta$ in $\{\mathtt{0},\mathtt{1}\}^N$. We define the corresponding rate as $R_N^{\max}(\delta) \equiv N^{-1}\log_2 \mathcal{N}_N^{\max}(\delta)$ and would like to compute this quantity in the infinite-dimensional limit

$$R^{\max}(\delta) \equiv \lim\sup_{N\to\infty} R_N^{\max}(\delta)\,. \tag{6.36}$$

The problem of determining the function $R^{\max}(\delta)$ is open: only upper and lower bounds are known. Here we shall derive the simplest of these bounds:

{pro:spheres}

Proposition 6.4

{eq:spack_propo}

$$1 - \mathcal{H}_2(2\delta) \le R^{\max}(\delta) \le 1 - \mathcal{H}_2(\delta) \tag{6.37}$$

The lower bound is often called the Gilbert-Varshamov bound, the upper bound is called the Hamming bound.

Proof: Lower bounds can be proved by analyzing good packing strategies. A simple such strategy consists in taking the sphere centers as 2^{NR} random points with uniform probability in the Hamming space. The minimum distance between any couple of points must be larger than $2N\delta$. It can be estimated by defining the distance enumerator $\mathcal{M}_2(d)$ which counts how many couples of points have distance d. It is straightforward to show that, if $d = 2N\delta$ and δ is kept fixed as $N \to \infty$:

$$\mathbb{E}\,\mathcal{M}_2(d) = \binom{2^{NR}}{2} 2^{-N} \binom{N}{d} \doteq 2^{N[2R-1+\mathcal{H}_2(2\delta)]}\,. \tag{6.38}$$

As long as $R < [1-\mathcal{H}_2(2\delta)]/2$, the exponent in the above expression is negative. Therefore, by Markov inequality, the probability of having any couple of centers ar a distance smaller than 2δ is exponentially small in the size. This implies that

$$R^{\max}(\delta) \ge \frac{1}{2}[1 - \mathcal{H}_2(2\delta)]\,. \tag{6.39}$$

A better lower bound can be obtained by a closer examination of the above (random) packing strategy. In Sec. 6.2 we derived the following result. If 2^{NR} points are chosen from the uniform distribution in the Hamming space $\{\mathtt{0},\mathtt{1}\}^N$, and one of them is considered, with high probability its closest neighbour is at a Hamming distance close to $N\delta_{\rm GV}(R)$. In other words, if we draw around each point a sphere of radius δ, with $\delta < \delta_{\rm GV}(R)/2$, and one of the spheres is selected randomly, with high probability it will not intersect any other sphere. This remark suggests the following trick (sometimes called **expurgation** in coding theory). Go through all the spheres one by one and check if it intersects any other one. If the answer is positive, simply eliminate the sphere. This reduces the cardinality of the packing, but only by a fraction approaching 0 as $N \to \infty$: the

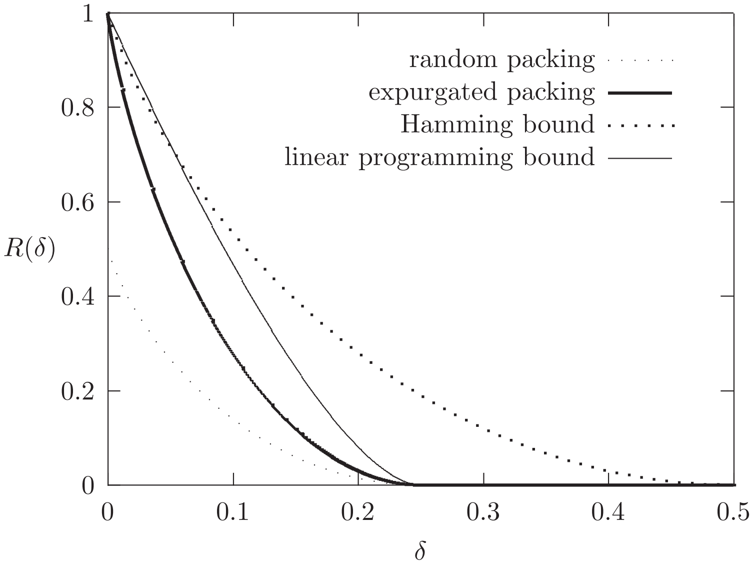


FIG. 6.6. Upper and lower bounds on the maximum packing rate $R^{\max}(\delta)$ of Hamming spheres of radius $N\delta$. Random packing and expurgated random packing provide lower bounds. The Hamming and linear programming bounds are upper bounds.

{fig:HammingSpheres}

packing rate is thus unchanged. As $\delta_{\rm GV}(R)$ is defined by $R = 1 - \mathcal{H}_2(\delta_{\rm GV}(R))$, this proves the lower bound in (6.37).

The upper bound can be obtained from the fact that the total volume occupied by the spheres is not larger than the volume of the hypercube. If we denote by $\Lambda_N(\delta)$ the volume of an N-dimensional Hamming sphere of radius $N\delta$, we get $\mathcal{N}_{\mathcal{S}}\,\Lambda_N(\delta) \leq 2^N$. Since $\Lambda_N(\delta) \doteq 2^{N\mathcal{H}_2(\delta)}$, this implies the upper bound in (6.37). □

Better upper bounds can be derived using more sophisticated mathematical tools. An important result of this type is the so-called *linear programming bound*:

$$R^{\max}(\delta) \leq \mathcal{H}_2(1/2 - \sqrt{2\delta(1-2\delta)})\,, \tag{6.40}$$

whose proof goes beyond our scope. On the other hand, no better lower bound than the Gilbert-Varshamov result is known. It is a widespread conjecture that this bound is indeed tight: in high dimension there is no better way to pack spheres than placing them randomly and expurgating the small fraction of them that are ‘squeezed’. The various bounds are shown in Fig. 6.6.

Exercise 6.6 Derive two simple alternative proofs of the Gilbert-Varshamov bound using the following hints:

1. Given a constant $\overline{\delta}$, let's look at all the 'dangerous' couples of points whose distance is smaller than $2N\overline{\delta}$. For each dangerous couple, we can expurgate one of its two points. The number of points expurgated is smaller or equal than the number of dangerous couples, which can be bounded using $\mathbb{E}\,\mathcal{M}_2(d)$. What is the largest value of $\overline{\delta}$ such that this expurgation procedure does not reduce the rate?
2. Construct a packing $\underline{x}_1 \dots \underline{x}_{\mathcal{N}}$ as follows. The first center $\underline{x}_1$ can be placed anywhere in $\{\mathtt{0},\mathtt{1}\}^N$. The second one is everywhere outside a sphere of radius $2N\delta$ centered in $\underline{x}_0$. In general the i-th center $\underline{x}_i$ can be at any point outside the spheres centered in $\underline{x}_1 \dots \underline{x}_{i-1}$. This procedures stops when the spheres of radius $2N\delta$ cover all the space $\{\mathtt{0},\mathtt{1}\}^N$, giving a packing of cardinality $\mathcal{N}$ equal to the number of steps and radius $N\delta$.

Let us now see the consequences of Proposition 6.4 for coding over the BSC. If the transmitted codeword in $\underline{x}^{(i)}$, the channel output will be (with high probability) at a distance close to Np from $\underline{x}^{(i)}$. Clearly $R \le R^{\max}(p)$ is a necessary and sufficient condition for existence of a code for the BSC which corrects *any* error pattern such that less than Np bits are flipped. Notice that this correction criterion is much stronger than requiring a vanishing (bit or block) error rate. The direct part of Shannon theorem shows the existence of codes with a vanishing (at $N \to \infty$) block error probability for $R < 1 - \mathcal{H}_2(p) = C_{\rm BSC}(p)$. As shown by the linear programming bound in Fig. 6.6 $C_{\rm BSC}(p)$ lies above $R^{\max}(p)$ for large enough p. Therefore, for such values of p, there is a non-vanishing interval of rates $R^{\max}(p) < R < C_{\rm BSC}(p)$ such that one can correct Np errors with high probability but one cannot correct *any* error pattern involving that many bits.

Let us show, for the BSC case, that the condition $R < 1 - \mathcal{H}_2(p)$ is actually a necessary one for achieving zero block error probability (this is nothing but the converse part of Shannon channel coding theorem **??**).

Define $\mathrm{P}_{\rm B}(k)$ the block error probability under the condition that k bits are flipped by the channel. If the codeword $\underline{x}^{(i)}$ is transmitted, the channel output lies on the border of a Hamming sphere of radius k centered in $\underline{x}^{(i)}$: $\partial B_i(k) \equiv \{\underline{z} \;:\; d(\underline{z}, \underline{x}^{(i)}) = k\}$. Therefore

$$\mathrm{P}_{\rm B}(k) = \frac{1}{2^{NR}} \sum_{i=1}^{2^{NR}} \left[1 - \frac{|\partial B_i(k) \cap \mathfrak{D}^{(i)}|}{|\partial B_i(k)|} \right] \ge \tag{6.41}$$

$$\ge 1 - \frac{1}{2^{NR}} \sum_{i=1}^{2^{NR}} \frac{|\mathfrak{D}^{(i)}|}{|\partial B_i(k)|} \,. \tag{6.42}$$

Since $\{\mathfrak{D}^{(i)}\}$ is a partition of $\{\mathtt{0},\mathtt{1}\}^N$, $\sum_i |\mathfrak{D}^{(i)}| = 2^N$. Moreover, for a typical channel realization k is close to Np, and $|\partial B_i(Np)| \doteq 2^{N\mathcal{H}_2(p)}$. We deduce that, for any $\varepsilon > 0$, and large enough N:

$$\mathrm{P_B} \geq 1 - 2^{N(1-R-\mathcal{H}_2(p)+\varepsilon)} \,, \tag{6.43}$$

and thus reliable communication is possible only if $R \leq 1 - \mathcal{H}_2(p)$.

6.6 Other random codes

A major drawback of the random code ensemble is that specifying a particular code (an element of the ensemble) requires $N2^{NR}$ bits. This information has to be stored somewhere when the code is used in practice and the memory requirement is soon beyond the hardware capabilities. A much more compact specification is possible for the **random linear code (RLC)** ensemble. In this case encoding is required to be a linear map, and any such map is equiprobable. Concretely, the code is fully specified by a $N \times M$ binary matrix $\mathbb{G} = \{G_{ij}\}$ (the **generating matrix**) and encoding is left multiplication by $\mathbb{G}$:

$$\underline{x} : \{\mathtt{0},\mathtt{1}\}^M \to \{\mathtt{0},\mathtt{1}\}^N \,, \tag{6.44}$$

$$\underline{z} \mapsto \mathbb{G}\,\underline{z} \,, \tag{6.45}$$

where the multiplication has to be carried modulo 2. Endowing the set of linear codes with uniform probability distribution is equivalent to assuming the entries of $\mathbb{G}$ to be i.i.d. random variables, with $G_{ij} = \mathtt{0}$ or $\mathtt{1}$ with probability 1/2. Notice that only MN bits are required for specifying an element of this ensemble.

Exercise 6.7 Consider a linear code with $N = 4$ and $|\mathfrak{C}| = 8$ defined by

$$\mathfrak{C} = \{(z_1 \oplus z_2,\, z_2 \oplus z_3,\, z_1 \oplus z_3,\, z_1 \oplus z_2 \oplus z_3) \mid\; z_1, z_2, z_3 \in \{\mathtt{0},\mathtt{1}\}\} \,, \tag{6.46}$$

where we denoted by $\oplus$ the sum modulo 2. For instance $(\mathtt{0110}) \in \mathfrak{C}$ because we can take $z_1 = \mathtt{1}$, $z_2 = \mathtt{1}$ and $z_3 = \mathtt{0}$, but $(\mathtt{0010}) \notin \mathfrak{C}$. Compute the distance enumerator for $\underline{x}_0 = (\mathtt{0110})$.

It turns out that the RLC has extremely good performances. As the original Shannon ensemble, it allows to communicate error-free below capacity. Moreover, the rate at which the block error probability $\mathrm{P_B}$ vanishes is faster for the RLC than for the RCE. This justifies the considerable effort devoted so far to the design and analysis of specific ensembles of linear codes satisfying additional computational requirements. We shall discuss some among the best ones in the following Chapters.

6.7 A remark on coding theory and disordered systems

{se:RCEConsiderations}

We would like to stress here the fundamental similarity between the analysis of random code ensembles and the statistical physics of disordered systems. As should be already clear, there are several sources of randomness in coding:

- First of all, the code used is chosen randomly from an ensemble. This was the original idea used by Shannon to prove the channel coding theorem.
- The codeword to be transmitted is chosen with uniform probability from the code. This hypothesis is supported by the source-channel separation theorem.
- The channel output is distributed, once the transmitted codeword is fixed, according to a probabilistic process which accounts for the channel noise.
- Once all the above elements are given, one is left with the decoding problem. As we have seen in Sec. 6.3.3, both classical MAP decoding strategies and finite-temperature decoding can be defined in a unified frame. The decoder constructs a probability distribution $P_\beta(\underline{x})$ over the possible channel inputs, and estimates its single bit marginals $P_\beta^{(i)}(x_i)$. The decision on the i-th bit depends upon the distribution $P_\beta^{(i)}(x_i)$.

The analysis of a particular coding system can therefore be regarded as the analysis of the properties of the distribution $P_\beta(\underline{x})$ when the code, the transmitted codeword and the noise realization are distributed as explained above.

In other words, we are distinguishing two levels of randomness[17]: on the first level we deal with the first three sources of randomness, and on the second level we use the distribution $P_\beta(\underline{x})$. The deep analogy with the theory of disordered system should be clear at this point. The code, channel input, and noise realization play the role of *quenched disorder* (the sample), while the distribution $P_\beta(\underline{x})$ is the analogous of the *Boltzmann distribution*. In both cases the problem consists in studying the properties of a probability distribution which is itself a random object.

Notes

The random code ensemble dates back to Shannon (Shannon, 1948) who used it (somehow implicitely) in his proof of the channel coding thorem. A more explicit (and complete) proof was provided by Gallager in (Gallager, 1965). The reader can find alternative proofs in standard textbooks such as (Cover and Thomas, 1991; Csiszár and Körner, 1981; Gallager, 1968).

The distance enumerator is a feature extensively investigated in coding theory. We refer for instance to (Csiszár and Körner, 1981; Gallager, 1968). A treatment of the random code ensemble in analogy with the random energy model was presented in (Montanari, 2001). More detailed results in the same spirit can be found in (Barg and G. David Forney, 2002). The analogy between coding theory and the statistical physics of disordered systems was put forward by Sourlas (Sourlas, 1989). Finite temperature decoding has been introduced in (Rujan, 1993).

[17] Further refinements of this point of view are possible. One could for instance argue that the code is not likely to be changed at each channel use, while the codeword and noise realization surely change. This remark is important, for instance, when dealing with finite-length effects

A key ingredient of our analysis was the assumption, already mentioned in Sec. 1.6.2, that any codeword is *a priori* equiprobable. The fundamental motivation for such an assumption is the source-channel separation theorem. In simple terms: one does not loose anything in constructing an encoding system in two blocks. First a source code compresses the data produced by the information source and outputs a sequence of i.i.d. unbiased bits. Then a channel code adds redundancy to this sequence in order to contrast the noise on the channel. The theory of error correcting codes (as well as the present Chapter) focuses on the design and analysis of this second block, leaving the first one to source coding. The interested reader may find a proofs of the separation theorem in (Cover and Thomas, 1991; Csiszár and Körner, 1981; Gallager, 1968).

Sphere packing is a classical problem in mathematics, with applications in various branches of science. The book by Conway and Sloane (Conway and Sloane, 1998) provides both a very good introduction and some far reaching results on this problem and its connections, in particular to coding theory. Finding the densest packing of spheres in $\mathbb{R}^n$ is an open problem when $n \geq 4$.

Appendix: Proof of Lemma 6.2

{se:apShannon1}

We estimate (to the leading exponential order in the large N limit) the probability $P_N(\varepsilon)$ for one of the incorrect codewords, $\underline{x}$, to have cost $E(\underline{x}) = N\varepsilon$. The channel output $\underline{y} = (y_1 \cdots y_N)$ is a sequence of N i.i.d. symbols distributed according to

$$Q(y) \equiv \sum_x Q(y|x) P(x)\,, \tag{6.47}$$

and the cost can be rewritten as:

$$E(\underline{x}) \equiv -\sum_{i=1}^{N} \log Q(y_i|x_i) = -N \sum_{x,y} Q(y) \log Q(y|x) \frac{1}{NQ(y)} \sum_{i=1}^{N} \mathbb{I}(x_i = x, y_i = y) \tag{6.48}$$

There are approximatively $NQ(y)$ positions i such that $y_i = y$, for $y \in \mathcal{Y}$. We assume that there are *exactly* $NQ(y)$ such positions, and that $NQ(y)$ is an integer (of course this hypothesis is in general false: it is a routine exercise, left to the reader , to show that it can be avoided with a small technical etour). Furthermore ⋆ we introduce

$$p_y(x) \equiv \frac{1}{NQ(y)} \sum_{i=1}^{N} \mathbb{I}(x_i = x,\ y_i = y)\,. \tag{6.49}$$

Under the above assumptions the function $p_y(x)$ is a probability distribution over $x \in \{\mathtt{0}, \mathtt{1}\}$ for each $y \in \mathcal{Y}$. Looking at the subsequence of positions i such that $y_i = y$, it counts the fraction of the x_i's such that $x_i = x$. In other words

$p_y(\cdot)$ is the type of the subsequence $\{x_i|y_i = y\}$. Because of Eq. (6.48), the cost is written in terms of these types as follows

$$E(\underline{x}) = -N \sum_{xy} Q(y) p_y(x) \log Q(y|x) \,. \tag{6.50}$$

Therefore $E(\underline{x})$ depends upon $\underline{x}$ uniquely through the types $\{p_y(\cdot) \,:\, y \in \mathcal{Y}\}$, and this dependence is linear in $p_y(x)$. Moreover, according to our definition of the RCE, $x_1, \dots, x_N$ are i.i.d. random variables with distribution $P(x)$. The probability $P(\varepsilon)$ that $E(\underline{x})/N = \varepsilon$ can therefore be deduced from the Corollary 4.5. To the leading exponential order, we get

$$P(\varepsilon) \doteq \exp\{-N\psi(\varepsilon) \log 2\} \,, \tag{6.51}$$

$$\psi(\varepsilon) \equiv \min_{p_y(\cdot)} \left[\sum_y Q(y) D(p_y||P) \;\text{ s.t. }\; \varepsilon = -\sum_{xy} Q(y) p_y(x) \log_2 Q(y|x) \right] \tag{6.52}$$

7

NUMBER PARTITIONING

{ch:number_part}

Number partitioning is one of the most basic optimization problems. It is very easy to state: “Given the values of N assets, is there a fair partition of them into two sets?”. Nevertheless it is very difficult to solve: it belongs to the NP-complete category, and the known heuristics are often not very good. It is also a problem with practical applications, for instance in multiprocessor scheduling.

In this Chapter, we shall pay special attention to the partitioning of a list of iid random numbers. It turns out that most heuristics perform poorly on this ensemble of instances. This motivates their use as a benchmark for new algorithms, as well as their analysis. On the other hand, it is relatively easy to characterize analytically the structure of random instances. The main result is that low cost configurations (the ones with a small unbalance between the two sets) can be seen as independent energy levels: the model behaves pretty much like the random energy model of Chap. 5.

7.1 A fair distribution into two groups?

{se:num_part_intro}

An instance of the number partitioning problem is a set of N positive integers $\mathcal{S} = \{a_1, \dots, a_N\}$ indexed by $i \in [N] \equiv \{1, \dots, N\}$. One would like to **partition** the integers in two subsets $\{a_i \,:\, i \in \mathcal{A}\}$ and $\{a_i \,:\, i \in \mathcal{B} \equiv [N]\setminus\mathcal{A}\}$ in such a way as to minimize the discrepancy among the sums of elements in the two subsets. In other words, a configuration is given by $\mathcal{A} \subseteq [N]$, and its cost is defined as

$$E_{\mathcal{A}} = \left| \left(\sum_{i\in\mathcal{A}} a_i \right) - \left(\sum_{i\in\mathcal{B}} a_i \right) \right| . \qquad (7.1)$$

{eq:numparcost}

A **perfect partition** is such that the total number in each subset equilibrate, which means $E_{\mathcal{A}} \leq 1$ (actually $E_{\mathcal{A}} = 0$ if $\sum_i a_i$ is even, or $E_{\mathcal{A}} = 1$ if $\sum_i a_i$ is odd). As usual, one can define several versions of the problem, among which: *i) The decision problem*: Does there exist a perfect partition? *ii) The optimization problem*: Find a partition of lowest cost.

There are also several variants of the problem. So far we have left free the size of $\mathcal{A}$. This is called the **unconstrained** version. On the other hand one can study a constrained version where one imposes that the cardinality difference $|\mathcal{A}| - |\mathcal{B}|$ of the two subsets is fixed to some number D. Here for simplicity we shall mainly keep to the unconstrained case.

{ex:8_warmup}

Exercise 7.1 As a small warm-up, the reader can show that (maybe writing a simple exhaustive search program):

The set $\mathcal{S}_1 = \{10, 13, 23, 6, 20\}$ has a perfect partition.

The set $\mathcal{S}_2 = \{6, 4, 9, 14, 12, 3, 15, 15\}$ has a perfect balanced partition.

In the set $\mathcal{S}_3 = \{93, 58, 141, 209, 179, 48, 225, 228\}$, the lowest possible cost is 5.

In the set $\mathcal{S}_4 = \{2474, 1129, 1388, 3752, 821, 2082, 201, 739\}$, the lowest possible cost is 48.

7.2 Algorithmic issues

7.2.1 *An NP-complete problem*

In order to understand the complexity of the problem, one must first measure its size. This is in turn given by the number of characters required for specifying a particular instance. In number partitioning, this depends crucially on how large the integers can be. Imagine that we restrict ourselves to the case:

{eq:np_sizedef}

$$a_i \in \{1, \dots, 2^M\} \qquad \forall\, i \in \{1, \dots, N\} \tag{7.2}$$

so that each of the N integers can be encoded with M bits. Then the entire instance can be encoded in $N\,M$ bits. It turns out that no known algorithm solves the number partitioning problem in a time upper bounded by a power of $N\,M$. Exhaustive search obviously finds a solution in 2^N operations for unbounded numbers (any M). For bounded numbers there is a simple algorithm running in ⋆ a time of order $N^2\,2^M$ (hint: look at all the integers between 1 and $N\,2^M$ and find recursively which of them can be obtained by summing the k first numbers in the set). In fact, number partitioning belongs to the class of NP-complete problems and is even considered as a fundamental problem in this class.

{se:np_KKalgo}

7.2.2 *A simple heuristic and a complete algorithm*

There is no good algorithm for the number partitioning problem. One of the best heuristics, due to Karmarkar and Karp (KK), uses the following idea. We start from a list $a_1, \dots, a_N$ which coincides with the original set of integers, and reduce it by erasing two elements a_i and a_j in the list, and replacing them by the difference $|a_i - a_j|$, if this difference is non-zero. This substitution means that a decision has been made to place a_i and a_j in two different subsets (but without fixing in which subset they are). One then iterates this procedure as long as the list contains two or more elements. If in the end one finds either an empty list or the list $\{1\}$, then there exists a perfect partitioning. In the opposite case, the remaining integer is the cost of a particular partitioning, but the problem could have better solutions. Of course, there is a lot of flexibility and ingenuity involved in the best choice of the elements a_i and a_j selected at each step. In the KK algorithm one picks up the two largest numbers.

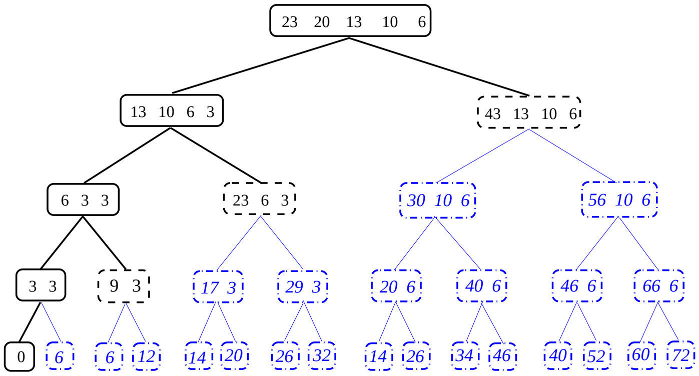


FIG. 7.1. A complete search algorithm: Starting from a list, one erases the two largest numbers a_i and a_j and generate two new lists: the left one contains $|a_i - a_j|$, the right one contains $a_i + a_j$. At the bottom of the tree, every leaf contains the cost of a valid partition. In the search for a perfect partition the tree can be pruned at the dashed leaves because the largest number is bigger than the sum of others: the dash-dotted lists are not generated. The KK heuristics picks up only the left branch. In this example it is successful and finds the unique perfect partition. {fig:numpart_ex}

Example 7.1 Let us see how it works on the first list of exercise 7.1: $\{10, 13, 23, 6, 20\}$. At the first iteration we substitute 23 and 20 by 3, giving the list $\{10, 13, 6, 3\}$. The next step gives $\{3, 6, 3\}$, then $\{3, 3\}$, then $\emptyset$, showing that there exists a perfect partition. The reader can find out how to systematically reconstruct the partition.

A modification due to Korf transforms the KK heuristic into a complete algorithm, which will return the best partitioning (eventually in exponential time). Each time one eliminates two elements a_i and a_j, two new lists are built: a 'left' list which contains $|a_i - a_j|$ (it corresponds to placing a_i and a_j in different groups) and a right one which contains $a_i + a_j$ (it corresponds to placing a_i and a_j in the same group). Iterating in this way one constructs a tree with 2^{N-1} terminal nodes, containing each the cost of a valid partition. Vice-versa, the cost of each possible partition is reported at one of the terminal nodes (notice that each of the 2^N possible partitions $\mathcal{A}$ is equivalent to its complement $[N] \setminus \mathcal{A}$). If one is interested only in the decision: 'is there a perfect partition?', the tree can be pruned as follows. Each time one encounters a list whose largest element is larger than the sum of all other elements plus 1, this list cannot lead to a perfect partition. One can therefore avoid to construct the sub-tree whose root is such a list. Figure 7.1 shows a simple example of application of this algorithm.

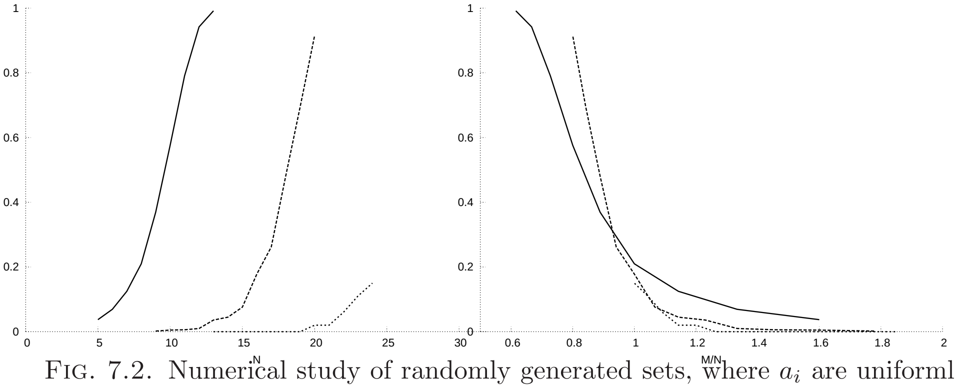


FIG. 7.2. Numerical study of randomly generated sets, where a_i are uniformly distributed in $\{1, \dots 2^M\}$, with $\sum_i a_i$ even. The fraction of samples with a perfect balanced partition is plotted versus N (left plot: from left to right $M = 8, 16, 24$), and versus $\kappa = M/N$ (right plot). In the limit $N \to \infty$ at fixed κ, it turns out that the probability becomes a step function, equal to 1 for $\kappa < 1$, to 0 for $\kappa > 1$ (see also Fig. 7.4).

{fig:nump_stat1}

7.3 Partition of a random list: experiments

{se:numpart_rand_exp}

A natural way to generate random instances of number partitioning is to choose the N input numbers a_i as iid. Here we will be interested in the case where they are uniformly distributed in the set $\{1, \dots, 2^M\}$. As we discussed in Chap. 3, one can use these random instances in order to test typical performances of algorithms, but we will also be interested in natural probabilistic issues, like the distribution of the optimal cost, in the limits where N and M go to ∞.

It is useful to first get an intuitive feeling of the respective roles of N (size of the set) and M (number of digits of each a_i - in base 2). Consider the instances $\mathcal{S}_2, \mathcal{S}_3, \mathcal{S}_4$ of example 1. Each of them contains $N = 8$ random numbers, but they are randomly generated with $M = 4, M = 8, M = 16$ respectively. Clearly, the larger M, the larger is the typical value of the a_i's, and the more difficult it is to distribute them fairly. Consider the costs of all possible partitions: it is reasonable to expect that in about half of the partitions, the most significant bit of the cost is 0. Among these, about one half should have the second significant bit equal to 0. The number of partitions is 2^{N-1}, this qualitative argument can thus be iterated roughly N times. This leads one to expect that, in a random instance with large N, there will be a significant chance of having a perfect partition if $N > M$. On the contrary, for $N < M$, the typical cost of the best partition should behave like 2^{M-N}.

This intuitive reasoning turns out to be essentially correct, as far as the leading exponential behavior in N and M is concerned. Here we first provide some numerical evidence, obtained with the complete algorithm of Sec. 7.2.2 for relatively small systems. In the next Section, we shall validate our conclusions by a sharper analytical argument.

Figure 7.2 shows a numerical estimate of the probability $p_{\rm perf}(N, M)$ that a

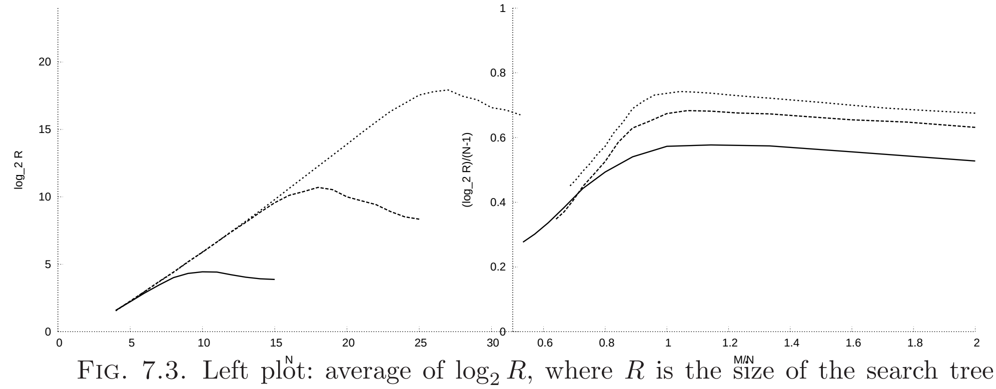


FIG. 7.3. Left plot: average of $\log_2 R$, where R is the size of the search tree. The three curves correspond to $M = 8, 16, 24$ (from left to right). The size grows exponentially with N, and reaches a maximum for $N \approx M$. Right plot: the average of $\log_2 R/(N-1)$ is plotted versus $\kappa = M/N$. {fig:nump_stat1bis}

randomly generated instance has a perfect partition, plotted versus N. This has been obtained by sampling $n_{\rm stat}$ instances of the problem for each considered pair N,M (here $n_{\rm stat} = 10000, 1000, 100$ when $M = 8, 16, 24$ respectively), and solving each instance by simple enumeration. The probability $p_{\rm perf}(N, M)$ was estimated as the fraction of the sampled instances for which a perfect partitioning was found. The standard deviation of such an estimate is $\sqrt{p_{\rm perf}(1 - p_{\rm perf})/n_{\rm stat}}$.

For a fixed value of M, $p_{\rm perf}(N, M)$ crosses over from a value close to 0 at small N to a value close to 1 at large N. The typical values of N where the crossover takes place seem to grow proportionally to M. It is useful to look at the same data from a slightly different perspective by defining the ratio

$$\kappa = \frac{M}{N}\,, \tag{7.3}$$ {eq:np_kappa_def}

and considering $p_{\rm perf}$ as a function of N and κ. The plot of $p_{\rm perf}(\kappa, N)$ versus κ at fixed N shows a very interesting behavior, cf. Fig. 7.2, right frame. A careful analysis of the numerical data [18] indicates that $\lim_{N\to\infty} p_{\rm perf}(\kappa, N) = 1$ for $\kappa < 1$, and $= 0$ for $\kappa > 1$. We stress that the limit $N \to \infty$ is taken with κ kept fixed (and therefore letting $M \to \infty$ proportionally to N). As we shall see in the following, we face here a typical example of a phase transition, in the sense introduced in Chap. 2. The behavior of a generic large instance changes completely when the control parameter κ crosses a critical value $\kappa_c \equiv 1$. For $\kappa < 1$ almost all instances of the problem have a perfect partition (in the large N limit), for $\kappa > 1$ almost none of them can be partitioned perfectly. This phenomenon has important consequences on the computational difficulty of the problem. A good measure of the performance of Korf's complete algorithm is the number R of lists generated in the tree before finding the optimal partition.

[18] In order to perform this analysis, guidance from the random cost model or from the exact results of the next sections is very useful.

In Fig. 7.3 we plot the quantity $\log_2 R$ averaged on the same instances which we had used for the estimation of $p_{\rm perf}$ in Fig. 7.2. The size of the search tree first grows exponentially with N and then reaches a maximum around $N \approx M$. Plotted as a function of κ, one sees a clear peak of $\log_2 R$, somewhere around $\kappa = \kappa_c = 1$: problems close to the critical point are the hardest ones for the algorithm considered. A similar behavior is found with other algorithms, and in fact we will encounter it in many other decision problems like e.g. satisfiability or coloring. When a class of random instances presents a phase transition as a function of one parameter, it is generally the case that the most difficult instances are found in the neighborhood of the phase transition.

{se:numpart_rand_th}

7.4 The random cost model

7.4.1 *Definition of the model*

Consider as before the probability space of random instances constructed by taking the numbers a_j to be iid uniformly distributed in $\{1, \dots, 2^M\}$. For a given partition $\mathcal{A}$, the cost $E_{\mathcal{A}}$ is a random variable with a probability distribution $P_{\mathcal{A}}$. Obviously, the costs of two partitions $\mathcal{A}$ and $\mathcal{A}'$ are correlated random variables. The random cost approximation consists in neglecting these correlations. Such an approximation can be applied to any kind of problem, but it is not always a good one. Remarkably, as discovered by Mertens, the random cost approximation turns out to be 'essentially exact' for the partitioning of iid random numbers.

In order to state precisely the above mentioned approximation, one defines a random cost model (RCM), which is similar to the REM of Chapter 5. A sample is defined by the costs of all the 2^{N-1} 'partitions' (here we identify the two complementary partitions $\mathcal{A}$ and $[N]\backslash\mathcal{A}$). The costs are supposed to be *iid random variables* drawn from the probability distribution $\mathcal{P}$. In order to mimic the random number partitioning problem, $\mathcal{P}$ is taken to be the same as the distribution of the cost of a random partition $\mathcal{A}$ in the original problem:

$$\mathcal{P} \equiv \frac{1}{2^{N-1}} \sum_{\mathcal{A}} \mathcal{P}_{\mathcal{A}} \, . \tag{7.4}$$

Here $\mathcal{P}_{\mathcal{A}}$ is the distribution of the cost of partition $\mathcal{A}$ in the original number partitioning problem.

Let us analyze the behavior of $\mathcal{P}$ for large N. We notice that the cost of a randomly chosen partition in the original problem is given by $|\sum_i \sigma_i a_i|$, where σ_i are iid variables taking value ± 1 with probability $1/2$. For large N, the distribution of $\sum_i \sigma_i a_i$ is characterized by the central limit theorem, and $\mathcal{P}$ is obtained by restricting it to the positive domain. In particular, the cost of a partition will be, with high probability, of order $\sqrt{N\alpha_M^2}$, where

$$\alpha_M^2 \equiv \mathbb{E}\, a^2 = \frac{1}{3}\, 2^{2M} + \frac{1}{2}\, 2^M + \frac{1}{6} \, . \tag{7.5}$$

Moreover, for any $0 \le x_1 < x_2$:

$$\mathcal{P}\left(\frac{E}{\sqrt{N\alpha_M^2}} \in [x_1, x_2]\right) \simeq \sqrt{\frac{2}{\pi}}\, \int_{x_1}^{x_2} e^{-x^2/2}\, dx\,.$$

Finally, the probability of a perfect partition $\mathcal{P}(E=0)$ is just the probability of return to the origin of a random walk with steps $\sigma_i a_i \in \{-2^M, \dots, -1\} \cup \{1, \dots, 2^M\}$. Assuming for simplicity that $\sum_i a_i$ is even, we get:

$$\mathcal{P}(0) \simeq 2\, \frac{1}{\sqrt{2\pi N\alpha_M^2}} \simeq \sqrt{\frac{6}{\pi N}} 2^{-M}\,, \tag{7.6}$$

where $1/\sqrt{2\pi N\alpha_M^2}$ is the density of a normal random variable of mean 0 and variance $N\alpha_M^2$ near the origin, and the extra factor of 2 comes from the fact that the random walk is on even integers only.

As we will show in the next Sections, the RCM is a good approximation for the original number partitioning problem. Some intuition for this property can be found in the exercise below.

{ex:8_remlike}

Exercise 7.2 Consider two random, uniformly distributed, independent partitions $\mathcal{A}$ and $\mathcal{A}'$. Let $\mathcal{P}(E, E')$ denote the joint probability of their energies when the numbers $\{a_i\}$ are iid and uniformly distributed over $\{1, \dots, 2^M\}$. Show that $\mathcal{P}(E, E') = \mathcal{P}(E)\mathcal{P}(E')[1 + o(1)]$ in the large N, M limit, if $E, E' < C\, 2^M$ for some fixed C.

7.4.2 *Phase transition*

We can now proceed with the analysis of the RCM. We shall first determine the phase transition, then study the phase $\kappa > 1$ where typically no perfect partition can be found, and finally study the phase $\kappa < 1$ where an exponential number of perfect partitions exist.

Consider a random instance of the RCM. The probability that *no* perfect partition exist is just the probability that each partition has a strictly positive cost. Since, within the RCM, the 2^{N-1} partitions have iid costs with distribution $\mathcal{P}$, we have:

$$1 - p_{\text{perf}}(\kappa, N) = [1 - \mathcal{P}(0)]^{2^{N-1}}\,. \tag{7.7}$$

In the large N limit with fixed κ, the zero cost probability is given by Eq. (7.6). In particular $\mathcal{P}(0) \ll 1$. Therefore:

$$p_{\text{perf}}(\kappa, N) = 1 - \exp[-2^{N-1}\mathcal{P}(0)] + o(1) = 1 - \exp\left[-\sqrt{\frac{3}{2\pi N}}\, 2^{N\,(1-\kappa)}\right] + o(1)\,. \tag{7.8}$$

{eq:pperf_pred}

This expression predicts a phase transition for the RCM at $\kappa_{\text{c}} = 1$. Notice in fact that $\lim_{N\to\infty} p_{\text{perf}}(\kappa, N) = 1$ if $\kappa < 1$, and $= 0$ if $\kappa > 1$. Moreover, it describes the precise behavior of $p_{\text{perf}}(\kappa, N)$ around the critical point κ_{c} for finite N: Let

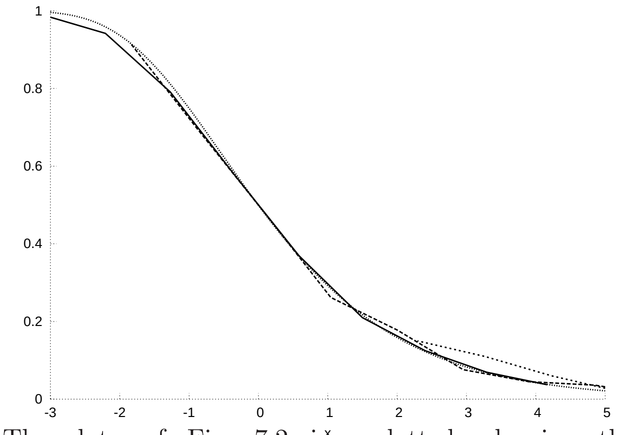


FIG. 7.4. The data of Fig. 7.2 is replotted, showing the (estimated) probability of perfect partition $p_{\text{perf}}(N, M)$ versus the rescaled variable $x = N(\kappa - \kappa_c) + (1/2)\log_2 N$. The agreement with the theoretical prediction (7.9) is very good.

{fig:nump_stat3}

us define the variable $x = N(\kappa - \kappa_c) + (1/2)\log_2 N$. In the limit $N \to \infty$ and $\kappa \to \kappa_c$ at *fixed* x, one finds the crossover behavior:

$$\lim_{\substack{N\to\infty \\ \kappa\to\kappa_c}} p_{\text{perf}}(\kappa, N) = 1 - \exp\left[-\sqrt{\frac{3}{2\pi}}\, 2^{-x}\right] . \tag{7.9}$$

{eq:NPfss}

This is an example of **finite-size scaling** behavior.

In order to compare the above prediction with our numerical results for the original number partitioning problem, we plot in Fig. 7.4 $p_{\text{perf}}(\kappa, N)$ versus the scaling variable x. Here we use the same data presented in Fig. 7.2, just changing the horizontal axis from N to x. The good collapse of the curves for various values of M provides evidence for the claim that the number partitioning problem is indeed asymptotically equivalent to the RCM and presents a phase transition at $\kappa = 1$.

{ex:8_oddeven}

Exercise 7.3 Notice that the argument before assume that $\sum_i a_i$ is even. This is the condition was imposed in the simulation whose results are presented in Fig. 7.4. How should one modify the estimate of $\mathcal{P}(0)$ in Eq. (7.6) when $\sum_i a_i$ is odd? Show that, in this case, if one keeps the definition $x = N(\kappa - \kappa_c) + (1/2)\log_2 N$, the scaling function becomes $1 - \exp\left[-2\sqrt{\frac{3}{2\pi}}\, 2^{-x}\right]$. Run a simulation to check this prediction.

7.4.3 *Study of the two phases*

Let us now study the minimum cost in the phase $\kappa > 1$. The probability that all configurations have a cost larger than E is:

{eq:FiniteNGroundState}

$$\mathbb{P}(\forall \mathcal{A}:\ E_{\mathcal{A}} > E) = \left(1 - \sum_{E'=0}^{E} \mathcal{P}(E')\right)^{2^{N-1}} . \tag{7.10}$$

This probability is non trivial (i.e. different form 0 or 1) if $\sum_{E'=0}^{E} \mathcal{P}(E') = O(2^{-N})$. It is easy to show that this sum can be estimated by substituting[19] $\mathcal{P}(E') \to \mathcal{P}(0)$, which gives the condition $E \sim 1/(\mathcal{P}(0)2^{N-1}) \sim 2^{M-N}\sqrt{N}$ We therefore get, from Eq. (7.10):

$$\lim_{N\to\infty} \mathbb{P}\left(\forall \mathcal{A}:\ E_{\mathcal{A}} > \frac{\varepsilon}{\mathcal{P}(0)2^{N-1}}\right) = e^{-\varepsilon}\ \mathbb{I}(\varepsilon \geq 0)\ . \tag{7.11}$$

In particular the mean of the distribution on the right hand side is equal to 1. This implies that the expectation of the lowest cost in the problem is $\mathbb{E}\, E_{\rm gs} = \sqrt{\frac{2\pi N}{3}}2^{N(\kappa-1)}$. These predictions also fit the numerical results for number partitioning very well.

{ex:8_extreme}

Exercise 7.4 Show that the probability density of the k-th lowest cost configuration, in the rescaled variable ε, is $\varepsilon^{k-1}/(k-1)!\ \exp(-\varepsilon)\ \mathbb{I}(\varepsilon > 0)$. This is a typical case of extreme value statistics for bounded iid variables.

In the phase $\kappa < 1$ we already know that, for almost all samples, there exists at least one configuration with zero cost. It is instructive to count the number of zero cost configurations. Since each configuration has zero cost independently with probability $\mathcal{P}(0)$, the number Z of zero cost configurations is a binomial random variable with distribution

$$P(Z) = \binom{2^{N-1}}{Z}\ \mathcal{P}(0)^Z\ \left[1 - \mathcal{P}(0)\right]^{2^{N-1}-Z} . \tag{7.12}$$

{eq:RCMdegeneracy}

In particular, for large N, Z concentrates around its average value $Z_{\rm av} \doteq 2^{N(1-\kappa)}$. One can define an entropy density of the ground state as:

$$s_{\rm gs} = \frac{1}{N}\log_2 Z\,. \tag{7.13}$$

{eq:rcm_entrop}

The RCM result (7.12) predicts that for $\kappa < 1$ the entropy density is close $1-\kappa$ with high probability. Once again, numerical simulations on the original number partitioning problem confirm this expectation.

[19] As the resulting value of E is much smaller than the scale over which $\mathcal{P}(E)$ varies significantly, cf. Eq. (7.6), the substitution of $\mathcal{P}(0)$ to $\mathcal{P}(E')$ is indeed consistent

{ex:8_integlog}

Exercise 7.5 Using the integral representation of the logarithm:

$$\log_2 x = \int_0^\infty \frac{dt}{t}\left(e^{-t\log 2} - e^{-tx}\right) \ , \tag{7.14}$$

{eq:log_int_rep}

compute $\mathbb{E}\, s_{\rm gs}$ directly. It will be useful to notice that the t integral is dominated by very small values of t, of order $1/(2^{N-1}\mathcal{P}(0))$. Then one easily finds $\mathbb{E}\, s_{\rm gs} \simeq (1/N)\log_2(2^{N-1}\mathcal{P}(0)) \simeq 1-\kappa$.

{se:nump_exact}

7.5 Partition of a random list: rigorous results

A detailed rigorous characterization of the phase diagram in the partitioning of random numbers has been obtained by Borgs, Chayes and Pittel. Basically it confirms the predictions of the RCM. We shall first state some of the exact results known for the balanced partitioning of N numbers. For definiteness we keep as before to the case where a_i are iid uniformly distributed in $\{1,\dots,2^M\}$, and both N and $\sum_{i=1}^N a_i$ are even. The following results hold in the 'thermodynamic limit' $N, M\to\infty$ with fixed $\kappa = M/N$,

{nump_th1}

Theorem 7.2 *There is a phase transition at $\kappa = 1$. For $\kappa < 1$, with high probability, a randomly chosen instance has a perfect balanced partition. For $\kappa > 1$, with high probability, a randomly chosen instance does not have a perfect balanced partition.*

{nump_th2}

Theorem 7.3 *In the phase $\kappa < 1$, the entropy density (7.13) of the number of perfect balanced partitions converges in probability to $s = 1-\kappa$.*

{nump_th3}

Theorem 7.4 *Define $\overline{E} = 2^{N(\kappa-1)}\ \sqrt{2\pi N/3}$ and let $E_1 \le \cdots \le E_k$ be the k lowest costs, with k fixed. Then the k-uple $\left(\varepsilon_1 = E_1/\overline{E},\dots,\varepsilon_k = E_k/\overline{E}\right)$ converges in distribution to $(W_1, W_1{+}W_2,\dots,W_1{+}\dots W_k)$, where W_i are iid random variables with distribution $P(W_i) = e^{-W_i}\ \mathbb{I}(W_i\ge 0)$. In particular the (rescaled) optimal cost distribution converges to $P(\varepsilon_1) = e^{-\varepsilon_1}\ \mathbb{I}(\varepsilon_1 \ge 0)$.*

Note that these results all agree with the RCM. In particular, Theorem 7.4 states that, for fixed k and $N\to\infty$, the lowest k costs are iid variables, as assumed in the RCM. This explains why the random cost approximation is so good.

The proofs of these theorems (and of more detailed results concerning the scaling in the neighborhood of the phase transition point $\kappa = 1$), are all based on the analysis of an integral representation for the number of partitions with a given cost which we will derive below. We shall then outline the general strategy by proving the existence of a phase transition, cf. Theorem 7.2, and we refer the reader to the original literature for the other proofs.

7.5.1 *Integral representation*

For simplicity we keep to the case where $\sum_i a_i$ is even, similar results can be obtained in the case of an odd sum (but the lowest cost is then equal to 1).

Proposition 7.5 *Given a set $\mathcal{S} = \{a_1, \dots, a_N\}$ with $\sum_i a_i$ even, the number Z of partitions with cost $E = 0$ can be written as:*

{eq:np_intrep}

$$Z = 2^{N-1} \int_{-\pi}^{\pi} \frac{dx}{2\pi} \prod_{j=1}^{N} \cos(a_j x) \,. \tag{7.15}$$

Proof: We represent the partition $\mathcal{A}$ by writing $\sigma_i = 1$ if $i \in \mathcal{A}$, and $\sigma_i = -1$ if $i \in \mathcal{B} = [N] \setminus \mathcal{A}$. One can write: $Z = \frac{1}{2} \sum_{\sigma_1,\dots,\sigma_N} \mathbb{I}\left(\sum_{j=1}^{N} \sigma_j a_j = 0\right)$, where the factor $1/2$ comes from the $\mathcal{A} - \mathcal{B}$ symmetry (the same partition is represented by the sequence $\sigma_1, \dots, \sigma_N$ and by $-\sigma_1, \dots, -\sigma_N$). We use the integral representations valid for any integer number a:

$$\mathbb{I}(a = 0) = \int_{-\pi}^{\pi} \frac{dx}{2\pi} e^{ixa} \ , \tag{7.16}$$

which gives:

$$Z = \frac{1}{2} \sum_{\sigma_1,\dots,\sigma_N} \int_{-\pi}^{\pi} \frac{dx}{2\pi} e^{ix(\sum_j \sigma_j a_j)} \ . \tag{7.17}$$

The sum over σ_i's gives the announced integral representation (7.15) □

{ex:8_highercost}

Exercise 7.6 Show that a similar representation holds for the number of partition with cost $E \geq 1$, with an extra factor $2\cos(Ex)$ in the integrand. For the case of balanced partitions, find a similar representation with a two-dimensional integral.

The integrand of (7.15) is typically exponential in N and oscillates wildly. It is thus tempting to compute the integral by the method of steepest descent. This strategy yields correct results in the phase $\kappa \leq 1$, but it is not easy to control it rigorously. Hereafter we use simple first and second moment estimates of the integral which are powerful enough to derive the main features of the phase diagram. Finer control gives more accurate predictions which go beyond this presentation.

7.5.2 *Moment estimates*

We start by evaluating the first two moments of the number of perfect partitions Z.

{propo:np_1}

Proposition 7.6 *In the thermodynamic limit the first moment of Z behaves as:*

$$\mathbb{E}\, Z = 2^{N(1-\kappa)} \sqrt{\frac{3}{2\pi N}} (1 + \Theta(1/N)) \tag{7.18}$$ {eq:np_mom1_res}

Proof: The expectation value is taken over choices of a_i where $\sum_i a_i$ is even. Let us use a modified expectation, denoted by $\mathbb{E}_i$, over all choices of $a_1, \dots, a_N$, without any parity constraint, so that a_i are iid. Clearly $\mathbb{E}_i Z = (1/2)\mathbb{E}\, Z$, because a perfect partition can be obtained only in the case where $\sum_i a_i$ is even, and this happens with probability 1/2.

Because of the independence of the a_i in the expectation $\mathbb{E}_i$, one gets from (7.15)

$$\mathbb{E}\, Z = 2\mathbb{E}_i Z = 2^N \int_{-\pi}^{\pi} \frac{dx}{2\pi}\, \left[\mathbb{E}_i \cos(a_1 x)\right]^N \, . \tag{7.19}$$

{eq:np_m1_1}

The expectation of the cosine is:

{eq:np_m1_2}

$$\mathbb{E}_i \cos(a_1 x) = 2^{-M} \cos\left(\frac{x}{2}(2^M+1)\right) \frac{\sin(2^M x/2)}{\sin(x/2)} \equiv g(x)\, . \tag{7.20}$$

A little thought shows that the integral in (7.19) is dominated in the thermodynamic limit by values of x very near to 0. Precisely we rescale the variable as $x = \hat{x}/(2^M\sqrt{N})$. Then one has $g(x) = 1 - \hat{x}^2/(6N) + \Theta(1/N^2)$. The leading behavior of the integral (7.20) at large N is thus given by:

$$\mathbb{E}\, Z = 2^{N-M} \frac{1}{\sqrt{N}} \int_{-\infty}^{\infty} \frac{d\hat{x}}{2\pi}\, \exp\left(-\frac{\hat{x}^2}{6}\right) = 2^{N-M}\sqrt{\frac{3}{2\pi N}}\, , \tag{7.21}$$

up to corrections of order $1/N$. □

{ex:8_thermod2}

Exercise 7.7 Show that, for E even, with $E \le C2^M$, for a fixed C, the number of partitions with cost E is also given by (7.18) in the thermodynamic limit.

{propo:np_2}

Proposition 7.7 *When $\kappa < 1$, the second moment of Z behaves in the thermodynamic limit as:*

{eq:np_mom2_res}

$$\mathbb{E}\, Z^2 = \left[\mathbb{E}\, Z\right]^2\, \left(1 + \Theta(1/N)\right)\, . \tag{7.22}$$

Proof: We again release the constraint of an even $\sum_i a_i$, so that:

$$\mathbb{E}\, Z^2 = 2^{2N-1} \int_{-\pi}^{\pi} \frac{dx_1}{2\pi} \int_{-\pi}^{\pi} \frac{dx_2}{2\pi} \left[\mathbb{E} \cos(a_1 x_1)\cos(a_1 x_2)\right]^N \tag{7.23}$$

The expectation of the product of the two cosines is:

{eq:np_m2_2}

$$\mathbb{E} \cos(a_1 x_1)\cos(a_1 x_2) = \frac{1}{2}\left[g(x_+) + g(x_-)\right]\, , \tag{7.24}$$

where $x_\pm = x_1 \pm x_2$. In order to find out which regions of the integration domain are important in the thermodynamic limit, one must be careful because the function $g(x)$ is 2π periodic. The double integral is performed in the square $[-\pi, +\pi]^2$. The region of this square where g can be very close to 1 are the 'center' where $x_1, x_2 = \Theta(1/(2^M\sqrt{N}))$, and the four corners, close to $(\pm\pi, \pm\pi)$, obtained from the center by a $\pm 2\pi$ shift in x_+ or in x_-. Because of the periodicity of $g(x)$, the total contribution of the four corners equals that of the center. Therefore one can first compute the integral near the center, using the change of variables

$x_{1(2)} = \hat{x}_{1(2)}/(2^M\sqrt{N})$. The correct value of $\mathbb{E}\,Z^2$ is equal to twice the result of this integral. The remaining part of the computation is straightforward, and gives indeed $\mathbb{E}Z^2 \simeq 2^{2N(1-\kappa)}\frac{3}{2\pi N}$.

In order for this argument to be correct, one must show that the contributions from outside the center are negligible in the thermodynamic limit. The leading correction comes from regions where $x_+ = \Theta(1/(2^M\sqrt{N}))$ while x_- is arbitrary. One can explicitly evaluate the integral in such a region by using the saddle point approximation. The result is of order $\Theta(2^{N(1-\kappa)}/N)$. Therefore, for $\kappa < 1$ the relative contributions from outside the center (or the corners) are exponentially small in N. A careful analysis of the above two-dimensional integral can be found in the literature. □

Propositions 7.6 and 7.7 above have the following important implications. For $\kappa > 1$, $\mathbb{E}\,Z$ is exponentially small in N. Since Z is a non-negative integer, this implies (first moment method) that, in most of the instances Z is indeed 0. For $\kappa < 1$, $\mathbb{E}\,Z$ is exponentially large. Moreover, the normalized random variable $Z/\mathbb{E}\,Z$ has a small second moment, and therefore small fluctuations. The second moment method then shows that Z is positive with high probability. We have thus proved the existence of a phase transition at $\kappa_{\rm c} = 1$, i.e. Theorem 7.2.

Exercise 7.8 Define as usual the partition function at inverse temperature β as $Z(\beta) = \sum_{\mathcal{A}} e^{-\beta E_{\mathcal{A}}}$. Using the integral representation

$$e^{-|U|} = \int_{-\infty}^{\infty} \frac{dx}{\pi}\frac{1}{1+x^2}e^{-ixU} \ , \tag{7.25}$$

and the relation $\sum_{k\in\mathbb{Z}} 1/(1+x^2k^2) = \pi/(x\tanh(\pi/x))$, show that the 'annealed average' for iid numbers a_i is

$$\mathbb{E}_i(Z) = 2^{N(1-\kappa)}\sqrt{\frac{3}{2\pi N}}\frac{1}{\tanh(\beta/2)}(1+\Theta(1/N)) \tag{7.26}$$

Notes

A nice elementary introduction to number partitioning is the paper by Hayes (Hayes, 2002). The NP-complete nature of the problem is a classical result which can be found in textbooks like (Papadimitriou, 1994; Garey and Johnson, 1979). The Karmarkar Karp algorithm was introduced in the technical report (Karmarkar and Karp, 1982). Korf's complete algorithm is in (Korf, 1998).

There has been a lot of work on the partitioning of random iid numbers. In particular, the large κ limit, after a rescaling of the costs by a factor 2^{-M}, deals with the case where a_i are real random numbers, iid on $[0, 1]$. The scaling of the cost of the optimal solution in this case was studied as soon as 1986 by Karmarkar, Karp, Lueker and Odlyzko (Karmarkar, Karp, Lueker and Odlyzko, 1986). On the algorithmic side this is a very challenging problem. As we have seen the optimal partition has a cost $O(\sqrt{N}2^{-N}$; however all known heuristics perform badly on this problem. For instance the KK heuristics finds solution with a cost $O(\exp\left[-.72(\log N)^2\right]$ which is very far from the optimal scaling (Yakir, 1996).

The phase transition was identified numerically by Gent and Walsh (Gent and Walsh, 1998), and studied through statistical physics methods by Ferreira and Fontanari (Ferreira and Fontanari, 1998) and Mertens (Mertens, 1998), who also introduced the random cost model (Mertens, 2000). His review paper (Mertens, 2001) provides a good summary of these works, and helps to solve the Exercises 7.2,7.4, and 7.7. The parity questions discussed in exercise 7.3 have been studied in (Bauke, 2002).

Elaborating on these statistical mechanics treatments, Borgs, Chayes and Pittel were able to establish very detailed rigorous results on the unconstrained problem (Borgs, Chayes and Pittel, 2001), and more recently, together with Mertens, on the constrained case (Borgs, Chayes, Mertens and Pittel, 2003). These result go much beyond the Theorems which we have stated here, and the interested reader is encouraged to study these papers. She will also find there all the technical details needed to fully control the integral representation used in Section 7.5, and the solutions to Exercises 7.5 and 7.6.

8

INTRODUCTION TO REPLICA THEORY

{ch:replicas_intro}

In the past 25 years the replica method has evolved into a rather sophisticated tool for attacking theoretical problems as diverse as spin glasses, protein folding, vortices in superconductors, combinatorial optimization, etc. In this book we adopt a different (but equivalent and, in our view, more concrete) approach: the so-called 'cavity method'. In fact, the reader can skip this Chapter without great harm concerning her understanding of the rest of this book.

It can be nevertheless instructive to have some knowledge of replicas: the replica method is an amazing construction which is incredibly powerful. It is not yet a rigorous method: it involves some formal manipulations, and a few prescriptions which may appear arbitrary. Nevertheless these prescriptions are fully specified, and the method can be regarded as an 'essentially automatic' analytic tool. Moreover, several of its most important predictions have been confirmed rigorously through alternative approaches. Among its most interesting aspects is the role played by 'overlaps' among replicas. It turns out that the subtle probabilistic structure of the systems under study are often most easily phrased in terms of such variables.

Here we shall take advantage of the simplicity of the Random Energy Model (REM) defined in Chapter 5 to introduce replicas. This is the topic of Sec. 8.1. A more complicated spin model is introduced and discussed in Sec. 8.2. In Sec. 8.3 we discuss the relationship between the simplest replica symmetry breaking scheme and the extreme value statistics. Finally, in the Appendix we briefly explain how to perform a local stability analysis in replica space. This is one of the most commonly used consistency checks in the replica method.

8.1 Replica solution of the Random Energy Model

{se:ReplicaREM}

As we saw in Sec. 5.1, a sample (or instance) of the REM is given by the values of 2^N energy levels E_j, with $j \in \{1, \dots, 2^N\}$. The energy levels are iid Gaussian random variables with mean 0 and variance $N/2$. A configuration of the REM is just the index j of one energy level. The partition function for a sample with energy levels $\{E_1 \dots, E_{2^N}\}$ is

$$Z = \sum_{j=1}^{2^N} \exp\left(-\beta E_j\right) \ , \tag{8.1}$$

{eq:rem_zdef}

and is itself a random variable (in the physicist language 'Z fluctuates from sample to sample'). In Chapter 5 we argued that intensive thermodynamic potentials are self-averaging, meaning that their distribution is sharply concentrated around the mean value in the large-N limit. Among these quantities, a prominent role is played by the free energy density $f = -1/(\beta N) \log Z$. Other potentials can in fact be computed from derivatives of the free energy. Unlike these quantities, the partition function has a broad distribution even for large sizes. In particular, its average is dominated (in the low temperature phase) by extremely rare samples. In order to have a fair description of the system, one has to compute the average of the log-partition function, $\mathbb{E} \log Z$, which, up to a constant, yields the average free energy density.

It turns out that computing integer moments of the partition function $\mathbb{E}\, Z^n$, with $n \in \mathbb{N}$, is much easier than computing the average log-partition function $\mathbb{E} \log Z$. This happens because Z is the sum of a large number of 'simple' terms.

If, on the other hand, we were able to compute $\mathbb{E}\, Z^n$ for any *real* n (or, at least, for n small enough), the average log-partition function could be determined using, for instance, the relation

{eq:replicalimit}

$$\mathbb{E} \log Z = \lim_{n \to 0} \frac{1}{n} \log(\mathbb{E}\, Z^n) \ . \tag{8.2}$$

The idea is to carry out the calculation of $\mathbb{E}\, Z^n$ 'as if' n were an integer. At a certain point (after having obtained a manageable enough expression), we shall 'remember' that n has indeed to be a real number and take this into account. As we shall see this whole line of approach has some flavor of an analytic continuation but in fact it has quite a few extra grains of salt...

The first step consists in noticing that Z^n can be written as an n-fold sum

{eq:Zngen}

$$Z^n = \sum_{i_1 \dots i_n = 1}^{2^N} \exp\left(-\beta E_{i_1} - \cdots - \beta E_{i_n}\right) \ . \tag{8.3}$$

This expression can be interpreted as the partition function of a new system. A configuration of this system is given by the n-uple $(i_1, \dots, i_n)$, with $i_a \in \{1, \dots, 2^N\}$, and its energy is $E_{i_1 \dots i_n} = E_{i_1} + \cdots + E_{i_n}$. In other words, the new system is formed of n statistically independent (in the physicist language: non-interacting) copies of the original one. We shall refer to such copies as **replicas**.

In order to evaluate the average of Eq. (8.3), it is useful to first rewrite it as:

$$Z^n = \sum_{i_1 \dots i_n = 1}^{2^N} \prod_{j=1}^{2^N} \exp\left[-\beta E_j \left(\sum_{a=1}^{n} \mathbb{I}(i_a = j)\right)\right] \ . \tag{8.4}$$

Exploiting the linearity of expectation, the independence of the E_j's, and their Gaussian distribution, one easily gets:

{eq:AverageReplicated}

$$\mathbb{E}\, Z^n = \sum_{i_1 \dots i_n = 1}^{2^N} \exp\left(\frac{\beta^2 N}{4} \sum_{a,b=1}^{n} \mathbb{I}(i_a = i_b)\right) \ . \tag{8.5}$$

$\mathbb{E}\, Z^n$ can also be interpreted as the partition function of a new 'replicated' system. As before, a configuration is given by the n-uple $(i_1, \dots, i_n)$, but now its energy is $E_{i_1 \dots i_n} = -N\beta/4 \sum_{a,b=1}^{n} \mathbb{I}(i_a = i_b)$.

This replicated system has several interesting properties. First of all, it is no longer a disordered system: the energy is a deterministic function of the configuration. Second, replicas do interact: the energy function cannot be written as a sum of single replica terms. The interaction amounts to an attraction between different replicas. In particular, the lowest energy configurations are obtained by setting $i_1 = \cdots = i_n$. Their energy is $E_{i_1 \dots i_n} = -N\beta n^2/4$. Third: the energy depends itself upon the temperature, although in a very simple fashion. Its effect will be stronger at low temperature.

The origin of the interaction among replicas is easily understood. For one given sample of the original problem, the Boltzmann distribution concentrates at low temperature ($\beta \gg 1$) on the lowest energy levels: all the replicas will tend to be in the same configuration with large probability. When averaging over the distribution of samples, we do not see any longer which configuration $i \in \{1 \dots 2^N\}$ has the lowest energy, but we still see that the replicas prefer to stay in the same state. There is no mystery in these remarks. The elements of the n-uple $(i_1 \dots i_n)$ are independent *conditional* on the sample, that is on realization of the energy levels E_j, $j \in \{1 \dots 2^N\}$. If we do not condition on the realization, $(i_1 \dots i_n)$ become dependent.

Given the replicas configurations $(i_1 \dots i_n)$, it is convenient to introduce the $n \times n$ matrix $Q_{ab} = \mathbb{I}(i_a = i_b)$, with elements in $\{0, 1\}$. We shall refer to this matrix as the **overlap matrix**. The summand in Eq. (8.5) depends upon the configuration $(i_1 \dots i_n)$ only through the overlap matrix. We can therefore rewrite the sum over configurations as:

$$\mathbb{E}\, Z^n = \sum_{Q} \mathcal{N}_N(Q) \exp\left(\frac{N\beta^2}{4} \sum_{a,b=1}^{n} Q_{ab} \right) . \tag{8.6}$$

Here $\mathcal{N}_N(Q)$ denotes the number of configurations $(i_1 \dots i_n)$ whose overlap matrix is $Q = \{Q_{ab}\}$, and the sum $\sum_Q$ runs over the symmetric $\{0, 1\}$ matrices with ones on the diagonal. The number of such matrices is $2^{n(n-1)/2}$, while the number of configurations of the replicated system is 2^{Nn}. It is therefore natural to guess that the number of configurations with a given overlap matrix satisfies a large deviation principle of the form $\mathcal{N}_N(Q) \doteq \exp(Ns(Q))$:

Exercise 8.1 Show that the overlap matrix always has the following form: There exists a partition $\mathcal{G}_1$, $\mathcal{G}_2$, $\dots$, $\mathcal{G}_{n_{\rm g}}$ of the n replicas (this means that $\mathcal{G}_1 \cup \mathcal{G}_2 \cup \cdots \cup \mathcal{G}_{n_{\rm g}} = \{1 \dots n\}$ and $\mathcal{G}_i \cap \mathcal{G}_j = \emptyset$) into $n_{\rm g}$ groups such that $Q_{ab} = 1$ if a and b belong to the same group, and $Q_{ab} = 0$ otherwise. Prove that $\mathcal{N}_N(Q)$ satisfies the large deviation principle described above, with $s(Q) = n_{\rm g} \log 2$.

Using this form of $\mathcal{N}_N(Q)$, the replicated partition function can be written as:

$$\mathbb{E}\, Z^n \doteq \sum_Q \exp\left(Ng(Q)\right) \qquad ; \qquad g(Q) \equiv \frac{\beta^2}{4} \sum_{a,b=1}^{n} Q_{ab} + s(Q)\,. \tag{8.7}$$

{eq:ReplicatedPartitionFuncti

The strategy of the replica method is to estimate the above sum using the saddle point method[20]. The 'extrapolation' to non-integer values of n is discussed afterward. Let us notice that this program is completely analogous to the treatment of the Curie-Weiss model in Sec. 2.5.2 (see also Sec. 4.3 for related background), with the extra step of extrapolating to non-integer n.

8.1.1 *Replica symmetric saddle point*

{se:ReplicaSymmetricREM}

The function $g(Q)$ is symmetric under permutation of replicas: Let $\pi \in S_n$ be a permutation of n objects, and denote by Q^π the matrix with elements $Q^\pi_{ab} = Q_{\pi(a)\pi(b)}$. Then $g(Q^\pi) = g(Q)$. This is a simple consequence of the fact that the n replicas were equivalent from the beginning. This symmetry is called the **replica symmetry**, and is a completely generic feature of the replica method.

When the dominant saddle point possesses this symmetry (i.e. when $Q^\pi = Q$ for any permutation π) one says that the system is **replica symmetric (RS)**. In the opposite case replica symmetry is spontaneously broken in the large N limit, in the same sense as we discussed in chapter 2 (see Sec. 2.5.2).

In view of this permutation symmetry, the simplest idea is to seek a replica symmetric saddle point. If Q is invariant under permutation, then necessarily $Q_{aa} = 1$, and $Q_{ab} = q_0$ for any couple $a \neq b$. We are left with two possibilities:

- The matrix $Q_{\rm RS,0}$ is defined by $q_0 = 0$. In this case $\mathcal{N}_N(Q_{\rm RS,0}) = 2^N(2^N - 1)\dots(2^N - n + 1)$, which yields $s(Q_{\rm RS,0}) = n \log 2$ and $g(Q_{\rm RS,0}) = n\left(\beta^2/4 + \log 2\right)$.
- The matrix $Q_{\rm RS,1}$ is defined by $q_0 = 1$. This means that $i_1 = \cdots = i_n$. There are of course $\mathcal{N}_N(Q_{\rm RS,1}) = 2^N$ choices of the n-uple $(i_1 \dots i_n)$ compatible with this constraint, which yields $s(Q_{\rm RS,1}) = \log 2$ and $g(Q_{\rm RS,1}) = n^2\beta^2/4 + \log 2$.

Keeping for the moment to these RS saddle points, one needs to find which one dominates the sum. In Figure 8.1 we plot the functions $g_0(n,\beta) \equiv g(Q_{\rm RS,0})$ and $g_1(n,\beta) \equiv g(Q_{\rm RS,1})$ for $n = 3$ and $n = 0.5$ as a functions of $T = 1/\beta$. Notice that the expressions we obtained for $g_0(n,\beta)$ and $g_1(n,\beta)$ are polynomials in n, which we can plot for non-integer values of n.

When $n > 1$, the situation is always qualitatively the same as the one shown in the $n = 3$ case. If we let $\beta_{\rm c}(n) = \sqrt{4\log 2/n}$, we have $g_1(\beta,n) > g_0(\beta,n)$ for $\beta > \beta_{\rm c}(n)$, while $g_1(\beta,n) < g_0(\beta,n)$ for $\beta < \beta_{\rm c}(n)$. Assuming for the moment that the sum in Eq. (8.7) is dominated by replica symmetric terms, we have $\mathbb{E}\, Z^n \doteq$

[20] Speaking of 'saddle points' is a bit sloppy in this case, since we are dealing with a *discrete* sum. By this, we mean that we aim at estimating the sum in Eq. (8.7) through a single 'dominant' term.

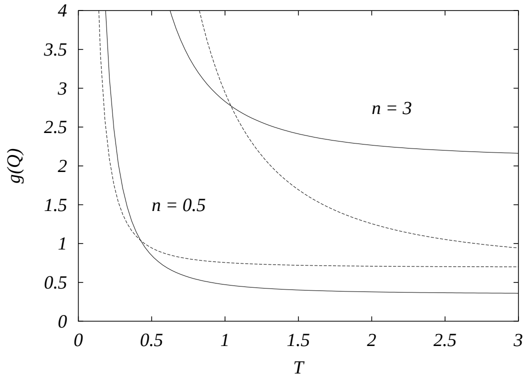


FIG. 8.1. Rate function $g(Q)$ for the REM, cf. Eq. (8.7) versus temperature. $g(Q)$ is evaluated here on the two replica-symmetric saddle points $Q_{\rm RS,0}$ (continuous curves) and $Q_{\rm RS,1}$ (dashed curves), in the cases $n = 3$ and $n = 0.5$. {fig:RemRSSaddlePoints}

$\exp\{N \max[g_0(\beta, n), g_1(\beta, n)]\}$. The point $\beta_{\rm c}(n)$ can therefore be interpreted as a phase transition in the n replicas system. At high temperatures ($\beta < \beta_{\rm c}(n)$) the $q_0 = 0$ saddle point dominates the sum: replicas are essentially independent. At low temperature the partition function is dominated by $q_0 = 1$: replicas are locked together. This fits nicely within our qualitative discussion of the replicated system in the previous Section.

The problems appear when considering the $n < 1$ situation. In this case we still have a phase transition at $\beta_{\rm c}(n) = \sqrt{4 \log 2/n}$, but the high and low temperature regimes exchange their roles. At low temperature ($\beta > \beta_{\rm c}(n)$) one has $g_1(\beta, n) < g_0(\beta, n)$, and at high temperature ($\beta < \beta_{\rm c}(n)$) one has $g_1(\beta, n) > g_0(\beta, n)$. If we applied the usual prescription and pick up the saddle point which maximizes $g(Q)$, we would obtain a nonsense, physically (replicas become independent at low temperatures, and correlated at high temperatures, contrarily to our general discussion) as well as mathematically (for $n \to 0$, the function $\mathbb{E}\, Z^n$ does not go to one, because $g_1(\beta, n)$ is not linear in n at small n). ⋆ As a matter of fact, the replica method prescribes that, in this regime $n < 1$, one must estimate the sum (8.7) using the *minimum* of $g(Q)$! There is no mathematical justification of this prescription in the present context. In the next example and the following Chapters we shall outline some of the arguments employed by physicists in order to rationalize this choice.

Example 8.1 In order to get some understanding of this claim, consider the following toy problem. We want to apply the replica recipe to the quantity $Z_{\rm toy}(n) = (2\pi/N)^{n(n-1)/4}$ (for a generic real n). For n integer, we have the following integral representation:

$$Z_{\rm toy}(n) = \int e^{-\frac{N}{2}\sum_{(ab)} Q_{ab}^2} \prod_{(ab)} dQ_{ab} \equiv \int e^{Ng(Q)} \prod_{(ab)} dQ_{ab} \, , \tag{8.8}$$

where (ab) runs over all the un-ordered couples of indices $a, b \in \{1 \dots n\}$ with $a \neq b$, and the integrals over Q_{ab} run over the real line. Now we try to evaluate the above integral by the saddle point method, and begin with the assumption that is dominated by a replica symmetric point $Q^*_{ab} = q_0$ for any $a \neq b$, yielding $g(Q^*) = -n(n-1)q_0^2/2$. Next, we have to fix the value of $q_0 \in \mathbb{R}$. It is clear that the correct result is recovered by setting $q_0 = 0$, which yields $Z_{\rm toy}(n) \doteq 1$. Moreover this is the unique choice such that $g(Q^*)$ is stationary. However, for $n < 1$, $q_0 = 0$ corresponds to a *minimum*, rather than to a maximum of $g(Q^*)$. A formal explanation of this odd behavior is that the number of degrees of freedom, the matrix elements Q_{ab} with $a \neq b$, becomes negative for $n < 1$.

This is one of the strangest aspects of the replica method, but it is unavoidable. Another puzzle which we shall discuss later concerns the exchange of order of the $N \to \infty$ and $n \to 0$ limits.

Let us therefore select the saddle point $q_0 = 0$, and use the trick (8.2) to evaluate the free energy density. Assuming that the $N \to \infty$ and $n \to 0$ limits commute, we get the RS free energy:

$$-\beta f \equiv \lim_{N\to\infty} \frac{1}{N} \mathbb{E} \log Z = \lim_{N\to\infty} \lim_{n\to 0} \frac{1}{Nn} \log(\mathbb{E}\, Z^n) = \lim_{n\to 0} \frac{1}{n}\, g_0(n,\beta) = \frac{\beta^2}{4} + \log 2 \, . \tag{8.9}$$

{eq:ReplicaSymmetricREM}

Comparing to the correct free energy density, cf. Eq. (5.15), we see that the RS result is correct, but only in the high temperature phase $\beta < \beta_{\rm c} = 2\sqrt{\log 2}$. It misses the phase transition. Within the RS framework, there is no way to get the correct solution for $\beta > \beta_{\rm c}$.

8.1.2 *One step replica symmetry breaking saddle point*

For $\beta > \beta_{\rm c}$, the sum (8.7) is dominated by matrices Q which are not replica symmetric. The problem is to find these new saddle points, and they must make sense in the $n \to 0$ limit. In order to improve over the RS result, one may try to enlarge the subspace of matrices to be optimized over (i.e. to weaken the requirement of replica symmetry). The **replica symmetry breaking (RSB)** scheme initially proposed by Parisi in the more complicated case of spin glass mean field theory, prescribes a recursive procedure for defining larger and larger spaces of Q matrices where to search for saddle points.

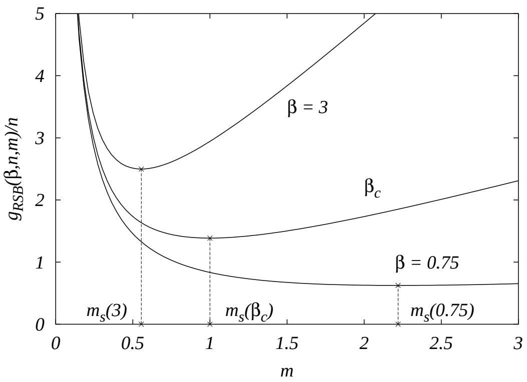


FIG. 8.2. The rate function $g(Q)$, cf. Eq. (8.7), evaluated on the one-step replica symmetry breaking point, as a function of the replica-symmetry breaking parameter m. {fig:RemRSB}

The first step of this procedure, is called **one step replica symmetry breaking (1RSB)**. In order to describe it, let us suppose that n is a multiple of m, and divide the n replicas into n/m groups of m elements each, and set:

$$\begin{aligned} &Q_{aa} = 1\,, \\ &Q_{ab} = q_1 \quad \text{if } a \text{ and } b \text{ are in the same group,} \\ &Q_{ab} = q_0 \quad \text{if } a \text{ and } b \text{ are in different groups.} \end{aligned} \tag{8.10}$$

Since in the case of the REM the matrix elements are in $\{0, 1\}$, this Ansatz is distinct from the RS one only if $q_1 = 1$ and $q_0 = 0$. This corresponds, after an eventual relabeling of the replica indices, to $i_1 = \cdots = i_m$, $i_{m+1} = \cdots = i_{2m}$, etc. The number of choices of $(i_1, \dots i_n)$ which satisfy these constraints is $\mathcal{N}_N(Q) = 2^N(2^N - 1)\cdots(2^N - n/m + 1)$, and therefore we get $s(Q) = (n/m)\log 2$. The rate function in Eq. (8.7) is given by $g(Q_{\rm RSB}) = g_{\rm RSB}(\beta, n, m)$:

$$g_{\rm RSB}(\beta, n, m) = \frac{\beta^2}{4}nm + \frac{n}{m}\log 2\ . \tag{8.11}$$ {eq:REMReplicaSymmetryBroken}

Following the discussion in the previous Section, we should minimize $g_{\rm RSB}(\beta, n, m)$ with respect to m, and then take the $n \to 0$ limit. Notice that Eq. (8.11) can be interpreted as an analytic function both in n and in $m \neq 0$. We shall therefore forget hereafter that n and m are integers with n a multiple of m. The first derivative of $g_{\rm RSB}(\beta, n, m)$ with respect to m, vanishes if $m = m_{\rm s}(\beta)$, where

$$m_{\rm s}(\beta) \equiv \frac{2\sqrt{\log 2}}{\beta} = \frac{\beta_{\rm c}}{\beta}\ . \tag{8.12}$$

Substituting in Eq. (8.11), and assuming again that we can commute the limits $n \to 0$ and $N \to \infty$, we get

$$-\beta f = \lim_{n \to 0} \frac{1}{n} \min_{m} g_{\text{RSB}}(\beta, n, m) = \beta \sqrt{\log 2}\,, \tag{8.13}$$

which is the correct result for $\beta > \beta_c$: $f = -\sqrt{\log 2}$. In fact we can recover the correct free energy of the REM in the whole temperature range if we accept that the inequality $1 \le m \le n$, valid for n, m integers, becomes $n = 0 \le m \le 1$ in the limit $n \to 0$ (we shall see later on other arguments supporting this prescription). If the minimization is constrained to $m \in [0, 1]$, we get a fully consistent answer: $m = \beta_c/\beta$ is the correct saddle point in the phase $\beta > \beta_c$, while for $\beta < \beta_c$ the parameter m sticks to the value $m = 1$. In Fig. 8.2 we sketch the function $g_{\text{RSB}}(\beta, n, m)/n$ for a few values of the temperature β.

8.1.3 *Comments on the replica solution*

One might think that the replica method is just a fancy way of reconstructing a probability distribution from its integer moments. We know how to compute the integer moments of the partition function $\mathbb{E}\, Z^n$, and we would like to infer the full distribution of Z, and in particular the value of $\mathbb{E} \log Z$. This is a standard topic in probability theory: the probability distribution can be reconstructed if its integer moments don't grow too fast as $n \to \infty$. A typical result is the following.

Theorem 8.2. (Carleman) *Let X be a real random variable with moments $\mu_n = \mathbb{E}\, X^n$ such that*

$$\sum_{n=1}^{\infty} \mu_{2n}^{-1/2n} = \infty\,. \tag{8.14}$$

Then any variable with the same moments is distributed identically to X.

For instance, if the moments don't grow faster than exponentially, $\mathbb{E}\, X^n \sim e^{\alpha n}$, their knowledge completely determines the distribution of X.

Let us try to apply the above result to the REM case treated in the previous pages. The replica symmetric calculation of Sec. 8.1.1 is easily turned into a lower bound:

$$\mathbb{E}\, Z^n \ge e^{n g(Q_{\text{RS},0})} \ge e^{N\beta^2 n^2/4}\,. \tag{8.15}$$

Therefore the sum in Eq. (8.14) converges and the distribution of Z is not necessarily fixed by its integer moments.

Exercise 8.2 Assume $Z = e^{-F}$, with F a Gaussian random variable, with probability density

$$p(F) = \frac{1}{\sqrt{2\pi}}\, e^{-F^2/2}\,. \tag{8.16}$$

Compute the integer moments of Z. Do they verify the hypothesis of Carleman Theorem? Show that the moments are unchanged if $p(F)$ is replaced by the density $p_a(F) = p(F)[1 + a \sin(2\pi F)]$, with $|a| < 1$ (from (Feller, 1968)).

In our replica approach, there exist several possible analytic continuations to non-integer n's, and the whole issue is to find the correct one. Parisi's Ansatz (and its generalization to higher order RSB that we will discuss below) gives a well defined class of analytic continuations, which turns out to be the correct one in many different problems.

The suspicious reader will notice that the moments of the REM partition function would not grow that rapidly if the energy levels had a distribution with bounded support. If for instance, we considered E_i to be Gaussian random variables conditioned to $E_i \in [-E_{\max}, E_{\max}]$, the partition function would be upper bounded by the constant $Z_{\max} = 2^N e^{\beta E_{\max}}$. Consequently, we would have $\mathbb{E}\, Z^n \leq Z_{\max}^n$, and the whole distribution of Z could be recovered from its integer moments. In order to achieve such a goal, we would however need to know exactly all the moments $1 \leq n < \infty$ at fixed N (the system size). What we are instead able to compute, in general, is the large N behavior at any fixed n. In most cases, this information is insufficient to insure a unique continuation to $n \to 0$.

In fact, one can think of the replica method as a procedure for computing the quantity

$$\psi(n) = \lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}\, Z^n \, , \tag{8.17}$$

whenever the limit exist. In the frequent case where $f = -\log Z/(\beta N)$ satisfies a large deviation principle of the form $P_N(f) \doteq \exp[-NI(f)]$, then we have

$$\mathbb{E}\, Z^n \doteq \int df \;\exp[-NI(f) - N\beta n f] \doteq \exp\{-N \inf[I(f) + \beta n f]\}\, . \tag{8.18}$$

Therefore $\psi(n) = -\inf[I(f)+\beta n f]$. In turns, the large deviation properties of f_N can be inferred from $\psi(n)$ through the Gärtner-Ellis theorem 4.12. The typical value of the free energy density is given by the location of the absolute minimum of $I(f)$. In order to compute it, one must in general use values of n which go to 0, and one cannot infer it from the integer values of n.

8.1.4 *Condensation*

{se:reprem_cond}

As we discussed in Chapter 5, the appearance of a low temperature 'glass' phase is associated with a condensation of the probability measure on few configurations. We described quantitatively this phenomenon by the participation ratio Y. For the REM we obtained $\lim_{N\to\infty} \mathbb{E}\, Y = 1 - \beta_{\rm c}/\beta$ for any $\beta > \beta_{\rm c}$ (see proposition 5.3). Let us see how this result can be recovered in just a few lines from a replica computation.

The participation ratio is defined by $Y = \sum_{j=1}^{2^N} p_j^2$, where $p_j = e^{-\beta E_j}/Z$ is Boltzmann's probability of the j'th energy level. Therefore:

$$
\begin{aligned}
\mathbb{E}\,Y &= \lim_{n\to 0}\ \mathbb{E}\left[Z^{n-2}\sum_{i=1}^{2^N} e^{-2\beta E_i}\right] && \text{[Definition of } Y\text{]}\\
&= \lim_{n\to 0}\mathbb{E}\left[\sum_{i_1\dots i_{n-2}} e^{-\beta(E_{i_1}+\cdots+E_{i_{n-2}})}\sum_{i=1}^{2^N} e^{-2\beta E_i}\right] && \text{[Assume } n\in\mathbb{N}\text{]}\\
&= \lim_{n\to 0}\mathbb{E}\left[\sum_{i_1\dots i_n} e^{-\beta(E_{i_1}+\cdots+E_{i_n})}\mathbb{I}(i_{n-1}=i_n)\right]\\
&= \lim_{n\to 0}\frac{1}{n(n-1)}\sum_{a\neq b}\mathbb{E}\left[\sum_{i_1\dots i_n} e^{-\beta(E_{i_1}+\cdots+E_{i_n})}\ \mathbb{I}(i_a=i_b)\right] && \text{[Symmetrize]}\\
&= \lim_{n\to 0}\frac{1}{n(n-1)}\sum_{a\neq b}\frac{\mathbb{E}\left[\sum_{i_1\dots i_n} e^{-\beta(E_{i_1}+\cdots+E_{i_n})}\ \mathbb{I}(i_a=i_b)\right]}{\mathbb{E}\left[\sum_{i_1\dots i_n} e^{-\beta(E_{i_1}+\cdots+E_{i_n})}\right]} && \text{[Denom. } \to 1\text{]}\\
&= \lim_{n\to 0}\frac{1}{n(n-1)}\sum_{a\neq b}\langle Q_{ab}\rangle_n\ , && (8.19)
\end{aligned}
$$

where the sums over the replica indices a, b run over $a, b \in \{1,\dots,n\}$, while the configuration indices i_a are summed over $\{1,\dots,2^N\}$. In the last step we introduced the notation

q:ExpectationReplicated}

$$
\langle f(Q)\rangle_n \equiv \frac{\sum_Q f(Q)\,\mathcal{N}_N(Q)e^{\frac{N\beta^2}{4}\sum_{a,b}Q_{ab}}}{\sum_Q \mathcal{N}_N(Q)e^{\frac{N\beta^2}{4}\sum_{a,b}Q_{ab}}}\,, \tag{8.20}
$$

and noticed that the sum over $i_1,\dots,i_n$ can be split into a sum over the overlap matrices Q and a sum over the n-uples $i_1\dots i_n$ having overlap matrix Q. Notice that $\langle\cdot\rangle_n$ can be interpreted as an expectation in the 'replicated system'.

In the large N limit $\mathcal{N}_N(Q) \doteq e^{Ns(Q)}$, and the expectation value (8.20) is given by a dominant[21] (saddle point) term: $\langle f(Q)\rangle_n \simeq f(Q^*)$. As argued in the previous Sections, in the low temperature phase $\beta > \beta_{\rm c}$, the saddle point matrix is given by the 1RSB expression (8.10).

[21] If the dominant term corresponds to a non-replica symmetric matrix Q^*, all the terms obtained by permuting the replica indices contribute with an equal weight. Because of this fact, it is a good idea to compute averages of symmetric functions $f(Q) = f(Q^\pi)$. This is what we have done in Eq. (8.19).

$$\begin{aligned}\mathbb{E}\,Y &= \lim_{n\to 0}\frac{1}{n(n-1)}\sum_{a\neq b} Q_{ab}^{\rm 1RSB} && \text{[Saddle point]}\\ &= \lim_{n\to 0}\frac{1}{n(n-1)}\, n[(n-m)q_0+(m-1)q_1] && \text{[Eq. (8.10)]}\\ &= 1-m = 1-\frac{\beta_{\rm c}}{\beta} && [q_0=0,\ q_1=1]\,. \qquad (8.21)\end{aligned}$$

This is exactly the result we found in proposition 5.3, using a direct combinatorial approach. It also confirms that the 1RSB Ansatz (8.10) makes sense only provided $0 \le m \le 1$ (the participation ratio Y is positive by definition). Compared to the computation in Sec. 5.3, the simplicity of the replica derivation is striking.

At first look, the manipulations in Eq. (8.19) seem to require new assumptions with respect to the free energy computation in the previous Sections. Replicas are introduced in order to write the Z^{-2} factor in the participation ratio, as the analytic continuation of a positive power Z^{n-2}. It turns out that this calculation is in fact equivalent to the one in (8.2). This follows from the basic observation that expectation values can be obtained as derivatives of $\log Z$ with respect to some parameters.

{ex:rem1}

Exercise 8.3 Using the replica method, show that, for $T < T_{\rm c}$:

$$\mathbb{E}\left(\sum_{j=1}^{2^N} p_j^r\right) = \frac{\Gamma(r-m)}{\Gamma(r)\Gamma(1-m)} = \frac{(r-1-m)(r-2-m)\dots(1-m)}{(r-1)(r-2)\dots(1)}\ , \qquad (8.22)$$

where $\Gamma(x)$ denotes Euler's Gamma function.

Exercise 8.4 Using the replica method, show that, for $T < T_{\rm c}$:

$$\mathbb{E}\left(Y^2\right) = \frac{3-5m+2m^2}{3}\ . \qquad (8.23)$$

8.2 The fully connected p-spin glass model

{se:PspinReplicas}

The replica method provides a compact and efficient way to compute –in a non rigorous way– the free energy density of the REM. The result proves to be exact, once replica symmetry breaking is used in the low temperature phase. However, its power can be better appreciated on more complicated problems which cannot be solved by direct combinatorial approaches. In this Section we shall apply the replica method to the so-called 'p-spin glass' model. This model has been invented in the theoretical study of spin glasses. Its distinguishing feature are interactions which involve groups p spins, with $p \ge 2$. It generalizes ordinary spin glass models, cf. Sec. 2.6, in which interactions involve couples of spins (i.e. $p = 2$). This provides an additional degree of freedom, the value of p, and different physical scenarios appear whether $p = 2$ or $p \geq 3$. Moreover, some pleasing simplifications show up for large p.

In the p**-spin model**, one considers the space of 2^N configurations of N Ising spins. The energy of a configuration $\sigma = \{\sigma_1, \dots, \sigma_N\}$ is defined as:

{eq:pspin_enedef}

$$E(\sigma) = -\sum_{i_1 < i_2 < \dots i_p} J_{i_1 \dots i_p} \sigma_{i_1} \cdots \sigma_{i_p} \tag{8.24}$$

where $\sigma_i \in \{\pm 1\}$. This is a disordered system: a sample is characterized by the set of all couplings $J_{i_1 \dots i_p}$, with $1 \leq i_1 < \cdots < i_p \leq N$. These are taken as iid Gaussian random variables with zero mean and variance $\mathbb{E}\, J^2_{i_1 \dots i_p} = p!/(2N^{p-1})$. Their probability density reads:

{eq:pspin_jdist}

$$P\left(J\right) = \sqrt{\frac{\pi p!}{N^{p-1}}}\ \exp\left(-\frac{N^{p-1}}{p!} J^2\right)\ ; \tag{8.25}$$

The p-spin model is a so-called **infinite range interaction** model: there is no notion of Euclidean distance between the positions of the spins. It is also called a **fully connected** model since each spin interacts directly with all the others. The last feature is at the origin of the special scaling of the variance of the J distribution in (8.25). A simple criterion for arguing that the proposed scaling is the correct one consists in requiring that a flip of a single spin generates an energy change of order 1 (i.e. finite when $N \to \infty$). More precisely, let $\sigma^{(i)}$ the configuration obtained from σ by reversing the spin i and define $\Delta_i \equiv [E(\sigma^{(i)}) - E(\sigma)]/2$. It is easy to see that $\Delta_i = \sum_{i_2 \dots i_p} J_{ii_1 \dots i_p} \sigma_i \sigma_{i_1} \cdots \sigma_{i_p}$. The sum is over $\Theta(N^{p-1})$ terms, and, if σ is a random configuration, the product $\sigma_i \sigma_{i_1} \cdots \sigma_{i_p}$ in each term is $+1$ or -1 with probability 1/2. The scaling in (8.25) insures that Δ_i is finite as $N \to \infty$ (in contrast, the $p!$ factor is just a matter of convention).

Why is it important that the Δ_i are of order 1? The intuition is that Δ_i estimates the interaction between a spin and the rest of the system. If Δ_i were much larger than 1, the spin σ_i would be completely frozen in the direction which makes Δ_i positive, and temperature wouldn't have any role. On the other hand, if Δ_i were much smaller than one, the spin i would be effectively independent from the others.

Exercise 8.5 An alternative argument can be obtained as follows. Show that, at high temperature $\beta \ll 1$: $Z = 2^N [1 + 2^{-1}\, \beta^2 \sum_{i_1 < \dots < i_p} J^2_{i_1 .. i_p} + O(\beta^3)]$. This implies $N^{-1} \mathbb{E} \log Z = \log 2 + C_N \beta^2/2 + O(\beta^3)$, with $C_N = 1$. What would happen with a different scaling of the variance? Which scaling is required in order for C_N to have a finite $N \to \infty$ limit?

The special case of $p = 2$ is the closest to the original spin glass problem and is known as the **Sherrington-Kirkpatrick** (or **SK**) model.

8.2.1 *The replica calculation*

Let us start by writing Z^n as the partition function for n non-interacting replicas σ_i^a, with $i \in \{1, \dots, N\}$, $a \in \{1, \dots, n\}$:

$$Z^n = \sum_{\{\sigma_i^a\}} \prod_{i_1 < \cdots < i_p} \exp\left(\beta J_{i_1 \dots i_p} \sum_{a=1}^{n} \sigma_{i_1}^a \dots \sigma_{i_p}^a \right) . \tag{8.26}$$

The average over the couplings $J_{i_1..i_p}$ is easily done by using their independence and the well known identity

$$\mathbb{E}\, e^{\lambda X} = e^{\frac{1}{2}\Delta\lambda^2} , \tag{8.27}$$

{eq:HubbardStrat}

holding for a Gaussian random variable X with zero mean and variance $\mathbb{E}\, X^2 = \Delta$. One gets:

$$\begin{aligned} \mathbb{E}\, Z^n &= \sum_{\{\sigma_i^a\}} \exp\left(\frac{\beta^2}{4} \frac{p\,!}{N^{p-1}} \sum_{i_1 < \cdots < i_p} \sum_{a,b} \sigma_{i_1}^a \sigma_{i_1}^b \; \sigma_{i_2}^a \sigma_{i_2}^b \cdots \sigma_{i_p}^a \sigma_{i_p}^b \right) \\ &\doteq \sum_{\{\sigma_i^a\}} \exp\left[\frac{\beta^2}{4} \frac{1}{N^{p-1}} \sum_{a,b} \left(\sum_i \sigma_i^a \sigma_i^b \right)^p \right] \end{aligned} \tag{8.28}$$

{eq:ReplicatedPspin}

where we have neglected corrections due to coincident indices $i_l = i_k$ in the first term, since they are irrelevant to the leading exponential order. We introduce for each $a < b$ the variables λ_{ab} and Q_{ab} by using the identity

$$1 = \int dQ_{ab}\, \delta\left(Q_{ab} - \frac{1}{N} \sum_{i=1}^{N} \sigma_i^a \sigma_i^b \right) = N \int dQ_{ab} \int \frac{d\lambda_{ab}}{2\pi}\, e^{-i\lambda_{ab}\left(NQ_{ab} - \sum_i \sigma_i^a \sigma_i^b \right)} , \tag{8.29}$$

with all the integrals running over the real line. Using it in Eq. (8.28), we get

$$\begin{aligned} \mathbb{E}\, Z^n &\doteq \int \prod_{a<b} dQ_{ab} \sum_{\{\sigma_i^a\}} \exp\left(\frac{N\beta^2}{4} n + \frac{N\beta^2}{2} \sum_{a<b} Q_{ab}^p \right) \delta\left(Q_{ab} - \frac{1}{N} \sum_{i=1}^{N} \sigma_i^a \sigma_i^b \right) \\ &\doteq \int \prod_{a<b} (dQ_{ab}\, d\lambda_{ab})\, e^{-NG(Q,\lambda)} \end{aligned} \tag{8.30}$$

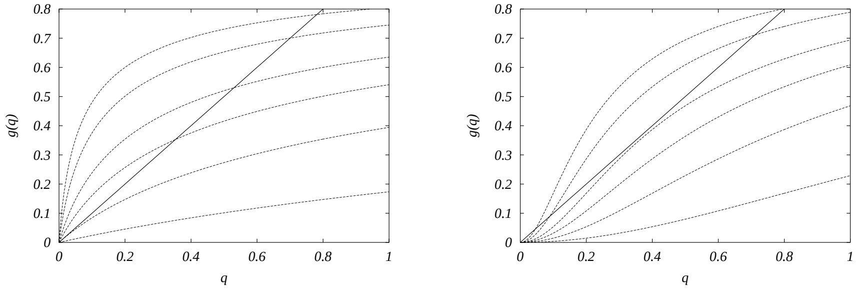
{eq:pspin_sp}

where we have introduced the function:

$$G(Q, \lambda) = -n\frac{\beta^2}{4} - \frac{\beta^2}{2} \sum_{a<b} Q_{ab}^p + \sum_{a<b} i\lambda_{ab} Q_{ab} - \log\left[\sum_{\{\sigma_a\}} e^{\sum_{a<b} i\lambda_{ab}\sigma_a\sigma_b} \right] , \tag{8.31}$$

{eq:PspinAction}

which depends upon the $n(n-1)/2 + n(n-1)/2$ variables Q_{ab}, λ_{ab}, $1 \le a < b \le n$.



{fig:PspinRS} FIG. 8.3. Graphical solution of the RS equations for the p-spin model, with $p = 2$ (SK model, left) and $p = 3$ (right). The various curves correspond to inverse temperatures $\beta = 4$, 3, 2, 1.5, 1, 0.5 (from top to bottom).

{AlternativeAction}

Exercise 8.6 An alternative route consists in noticing that the right hand side of Eq. (8.28) depends upon the spin configuration only through the overlap matrix $Q_{ab} = N^{-1} \sum_i \sigma_i^a \sigma_i^b$, with $a < b$. The sum can be therefore decomposed into a sum over the overlap matrices and a sum over configurations with a given overlap matrix:

$$\mathbb{E}\, Z^n \doteq \sum_Q \mathcal{N}_N(Q) \, \exp\left(\frac{N\beta^2}{4} n + \frac{N\beta^2}{2} \sum_{a<b} Q_{ab}^p \right) . \tag{8.32}$$

Here $\mathcal{N}_N(Q)$ is the number of spin configurations with a given overlap matrix Q. In analogy to the REM case, it is natural to guess a large deviations principle of the form $\mathcal{N}_N(Q) \doteq \exp[Ns(Q)]$. Use the Gärtner-Ellis theorem 4.12 to obtain an expression for the 'entropic' factor $s(Q)$. Compare the resulting formula for $\mathbb{E}\, Z^n$ with Eq. (8.28).

Following our general approach, we shall estimate the integral (8.30) at large N by the saddle point method. The stationarity conditions of G are most easily written in terms of the variables $\mu_{ab} = i\lambda_{ab}$. By differentiating Eq. (8.31) with respect to its arguments, we get $\forall a < b$

{eq:pspin_speq}

$$\mu_{ab} = \frac{1}{2} p\beta^2 \, Q_{ab}^{p-1} \, , \qquad\qquad Q_{ab} = \langle \sigma_a \sigma_b \rangle_n \, , \tag{8.33}$$

where we have introduced the average within the replicated system

{eq:rep_onesitez}

$$\langle f(\sigma) \rangle_n \equiv \frac{1}{z(\mu)} \sum_{\{\sigma^a\}} f(\sigma) \, \exp\left(\sum_{a<b} \mu_{ab} \sigma_a \sigma_b \right) \, , \quad z(\mu) \equiv \sum_{\{\sigma^a\}} \exp\left(\sum_{a<b} \mu_{ab} \, \sigma_a \sigma_b \right) , \tag{8.34}$$

for any function $f(\sigma) = f(\sigma^1 \dots \sigma^n)$.

We start by considering a RS saddle point: $Q_{ab} = q$; $\mu_{ab} = \mu$ for any $a \neq b$. Using the Gaussian identity (8.27), one finds that the saddle point equations (8.33) become:

{eq:ps_speq_rs}

$$\mu = \frac{1}{2}\, p\beta^2\, q^{p-1}\,, \qquad q = \mathsf{E}_z \tanh^2\left(z\sqrt{\mu}\right)\,, \tag{8.35}$$

where E_z denotes the expectation with respect to a Gaussian random variable z of zero mean and unit variance. Eliminating μ, we obtain an equation for the overlap parameter: $q = r(q)$, with $r(q) \equiv \mathsf{E}_z \tanh^2(z\sqrt{p\beta^2\, q^{p-1}/2})$. In Fig. 8.3 we plot the function $r(q)$ for $p = 2, 3$ and various temperatures. The equations (8.35) always admit the solution $q = \mu = 0$. Substituting into Eq. (8.31), and using the trick (8.2) this solution would yield a free energy density

$$f_{\rm RS} = \lim_{n\to 0} \frac{1}{\beta n} G(Q^{\rm RS}, \lambda^{\rm RS}) = -\beta/4 - (1/\beta)\log 2\,. \tag{8.36}$$

At low enough temperature, other RS solutions appear. For $p = 2$, a single such solution departs continuously from 0 at $\beta_{\rm c} = 1$, cf. Fig. 8.3, left frame. For $p \geq 3$ a couple of non-vanishing solutions appear discontinuously for $\beta \geq \beta_*(p)$ and merge as $\beta \downarrow \beta_*(p)$, cf. Fig. 8.3, right frame. However two arguments allow to discard these saddle points:

- Stability argument: One can compute the Taylor expansion of $G(Q, \lambda)$ around such RS saddle points. The saddle point method can be applied only if the matrix of second derivatives has a defined sign. As discussed in the Appendix, this condition does not hold for the non-vanishing RS saddle points.
- Positivity of the entropy: As explained in Chap. 2, because of the positivity of the entropy, the free energy of a physical system with discrete degrees of freedom must be a decreasing function of the temperature. Once again, one can show that this condition is not satisfied by the non-vanishing RS saddle points.

  On the other hand, the $q = 0$ saddle point also violates this condition at low enough temperature (as the reader can show from Eq. (8.36)). ⋆

The above arguments are very general. The second condition, in particular, is straightforward to be checked and must always be satisfied by the correct saddle point. The conclusion is that none of the RS saddle points is correct at low temperatures. This motivates us to look for 1RSB saddle points. We partition the set of n replicas into n/m groups of m replicas each and seek a saddle point of the following 1RSB form:

$$\begin{aligned} &Q_{ab} = q_1\,, \quad \mu_{ab} = \mu_1\,, \quad \text{if } a \text{ and } b \text{ belong to the same group,}\\ &Q_{ab} = q_0\,, \quad \mu_{ab} = \mu_0\,, \quad \text{if } a \text{ and } b \text{ belong to different groups.} \end{aligned} \tag{8.37}$$

{eq:1RSBAnsatzPspin}

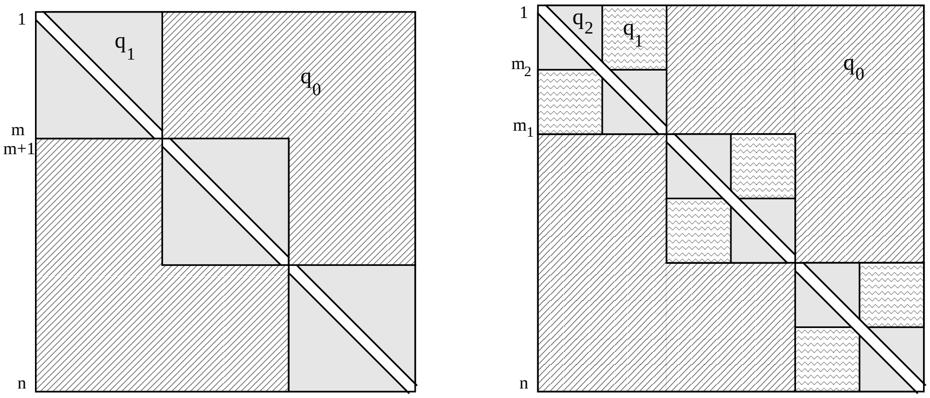


FIG. 8.4. Structure of the Q_{ab} matrix when replica symmetry is broken. Left: 1RSB Ansatz. The $n(n-1)/2$ values of Q_{ab} are the non diagonal elements of a symmetric $n \times n$ matrix. The n replicas are divided into n/m blocks of size m. When a and b are in the same block, $Q_{ab} = q_1$, otherwise $Q_{ab} = q_0$. Right: 2RSB Ansatz: an example with $n/m_1 = 3$ and $m_1/m_2 = 2$.

{fig:pspin_1rsb_Ansatz}

In practice one can relabel the replicas in such a way that the groups are formed by successive indices $\{1 \dots m\}$, $\{m+1 \dots 2m\}$, $\dots$, $\{n-m+1 \dots n\}$ (see Fig. 8.4)[22].

The computation of $G(Q, \lambda)$ on this saddle point makes repeated use of the ⋆ identity (8.27) and is left as an exercise.One gets:

$$G(Q^{1\text{RSB}}, \lambda^{1\text{RSB}}) = -n\frac{\beta^2}{4} + n\frac{\beta^2}{4}\left[(1-m)q_1^p + mq_0^p\right] - \frac{n}{2}\left[(1-m)q_1\mu_1 + mq_0\mu_0\right] \\ + \frac{n}{2}\mu_1 - \log\left\{\mathsf{E}_{z_0}\left[\mathsf{E}_{z_1}(2\cosh(\sqrt{\mu_0}\, z_0 + \sqrt{\mu_1 - \mu_0}\, z_1))^m\right]^{n/m}\right\} \tag{8.38}$$

{eq:1RSBFreeEnergy}

where E_{z_0} and E_{z_1} denote expectations with respect to the independent Gaussian random variables z_0 and z_1 with zero mean and unit variance.

[22] Some of the other labellings of the replicas give distinct 1RSB saddle points with the same value of $G(Q, \lambda)$. This is a general feature of RSB saddle points, that we already encountered when studying the REM, cf. Sec. 8.1.4.

Exercise 8.7 Show that the limit $G_{\rm 1RSB}(q,\mu;m) = \lim_{n\to 0} n^{-1}\, G(Q^{\rm 1RSB},\lambda^{\rm 1RSB})$ exists, and compute the function $G_{\rm 1RSB}(q,\mu;m)$. Determine the stationarity condition for the parameters q_1, q_0, μ_1, μ_0 and m by computing the partial derivatives of $G_{\rm 1RSB}(q,\mu;m)$ with respect to its arguments and setting them to 0. Show that these equations are always consistent with $q_0 = \mu_0 = 0$, and that

$$
\begin{aligned}
G_{\rm 1RSB}|_{q_0,\mu_0=0} = & -\frac{1}{4}\beta^2[1-(1-m)q_1^p] + \frac{1}{2}\mu_1[1-(1-m)q_1] \\
& -\frac{1}{m}\log \mathsf{E}_z\left[(2\cosh(\sqrt{\mu_1}\, z))^m\right]\,. \qquad (8.39)
\end{aligned}
$$

Picking up the solution $q_0 = \mu_0 = 0$, the stationarity conditions[23] for the remaining parameters q_1 and μ_1 read

$$
\mu_1 = \frac{1}{2}\, p\beta^2\, q_1^{p-1}\,, \qquad q_1 = \frac{\mathsf{E}_z\left[(2\cosh(\sqrt{\mu_1}\, z))^m (\tanh(\sqrt{\mu_1}\, z))^2\right]}{\mathsf{E}_z\left[(2\cosh(\sqrt{\mu_1}\, z))^m\right]}\,. \qquad (8.40)
$$

These equations always admit the solution $q_1 = \mu_1 = 0$: this choice reduces in fact to a replica symmetric Ansatz, as can be seen from Eq. (8.37). Let us now consider the $p \geq 3$ case. At low enough temperature two non-vanishing solutions appear. A local stability analysis shows that the largest one, let us call it $mu_1^{\rm sp}$, $q_1^{\rm sp}$, must be chosen.

The next step consists in optimizing $G_{\rm 1RSB}(q^{\rm sp},\mu^{\rm sp};m)$ with respect to $m \in [0,1]$ (notice that $G_{\rm 1RSB}$ depends on m both explicitly and through $q^{\rm sp},\mu^{\rm sp}$). It turns out that a unique stationary point $m_{\rm s}(\beta)$ exists, but $m_{\rm s}(\beta) \in [0,1]$ only at low enough temperature $\beta > \beta_{\rm c}(p)$. We refer to the literature for an explicit characterization of $\beta_{\rm c}(p)$. At the transition temperature $\beta_{\rm c}(p)$, the free energy of the 1RSB solution becomes equal to that of the RS one. There is a phase transition from a RS phase for $\beta < \beta_{\rm c}(p)$ to a 1RSB phase for $\beta > \beta_{\rm c}(p)$.

These calculations are greatly simplified (and can be carried out analytically) in the large p limit. The leading terms in a large p expansion are:

$$
\beta_{\rm c}(p) = 2\sqrt{\log 2} + e^{-\Theta(p)}\,, \quad m_{\rm s}(\beta) = \frac{\beta_{\rm c}(p)}{\beta} + e^{-\Theta(p)}\,, \quad q_1 = 1 - e^{-\Theta(p)}\,.(8.41)
$$

The corresponding free energy density is constant in the whole low temperature phase, equal to $-\sqrt{\log 2}$. The reader will notice that several features of the REM are recovered in this large p limit. One can get a hint that this should be the case from the following exercise:

[23] They are most easily obtained by differentiating Eq. (8.39) with respect to q_1 and μ_1.

Exercise 8.8 Consider a p-spin glass problem, and take an arbitrary configuration $\sigma = \{\sigma_1, \dots, \sigma_N\}$. Let $P_\sigma(E)$ denote the probability that this configuration has energy E, when a sample (i.e. a choice of couplings $J_{i_1 \dots i_p}$) is chosen at random with distribution (8.25). Show that $P_\sigma(E)$ is independent of σ, and is a Gaussian distribution of mean 0 and variance $N/2$. Now take two configurations σ and σ', and show that the joint probability distribution of their energies, respectively E and E', in a randomly chosen sample, is:

$$P_{\sigma,\sigma'}(E, E') = C \exp\left[-\frac{(E+E')^2}{2N(1+x^p)} - \frac{(E-E')^2}{2N(1-x^p)}\right] \tag{8.42}$$

where $x = (1/N)\sum_i \sigma_i \sigma_i'$, and C is a normalization constant. When $|x| < 1$ the energies of the two configurations become uncorrelated as $p \to \infty$, (i.e. $\lim_{p\to\infty} P_{\sigma,\sigma'}(E, E') = P_\sigma(E) P_{\sigma'}(E')$), suggesting a REM-like behavior.

In order to know if the 1RSB solution which we have just found is the correct one, one should first check its stability by verifying that the eigenvalues of the Hessian (i.e. the matrix of second derivatives of $G(Q, \lambda)$ with respect to its arguments) have the correct sign. Although straightforward in principle, this computation becomes rather cumbersome and we shall just give the result, due Elizabeth Gardner. The 1RSB solution is stable only in some intermediate phase $\beta_{\rm c}(p) < \beta < \beta_{\rm u}(p)$. At the inverse temperature $\beta_{\rm u}(p)$ there is a second transition to a new phase which involves a more complex replica symmetry breaking scheme.

The 1RSB solution was generalized by Parisi to higher orders of RSB. His construction is a hierarchical one. In order to define the structure of the Q_{ab} matrix with two steps of replica symmetry breaking (**2RSB**), one starts from the 1RSB matrix of Fig. 8.4 (left panel). The off diagonal blocks with matrix elements q_0 are left unchanged. The diagonal blocks are changed: take any diagonal block of size $m_1 \times m_1$ (we now call $m = m_1$). In the 1RSB case all its matrix elements are equal to q_1. In the 2RSB case the m_1 replicas are split into m_1/m_2 blocks of m_2 replicas each. The matrix elements in the off diagonal blocks remain equal to q_1. The ones in the diagonal blocks become equal to a new number q_2 (see Fig. 8.4, right panel). The matrix is parametrized by 5 numbers: q_0, q_1, q_2, m_1, m_2. This construction can obviously be generalized by splitting the diagonal blocks again, grouping m_2 replicas into m_2/m_3 groups of m_3 replicas. The so-called **full replica symmetry breaking Ansatz (FRSB)** Ansatz corresponds to iterating this procedure R times, and eventually taking R to infinity. Notice that, while the construction makes sense, for n integer, only when $n \geq m_1 \geq m_2 \geq \cdots \geq m_R \geq 1$, in the $n \to 0$ limit this order is reversed to $0 \leq m_1 \leq m_2 \leq \cdots \leq m_R < 1$. Once one assumes a R-RSB Ansatz, computing the rate function G and solving the saddle point equations is a matter of calculus (special tricks have been developed for $R \to \infty$). It turns out that, in order to find a stable solution in the phase $\beta > \beta_{\rm u}(p)$, a FRSB Ansatz is required. This same situation is also encountered in the case of the SK model, in the whole phase $\beta > 1$, but its description would take us too far.

8.2.2 *Overlap distribution*

Replica symmetry breaking appeared in the previous Sections as a formal trick for computing certain partition functions. One of the fascinating features of spin-glass theory is that RSB has a very concrete physical (as well as probabilistic) interpretation. One of the main characteristics of a system displaying RSB is the existence, in a typical sample, of some spin configurations which are very different from the lowest energy (ground state) configuration, but are very close to it in energy. One gets a measure of this property through the distribution of overlaps between configurations. Given two spin configurations $\sigma = \{\sigma_1, \dots, \sigma_N\}$ and $\sigma' = \{\sigma'_1, \dots, \sigma'_N\}$, the **overlap** between σ and σ' is:

$$q_{\sigma\sigma'} = \frac{1}{N}\sum_{i=1}^{N} \sigma_i \sigma'_i \ , \tag{8.43}$$

so that $N(1-q_{\sigma\sigma'})/2$ is the Hamming distance between σ and σ'. For a given sample of the p-spin glass model, which we denote by J, the **overlap distribution** $P_J(q)$ is the probability density that two configuration, randomly chosen with the Boltzmann distribution, have overlap q:

$$\int_{-1}^{q} P_J(q')\, dq' = \frac{1}{Z^2}\sum_{\sigma,\sigma'} \exp\left[-\beta E(\sigma) - \beta E(\sigma')\right] \mathbb{I}\left(q_{\sigma\sigma'} \le q\right) \tag{8.44}$$

Let us compute the expectation of $P_J(q)$ in the thermodynamic limit:

$$P\left(q\right) \equiv \lim_{N\to\infty} \mathbb{E}\, P_J\left(q\right) \tag{8.45}$$

using replicas. One finds:

$$\int_{-1}^{q} P(q')\, dq' = \lim_{n\to 0} \sum_{\sigma^1 \dots \sigma^n} \mathbb{E}\left[\exp\left(-\beta \sum_a E(\sigma^a)\right)\right] \mathbb{I}\left(q_{\sigma^1\sigma^2} \le q\right) \tag{8.46}$$

The calculation is very similar to the one of $\mathbb{E}\left(Z^n\right)$, the only difference is that now the overlap between replicas 1 and 2 is fixed to be $\le q$. Following the same steps as before, one obtains the expression of $P(q)$ in terms of the saddle point matrix $Q^{\rm sp}_{ab}$. The only delicate point is that there may be several RSB saddle points related by a permutation of the replica indices. If $Q = \{Q_{ab}\}$ is a saddle point, any matrix $\left(Q^\pi\right)_{ab} = Q_{\pi(a),\pi(b)}$ (with π a permutation in S_n) is also a saddle point, with the same weight: $G(Q^\pi) = G(Q)$. When computing $P\left(q\right)$, we need to sum up over all the equivalent distinct saddle points, which gives in the end:

$$\int_{-1}^{q} P(q')\, dq' = \lim_{n\to 0} \frac{1}{n\left(n-1\right)} \sum_{a\ne b} \mathbb{I}\left(Q^{\rm sp}_{ab} \le q\right) \ . \tag{8.47}$$

In case of a RS solution one has:

$$\int_{-1}^{q} P(q')\, dq' = \mathbb{I}\left(q^{\rm RS} \le q\right) \,, \tag{8.48}$$

with $q^{\rm RS}$ the solution of the saddle point equations (8.35). In words: if two configurations σ and σ' are drawn according to the Boltzmann distribution, their overlap will be $q^{\rm RS}$ with high probability. Since the overlap is the sum of many 'simple' terms, the fact that its distribution concentrates around a typical value is somehow expected.

In a 1RSB phase characterized by the numbers $q_0, q_1, \lambda_0, \lambda_1, m$, one finds:

$$\int_{-1}^{q} P(q')\, dq' = (1-m)\, \mathbb{I}\left(q_1 \le q\right) + m\, \mathbb{I}\left(q_0 \le q\right) \;. \tag{8.49}$$

The overlap can take with finite probability two values: q_0 or q_1. This has a very nice geometrical interpretation. When sampling configurations randomly chosen with the Boltzmann probability, at an inverse temperature $\beta > \beta_{\rm c}(p)$, the configurations will typically be grouped into clusters, such that any two configurations in the same cluster have an overlap q_1, while configurations in different clusters have an overlap $q_0 < q_1$, and thus a larger Hamming distance. When picking at random two configurations, the probability that they fall in the same cluster is equal to $1-m$. The clustering property is a rather non-trivial one: it would have been difficult to anticipate it without a detailed calculation. We shall encounter later several other models where it also occurs. Although the replica derivation presented here is non rigorous, the clustering phenomenon can be proved rigorously.

In a solution with higher order RSB the $P(q)$ function develops new peaks. The geometrical interpretation is that clusters contain some sub-clusters, which themselves contain sub-clusters etc... this hierarchical structure leads to the property of **ultrametricity**. Consider the triangle formed by three independent configurations drawn from the Boltzmann distribution, and let the lengths of its sides be measured using to the Hamming distance. With high probability, such a triangle will be either equilateral, or isosceles with the two equal sides larger than the third one. In the case of full RSB, $P(q)$ has a continuous part, showing that the clustering property is not as sharp, because clusters are no longer well separated; but ultrametricity still holds.

{ex:Yuniversal}

Exercise 8.9 For a given sample of a p-spin glass in its 1RSB phase, define Y as the probability that two configurations fall into the same cluster. More precisely: $Y = \int_q^1 P_J(q')\, dq'$, where $q_0 < q < q_1$. The previous analysis shows that $\lim_{N\to\infty} \mathbb{E}\, Y = 1-m$. Show that, in the large N limit, $\mathbb{E}\left(Y^2\right) = \frac{3-5m+2m^2}{3}$, as in the REM. Show that all moments of Y are identical to those of the REM. The result depends only on the 1RSB structure of the saddle point, not on any of its details.

{se:ReplicaExtreme}

8.3 Extreme value statistics and the REM

Exercise 8.9 suggests that there exist universal properties which hold in the glass phase, independently of the details of the model.

In systems with a 1RSB phase, this universality is related to the universality of extreme value statistics. In order to clarify this point, we shall consider in this Section a slightly generalized version of the REM. Here we assume the energy levels to be $M = 2^N$ iid random variables admitting a probability density function (pdf) $P(E)$ with the following properties:

1. $P(E)$ is continuous.
2. $P(E)$ is strictly positive on a semi-infinite domain $-\infty < E \le E_0$.
3. In the $E \to -\infty$ limit, $P(E)$ vanishes more rapidly than any power law. We shall keep here to the simple case in which

$$P(E) \simeq A \exp\left(-B|E|^{\delta}\right) \qquad \text{a}s \quad E \to -\infty\,, \tag{8.50}$$

{eq:gumbel_hyp}

   for some positive constants A, B, δ.

We allow for such a general probability distribution because we want to check which properties of the corresponding REM are universal.

As we have seen in Chap. 5, the low temperature phase of the REM is controlled by a few low-energy levels. Let us therefore begin by computing the distribution of the lowest energy level among $E_1, \dots, E_M$ (we call it $E_{\rm gs}$). Clearly,

$$\mathbb{P}[E_{\rm gs} > E] = \left[\int_E^{\infty} P(x)\; dx\right]^M \;. \tag{8.51}$$

Let $E^*(M)$ be the value of E such that $\mathbb{P}[E_i < E] = 1/M$ for one of the energy levels E_i. For $M \to \infty$, one gets

$$|E^*(M)|^{\delta} = \frac{\log M}{B} + O(\log\log M)\,. \tag{8.52}$$

Let's focus on energies close to $E^*(M)$, such that $E = E^*(M) + \varepsilon/(B\delta|E^*(M)|^{\delta-1})$, and consider the limit $M \to \infty$ with ε fixed. Then:

$$\begin{aligned}\mathbb{P}[E_i > E] &= 1 - \frac{A}{B\delta|E|^{\delta-1}} e^{-B|E|^{\delta}}\,[1+o(1)] = \\ &= 1 - \frac{1}{M}\, e^{\varepsilon}\,[1+o(1)]\;. \end{aligned} \tag{8.53}$$

Therefore, if we define the rescaled ground state energy through $E_{\rm gs} = E^*(M) + \varepsilon_{\rm gs}/(B\delta|E^*(M)|^{\delta-1})$, we get

$$\lim_{N\to\infty} \mathbb{P}[\varepsilon_{\rm gs} > \varepsilon] = \exp\left(-e^{\varepsilon}\right)\;. \tag{8.54}$$

In other words, the pdf of the rescaled ground state energy converges to $P_1(\varepsilon) = \exp(\varepsilon - e^{\varepsilon})$. This limit distribution, known as Gumbel's distribution, is *universal*. The form of the energy level distribution $P(E)$ only enters in the values of the shift and the scale, but not in the form of $P_1(\varepsilon)$. The following exercises show that several other properties of the glass phase in the REM are also universal.

Exercise 8.10 Let $E_1 \le E_2 \le \cdots \le E_k$ be the k lowest energies. Show that universality also applies to the joint distribution of these energies, in the limit $M \to \infty$ at fixed k. More precisely, define the rescaled energies $\varepsilon_1 \le \cdots \le \varepsilon_k$ through $E_i = E^*(M) + \frac{\varepsilon_i}{B\delta|E^*(M)|^{\delta-1}}$. Prove that the joint distribution of $\varepsilon_1, \dots, \varepsilon_k$ admits a density which converges (as $M \to \infty$) to

$$P_k(\varepsilon_1, \dots, \varepsilon_k) = \exp\left(\varepsilon_1 + \cdots + \varepsilon_k - e^{\varepsilon_k}\right)\, \mathbb{I}\left(\varepsilon_1 \le \cdots \le \varepsilon_k\right)\,. \tag{8.55}$$

Exercise 8.11 Consider a REM where the pdf of the energies satisfies the hypotheses 1-3 above, and $M = 2^N$. Show that, in order for the ground state energy to be extensive (i.e. $E_1 \sim N$ in the large N limit), one must have $B \sim N^{1-\delta}$. Show that the system has a phase transition at the critical temperature $T_{\rm c} = \delta\, (\log 2)^{(\delta-1)/\delta}$.

Define the participation ratios $Y_r \equiv \sum_{j=1}^{2^N} p_j^r$. Prove that, for $T < T_{\rm c}$, these quantities signal a condensation phenomenon. More precisely:

$$\lim_{N\to\infty} \mathbb{E}\, Y_r = \frac{\Gamma(r-m)}{\Gamma(r)\Gamma(1-m)}\,, \tag{8.56}$$

where $m = (T/T_{\rm c})\,\min\{\delta, 1\}$, as in the standard REM (see Sec. 8.3). (Hint: One can prove this equality by direct probabilistic means using the methods of Sec. 5.3. For $\delta > 1$, one can also use the replica approach of Sec. 8.1.4).

In the condensed phase only the configurations with low energies count, and because of the universality of their distribution, the moments of the Boltzmann probabilities p_j are universal. These universal properties are also captured by the 1RSB approach. This explains the success of this 1RSB in many systems with a glass phase.

A natural (and fascinating) hypothesis is that higher orders of RSB correspond to different universality classes of extreme values statistics for correlated variables. The mathematical definition of these universality classes have not yet been studied in the mathematical literature, to our knowledge.

{se:repli_app}

8.4 Appendix: Stability of the RS saddle point

In order to establish if a replica saddle point is correct, one widely used criterion is its local stability. In order to explain the basic idea, let us move a step backward and express the replicated free energy as an integral over uniquely the overlap parameters

$$\mathbb{E}\, Z^n \doteq \sum_{Q} e^{N\widehat{G}(Q)}\,. \tag{8.57}$$

Such an expression can either be obtained from Eq. (8.30) by integrating over $\{\lambda_{ab}\}$, or as described in Exercise 8.6. Following the last approach, we get

$$\widehat{G}(Q) = -n\frac{\beta^2}{4} - \frac{\beta^2}{2}\sum_{a<b} Q_{ab}^p - s(Q)\,, \tag{8.58}$$

where

$$s(Q) = -\sum_{a<b}\mu_{ab}Q_{ab} + \psi(\mu)\Big|_{\mu=\mu^*(Q)}\,,\;\; \psi(\mu) = \log\left[\sum_{\{\sigma_a\}} e^{\sum_{a<b}\mu_{ab}\sigma_a\sigma_b}\right] \,, \tag{8.59}$$

and $\mu^*(Q)$ solves the equation $Q_{ab} = \frac{\partial\psi(\mu)}{\partial\mu_{ab}}$. In other words $s(Q)$ is the Legendre transform of $\psi(\mu)$ (apart from an overall minus sign). An explicit expression of $s(Q)$ is not available but we shall only need the following well known property of Legendre transforms

$$\frac{\partial^2 s(Q)}{\partial Q_{ab}\partial Q_{cd}} = -C^{-1}_{(ab)(cd)}\,, \qquad C_{(ab)(cd)} \equiv \frac{\partial^2\psi(\mu)}{\partial\mu_{ab}\partial\mu_{cd}}\Big|_{\mu=\mu^*(Q)}\,, \tag{8.60}$$

where C^{-1} is the inverse of C in matrix sense. The right hand side is in turn easily written down in terms of averages over the replicated system, cf. Eq. (8.34):

$$C_{(ab)(cd)} = \langle\sigma_a\sigma_b\sigma_c\sigma_d\rangle_n - \langle\sigma_a\sigma_b\rangle_n\langle\sigma_c\sigma_d\rangle_n\,. \tag{8.61}$$

Assume now that $(Q^{\rm sp}, \lambda^{\rm sp})$ is a stationary point of $G(Q,\lambda)$. This is equivalent to say that $Q^{\rm sp}$ is a stationary point of $\widehat{G}(Q)$ (the corresponding value of μ coincides with $i\lambda^{\rm sp}$). We would like to estimate the sum (8.57) as $\mathbb{E}Z^n \doteq e^{N\widehat{G}(Q^{\rm sp})}$. A necessary condition for this to be correct is that the matrix of second derivatives of $\widehat{G}(Q)$ is positive semidefinite at $Q = Q^{\rm sp}$. This is referred to as the **local stability** condition. Using Eqs. (8.58) and (8.61), we get the explicit condition

$$M_{(ab)(cd)} \equiv \left[-\frac{1}{2}\beta^2 p(p-1)Q_{ab}^{p-2}\,\delta_{(ab),(cd)} + C^{-1}_{(ab)(cd)}\right] \succeq 0\,, \tag{8.62}$$

where we use the symbol $A \succeq 0$ to denote that the matrix A is positive semidefinite.

In this technical appendix we sketch this computation in two simple cases: the stability of the RS saddle point for the general p-spin glass in zero magnetic field, and the SK model in a field.

We consider first the RS saddle point $Q_{ab} = 0,\ \lambda_{ab} = 0$ in the p-spin glass. In this case

$$\langle f(\sigma)\rangle_n = \frac{1}{2^n}\sum_{\{\sigma_a\}} f(\sigma)\,. \tag{8.63}$$

It is then easy to show that $M_{(ab)(cd)} = \delta_{(ab),(cd)}$ for $p \geq 3$ and $M_{(ab)(cd)} = (1-\beta^2)\delta_{(ab),(cd)}$ for $p = 2$. The situations for $p = 2$ and $p \geq 3$ are very different:

- If $p = 2$ (the SK model) the RS solution is stable for $\beta < 1$, and unstable for $\beta > 1$.
- When $p \geq 3$, the RS solution is always stable.

Let us now look at the SK model in a magnetic field. This is the $p = 2$ case but with an extra term $-B \sum_i \sigma_i$ added to the energy (8.24). It is straightforward to ⋆ repeat all the replica computations with this extra term. The results are formally identical if the average within the replicated system (8.34) is changed to:

$$\langle f(\sigma) \rangle_{n,B} \equiv \frac{1}{z(\mu)} \sum_{\{\sigma^a\}} f(\sigma) \; \exp\left(\sum_{a<b} \mu_{ab}\, \sigma_a \sigma_b + \beta B \sum_a \sigma_a \right) \tag{8.64}$$

$$z(\mu) \equiv \sum_{\{\sigma^a\}} \; \exp\left(\sum_{a<b} \mu_{ab}\, \sigma_a \sigma_b + \beta B \sum_a \sigma_a \right) . \tag{8.65}$$

The RS saddle point equations (8.35) are changed to:

{eq:sk_speq_rs}
$$\mu = \beta^2 \, q \, , \qquad q = \mathsf{E}_z \tanh^2 (z\sqrt{\mu} + \beta B) \; . \tag{8.66}$$

and the values of q, μ are non-zero at any positive β, when $B \neq 0$. This complicates the stability analysis.

Since $p = 2$, we have $M_{(ab)(cd)} = -\beta^2 \delta_{(ab)(cd)} + C^{-1}_{(ab)(cd)}$. Let $\{\lambda_j\}$ be the eigenvalues of $C_{(ab)(cd)}$. Since $C \succeq 0$, the condition $M \succeq 0$ is in fact equivalent to $1 - \beta^2 \lambda_j \geq 0$, for all the eigenvalues λ_j.

The matrix elements $C_{(ab)(cd)}$ take three different forms, depending on the number of common indices in the two pairs (ab), (cd):

$$\begin{aligned}
C_{(ab)(ab)} &= 1 - \left[\mathsf{E}_z \tanh^2 (z\sqrt{\mu} + \beta B)\right]^2 \equiv U \\
C_{(ab)(ac)} &= \mathsf{E}_z \tanh^2 (z\sqrt{\mu} + \beta B) - \left[\mathsf{E}_z \tanh^2 (z\sqrt{\mu} + \beta B)\right]^2 \equiv V \\
C_{(ab)(cd)} &= \mathsf{E}_z \tanh^4 (z\sqrt{\mu} + \beta B) - \left[\mathsf{E}_z \tanh^2 (z\sqrt{\mu} + \beta B)\right]^2 \equiv W \; ,
\end{aligned}$$

where $b \neq c$ is assumed in the second line, and all indices are distinct in the last line. We want to solve the eigenvalue equation $\sum_{(cd)} C_{(ab)(cd)} x_{cd} = \lambda x_{(ab)}$.

A first eigenvector is the uniform vector $x_{(ab)} = x$. Its eigenvalue is $\lambda_1 = U + 2(n-2)V + (n-2)(n-3)/2W$. Next we consider eigenvectors which depend on one special value θ of the replica index in the form: $x_{(ab)} = x$ if $a = \theta$ or $b = \theta$, and $x_{(ab)} = y$ in all other cases. Orthogonality to the uniform vector is enforced by choosing $x = (1 - n/2)y$, and one finds the eigenvalue $\lambda_2 = U + (n-4)V + (3-n)W$. This eigenvalue has degeneracy $n - 1$. Finally we consider eigenvectors which depend on two special values θ, ν of the replica index: $x_{(\theta,\nu)} = x$, $x_{(\theta,a)} = x_{(\nu,a)} = y$, $x_{(ab)} = z$, where a and b are distinct form θ, ν. Orthogonality to the previously found eigenvectors imposes $x = (2-n)y$ and $y = [(3-n)/2]z$. Plugging this into the eigenvalue equation, one gets the eigenvalue $\lambda_3 = U - 2V + W$, with degeneracy $n(n-3)/2$.

In the limit $n \to 0$, the matrix C has two distinct eigenvalues: $\lambda_1 = \lambda_2 = U - 4V + 3W$ and $\lambda_3 = U - 2V + W$. Since $V \geq W$, the most dangerous eigenvalue is λ_3 (called the **replicon eigenvalue**). This implies that the RS solution of the SK model is locally stable if and only if

$$\mathsf{E}_z \left[1 - \tanh^2\left(z\sqrt{\mu} + \beta B\right)\right]^2 \leq T^2 \tag{8.67}$$

The inequality is saturated on line in the plane T, B, called the **AT line**. which behaves like $T = 1 - \left(\frac{3}{4}\right)^{2/3} B^{2/3} + o(B^{2/3})$ for $B \to 0$ and like $T \simeq \frac{4}{3\sqrt{2\pi}} e^{-B^2/2}$ for $B \gg 1$.

{ex:SK_J0}

Exercise 8.12 The reader who wants to test her understanding of these replica computations computation can study the SK model in zero field ($B = 0$), but in the case where the couplings have a ferromagnetic bias: J_{ij} are iid Gaussian distributed, with mean J_0/N and variance $1/N$.

(i) Show that the RS equations (8.35) are modified to:

$$\mu = \beta^2\, q \;\; ; \;\; q = \mathsf{E}_z \tanh^2\left(z\sqrt{\mu} + \beta J_0 m\right) \;\; ; \;\; m = \mathsf{E}_z \tanh\left(z\sqrt{\mu} + \beta J_0 m\right) \tag{8.68}$$

{eq:SK_J0_RS_SP}

(ii) Solve numerically these equations. Notice that, depending on the values of T and J_0, three types of solutions can be found: (1) a paramagnetic solution $m = 0, q = 0$, (2) a ferromagnetic solution $m > 0, q > 0$, (3) a spin glass solution $m = 0, q > 0$.

(iii) Show that the AT stability condition becomes:

$$\mathsf{E}_z \left[1 - \tanh^2\left(z\sqrt{\mu} + \beta J_0 m\right)\right]^2 < T^2 \tag{8.69}$$

{eq:SK_J0_RS_AT}

and deduce that the RS solution found in (i), (ii) is stable only in the paramagnetic phase and in a part of the ferromagnetic phase.

Notes

The replica solution of the REM was derived in the original work of Derrida introducing the model (Derrida, 1980; Derrida, 1981). His motivation for introducing the REM came actually from the large p limit of p-spin glasses.

The problem of moments is studied for instance in (Shohat and Tamarkin, 1943).

The first universally accepted model of spin glasses was introduced by Edwards and Anderson (Edwards and Anderson, 1975). The mean field theory was defined by Sherrington and Kirkpatrick (Sherrington and Kirkpatrick, 1975; Kirkpatrick and Sherrington, 1978), who considered the RS solution. The instability of this solution in the $p = 2$ case was found by de Almeida and Thouless (de Almeida and Thouless, 1978), who first computed the location of the AT line. The solution to exercise (8.12) can be found in (Kirkpatrick and Sherrington, 1978; de Almeida and Thouless, 1978).

Parisi's Ansatz was introduced in a couple of very inspired works starting in 1979 (Parisi, 1979; Parisi, 1980*b*; Parisi, 1980*a*). His original motivation came from his reflection on the meaning of the permutation group S_n when $n < 1$, and particularly in the $n \to 0$ limit. Unfortunately there has not been any mathematical developments along these lines. The replica method, in the presence of RSB, is still waiting for a proper mathematical framework. On the other hand it is a very well defined computational scheme, which applies to a wide variety of problems. The physical interpretation of RSB in terms of condensation was found by Parisi (Parisi, 1983), and developed in (Mézard, Parisi, Sourlas, Toulouse and Virasoro, 1985), which discussed the distribution of weights in the glass phase and its ultrametric organization. The p-spin model has been analyzed at large p with replicas in (Gross and Mézard, 1984). The clustering phenomenon has been discovered in this work. The finite p case was later studied in (Gardner, 1985). A rigorous treatment of the clustering effect in the p-spin glass model was developed by Talagrand (Talagrand, 2000) and can be found in his book (Talagrand, 2003).

The connection between 1RSB and Gumbel's statistics of extremes is discussed in (Bouchaud and Mézard, 1997). A more detailed presentation of the replica method, together with some reprints of most of these papers, can be found in (Mézard, Parisi and Virasoro, 1987).