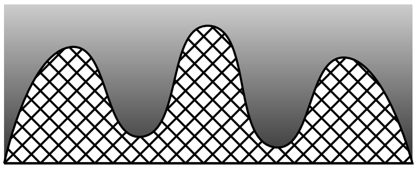
4 Frozen disorder in physical systems

A glass is essentially a frozen fluid in which the atoms are disordered as in a liquid, but they do not move, as in a solid.

G. Parisi[1]

4a Atmosphere over a random relief



The density of the air on the sea-level is about 1.2 kg/m^3, but 1 km higher it is about 1.1 kg/m^3. All that results from an equilibrium between the pressure and the weight of the air. We use, on one hand, the equation of state $p = C\rho$, where p is the pressure, ρ is the density, and C is a constant (about $83 \cdot 10^3$ m^2/sec^2); on the other hand, the equilibrium equation $g\rho(z) = -p'(z)$, where z is the altitude, and g the acceleration of gravity on Earth (about 9.8 m/sec^2). We get the differential equation $\rho'(z) = -(g/C)\rho(z)$ and its solution $\rho(z) = \text{const} \cdot \mathrm{e}^{-(g/C)z}$. We may use C/g (about 8.4 km) as the unit of altitude, then[2]

$$\rho(z) = \text{const} \cdot \mathrm{e}^{-z} .$$

However, it holds in the domain

$$z > H(x, y) ;$$

[1]See page 298 of the book "Field theory, disorder and simulations", World Scientific, 1992.

[2]In fact, a crude approximation, because the temperature of the air depends on the altitude.

the function H describes the relief.

Projecting the 3-dimensional distribution of (the mass of) the air to the horizontal plane (x, y) we get the 2-dimensional measure

$$\mathrm{const} \cdot \mathrm{e}^{-H(x,y)} \, \mathrm{d}x\mathrm{d}y \, .$$

The constant is determined by the relation $\mathrm{const} \cdot \iint \mathrm{e}^{-H(x,y)} \, \mathrm{d}x\mathrm{d}y = m$, where m is the total mass of the air. In order to keep m finite we either restrict ourselves to a bounded region of the plane or assume that $H(x, y) \to \infty$ (not too slowly) as $x^2 + y^2 \to \infty$. Thus, we consider the measure

$$\text{(4a1)} \qquad \begin{gathered} \frac{m}{Z_H} \mathrm{e}^{-H(x,y)} \, \mathrm{d}x\mathrm{d}y = m \mathrm{e}^{-(H(x,y)-F_H)} \, \mathrm{d}x\mathrm{d}y \, , \\ \text{where} \quad Z_H = \iint \mathrm{e}^{-H(x,y)} \, \mathrm{d}x\mathrm{d}y \quad \text{and} \quad F_H = -\ln Z_H \, . \end{gathered}$$

4a2 Exercise. Prove the Lipschitz property of F,

$$|F_{H_1} - F_{H_2}| \le \sup_{x,y} |H_1(x, y) - H_2(x, y)| \, .$$

Hint: recall 2a3 and 2d7.

4a3 Exercise. Let H be a finite-dimensional Gaussian random function,

$$H(x, y, \omega) = f_0(x, y) + \zeta_1(\omega) f_1(x, y) + \dots + \zeta_n(\omega) f_n(x, y) \, ,$$

where $(\zeta_1, \dots, \zeta_n)$ is an orthogaussian sequence, and $f_0, \dots, f_n$ are bounded continuous functions on a bounded region of $\mathbb{R}^2$. Prove that the random variable $(1/\sigma_{\max}) F_H$ is more concentrated than $\mathrm{N}(0, 1)$; here

$$\sigma_{\max} = \sup_{x,y} \sqrt{f_1^2(x, y) + \dots + f_n^2(x, y)} \, .$$

Hint: recall the phrase after 2d7; use 4a2.

4b A molecule over a random relief

The air is a mix of gases, but still, it is convenient to speak about a 'molecule of air' of the averaged mass $m_1 = 48 \cdot 10^{-27}\,\mathrm{kg}$ (29 atomic mass units). A randomly chosen molecule of air is distributed in the domain $\{(x, y, z) : z > H(x, y)\}$ according to the measure $\frac{1}{Z_H} \mathrm{e}^{-z} \, \mathrm{d}x\mathrm{d}y\mathrm{d}z = \mathrm{e}^{-(z-F_H)} \, \mathrm{d}x\mathrm{d}y\mathrm{d}z$, which follows from formulas of 4a derived for the air treated as a continuous medium (rather than a system of molecules). Recall that our unit of the altitude z is 8.4 km. Being at this altitude, the molecule has the potential energy $m_1 gh = 48 \cdot 10^{-27}\,\mathrm{kg} \cdot 9.8\,\mathrm{m/sec}^2 \cdot 8.4 \cdot 10^3\,\mathrm{m} = 4.0 \cdot 10^{-21}\,\mathrm{J}$ equal to $k_{\mathrm{B}}T$, where $T = 288\,\mathrm{K}$ is the temperature and $k_{\mathrm{B}} = 1.38 \cdot 10^{-23}\,\mathrm{J/K}$ is the so-called Boltzmann constant. That is not a coincidence! Thermal motion endows each degree of freedom with the energy $k_{\mathrm{B}}T$ (on the average).[1] No need to choose the molecule at random; each molecule moves randomly and visits all locations according to the distribution $\frac{1}{Z_H}\mathrm{e}^{-z}\,\mathrm{d}x\mathrm{d}y\mathrm{d}z$.

The Boltzmann constant is not specific for the air (nor even for gases), it is one of the universal physical constants. Nevertheless we can get our special altitude 8.4 km as $(k_{\mathrm{B}}T)/(m_1 g)$. Here we use a parameter of the molecule (m_1), Earth (g) and the air (T). However, the temperature is the only relevant parameter of the air! The constant $C = 83 \cdot 10^3\,\mathrm{m}^2/\mathrm{sec}^2$ used in 4a for describing the air as a continuous medium is no more needed. What does it mean?

It means that we face a very general physical principle. The air is just a 'heat bath' (reservoir), — a large system at some temperature T. The molecule is in (thermal) equilibrium with the heat bath. This is enough in order to determine uniquely statistical properties of the molecule! We could put the single molecule to a high vessel at the temperature $T = 288\,\mathrm{K}$ and get the same distribution ($\mathrm{const} \cdot \mathrm{e}^{-z}\,\mathrm{d}x\mathrm{d}y\mathrm{d}z$).[2]

4c The meaning of temperature

A finite quantum system (for instance, a finite system of interacting spins) is described by a Hermitian operator H (the Hamiltonian) on an n-dimensional Hilbert space; its eigenvalues $E_1, \dots, E_n \in \mathbb{R}$ are energy levels.[3] The tiny portion of quantum theory, needed here, deals only with commuting operators that may be represented by diagonal matrices,

$$H = \mathrm{diag}(E_1, \dots, E_n)\,.$$

The very general physical principle mentioned in 4b states the following. If the system is in thermal equilibrium with a heat bath at temperature T, then its energy takes on each value E_k with the corresponding probability[4]

$$\mathrm{const} \cdot \exp\Big(-\frac{E_k}{k_{\mathrm{B}}T} \Big)\,;$$

[1] However, see 4c9, 4c10.

[2] The vessel is void of gas, but full of infrared radiation.

[3] The time evolution of a state vector (of isolated system) is described by the Schrödinger equation $\mathrm{i}\hbar\psi'(t) = H\psi(t)$ and its solution $\psi(t) = \exp(-\mathrm{i}tH/\hbar)\psi(0)$, but we do not need it.

[4] In the language of operators, the (mixed) state of the system is described by the density matrix $\mathrm{const} \cdot \exp\big(-\frac{H}{k_{\mathrm{B}}T}\big)$, but we do not need it.

of course, $1/\text{const} = \sum_k \exp\bigl(-\frac{E_k}{k_B T}\bigr)$.

In other words, the energy distribution is

$$\text{const} \cdot \exp\Bigl(-\frac{E}{k_B T}\Bigr)\mu(\mathrm{d}E)\,,$$

where $\mu\bigl(\{E\}\bigr) = \#\{k : E_k = E\}$ or, if you prefer, $\mu\bigl(\{E\}\bigr) = \#\{k : E_k = E\}/n$ (the coefficient is anyway absorbed by 'const'). For large quantum systems, the discrete energy spectrum μ is often approximated by a continuous measure. Especially, classical systems are treated in this way.

4c1 Definition. Let μ be a finite or locally finite measure on $[0,\infty)$ such that $\int \exp(-\varepsilon E)\,\mu(\mathrm{d}E) < \infty$ for all $\varepsilon > 0$. For any $T \in (0,\infty)$, the *Gibbs measure* $\tilde G$ on $[0,\infty)$ (corresponding to T and μ) is[1]

$$\tilde G(\mathrm{d}E) = \frac{1}{Z}\exp\Bigl(-\frac{E}{k_B T}\Bigr)\mu(\mathrm{d}E)\,,$$
$$\text{where}\quad Z = \int \exp\Bigl(-\frac{E}{k_B T}\Bigr)\mu(\mathrm{d}E)\,.$$

Often μ is the image of another measure ν on another space X under a map $H : X \to [0,\infty)$. Then the measure

$$\frac{1}{Z}\exp\Bigl(-\frac{H(x)}{k_B T}\Bigr)\nu(\mathrm{d}x)$$

on X is also called the Gibbs measure. (Its image under H is the measure $\tilde G$ of 4c1.)

Instead of T it is convenient to use the *inverse temperature*

$$\beta = \frac{1}{k_B T}\,.$$

Thus, the so-called *partition function*

$$Z_\beta = \int \mathrm{e}^{-\beta E}\mu(\mathrm{d}E) = \int \mathrm{e}^{-\beta H}\,\mathrm{d}\nu$$

is the Laplace transform of μ (therefore, a smooth decreasing[2] function of β), and the Gibbs measure is[3]

$$G_\beta = \frac{1}{Z_\beta}\mathrm{e}^{-\beta H}\nu = \mathrm{e}^{-\beta(H-F_\beta)}\nu\,,\quad \tilde G_\beta(\mathrm{d}E) = \frac{1}{Z_\beta}\mathrm{e}^{-\beta E}\mu(\mathrm{d}E) = \mathrm{e}^{-\beta(E-F_\beta)}\mu(\mathrm{d}E)\,,$$
$$\text{where}\quad F_\beta = -\frac{1}{\beta}\ln Z_\beta\,.$$

[1] Later it will be denoted $\tilde G_\beta$.

[2] And logarithmically convex, see 4c7.

[3] F_β is called *free energy* (do not ask me, why).

Note especially the original case:

$$
\begin{gathered}
X = \{1, \dots, n\}, \quad H(k) = E_k, \quad \nu(\{k\}) = 1; \\
G_\beta(\{k\}) = \frac{1}{Z_\beta} \mathrm{e}^{-\beta E_k}, \quad Z_\beta = \sum_k \mathrm{e}^{-\beta E_k}; \\
\mu(\{E\}) = \#\{k : E_k = E\}; \quad \tilde G_\beta(\{E\}) = \frac{1}{Z_\beta} \mathrm{e}^{-\beta E} \mu(\{E\}).
\end{gathered}
\tag{4c2}
$$

Derivatives of Z_β (or F_β) are of interest, as we will see now.

4c3 Exercise. The *mean energy* (called also internal energy)

$$
U_\beta = \int E\, \tilde G_\beta(\mathrm{d}E) = \int H \,\mathrm{d}G_\beta
$$

may be calculated as

$$
U_\beta = -\frac{\mathrm{d}}{\mathrm{d}\beta} \ln Z_\beta = \frac{\mathrm{d}}{\mathrm{d}\beta} (\beta F_\beta).
$$

Prove it.

The *specific heat,* defined by[1]

$$
c_\beta = \frac{\mathrm{d}}{\mathrm{d}T} U_{1/(k_{\mathrm{B}} T)},
$$

may be calculated as

$$
c_\beta = -k_{\mathrm{B}} \beta^2 \frac{\mathrm{d}}{\mathrm{d}\beta} U_\beta = k_{\mathrm{B}} \beta^2 \frac{\mathrm{d}^2}{\mathrm{d}\beta^2} \ln Z_\beta = -k_{\mathrm{B}} \beta^2 \frac{\mathrm{d}^2}{\mathrm{d}\beta^2} (\beta F_\beta).
$$

4c4 Exercise. Prove that

$$
\int (H - U_\beta)^2 \,\mathrm{d}G_\beta = \frac{c_\beta}{k_{\mathrm{B}} \beta^2}.
$$

Hints: first, $\int (H - U_\beta)^2 \,\mathrm{d}G_\beta = \int H^2 \,\mathrm{d}G_\beta - \bigl(\int H \,\mathrm{d}G_\beta \bigr)^2$; second, $\frac{\mathrm{d}}{\mathrm{d}\beta} U_\beta = \frac{\mathrm{d}}{\mathrm{d}\beta} \frac{\int H \mathrm{e}^{-\beta H} \,\mathrm{d}\nu}{\int \mathrm{e}^{-\beta H} \,\mathrm{d}\nu} = \dots$

In the original case (4c2), the *entropy* of G_β is

$$
S_\beta = -\sum_{k=1}^n G_\beta(\{k\}) \ln G_\beta(\{k\}) = -\sum_{E : \exists k\, E_k = E} \tilde G_\beta(\{E\}) \ln \frac{\tilde G_\beta(\{E\})}{\mu(\{E\})}.
$$

[1] In physics it is often denoted c_V.

In general we may define the entropy by

$$S_\beta = -\int_X \Big(\ln \frac{\mathrm{d}G_\beta}{\mathrm{d}\nu} \Big) \, \mathrm{d}G_\beta = \int_X (\beta H + \ln Z_\beta) \, \mathrm{d}G_\beta = \\ = \int_{[0,\infty)} (\beta E + \ln Z_\beta) \, \tilde G_\beta(\mathrm{d}E) = -\int_{[0,\infty)} \Big(\ln \frac{\mathrm{d}\tilde G_\beta}{\mathrm{d}\mu} \Big) \, \mathrm{d}\tilde G_\beta \, .$$

4c5 Exercise. Prove that[1]

$$S_\beta = \beta(U_\beta - F_\beta) = \frac{U_\beta - F_\beta}{k_\mathrm{B} T} \, .$$

Hint: $S_\beta = Z_\beta^{-1} \int (\beta H + \ln Z_\beta) \mathrm{e}^{-\beta H} \, \mathrm{d}\nu = \dots$

By the way, if we replace μ, ν with $c\mu, c\nu$ for some $c \in (0,\infty)$ then $\tilde G_\beta$, G_β, U_β and c_β remain intact, but Z_β, F_β and S_β change. On the other hand, an energy shift $H \mapsto H + \mathrm{const}$ leads to the same shift of F_β and U_β, but leaves intact $\tilde G_\beta$, G_β, S_β and c_β.

4c6 Exercise. For every $\beta \in (0,\infty)$,

$$\nu\{x : H(x) \le U_\beta\} \le \exp S_\beta \, .$$

Prove it.

Hint: $\nu\{x : H(x) \le E\} \le \big(\int \mathrm{e}^{-\beta H} \, \mathrm{d}\nu \big) / \big(\mathrm{e}^{-\beta E} \big)$.

4c7 Exercise. The function $\beta \mapsto \ln Z_\beta$ is convex.

Prove it.

Hint: $\int \sqrt{\mathrm{e}^{-\beta_1 H}} \sqrt{\mathrm{e}^{-\beta_2 H}} \, \mathrm{d}\nu \le \sqrt{\int \mathrm{e}^{-\beta_1 H} \, \mathrm{d}\nu} \sqrt{\int \mathrm{e}^{-\beta_2 H} \, \mathrm{d}\nu}$.

It follows immediately that

$$\begin{gathered} \text{the function } \beta \mapsto \beta F_\beta \text{ is concave;} \\ \text{the function } \beta \mapsto U_\beta \text{ is decreasing;} \\ c_\beta \ge 0 \, . \end{gathered} \tag{4c8}$$

4c9 Exercise. [2] (a) Let $X = [0,\infty)^n$, $H(x_1,\dots,x_n) = x_1 + \dots + x_n$, $\nu = \mathrm{mes}_n |_X$. Prove that $U_\beta = n/\beta = n k_\mathrm{B} T$ and $c_\beta = n k_\mathrm{B}$.

(b) Generalize it for any strictly positive linear form H on $[0,\infty)^n$.

Hint: reduce the general case to $n = 1$.

[1] In physics, the entropy is multiplied by k_B; it is rather $(U - F)/T$.

[2] Think about the altitudes of the molecules.

4c10 Exercise. [1] (a) Let $X = \mathbb{R}^n$, $H(x_1, \dots, x_n) = \frac{1}{2}(x_1^2 + \dots + x_n^2)$, $\nu = \operatorname{mes}_n$. Prove that $U_\beta = \frac{n}{2\beta} = \frac{n}{2} k_{\mathrm{B}} T$ and $c_\beta = \frac{n}{2} k_{\mathrm{B}}$.

(b) Generalize it for any strictly positive quadratic form H on $\mathbb{R}^n$.

Hint: reduce (b) to (a), and (a) to $n = 1$. Or alternatively: G_β is a Gaussian measure, $Z_\beta = (\mathrm{const}/\beta)^{n/2}$.

(You may also try $H(x) = |x|^\alpha$ for $x \in \mathbb{R}$.)

4c11 Exercise. [2] Let $X = \{0, 1, 2, \dots\}$, $H(x) = x$, $\nu(\{x\}) = 1$ for all x. Prove that $U_\beta = 1/(\mathrm{e}^\beta - 1)$ and $c_\beta = k_{\mathrm{B}} \beta^2 \mathrm{e}^\beta / (\mathrm{e}^\beta - 1)^2$. Generalize it for $X = \{0, 1, 2, \dots\}^n$.

Note that $c_\beta \to n k_{\mathrm{B}}$ as $\beta \to 0$ (high temperature), but $c_\beta \to 0$ as $\beta \to \infty$ (low temperature); compare it with 4c9(a).

4c12 Exercise. [3] Let $X = \{0, 1\}$, $H(x) = x$, $\nu(\{x\}) = 1$ for $x = 0, 1$. Prove that $U_\beta = 1/(\mathrm{e}^\beta + 1)$ and $c_\beta = k_{\mathrm{B}} \beta^2 \mathrm{e}^\beta / (\mathrm{e}^\beta + 1)^2$. Generalize it for $X = \{0, 1\}^n$.

This time, $c_\beta \to 0$ in both limits (high and low temperature).

First and second moments of $\tilde{G}_\beta$ are used in 4c3–4c5; more generally, Z_β gives us all moments of $\tilde{G}_\beta$ via the moment generating function, as follows.

4c13 Exercise. (a) For every $\lambda \in (-\infty, \beta)$,

$$\int \mathrm{e}^{\lambda H} \, \mathrm{d}G_\beta = \int \mathrm{e}^{\lambda E} \, \tilde{G}_\beta(\mathrm{d}E) = \frac{Z_{\beta - \lambda}}{Z_\beta} \, .$$

(b) For every $\beta > 0$, the measure $\tilde{G}_\beta$ has all moments, and

$$\int H^m \, \mathrm{d}G_\beta = \int E^m \, \tilde{G}_\beta(\mathrm{d}E) = \frac{(-1)^m}{Z_\beta} \frac{\mathrm{d}^m}{\mathrm{d}\beta^m} Z_\beta \, .$$

(c) Basically, 4c3 and 4c4 are special cases of (b) for $m = 1, 2$.

Prove it. Does (b) hold for $\beta = 0$?

4d The random energy model

The simplest nontrivial Gaussian process H consists of independent random variables:

$H(1), \dots, H(n)$ are orthogaussian; $\quad X = \{1, \dots, n\}, \quad \nu(\{x\}) = \frac{1}{n}$ for all x.

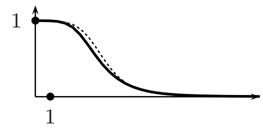
[1] Applicable to classical harmonic oscillators.

[2] Applicable to quantum harmonic oscillators.

[3] Applicable to quantum spins.

Negative values of H do not fit to the framework of 4c, which is harmless as far as H is bounded from below; recall that an energy shift $H \mapsto H + \text{const}$ does not influence G_β and c_β.

For large n we have $\mu \approx \gamma^1 = \mathrm{N}(0,1)$, thus we may expect $Z_\beta \approx \int_{-\infty}^{+\infty} \mathrm{e}^{-\beta E} \, \gamma^1(\mathrm{d}E) = \exp\bigl(\frac12 \beta^2\bigr)$, $\tilde G_\beta \approx \mathrm{N}(-\beta, 1)$, $U_\beta \approx -\beta$, $-\frac{\mathrm{d}}{\mathrm{d}\beta} U_\beta \approx 1$ and $c_\beta \approx k_{\mathrm{B}} \beta^2$. Here is a simulation:



the graph of $\beta \mapsto -\frac{\mathrm{d}}{\mathrm{d}\beta} U_\beta$
$n = 1\,000\,000$
two samples

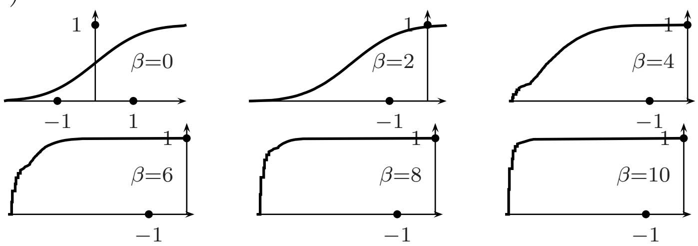
The naive expectation is confirmed for small β (high temperature) but refuted for large β (low temperature). Why?

Taking into account that

$$-\frac{\mathrm{d}}{\mathrm{d}\beta} U_\beta = \int (E - U_\beta)^2 \, \tilde G_\beta(\mathrm{d}E) \tag{4d1}$$

we want to look at $\tilde G_\beta$.

(4d2)



the c.d.f. of $\tilde G_\beta$
$n = 1\,000\,000$
one sample

For small β we see $\tilde G_\beta \approx \mathrm{N}(-\beta, 1)$, indeed. However, for large β this cannot happen, since $\tilde G_\beta$ cannot pass the left endpoint of the sample!

The calculation

$$\begin{aligned} \mathbb{E}\, Z_\beta &= \tfrac1n \mathbb{E}\left(\mathrm{e}^{-\beta H(1)} + \dots + \mathrm{e}^{-\beta H(n)}\right) = \mathbb{E}\, \mathrm{e}^{-\beta H(1)} = \\ &= \frac{1}{\sqrt{2\pi}} \int \mathrm{e}^{-\beta E} \mathrm{e}^{-E^2/2} \, \mathrm{d}E = \mathrm{e}^{\beta^2/2} \underbrace{\frac{1}{\sqrt{2\pi}} \int \exp\bigl(-\tfrac12 (E+\beta)^2\bigr) \, \mathrm{d}E}_{=1} \end{aligned}$$

gives us more than just the value of $\mathbb{E}\, Z_\beta$. It also shows that only values $E = -\beta + O(1)$ are essential. Roughly, $Z_\beta \approx \frac1n \mathrm{e}^{\beta^2} K$, where $K = \#\{k : H(k) = -\beta + O(1)\}$ is distributed binomially, $\mathrm{Binom}(n,p)$, $p \approx \mathrm{e}^{-\beta^2/2}$. Two cases are very different: $np \gg 1$ and $np \ll 1$.

The first case: $np \gg 1$. Here, $K \approx np$, since the mean square deviation $\sqrt{np(1-p)}$ of K is much less than np. Thus, $Z_\beta \approx \frac1n \mathrm{e}^{\beta^2} n \mathrm{e}^{-\beta^2/2} = \mathrm{e}^{\beta^2/2}$ (not just in the mean).

The second case: $np \ll 1$. Here, with probability close to 1, $K = 0$ and $Z_\beta \approx 0$. The large mean value $\mathbb{E}\, Z_\beta = \mathrm{e}^{\beta^2/2}$ is supported by rare events of very large Z_β.

Unfortunately, these arguments are sloppy. The interval $-\beta + O(1)$ is too large; both functions, $\mathrm{e}^{-\beta E}$ and $\Phi'(E)$ are too far from being constant on this interval. See below for a correct version. First we deal with the case $n\mathrm{e}^{-\beta^2/2} \gg 1$.

4d3 Exercise. For any $E \in \mathbb{R}$, the random variable

$$Z_\beta(-E,\infty) = \frac{1}{n} \sum_{k:H(k)\ge -E} \mathrm{e}^{-\beta H(k)}$$

has the following expectation and mean square deviation:

$$\mathbb{E}\, Z_\beta(-E,\infty) = \mathrm{e}^{\beta^2/2}\Phi(E-\beta)\,,$$

$$\sigma\big(Z_\beta(-E,\infty)\big) = \frac{1}{\sqrt{n}}\sqrt{\mathrm{e}^{2\beta^2}\Phi(E-2\beta) - \mathrm{e}^{\beta^2}\Phi^2(E-\beta)} \le \frac{1}{\sqrt{n}}\mathrm{e}^{\beta^2}\sqrt{\Phi(E-2\beta)}\,;$$

therefore

$$\frac{\sigma\big(Z_\beta(-E,\infty)\big)}{\mathbb{E}\, Z_\beta(-E,\infty)} \le \frac{\mathrm{e}^{\beta^2/2}}{\sqrt{n}} \frac{\sqrt{\Phi(E-2\beta)}}{\Phi(E-\beta)}\,.$$

Prove it.

Hint: $Z_\beta(-E,\infty)$ is the sum of n independent random variables; reduce the general case to $n = 1$, then calculate the integrals.

Taking $E \to +\infty$ we see that Z_β is nearly non-random provided that $\mathrm{e}^{\beta^2} \ll n$. However, $\mathrm{e}^{\beta^2/2} \ll n$ should be enough! Truncation will help.

4d4 Exercise. Let $\beta_1, \beta_2, \dots$ satisfy

$$\frac{\beta_n}{\sqrt{2\ln n}} \to 1 \quad \text{and} \quad \sqrt{2\ln n} - \beta_n \to +\infty\,.$$

Then

$$\frac{\sigma\big(Z_{\beta_n}(-\beta_n - a,\infty)\big)}{\mathbb{E}\, Z_{\beta_n}(-\beta_n - a,\infty)} \to 0$$

for any $a \in [0,\infty)$. Moreover,

$$\frac{\sigma\big(Z_{\beta_n}(-\beta_n - a_n,\infty)\big)}{\mathbb{E}\, Z_{\beta_n}(-\beta_n - a_n,\infty)} \to 0$$

for any $a_1, a_2, \dots \in [0,\infty)$ satisfying $a_n = o(\sqrt{2\ln n} - \beta_n)$.

Prove it.

Hint. Let $c_n = \sqrt{2\ln n} - \beta_n$, then $\frac{1}{n}\mathrm{e}^{\beta_n^2}\Phi(a_n - \beta_n) \ll \exp\big((\sqrt{2\ln n} - c_n)^2 - \frac{1}{2}(\sqrt{2\ln n} - c_n - a_n)^2 - \ln n\big) = \exp\big(-c_n\sqrt{2\ln n} + o(c_n\sqrt{2\ln n})\big) \to 0$.

The condition $\beta_n \sim \sqrt{2 \ln n}$ is not essential.

4d5 Exercise. Let $\beta_1, \beta_2, \dots$ satisfy

$$\sqrt{2 \ln n} - \beta_n \to +\infty \,. \tag{4d6}$$

Then

$$\frac{\sigma\big(Z_{\beta_n}(-\beta_n - a_n, \infty) \big)}{\mathbb{E}\, Z_{\beta_n}(-\beta_n - a_n, \infty)} \to 0$$

for any $a_1, a_2, \dots \in [0, \infty)$ satisfying $a_n = o(\sqrt{2 \ln n} - \beta_n)$.

Prove it.

Hint. The function $\beta \mapsto \beta^2 - \frac12 (\beta - a)^2$ is increasing; choose c_n such that $\beta_n \le \sqrt{2 \ln n} - c_n$, $c_n \to \infty$, $c_n = o(\sqrt{\ln n})$ and $a_n = o(c_n)$. (Try $c_n = \sqrt{(1 + a_n)(\sqrt{2 \ln n} - \beta_n)}$.)

It follows that

$$\Bigg\| \frac{Z_{\beta_n}(-\beta_n - a_n, \infty)}{\mathbb{E}\, Z_{\beta_n}(-\beta_n - a_n, \infty)} - 1 \Bigg\|_2 \to 0 \,.$$

4d7 Exercise. Let $\beta_1, \beta_2, \dots$ satisfy (4d6). Then

$$\bigg\| \frac{Z_{\beta_n}}{\mathbb{E}\, Z_{\beta_n}} - 1 \bigg\|_1 \to 0 \,, \quad \text{that is,} \quad \| \mathrm{e}^{-\beta_n^2/2} Z_{\beta_n} - 1 \|_1 \to 0 \,.$$

Prove it.

Hint: choose $a_n \to \infty$ such that $0 \le a_n = o(\sqrt{2 \ln n} - \beta_n)$, then

$$\bigg\| \frac{Z_{\beta_n} - Z_{\beta_n}(-\beta_n - a_n, \infty)}{\mathbb{E}\, Z_{\beta_n}} \bigg\|_1 = \Phi(-a_n) \to 0 \,;$$

use 4d5.

4d8 Exercise. Prove that

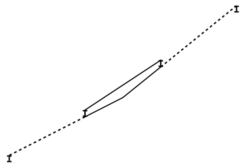
$$\frac{1}{\ln n} \ln Z_{\alpha \sqrt{2 \ln n}} \to \alpha^2 \quad \text{in probability as } n \to \infty$$

for every $\alpha \in [0, 1)$.

Hint: 4d7 gives $\mathrm{e}^{-\beta_n^2/2} Z_{\beta_n} \to 1$ in probability for $\beta_n = \alpha \sqrt{2 \ln n}$; take the logarithm.

Note that 4d7 was weaken thrice on the way toward 4d8. First, convergence in L_1 was replaced with convergence in probability. Second, $\sqrt{2\ln n}-\beta_n$ could be $o(\sqrt{\ln n})$. Third, the rate of convergence $o(1/\ln n)$ could be claimed. However, the weaker statement 4d8 is easier to grasp.

The (random) function $\beta \mapsto \ln Z_\beta$ is (almost sure) convex by 4c7, therefore the function $\alpha \mapsto \frac{1}{\ln n} \ln Z_{\alpha\sqrt{2\ln n}}$ is also convex. Thus, convergence in probability at every point (separately) implies uniform convergence in probability on $[\varepsilon, 1-\varepsilon]$ and moreover, convergence of derivatives.

(4d9) 

We turn to the case $\alpha > 1$. Recall that $-\frac{\mathrm{d}}{\mathrm{d}\beta} \ln Z_\beta = U_\beta = \int E\,\tilde G_\beta(\mathrm{d}E)$. The latter cannot be less than the left endpoint of the sample,

$$-\frac{\mathrm{d}}{\mathrm{d}\beta} \ln Z_\beta \ge \min_k H(k)\,.$$

Thus,

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}\frac{1}{\ln n} \ln Z_{\alpha\sqrt{2\ln n}} \le \sqrt{\frac{2}{\ln n}} \max_k(-H(k))\,. \tag{4d10}$$

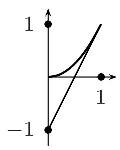
The right-hand side converges in probability (as $n \to \infty$) to 2, which was basically seen in 2c, but can also be proven via the arguments of this section. Namely, in the limit the derivative cannot be less than $\frac{\mathrm{d}}{\mathrm{d}\alpha}\big|_{\alpha=1} \alpha^2 = 2$, therefore

$$\mathbb{P}\Big(\max_k(-H(k)) \ge (1-\varepsilon)\sqrt{2\ln n}\Big) \to 1 \quad \text{as } n \to \infty$$

for every $\varepsilon > 0$. On the other hand,

$$Z_\beta = \frac{1}{n}\sum_k \mathrm{e}^{-\beta H(k)} \ge \frac{1}{n}\exp\big(-\beta \min_k H(k)\big) = \frac{1}{n}\exp\big(\beta\max_k(-H(k))\big)\,;$$

$$\underbrace{\frac{1}{\ln n}\ln Z_{\alpha\sqrt{2\ln n}}}_{\approx\alpha^2} \ge \alpha\sqrt{\frac{2}{\ln n}}\max_k(-H(k)) - 1\,,$$



therefore

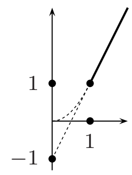
$$\mathbb{P}\Big(\max_k(-H(k)) \le (1+\varepsilon)\sqrt{2\ln n}\Big) \to 1 \quad \text{as } n \to \infty$$

for every $\varepsilon > 0$. We see that

$$\text{(4d11)} \qquad \frac{1}{\sqrt{2\ln n}} \max_k (-H(k)) \to 1 \quad \text{in probability as } n \to \infty\,.$$

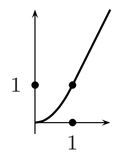
The derivative (4d10) cannot exceed 2 in the limit; but we saw that it cannot be less than 2 in the limit. It means that

$$\frac{1}{\ln n} \ln Z_{\alpha\sqrt{2\ln n}} \to 2\alpha - 1 \quad \text{in probability as } n \to \infty$$



for each $\alpha \in [1, \infty)$. We summarize the two cases:

$$\lim_{n\to\infty} \frac{1}{\ln n} \ln Z_{\alpha\sqrt{2\ln n}} = \begin{cases} \alpha^2 & \text{for } \alpha \in [0,1], \\ 2\alpha - 1 & \text{for } \alpha \in [1,\infty); \end{cases}$$



convergence in probability is meant.

As was noted, convergence of *convex* functions implies convergence of their first derivatives. By 4c3,

$$\frac{\mathrm{d}}{\mathrm{d}\alpha} \frac{1}{\ln n} \ln Z_{\alpha\sqrt{2\ln n}} = -\sqrt{\frac{2}{\ln n}}\, U_{\alpha\sqrt{2\ln n}}\,,$$

thus,

$$\text{(4d12)} \qquad \lim_{n\to\infty} \frac{1}{\sqrt{2\ln n}} U_{\alpha\sqrt{2\ln n}} = \begin{cases} -\alpha & \text{for } \alpha \in [0,1], \\ -1 & \text{for } \alpha \in [1,\infty) \end{cases}$$

(convergence in probability, as before). Compare the mean energy with the left endpoint of the sample!

Similarly,

$$\lim_{n\to\infty} \frac{1}{\sqrt{2\ln n}} F_{\alpha\sqrt{2\ln n}} = \begin{cases} -\frac{\alpha}{2} & \text{for } \alpha \in [0,1], \\ -1 + \frac{1}{2\alpha} & \text{for } \alpha \in [1,\infty); \end{cases}$$

$$\lim_{n\to\infty} \frac{1}{\ln n} S_{\alpha\sqrt{2\ln n}} = \begin{cases} -\alpha^2 & \text{for } \alpha \in [0,1], \\ -1 & \text{for } \alpha \in [1,\infty). \end{cases}$$

4d13 Exercise. Prove that for any $\alpha \in (0,1)$ and $\varepsilon > 0$ the probability of the following event tends to 1 as $n \to \infty$:

$$\nu\{x : H(x) \le -(\alpha+\varepsilon)\sqrt{2\ln n}\} \le \exp(-\alpha^2 \ln n)\,.$$

Hint: use 4c6.

The entropy used above is the entropy of G_β (or $\tilde G_\beta$) w.r.t. the uniform *probability* measure ν (or μ). The entropy w.r.t. the *counting* measure is higher by $\ln n$;

$$\lim_{n\to\infty} \frac{1}{\ln n}\big(S_{\alpha\sqrt{2\ln n}} + \ln n\big) = \begin{cases} 1-\alpha^2 & \text{for } \alpha \in [0,1], \\ 0 & \text{for } \alpha \in [1,\infty). \end{cases}$$

It does not mean that the entropy is small for $\alpha > 1$, it only means that it is $o(\ln n)$. In fact, it has a nontrivial limiting distribution (as $n \to \infty$). The order statistics $E^{(1)} < E^{(2)} < \dots$ are approximately

$$E^{(k)} = -c_n + \frac{\ln T_k}{\sqrt{2\ln n}}\,,$$

where $c_n \sim \sqrt{2\ln n}$ is defined by $\Phi(-c_n) = 1/n$ (compare it with a_n of 2c), and $T_1 < T_2 < \dots$ are a Poisson point process on $[0,\infty)$; in other words, random variables $T_1, T_2 - T_1, T_3 - T_2, \dots$ are independent, distributed Exp(1) each. Roughly, $T_k \approx k$, thus

$$\exp\big(-\alpha\sqrt{2\ln n}\,E^{(k)}\big) \approx \underbrace{\exp\big(\alpha\sqrt{2\ln n}\,c_n\big)}_{=\text{const}} \cdot \underbrace{\exp(-\alpha\ln T_k)}_{=T_k^{-\alpha}\approx k^{-\alpha}};$$

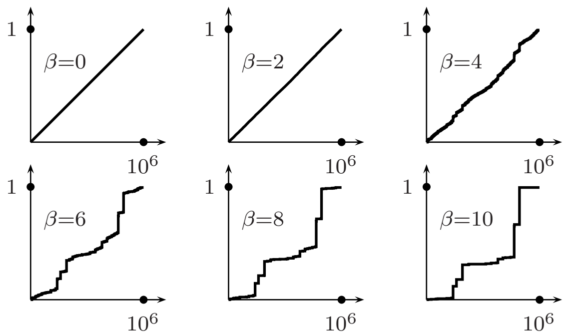
for $\alpha > 1$ the series converges, and the Gibbs measure $\tilde G_{\alpha\sqrt{2\ln n}}$ is approximately

$$\tilde G_{\alpha\sqrt{2\ln n}}\big(\{E^{(k)}\}\big) \approx \frac{T_k^{-\alpha}}{T_1^{-\alpha} + T_2^{-\alpha} + \dots}\,.$$

The (finite) entropy of the right-hand side gives us the limit (in distribution) of $S_{\alpha\sqrt{2\ln n}} + \ln n$ as $n \to \infty$. In contrast, for $\alpha < 1$,

$$\tilde G_{\alpha\sqrt{2\ln n}}\big(\{E^{(k)}\}\big) \to 0 \quad \text{in probability as } n \to \infty\,.$$

(I do not prove these facts.) In this sense, in the limit ($n \to \infty$) the Gibbs measure $G_{\alpha\sqrt{2\ln n}}$ gets continuous (nonatomic) if $\alpha < 1$ but discrete (purely atomic) if $\alpha > 1$.



the c.d.f. of G_β
$n = 1\,000\,000$
one sample

The shape of $\tilde G_{\alpha\sqrt{2\ln n}}$ for $\alpha < 1$ can be found via 4d3–4d7. Indeed, for every $u \in \mathbb{R}$

$$\tilde G_{\alpha\sqrt{2\ln n}}\big([-\alpha\sqrt{2\ln n}+u,\infty)\big) = \frac{Z_\beta(-\alpha\sqrt{2\ln n}+u,\infty)}{Z_\beta} \sim \\ \sim \mathrm{e}^{-\alpha^2 \ln n} Z_\beta(-\alpha\sqrt{2\ln n}+u,\infty)$$

by 4d7, thus

$$\mathbb{E}\, \tilde G_{\alpha\sqrt{2\ln n}}\big([-\alpha\sqrt{2\ln n}+u,\infty)\big) \sim \Phi(-u)$$

by 4d3. By 4d5,

$$\text{(4d14)} \quad \tilde G_{\alpha\sqrt{2\ln n}}\big([-\alpha\sqrt{2\ln n}+u,\infty)\big) \to \Phi(-u) \quad \text{in probability as } n \to \infty$$

for any $\alpha \in [0,1)$. It means that the shape is normal,

$$\tilde G_{\alpha\sqrt{2\ln n}} \approx \mathrm{N}(-\alpha\sqrt{2\ln n}, 1)\,.$$

Compare it with (4d2) and (4d12). Note however that (4d12) does not follow from (4d14) (tails could contribute too much). Similarly, (4d14) does not imply convergence of second moments, but still hints that (4d1) converges to 1, that is (recall 4c4),

$$\text{(4d15)} \qquad \frac{1}{\ln n} c_{\alpha\sqrt{2\ln n}} \to 2k_{\mathrm{B}}\alpha^2 \quad \text{in probability as } n \to \infty$$

for any $\alpha \in [0,1)$. How to prove it? Convexity does not help, since convergence of convex functions does not imply convergence of *second* derivatives.[1]

We consider the moment generating function f_n of the centered $\tilde G_{\beta_n}$,

$$f_n(\lambda) = \int \mathrm{e}^{\lambda(H+\beta_n)}\, \mathrm{d}G_{\beta_n} = \int \mathrm{e}^{\lambda(E+\beta_n)}\, \tilde G_{\beta_n}(\mathrm{d}E)\,.$$

[1]A smooth convex function can be approximated by piecewise linear convex functions. . .

By 4c13(a),

$$f_n(\lambda) = \mathrm{e}^{\lambda\beta_n} \frac{Z_{\beta_n-\lambda}}{Z_{\beta_n}} .$$

Let β_n satisfy (4d6), then $Z_{\beta_n} \sim \mathrm{e}^{\beta_n^2/2}$ and $Z_{\beta_n-\lambda} \sim \mathrm{e}^{(\beta_n-\lambda)^2/2}$ by 4d7, therefore

$$f_n(\lambda) \to \mathrm{e}^{\lambda^2/2} \quad \text{in probability as } n \to \infty$$

for every $\lambda \in \mathbb{R}$. This is more than enough in order to ensure convergence (in probability) of all moments in (4d14).[1] Moreover, it is enough in order to get it without (4d14). Indeed, convergence in probability of *convex* functions $\ln f_n(\lambda)$ to $\lambda^2/2$ at each λ (separately) implies uniform convergence in probability on every bounded interval (recall (4d9)), which ensures convergence of all moments to the moments of $\mathrm{N}(0,1)$. Relation (4d15) is thus proven.

4e The spherical model

Now the Gaussian process H is a random quadratic form on a sphere,

$$X = \{x \in \mathbb{R}^n : |x| = 1\} , \qquad (= \mathrm{S}^{n-1})$$
$$\nu \text{ is the uniform probability measure on } X,$$
$$H(x) = \sqrt{n}\langle Ax, x\rangle \quad \text{for } x \in X;$$

here A is the Gaussian random matrix of Sect. 3 (recall (3a2)). The coefficient $\sqrt{n}$ ensures that

$$H(x) \sim \mathrm{N}(0,1) \quad \text{for each } x \in X,$$

similarly to 4d; however, the random variables $H(x)$ are correlated. As always, the Gaussian process is described by its covariation function.

4e1 Exercise. Prove that

$$\langle H(x), H(y)\rangle = \mathbb{E}\, H(x)H(y) = \langle x, y\rangle^2 \quad \text{for } x, y \in X.$$

Hint: reduce the general case to $x = e_1$ (and $n = 2$, if you want).

In fact, $\|A\|$ is close to $\sqrt{2}$ (highly probably for large n, see the note after 3e8), thus $\sup H(x)$ is close to $\sqrt{2n}$ and, more importantly,

$$\text{(4e2)} \qquad \frac{1}{\sqrt{2n}} \min_{x\in X} H(x) \to -1 \quad \text{in probability as } n \to \infty.$$

[1] It is enough that $\sup_n f_n(\pm\varepsilon) < \infty$ for a single $\varepsilon > 0$. Convergence in probability does not imply convergence a.s.; however, every subsequence contains a subsequence converging a.s.

Compare it with (4d11); the n-dimensional spherical model looks similar to the e^n-dimensional random energy model of 4d (but uses only $\frac{1}{2}n(n+1)$ orthogaussian random variables). The latter has a threshold at $\beta = \sqrt{2\ln(\mathrm{e}^n)} = \sqrt{2n}$; we may expect the same for the spherical model (but in fact, we will get a threshold at $\beta = \sqrt{n/2}$).

The question is, how to calculate $\int \mathrm{e}^{-\beta H}\,\mathrm{d}\nu$; it is difficult to integrate over the sphere!

DIGRESSION: INTEGRATING OVER THE SPHERE

The Gaussian measure γ^n lives near the sphere (of radius $\sqrt{n}$ rather than 1, which is not a problem) and allows for a simple integration.

4e3 Exercise. Prove that

$$\int \exp(a_1 x_1^2 + \dots + a_n x_n^2)\,\gamma^n(\mathrm{d}x) = \prod_{k=1}^n (1-2a_k)^{-1/2}$$

for all $a_1,\dots,a_n \in (-\infty, 1/2)$.

Hint: reduce to $n=1$; note that $\int \mathrm{e}^{-u^2/2}\,\mathrm{d}u = \sqrt{2\pi}$ and $\int \exp\bigl(-\frac{u^2}{2\sigma^2}\bigr)\,\mathrm{d}u = \sqrt{2\pi}\sigma$.

Still, we need to integrate over the sphere. We have $\exp(a_1 x_1^2 + \dots + a_n x_n^2)\,\gamma^n(\mathrm{d}x) = \mathrm{const}\cdot\gamma(\mathrm{d}x)$, where const is the integral calculated above, and γ is another Gaussian measure,

$$\text{(4e4)}\qquad \frac{\gamma(\mathrm{d}x)}{\mathrm{d}x} = \prod_{k=1}^n \frac{1}{\sqrt{2\pi}\sigma_k}\exp\Bigl(-\frac{x_k^2}{2\sigma_k^2}\Bigr), \qquad \sigma_k = (1-2a_k)^{-1/2}\,.$$

The distribution of $x \mapsto |x|^2$ w.r.t. γ is the distribution of $(\sigma_1\zeta_1)^2 + \dots + (\sigma_n\zeta_n)^2$ for orthogaussian $\zeta_1,\dots,\zeta_n$, it is the convolution of n special distributions (well-known as gamma distributions). The density of the distribution at a point r^2 is proportional to the integral of $\gamma(\mathrm{d}x)/\mathrm{d}x$ over the r-sphere (think, why). The coefficient depends on r but not γ. We may get rid of the coefficient by comparing γ with γ^n (its density is constant on the sphere),

$$\text{(4e5)}\quad \frac{(\text{density of } |x|^2 \text{ w.r.t. } \gamma \text{ at } r^2)}{(\text{density of } |x|^2 \text{ w.r.t. } \gamma^n \text{ at } r^2)} = \biggl(\text{mean of } \frac{\gamma(\mathrm{d}x)}{\gamma^n(\mathrm{d}x)} \text{ over } r\text{-sphere}\biggr)$$

for every nondegenerate Gaussian measure γ on $\mathbb{R}^n$ (irrespective of (4e4)). The distribution of $|x|^2$ w.r.t. γ^n is well-known (a gamma distribution) and its density is easy to write down explicitly, which cannot be said about γ. Two clever ideas help a lot. First, we may restrict ourselves to logarithmic asymptotics; in other words, we may calculate up to factors $\mathrm{e}^{o(n)}$. Second, if the mean value of $|x|^2$ w.r.t. γ is equal to 1, then the density at 1 is neither too small nor too large, it must be $\mathrm{e}^{o(n)}$, uniformly over all γ satisfying the restriction $\int |x|^2\,\gamma(\mathrm{d}x) = 1$. (I omit the proof.) By rescaling,

$$\Bigl(\text{density of } |x|^2 \text{ w.r.t. } \gamma \text{ at } \textstyle\int |x|^2\,\gamma(\mathrm{d}x)\Bigr) = \frac{\mathrm{e}^{o(n)}}{\int |x|^2\,\gamma(\mathrm{d}x)}\,,$$

uniformly over all nondegenerate Gaussian measures γ on $\mathbb{R}^n$. Thus,

$$\frac{(\text{density of } |x|^2 \text{ w.r.t. } \gamma \text{ at } n)}{(\text{density of } |x|^2 \text{ w.r.t. } \gamma^n \text{ at } n)} = \mathrm{e}^{o(n)}$$

uniformly over all γ such that $\int |x|^2\,\gamma(\mathrm{d}x) = n$. By (4e5),

$$\Bigl(\text{mean of } \frac{\gamma(\mathrm{d}x)}{\gamma^n(\mathrm{d}x)} \text{ over } \sqrt{n}\text{-sphere}\Bigr) = \mathrm{e}^{o(n)}$$

for these γ. Taking into account that

$$\frac{\gamma^n(\mathrm{d}x)}{\mathrm{d}x} = (2\pi\mathrm{e})^{-n/2} \quad \text{on the } \sqrt{n}\text{-sphere}$$

(check it), we get

$$\Bigl(\text{mean of } \frac{\gamma(\mathrm{d}x)}{\mathrm{d}x} \text{ over } \sqrt{n}\text{-sphere}\Bigr) = (2\pi\mathrm{e})^{-n/2}\mathrm{e}^{o(n)}$$

for γ such that $\int |x|^2\,\gamma(\mathrm{d}x) = n$. By rescaling,

$$\Bigl(\text{mean of } \frac{\gamma(\mathrm{d}x)}{\mathrm{d}x} \text{ over } r\text{-sphere}\Bigr) = \Bigl(\frac{n}{2\pi\mathrm{e}r^2}\Bigr)^{n/2}\mathrm{e}^{o(n)} \quad \text{where } r^2 = \int |x|^2\,\gamma(\mathrm{d}x)\,,$$

uniformly over all nondegenerate Gaussian measures γ on $\mathbb{R}^n$.

Returning to γ defined by (4e4) we have

$$\Bigl(\text{mean of } \exp\bigl(\sum_k a_k x_k^2\bigr) \text{ over } r\text{-sphere}\Bigr) = \\ = \Bigl(\prod_k (1-2a_k)\Bigr)^{-1/2} \exp\Bigl(\frac{r^2}{2}\Bigr)\Bigl(\frac{n}{\mathrm{e}r^2}\Bigr)^{n/2}\mathrm{e}^{o(n)}$$

uniformly over $a_1,\dots,a_n \in (-\infty, 1/2)$; here

$$r^2 = \int |x|^2\,\gamma(\mathrm{d}x) = \sum_k \sigma_k^2 = \sum_k (1-2a_k)^{-1}\,.$$

By rescaling,

$$\Big(\text{mean of } \exp\big(r^2 \textstyle\sum_k a_k x_k^2\big) \text{ over } r\text{-sphere}\Big) = \text{the same}\,.$$

We want to get $r^2 a_k = b_k$ for given numbers $b_k \in \mathbb{R}$. Thus, r must satisfy $r^2 = \sum_k (1 - 2b_k r^{-2})^{-1}$, that is,

$$\sum_k (r^2 - 2b_k)^{-1} = 1\,. \tag{4e6}$$

The sum is a continuous, strictly decreasing function of $r \in (\sqrt{2b_n}, \infty)$ (assuming $b_n = \max(b_1, \dots, b_n)$), it tends to $+\infty$ as $r \to \sqrt{2b_n}$ and to 0 as $r \to \infty$. Therefore there exists exactly one solution r of the equation. For this r we have

$$\begin{aligned}\Big(\text{mean of } \exp\big(\sum_k b_k x_k^2\big) \text{ over 1-sphere}\Big) &= \\ &= \Big(\prod_k (r^2 - 2b_k) \Big)^{-1/2} \exp\Big(\frac{r^2}{2}\Big) \Big(\frac{n}{\mathrm{e}}\Big)^{n/2} \mathrm{e}^{o(n)}\,.\end{aligned} \tag{4e7}$$

END OF DIGRESSION

Now we can calculate $\int \mathrm{e}^{-\beta H} \, \mathrm{d}\nu$ up to $\mathrm{e}^{o(n)}$, using *Wigner's semi-circle law* (Theorem 3e6) that describes the spectrum $(\lambda_1, \dots, \lambda_n)$ of the (random) matrix A via the measure $\frac{1}{\pi}\sqrt{2 - \lambda^2} \, \mathrm{d}\lambda$.

As was said, we may expect a threshold at $\beta \asymp \sqrt{n}$; thus we take

$$\beta = \alpha \sqrt{n}$$

and consider small and large α separately.

First, let α be small. We have to solve Equation 4e6 for $b_k = -\beta\sqrt{n}\lambda_k = -\alpha n \lambda_k$;

$$\begin{aligned}\sum_k (r^2 - 2b_k)^{-1} = \sum_k (r^2 + 2\alpha n \lambda_k)^{-1} = \frac{1}{n} \sum_k \Big(\frac{r^2}{n} + 2\alpha\lambda_k \Big)^{-1} &= \\ = \int_{-\sqrt{2}}^{+\sqrt{2}} (u + 2\alpha\lambda)^{-1} \tfrac{1}{\pi} \sqrt{2 - \lambda^2} \, \mathrm{d}\lambda + o(1)\,,\end{aligned}$$

assuming that $r^2 = nu$ and $2\sqrt{2}\,\alpha < u$. The general formula

$$\int_{-\sqrt{2}}^{+\sqrt{2}} (u + v\lambda)^{-1} \tfrac{1}{\pi} \sqrt{2 - \lambda^2} \, \mathrm{d}\lambda = \frac{2}{u + \sqrt{u^2 - 2v^2}}$$

(valid for $0 < v < u/\sqrt{2}$) gives us

$$\sum_k (r^2 - 2b_k)^{-1} = \frac{2}{u + \sqrt{u^2 - 8\alpha^2}} + o(1)\,.$$

We solve the equation (ignoring $o(1)$): $u + \sqrt{u^2 - 8\alpha^2} = 2$; $u \le 2$, and $u^2 - 8\alpha^2 = (2-u)^2$; $u = 1 + 2\alpha^2$, and get

$$r^2 = n(1 + 2\alpha^2) + o(n) \tag{4e8}$$

provided that $\alpha < 1/\sqrt{2}$ (otherwise, for $\alpha > 1/\sqrt{2}$ the equation $u + \sqrt{u^2 - 8\alpha^2} = 2$ has no solutions, and for $\alpha = 1/\sqrt{2}$ the solution violates the restriction $2\sqrt{2}\,\alpha < u$). We put this r into (4e7);

$$\begin{aligned}
\text{(4e9)}\quad n^{n/2}\Big(\prod_k (r^2 - 2b_k) \Big)^{-1/2} &= \bigg(\prod_k \frac{r^2 - 2b_k}{n} \bigg)^{-1/2} = \\
&= \bigg(\prod_k \big(1 + 2\alpha^2 + 2\alpha\lambda_k + o(1) \big) \bigg)^{-1/2} = \\
&= \exp\Big(-\frac{n}{2} \cdot \frac{1}{n} \sum_k \ln(1 + 2\alpha^2 + 2\alpha\lambda_k) + o(n) \Big) = \\
&= \exp\Big(-\frac{n}{2} \int_{-\sqrt{2}}^{+\sqrt{2}} \ln(1 + 2\alpha^2 + 2\alpha\lambda) \tfrac{1}{\pi}\sqrt{2 - \lambda^2}\,\mathrm{d}\lambda \Big) \mathrm{e}^{o(n)}\,.
\end{aligned}$$

We apply the general formula

$$\int_{-\sqrt{2}}^{+\sqrt{2}} \ln(u + v\lambda) \tfrac{1}{\pi}\sqrt{2 - \lambda^2}\,\mathrm{d}\lambda = \ln\frac{u + \sqrt{u^2 - 2v^2}}{2} + \Big(\frac{v}{u + \sqrt{u^2 - 2v^2}} \Big)^2 \tag{4e10}$$

(valid when $0 < v \le u/\sqrt{2}$) for $u = 1 + 2\alpha^2$, $v = 2\alpha$, note that $\sqrt{u^2 - 2v^2} = 1 - 2\alpha^2$ and get just

$$n^{n/2}\Big(\prod_k (r^2 - 2b_k) \Big)^{-1/2} = \mathrm{e}^{-n\alpha^2/2}\mathrm{e}^{o(n)}\,.$$

By (4e7),

$$\begin{aligned}
Z_\beta = \int \mathrm{e}^{-\beta H}\,\mathrm{d}\nu &= \mathrm{e}^{-n\alpha^2/2}\mathrm{e}^{r^2/2}\mathrm{e}^{-n/2}\mathrm{e}^{o(n)} = \\
&= \exp\Big(-\frac{n\alpha^2}{2} + n\frac{1 + 2\alpha^2}{2} - \frac{n}{2} \Big)\mathrm{e}^{o(n)} = \mathrm{e}^{n\alpha^2/2}\mathrm{e}^{o(n)} = \mathrm{e}^{\beta^2/2}\mathrm{e}^{o(n)}\,,
\end{aligned}$$

that is,

$$\frac{1}{n}\ln Z_{\alpha\sqrt{n}} \to \frac{\alpha^2}{2} \quad \text{in probability as } n \to \infty$$

for $\alpha \in [0, 1/\sqrt{2})$.

We turn to the case $\alpha > 1/\sqrt{2}$.

4e11 Exercise. The solution r of Equation (4e6) satisfies

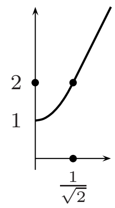
$$\frac{r^2}{n} \to 2\sqrt{2}\alpha \quad \text{in probability as } n \to \infty.$$

Prove it.

Hint: on one hand, if $u > 2\sqrt{2}\alpha > 2$ then $u + \sqrt{u^2 - 8\alpha^2} > 2$; on the other hand, $\min_k \lambda_k \to -\sqrt{2}$ in probability as $n \to \infty$ (take it for granted), thus $\max b_k \sim \sqrt{2}\alpha n$.

Compare it with (4e8):

$$\lim_{n\to\infty} \frac{r^2}{n} = \begin{cases} 1 + 2\alpha^2 & \text{for } \alpha \in [0, 1/\sqrt{2}], \\ 2\sqrt{2}\alpha & \text{for } \alpha \in [1/\sqrt{2}, \infty). \end{cases}$$



convergence in probability is meant. Similarly to (4e9),

$$n^{n/2}\Big(\prod_k (r^2 - 2b_k)\Big)^{-1/2} = \Big(\prod_k \big(2\sqrt{2}\alpha + 2\alpha\lambda_k + o(1)\big)\Big)^{-1/2} =$$
$$= \exp\Big(-\frac{n}{2}\int_{-\sqrt{2}}^{+\sqrt{2}} \ln(2\sqrt{2}\alpha + 2\alpha\lambda)\tfrac{1}{\pi}\sqrt{2-\lambda^2}\,\mathrm{d}\lambda\Big)\mathrm{e}^{o(n)},$$

which, unfortunately, does not follow from Theorem 3e6, since the logarithm is not bounded on $(-\sqrt{2}, \sqrt{2})$. More detailed information about eigenvalues near $-\sqrt{2}$ is needed, and the small gap $\frac{r^2}{n} - 2\sqrt{2}\alpha$ should be taken into account. (I omit the proof.) Anyway, the integral converges; it is (4e10) for $u = 2\sqrt{2}\alpha$, $v = 2\alpha$ (thus, $\sqrt{u^2 - 2v^2} = 0$);

$$n^{n/2}\Big(\prod_k (r^2 - 2b_k)\Big)^{-1/2} =$$
$$= \exp\Big(-\frac{n}{2}\Big(\ln(\alpha\sqrt{2}) + \frac{1}{2}\Big)\Big)\mathrm{e}^{o(n)} = (\alpha\sqrt{2})^{-n/2}\mathrm{e}^{-n/4}\mathrm{e}^{o(n)}.$$

By (4e7),

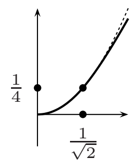
$$Z_\beta = \int \mathrm{e}^{-\beta H} \,\mathrm{d}\nu = (\alpha\sqrt{2})^{-n/2} \mathrm{e}^{-n/4} \mathrm{e}^{r^2/2} \mathrm{e}^{-n/2} \mathrm{e}^{o(n)} =$$
$$= (\alpha\sqrt{2})^{-n/2} \exp\Big(\alpha n \sqrt{2} - \frac{3n}{4} \Big) \mathrm{e}^{o(n)} \,,$$

that is,

$$\frac{1}{n} \ln Z_{\alpha\sqrt{n}} \to \alpha\sqrt{2} - \frac12 \ln(\alpha\sqrt{2}) - \frac34 \quad \text{in probability as } n \to \infty$$

for $\alpha \in (1/\sqrt{2}, \infty)$. We summarize the two cases:
(4e12)

$$\lim_{n\to\infty} \frac1n \ln Z_{\alpha\sqrt{n}} = \begin{cases} \alpha^2/2 & \text{for } \alpha \in [0, 1/\sqrt{2}], \\ \alpha\sqrt{2} - \frac12 \ln(\alpha\sqrt{2}) - \frac34 & \text{for } \alpha \in [1/\sqrt{2}, \infty); \end{cases}$$



convergence in probability is meant.

4e13 Exercise. Prove that

$$\lim_{n\to\infty} \frac{1}{\sqrt{n}} U_{\alpha\sqrt{n}} = \begin{cases} -\alpha & \text{for } \alpha \in (0, 1/\sqrt{2}], \\ -\sqrt{2} + \frac{1}{2\alpha} & \text{for } \alpha \in [1/\sqrt{2}, \infty); \end{cases}$$
$$\lim_{n\to\infty} \frac{1}{\sqrt{n}} F_{\alpha\sqrt{n}} = \begin{cases} -\frac{\alpha}{2} & \text{for } \alpha \in (0, 1/\sqrt{2}], \\ -\sqrt{2} + \frac{1}{2\alpha} \ln(\alpha\sqrt{2}) + \frac{3}{4\alpha} & \text{for } \alpha \in [1/\sqrt{2}, \infty); \end{cases}$$
$$\lim_{n\to\infty} \frac{1}{n} S_{\alpha\sqrt{n}} = \begin{cases} -\frac{\alpha^2}{2} & \text{for } \alpha \in (0, 1/\sqrt{2}], \\ -\frac12 \ln\alpha - \frac14 (1 + \ln 2) & \text{for } \alpha \in [1/\sqrt{2}, \infty) \end{cases}$$

(convergence in probability is meant).
Hint: similar to (4d12).

4e14 Exercise. Prove that for any $\alpha \in (0,1)$ and $\theta \in (0,1)$ the probability of the following event tends to 1 as $n \to \infty$:

$$\nu\{x : H(x) \le -(\theta+\varepsilon)\sqrt{2n}\} \le \begin{cases} \exp(-n\theta^2) & \text{for } \theta \le 1/2, \\ (2-2\theta)^{n/2} \mathrm{e}^{-n/4} & \text{for } \theta \ge 1/2. \end{cases}$$

Hint: similar to 4d13.

The threshold at $\pm\sqrt{2n}$ is evident, but one more threshold at $\pm\frac12\sqrt{2n}$ is mysterious!

4f The Sherrington-Kirkpatrick model

We return to the model touched upon in Sect. 1g,

$$X = \{-1,1\}^n \,, \qquad \nu \text{ is the uniform probability measure on } X,$$
$$H(\sigma_1,\dots,\sigma_n) = -\frac{1}{\sqrt{n}} \sum_{k<l} \zeta_{k,l}\sigma_k\sigma_l \quad \text{for } (\sigma_1,\dots,\sigma_n) \in X,$$

where $(\zeta_{k,l})_{k<l}$ is a family of $n(n-1)/2$ orthogaussian functions. They model disorder in a system of n spins $\sigma_1,\dots,\sigma_n = \pm 1$; namely, $\Xi(\sigma_1,\dots,\sigma_n)$ is the energy of the spin configuration $\sigma_1,\dots,\sigma_n$. That is the Sherrington-Kirkpatrick model for spin glasses, well-known in statistical physics (SK model). It ignores the geometric location of atoms, assuming that all pairs interact in the same way ('mean field approximation').

The SK model is related to the spherical model as vertices of the cube to the sphere, in the following sense.

4f1 Exercise. Prove that

$$H(\sigma) = -\frac{1}{\sqrt{2}} \big(\langle A\sigma, \sigma \rangle - \operatorname{trace} A \big) \,,$$

where A is the Gaussian random matrix of Sect. 3 (used in 4e).

The trace term is of little importance.

4f2 Exercise. Prove that

$$\int \exp\Big(\frac{\beta}{\sqrt{2}} \langle A\sigma, \sigma \rangle \Big) \, \nu(\mathrm{d}\sigma) = Z_\beta \exp\Big(\frac{\beta}{\sqrt{2}} \operatorname{trace} A \Big) \,,$$

the two factors in the right hand side being independent; and $\operatorname{trace} A$ is distributed $\mathrm{N}(0,1)$.

4f3 Exercise. Derive from (4e2) that for every $\varepsilon > 0$ the probability of the following event tends to 1 as $n \to \infty$:

$$\max_\sigma \big(-H(\sigma) \big) \le (1+\varepsilon) n \,.$$

Computer simulations seem to show that

$$\lim_{n\to\infty} \frac{1}{n} \max_\sigma \big(-H(\sigma) \big) = 0.7366\dots$$

see [2, (1.2)].

4f4 Exercise. (See also [1, Prop. 2.2]) For each β,

$$\sqrt{\frac{2}{n}}\,F_\beta = -\sqrt{\frac{2}{n}}\,\frac{1}{\beta}\ln Z_\beta \quad \text{is more concentrated than } \mathrm{N}(0,1).$$

Prove it.
Hint: similar to 4a3.

4f5 Exercise. For each β,

$$\mathbb{E}\, Z_\beta = \mathrm{e}^{n\beta^2/4}\,.$$

Prove it.
Hint: $\mathbb{E} \int \mathrm{e}^{-\beta H}\,\mathrm{d}\nu = \int (\mathbb{E}\,\mathrm{e}^{-\beta H})\,\mathrm{d}\nu$.

By Jensen's inequality, $\mathbb{E}\,\ln Z_\beta \le \ln \mathbb{E}\, Z_\beta$, thus

$$\mathbb{E}\,\frac{1}{n}\ln Z_\beta \le \frac{\beta^2}{4}\,. \tag{4f6}$$

4f7 Theorem. (Aizenman, Lebowitz, Ruelle; see [1, Th. 3.3]) For each $\beta < 1$,

$$\mathbb{E}\,\frac{1}{n}\ln Z_\beta \to \frac{\beta^2}{4} \quad \text{as } n \to \infty.$$

Feel free to use Theorem 4f7 even though I give no proof.

4f8 Exercise. For each $\beta \in (0,1)$,

$$\frac{1}{n}\ln Z_\beta \to \frac{\beta^2}{4} \quad \text{in probability as } n \to \infty.$$

Prove it.
Hint: combine 4f7 and 4f4.

4f9 Exercise. For each $\beta \in (0,1)$,

$$\lim_{n\to\infty} \frac{1}{n} U_\beta = -\frac{1}{2}\beta\,; \quad \lim_{n\to\infty} \frac{1}{n} F_\beta = -\frac{1}{4}\beta\,; \quad \lim_{n\to\infty} \frac{1}{n} S_\beta = -\frac{1}{4}\beta^2\,;$$

convergence in probability is meant.
Prove it.
Hint: similar to 4e13.

4f10 Exercise. For each $\beta \in (0,1)$ and $\varepsilon > 0$ the probability of the following event tends to 1 as $n \to \infty$:

$$\frac{\#\{\sigma : H(\sigma) \le -n(\beta+\varepsilon)/2\}}{2^n} \le \mathrm{e}^{-n\beta^2/4} \, .$$

Prove it.

Hint: similar to 4d15, 4e14.

"What really happens for $\beta > 1$? The physicists have proposed an entire theory, of great complexity. It seems so much out of reach of the current rigorous methods that there is no point to even discuss it." Talagrand [1, p. 211].

However, we know that the threshold at 1 really exists.

4f11 Theorem. (Comets [3]; see also [1, Th. 3.13]) For all $\beta \in (1, \infty)$,

$$\limsup_{n\to\infty} \mathbb{E} \, \frac{1}{n} \ln Z_\beta \le \beta - \frac{1}{2} \ln \beta - \frac{3}{4} < \frac{\beta^2}{4} \, .$$

Compare the right hand side with (4e12).

Theorem 4f11 follows immediately from (4e12) and the following inequality, well-known as 'domination of the SK model by the spherical model':

$$\mathbb{E} \, \ln Z_\beta^{\mathrm{SK}} \le \mathbb{E} \, \ln Z^{\mathrm{sphere}}_{\beta\sqrt{n/2}} \quad \text{for } \beta \in [0,\infty), \tag{4f12}$$

where Z_β^{SK} is Z_β of this section, while $Z_\beta^{\mathrm{sphere}}$ stands for Z_β of Section 4e. The domination is obtained by averaging over rotations $O \in \mathrm{O}(n)$. Below, $\mathrm{d}O$ denotes the uniform probability measure on $\mathrm{O}(n)$, and ν^{sphere} — the uniform probability measure on the 1-sphere of $\mathbb{R}^n$ (that is, ν of Sect. 4e). We extend H from $X = \{-1,1\}^n$ to the whole $\mathbb{R}^n$ according to 4f1.

4f13 Exercise. For every $\sigma \in \{-1,1\}^n$ and $\beta \in [0,\infty)$,

$$\int \mathrm{e}^{-\beta H(O\sigma)} \, \mathrm{d}O = \int \exp\Big(\frac{\beta n}{\sqrt{2}} \langle Ax, x\rangle - \frac{\beta}{\sqrt{2}} \operatorname{trace} A \Big) \, \nu^{\mathrm{sphere}}(\mathrm{d}x) \, .$$

Prove it.

Hint: $O\sigma$ is distributed uniformly on the $\sqrt{n}$-sphere.

4f14 Exercise. For every $\beta \in [0,\infty)$,

$$\int \ln \Big(\int \mathrm{e}^{-\beta H(O\sigma)} \, \nu(\mathrm{d}\sigma) \Big) \, \mathrm{d}O \le$$

$$\le -\frac{\beta}{\sqrt{2}} \operatorname{trace} A + \ln \int \exp\Big(\frac{\beta n}{\sqrt{2}} \langle Ax, x\rangle \Big) \, \nu^{\mathrm{sphere}}(\mathrm{d}x) \, .$$

Prove it.

Hint: $\int \ln(\dots) \le \ln \int(\dots)$; use 4f13.

4f15 Exercise. Prove (4f12).

Hint: take the expectation of 4f14 and note that the expectation (and moreover, the distribution) of $\ln\bigl(\int \mathrm{e}^{-\beta H(O\sigma)}\,\nu(\mathrm{d}\sigma)\bigr)$ does not depend on O (since $O^{-1}AO$ is distributed like A).

References

The content of my Sect. 4 is only a tip of the iceberg. See [1], [2] for some more realistic but more complicated models and finer arguments, and [2] for related topics in neural networks.

[1] M. Talagrand, *Mean field models for spin glasses: a first course,* Lecture Notes in Mathematics **1816**, 181–285 (2003).

[2] M. Talagrand, *Huge random structures and mean field models for spin glasses,* Proc. Intern. Congress Math. 1998, vol. 1, 507–536.

[3] F. Comets, *A spherical bound for the Sherrington-Kirkpatrick model,* Astérisque **236**, 103–108 (1996).

Index