Short Course on the Sherrington Kirkpatrick Model : Lecture 1

Shannon Starr

20 October 2003

Material which is not otherwise attributed in these lectures is motivated by joint work with Michael Aizenman and Robert Sims, or, in other cases, is material communicated to me in private conversations by Michael Aizenman.

1 Gaussian Processes

By a Gaussian process, we will always mean a *centered* Gaussian process.

1.1 Gaussian random variables and Wick's rule

Consider first Gaussian random variables. A Gaussian random variable with covariance 1 is a real random variable X such that the probability distribution on $\mathbb{R}$ given by $\mu(A) := \mathbb{E}\{X \in A\}$ is equal to $d\mu(z) = e^{-z^2/2}\,dz/\sqrt{2\pi}$. We will use the notation, $d\mu(z)$ to mean $e^{-z^2/2}\,dz/\sqrt{2\pi}$ henceforth. For any $\sigma \in \mathbb{R}$, a random variable Y is said to be a Gaussian random variable with covariance σ^2, if there is a Gaussian random variable with covariance 1 such that $Y = \sigma X$. We call the Gaussian random variable with covariance 1 "normal". We denote the class of Gaussian random variables with covariance σ^2 by $\mathcal{N}(0,\sigma^2)$. E.g., X is a normal Gaussian random variable iff $X \in \mathcal{N}(0,1)$. Defining $\mu_\sigma(A) = \mathbb{E}\{Y \in A\}$ to be the probability distribution for Y, one sees that, if $\sigma \neq 0$, then $d\mu_\sigma(z) = d\mu(z/\sigma)$. If $\sigma = 0$ then $\mu_0 = \delta_0$, the Dirac measure concentrated at 0. For any continuous function of compact support ϕ, the map

$$\sigma \mapsto \int_{\mathbb{R}} \phi(z) d\mu_\sigma(z) = \int_{\mathbb{R}} \phi(\sigma z) d\mu(z) \tag{1}$$

is continuous, which means the map $\sigma \mapsto \mu_\sigma$ is continuous with respect to the vague topology on $\mathcal{M}_1(\mathbb{R})$, where the latter is the set of all Borel probability measures on $\mathbb{R}$. The vague topology is the weak-* topology with respect to $\mathcal{C}_0(\mathbb{R})$. The only thing which is the slightest bit tricky conceptually (but it is also one of the first points in a basic probability course) is that probability measures which are absolutely continuous with respect to Lebesgue measure, such as μ_σ for $\sigma \neq 0$ can converge vaguely to a probability measure which is singular with respect to Lebesgue measure, such as δ_0. While this is a conceptual hurdle, it is a mathematical boon because it means that the absolutely continuous probability measures are dense in $\mathcal{M}_1(\mathbb{R})$ with respect to the vague topology.

The Gaussian process is a remarkable object. Because of the central limit theorem, it arises over and over again in probability theory. In addition it is crucially important in analysis, for example as the optimizer in many important inequalities. (See for example [7].) It is an advantage to have another characterization of the Gaussian process which is a little more algebraic.

Lemma 1.1 *(Wick's rule) For any $\sigma \in \mathbb{R}$, if $X \in \mathcal{N}(0,\sigma^2)$ then $\mathbb{E}\{X\} = 0$ and $\mathbb{E}\{X^2\} = \sigma^2$, and for any polynomial $f \in \mathbb{C}[z]$,*

$$\mathbb{E}\{X \cdot f(X)\} = \mathbb{E}\{f'(X)\} \cdot \mathbb{E}\{X^2\}\,. \tag{2}$$

If Y is any random variable satisfying $\mathbb{E}\{Y\} = 0$, $\mathbb{E}\{Y^2\} = \sigma^2$ and

$$\mathbb{E}\{Y \cdot f(Y)\} = \mathbb{E}\{f'(Y)\} \cdot \mathbb{E}\{Y^2\}\,. \tag{3}$$

for every polynomial $f \in \mathbb{C}[z]$, then $Y \in \mathcal{N}(0,\sigma^2)$.

Homework 1. *Check that if $X \in \mathcal{N}(0,\sigma^2)$, then (2) is satisfied, by integrating-by-parts with respect to the Gaussian distribution μ_σ. In fact it is satisfied as long as $f \in \mathcal{C}^1(\mathbb{R})$ and $zf(z)$, $f'(z)$ are in $L^1(\mathbb{R},\mu_\sigma)$.*

The homework problem proves the first half of the lemma, in fact the important half for our purposes. (It is a general principle that in classification problems the more frequently used half of the classification is the easier one.) The subtler half of the lemma will follow from the determinacy of the classical moment problem, which we will explain next.

1.2 Digression on the classical moment problem

Homework 2. *Using (3) calculate all the moments of Y. Prove that $\mathbb{E}\{Y^{2k+1}\} = 0$ for all $k \in \mathbb{N}$, while $\mathbb{E}\{Y^{2k}\} = \alpha_k \sigma^{2k}$ where α_k is the number of perfect matchings between $2k$ vertices, numerically $\alpha_k = (2k)!/(2^k k!)$.*

From the homework problem, we know that all the integer moments of Y are equal to all the integer moments of X. There is a problem of great importance in the development of classical analysis called the "classical moment problem". It asks, under which conditions a sequence of numbers ($\gamma_n : n \in \mathbb{N} = \{0,1,\dots\}$) can equal the integer moments of a measure on the real line $\gamma_n = \int_{\mathbb{R}} z^n \, d\alpha(z)$, and under which conditions the measure is uniquely determined. (More specifically, it is called the Hamburger moment problem. Another question is the Stieltjes moment problem, which is analogous to the Hamburger moment problem but restricting to measures supported on $\mathbb{R}_+ = [0,\infty)$.) A good reference is [9]. A necessary and sufficient condition for existence is that for any $n \in \mathbb{N}^* = \{1,2,\dots\}$ and any $\lambda_0,\dots,\lambda_{n-1} \in \mathbb{R}$, one has

$$\sum_{i,j=0}^{n-1} \lambda_i \lambda_j \gamma_{i+j} \geq 0\,. \tag{4}$$

This is clearly necessary because the double sum equals $\int_{\mathbb{R}} (\sum_{i=0}^{n-1} \lambda_i z^i)\, d\alpha(z)$ in case α exists. For sufficiency, see [9].

The question of uniqueness, which in this context is called determinacy, is more subtle. Basically, a complete answer can be given in terms of whether or not a certain symmetric unbounded operator is essentially selfadjoint. Given the integer moments, one knows how to evaluate $\mathbb{E}\{f(Y)\}$ for any polynomial $f \in \mathbb{C}[z]$, and sense all the moments are assumed finite, $\mathbb{E}\{f(Y)\}$ is always finite, moreover it is independent of which of the possibly nonunique measure α solving the moment problem we choose. (Here we implicitly assume at least one solution exists, i.e., we assume that the condition for existence has been satisfied. It is clear, though more cumbersome, how to define the inner product in terms of the sequence ($\gamma_n : n \in \mathbb{N}$), alone.) One can complete $\mathbb{C}[z]$ to a Hilbert space $\mathcal{H}_\gamma$ with respect to the inner product $\langle f, g\rangle = \mathbb{E}\{f(Y)g^*(Y)\}$ where $g^*(z) = \overline{g(\overline{z})}$. Then there is a densely defined operator A on $\mathcal{H}_\gamma$ with domain $D(A) = \mathbb{C}[z]$ given by $A[f(z)] = zf(z)$. A necessary and sufficient condition for uniqueness of α solving the classical moment problem is that A is essentially selfadjoint.

Though this gives an elegant solution to the problem, it is not entirely trivial to determine when A is essentially selfadjoint. One condition is the following: if the set of moments satisfy

$$|\gamma_n| \leq C R^n n! \tag{5}$$

for some $C, R \in \mathbb{R}_+$, then A is essentially selfadjoint. On the opposite side, there is the following result: suppose there is one solution of the moment problem α such that α is absolutely continuous with respect to Lebesgue measure and $\frac{d\alpha(z)}{dz} = F(z) \in (0,1]$ for all $z \in \mathbb{R}$ and

$$\int_{\mathbb{R}} \frac{-\log(F(z))}{1+z^2}\, dz < \infty\,, \tag{6}$$

then the classical moment problem is indeterminate. Incidentally, whenever the classical moment problem is indeterminate (uniqueness does not hold) then there are infinitely many solutions, and much more can be said about the structure of the simplex of all solutions. Neither the positive nor negative criterion for determinacy is sharp, but they are rather useful in practice. One really should consult [9] for more information. Since the question of uniqueness of analytic functions as determined from their values at the integers will arise again, it is interesting to have such explicit criteria in this context.

For now, we observe that $\alpha_k \sigma^{2k} \leq (2k)! \sigma^{2k}$, so that the positive criterion for determinacy 5 is satisfied. Therefore, $\nu(A) = \mathbb{E}\{Y \in A\}$ has a unique solution. Since we know the Gaussian process gives one solution, it is the unique solution.

1.3 Gaussian processes

Given an index set $\mathcal{A}$, a kernel on $\mathcal{A}$ is a function $Q : \mathcal{A} \times \mathcal{A} \to \mathbb{C}$. We will always consider real kernels, so that the range is $\mathbb{R}$ not $\mathbb{C}$. Such a kernel is positive semidefinite iff

1. It is symmetric; i.e., $Q(\alpha, \alpha') = Q(\alpha', \alpha)$ for all $\alpha, \alpha' \in \mathcal{A}$;
2. For any $n \in \mathbb{N} = \{1, 2, \dots\}$ and for any $\alpha_1, \dots, \alpha_n \in \mathcal{A}$ and for any $\gamma_1, \dots, \gamma_n \in \mathbb{R}$, one has $\sum_{i,j=1}^n Q(\alpha_i, \alpha_j)\gamma_i \gamma_j \geq 0$.

In the case that $\mathcal{A}$ is finite, which is what we should consider first, this is equivalent to the fact that the matrix $(Q(\alpha, \alpha') : \alpha, \alpha' \in \mathcal{A})$ is positive semidefinite, which simply means that the matrix is real, symmetric and all its eigenvalues are in $\mathbb{R}_+ = [0, \infty)$.

A random process $(Z_\alpha : \alpha \in \mathcal{A})$ is said to be a Gaussian process[1] if, for any $n \in \mathbb{N}$ and any $\alpha_1, \dots, \alpha_n \in \mathcal{A}$ and any $\gamma = (\gamma_1, \dots, \gamma_n) \in \mathbb{R}$, the random variable $Z = \sum_{i=1}^n \gamma_i Z_{\alpha_i}$ is a Gaussian random variable. Note that the covariance kernel $Q(\alpha, \alpha') = \mathbb{E}\{Z_\alpha Z_{\alpha'}\}$ is positive semidefinite, for any stochastic process. But for a Gaussian process it uniquely determines the process because it uniquely determines the variance of Z,

$$\sigma^2 = \sum_{i,j=1}^n Q(\alpha_i, \alpha_j)\gamma_i \gamma_j \,. \tag{7}$$

If $Q : \mathcal{A} \times \mathcal{A} \to \mathbb{R}$ is any positive semidefinite kernel, then there is a unique (modulo distributional equivalence) Gaussian process with this as its covariance matrix. For example, if $\mathcal{A} = \{1, \dots, n\}$ then one can allow $(X_1, \dots, X_n)$ to be independent Gaussians in $\mathcal{N}(0,1)$ and take $Z_i = \sum_{j=1}^n [Q^{1/2}]_{i,j} X_j$ for $i = 1, \dots, n$. It is sometimes useful to prove that a kernel is positive semidefinite by showing that it is the covariance of a Gaussian process. Notationally, if the covariance kernel is $Q(\alpha, \alpha') = \delta_{\alpha,\alpha'}$ we will call this an iid (independent, identically distributed) $\mathcal{N}(0,1)$ process.

The fundamental property of Gaussian processes is again Wick's rule. In this case, if $f \in \mathbb{C}[z_\alpha : \alpha \in \mathcal{A}]$ is any polynomial (in finitely many coordinates), then Wick's rule states that for any $\alpha_1 \in \mathcal{A}$,

$$\mathbb{E}\{Z_{\alpha_1} f(Z_\alpha : \alpha \in \mathcal{A})\} = \sum_{\alpha_2 \in \mathcal{A}} Q(\alpha_1, \alpha_2) \mathbb{E}\left\{\frac{\partial f}{\partial z_(\alpha_2)}(Z_\alpha : \alpha \in \mathcal{A})\right\} \,. \tag{8}$$

One can define the L^2-norm on the set of polynomials by $\|f\|_2 = \mathbb{E}\{f(Z)^2\}^{1/2}$. In this case statistical independence and perpendicularity are the same. Taking the closure of the polynomials gives a Hilbert space of functions $f : \mathbb{R}^{\mathcal{A}} \to \mathbb{C}$. In this framework one can extend Wick's rule to the appropriate Sobolev space, and we will call this *the generalized Wick's rule.* (In practice it will always be obvious how to approximate by polynomials.)

Homework 3. Prove that for any $n \in \mathbb{N}$, and any $\alpha_1, \dots, \alpha_n \in \mathcal{A}$, one has

$$\mathbb{E}\{Z_{\alpha_1}, \dots, Z_{\alpha_n}\} = \begin{cases} 0 & n \text{ is odd;} \\ \sum'_{\pi \in S_n} \prod_{k=1}^{n/2} Q(\alpha_{\pi(2k-1)}, \alpha_{\pi(2k)}) & n \text{ is even.} \end{cases} \tag{9}$$

where the primed sum means to restrict to those permutations satistifying $\pi(1) < \pi(3) < \cdots < \pi(n-1)$ and $\pi(2k-1) < \pi(2k)$ for every $k = 1, \dots, n/2$.

[1] I apologize for using Z for a Gaussian process here, when shortly the symbol Z will have a different meaning, as the partition function.

2 Slepian's lemma

The following result is key for Lectures 3 on. But it also arises today, so we mention it now.

Lemma 2.1 *Let $(X_\alpha : \alpha \in \mathcal{A})$ and $(Y_\alpha : \alpha \in \mathcal{A})$ be Gaussian processes with covariance kernels Q and P, respectively. Suppose, furthermore, that $Q(\alpha, \alpha') \leq P(\alpha, \alpha')$ for every $\alpha, \alpha' \in \mathcal{A}$. Let f be a function (in the intersection of the appropriate Sobolev spaces) such that*

$$\forall \alpha, \alpha' \in \mathcal{A} \text{ such that } Q(\alpha, \alpha') < P(\alpha, \alpha'), \ \frac{\partial^2 f}{\partial z_\alpha \partial z_{\alpha'}} \geq 0 \,. \tag{10}$$

Then $\mathbb{E}\{f(X_\alpha : \alpha \in \mathcal{A})\} \leq \mathbb{E}\{f(Y_\alpha : \alpha \in \mathcal{A})\}$.

In this generality, the lemma was first proved in [4], along with other generalizations to non-Gaussian processes. This reference is available online, and there is a pointer to it on the references section of the class webpage. Talagrand refers to [6], although his proof seems to come later. The proof we give is essentially that of [4]. This procedure is called "quadratic replica interpolation" by Guerra and Toninelli, and the "smart path" method by Talagrand.

Proof. The result will follow by approximation if we prove it for all polynomials. In this case we may assume that $\mathcal{A} = \{1, \dots, n\}$ for some $n \in N$ since a polynomial depends on only finitely many variables. Take an independent coupling of $X = (X_1, \dots, X_n)$ and $Y = (Y_1, \dots, Y_n)$, and for every $t \in [0, 1]$ define $Z_{i,t} = \sqrt{1-t} X_i + \sqrt{t} Y_i$ for $i = 1, \dots, n$ Then $Z = (Z_{1,t}, \dots, Z_{n,t})$ is a Gaussian process for all $t \in [0, 1]$, and for any polynomial g, the map $t \mapsto \mathbb{E}\{g(Z_1, \dots, Z_n)\}$ is continuous. Also, for $0 < t < 1$, observing that

$$\partial_t Z_{i,t} = \frac{-1}{2\sqrt{1-t}} X_i + \frac{1}{2\sqrt{t}} Y_i \tag{11}$$

we see that

$$(Z_t; \partial_t Z_t) = (Z_{1,t}, \dots, Z_{n,t}, \partial_t Z_{1,t}, \dots, \partial_t Z_{n,t}) \tag{12}$$

is another Gaussian process. The important formula is that

$$\mathbb{E}\{Z_{i,t} \partial_t Z_{j,t}\} = \frac{1}{2}(P_{ij} - Q_{ij}) \geq 0 \,. \tag{13}$$

Now, let $E_t = \mathbb{E}\{f(Z_t)\}$. For $0 < t < 1$, this is evidently differentiable with

$$\partial_t E_t = \mathbb{E}\{\partial_t f(Z_t)\} = \sum_{i=1}^{n} \mathbb{E}\left\{\frac{\partial f}{\partial z_i}(Z_t) \partial_t Z_{i,t}\right\} \tag{14}$$

by the chain-rule. Then, using Wick's rule,

$$\partial_t E_t = \sum_{i,j=1}^{n} \mathbb{E}\left\{\frac{\partial^2 f}{\partial z_i \partial z_j}\right\} \mathbb{E}\{Z_{j,t} \partial_t Z_{i,t}\} \,. \tag{15}$$

Using equation (13) and (10) we see that this is nonnegative. Therefore, $E_0 \leq E_1$, and that is what we wanted to prove.□

3 Introduction to the Sherrington-Kirkpatrick Model

For $N \in \mathbb{N} = \{1, 2, \dots, \}$ let $(J_{ij} : i, j \in \{1, \dots, N\})$ be an iid $\mathcal{N}(0, 1)$ Gaussian process. Define $\Omega_N = \{+1, -1\}^N = \{\sigma = (\sigma_1, \dots, \sigma_N)\}$, where $\sigma_1, \dots, \sigma_N$ are Ising variables. The Sherrington-Kirkpatrick is a mean-field spin glass, determined by the random Hamiltonian function $H_N : \Omega_N \times \mathbb{R} \to \mathbb{R}$ where

$$H_N(\sigma, h) = \frac{-1}{\sqrt{2N}} \sum_{i,j=1}^{N} J_{ij} \sigma_i \sigma_j - h \sum_{i=1}^{N} \sigma_i \,. \tag{16}$$

Here the parameter h is the external magnetic field. It is introduced primarily to break spin-flip symmetry but it plays a second important role which will become apparent when we talk about the *cavity method* in Lectures 3 and 4. (By this we mean what the physicists do when they use this term, not what Talagrand does). Note that

$$K_N(\sigma) = \frac{-1}{\sqrt{2N}} \sum_{i,j=1}^{N} J_{ij}\sigma_i\sigma_j \tag{17}$$

defines a Gaussian process with

$$\begin{aligned}
\mathbb{E}\{K_N(\sigma)K_N(\sigma')\} &= \frac{1}{2N}\sum_{i,j=1}^{N}\sum_{k,l=1}^{N}\mathbb{E}\{J_{ij}J_{kl}\}\sigma_i\sigma_j\sigma'_k\sigma'_l \\
&= \frac{1}{2N}\sum_{i,j=1}^{N}\sum_{k,l=1}^{N}\delta_{ik}\delta_{jl}\sigma_i\sigma_j\sigma'_k\sigma'_l \\
&= \frac{1}{2N}\sum_{i,j=1}^{N}\sigma_i\sigma_j\sigma'_i\sigma'_j \\
&= \frac{N}{2}q^2_{\sigma,\sigma'}
\end{aligned} \tag{18}$$

where $q_{\sigma,\sigma'}$ is the *overlap*

$$q_{\sigma,\sigma'} := \frac{1}{N}\sum_{i=1}^{N}\sigma_i\sigma'_i\,. \tag{19}$$

Note that $q_{\sigma,\sigma'} \in \{-1, -1+\frac{1}{N}, \dots, 1\}$. Also, note that $q^2_{\sigma,\sigma'}$ is a positive semidefinite kernel because we have realized it as the covariance of a Gaussian process. Indeed, $q_{\sigma,\sigma'}$ is a positive semidefinite kernel because it is the covariance of the Gaussian process $(H'_N(\sigma) : \sigma \in \Omega_N)$, where

$$H'_N(\sigma) = \frac{1}{\sqrt{N}}\sum_{i=1}^{N} J_i\sigma_i \tag{20}$$

and $J_1, \dots, J_N$ are iid $\mathcal{N}(0,1)$ random variables. This will be important later on.

3.1 Physical Motivation

We will skip the physical motivation which was explained in class. For a good description of spin glasses, consult the magnetism book by Mattis.

The physical quantity one studies for a finite spin system is the Gibbs state. This is a measure on Ω_N such that the measure of a particular spin configuration σ is $e^{-\beta H_N(\sigma,h)}/Z_N(\beta,h)$, where $Z_N(\beta,h)$ is a normalization called the partition function

$$Z_N(\beta,h) = \sum_{\sigma\in\Omega_N} e^{-\beta H_N(\sigma,h)}\,. \tag{21}$$

We will write expectations with respect to the Gibbs measure using angle brackets,

$$\langle f\rangle_{\beta,H_N} = \sum_{\sigma\in\Omega_N} Z_N(\beta,h)^{-1}e^{-\beta H_N(\sigma,h)}f(\sigma)\,. \tag{22}$$

Be warned that these numbers are still random since we have not yet taken the expectation with respect to the random couplings.

It turns out that quite a bit of information is contained in the partition function, itself. For example, the free energy is

$$A_N(\beta,h) = -\frac{1}{\beta}\log Z_N(\beta,h)\,, \tag{23}$$

(this is extensive) and the pressure (which is intrinsic) is

$$P_N(\beta, h) = -\frac{1}{N} \log Z_N(\beta, h)\,. \tag{24}$$

Note that $P_N(\beta, h)$ is jointly convex in β and βh. Two examples of physical quantities which one can calculate from the pressure alone are the average magnetization

$$M(\beta, h) = \langle \frac{1}{N} \sum_{i=1}^{N} \sigma_i \rangle_{beta, H_N} = \frac{1}{\beta} \frac{\partial}{\partial h} P_N(\beta, h) \tag{25}$$

and the average spin-spin correlation

$$\frac{1}{N^2} \sum_{i,j=1}^{N} (\langle \sigma_i \sigma_j \rangle - \langle \sigma_i \rangle \langle \sigma_j \rangle) = \frac{1}{\beta^2} \frac{\partial^2}{\partial h^2} P_N(\beta, h)\,. \tag{26}$$

Note that for $h = 0$, the latter quantity has the physical significance of the magnetic susceptibility. It is actually the pressure and the free energy which we will be concerned with, from here on.

Using the martingale method, Pastur and Scherbina [8], proved that

$$\lim_{N \to \infty} \mathbb{E}\{(P_N(\beta, h) - \mathbb{E}\{P_N(\beta, h)\})^2\} = 0\,, \tag{27}$$

which is called "concentration of measure". However, somewhat surprisingly, it was only proved last year that

$$\lim_{N \to \infty} \mathbb{E}\{P_N(\beta, h)\} \tag{28}$$

exists for all β and h. This is the subject of Lecture 3.

For any physical model, it is helpful if there is some early success in its analysis, even if it turns out later that the success was bogus. This is the case with Sherrington and Kirkpatrick's replica method. Supposedly their method allows one to calculate the quenched pressure

$$p_N(\beta, h) = \mathbb{E}\{P_N(\beta, h)\} \tag{29}$$

for all β and h.

4 The Replica Method

Rather than studying Sherrington and Kirkpatrick's wrong paper, we prefer to describe van Hemmen and Palmer's paper [2], which is mathematically corrrect.

Begin with the following premise. We want to calculate $p_N(\beta, h) = \mathbb{E}\{\frac{1}{N} \log Z_N(\beta, h)\}$. It is a good idea is to analyze the distribution of $\log Z_N(\beta, h)$ through its Laplace transform. For $\lambda \in \mathbb{R}$, define

$$\psi_N(\beta, h; \lambda) = \mathbb{E}\{e^{\lambda \log Z_N(\beta, h)}\} = \mathbb{E}\{Z_N(\beta, h)\}\,. \tag{30}$$

The derivative of the Laplace transform at 0 is the expecation. Rather than study this quantity, we prefer to consider

$$\phi_N(\beta, h; \lambda) = \frac{1}{N\lambda} \log\, \phi_N(\lambda). \tag{31}$$

For $\lambda = 0$, define $\phi_N(\beta, h; 0) = p_N(\beta, h)$ which makes $\phi_N(\beta, h; \lambda)$ continuous in λ, β and h. We can prove that

$$\phi(\beta, h; \lambda) = \lim_{N \to \infty} \phi_N(\beta, h; \lambda) \tag{32}$$

exists for all β, h and λ. In fact, this is the following homework.

Homework 4. (a) *Prove that for $\lambda \neq 0$ and $\binom{\lambda}{2} \geq 0$, the sequence $(N\phi_N(\lambda))_{N=1}^{\infty}$ is subadditive.*
(b) *Prove that for $\lambda \neq 0$ and $\binom{\lambda}{2} \leq 0$, the sequence $(N\phi_N(\lambda))_{N=1}^{\infty}$ is superadditive.*

(This homework is a little too difficult as it stands, but it will be easy after lecture 3. Until then: think Slepian's lemma.)

It is rather simple to obtain a formula for $\phi(n)$, when $n \in \mathbb{N}$ as a straightforward optimization problem. We state this, here.

Theorem 4.1 *(van Hemmen and Palmer) For $n \in \mathbb{N}$, $\phi(n) = \frac{\beta^2}{4} + \frac{1}{n}\sup_{\boldsymbol{Q}} \mathcal{F}_n(\boldsymbol{Q})$ where $\boldsymbol{Q} = (Q_{\alpha,\alpha'} : 1 \leq \alpha < \alpha' \leq n)$ is a real triangular array, and*

$$\mathcal{F}_n(\boldsymbol{Q}) = -\frac{\beta^2}{2}\sum_{\alpha<\alpha'} Q_{\alpha,\alpha'}^2 + \log\left(\sum_{S\in\Omega_n} e^{\beta h \sum_{\alpha=1}^n S_\alpha + \beta^2 \sum_{\alpha<\alpha'} Q_{\alpha,\alpha'} S_\alpha S_{\alpha'}}\right) \tag{33}$$

Theorem 4.2 *(van Hemmen, Palmer and Lieb) If $h = 0$, the optimization problem listed above has at least one maximizer with $Q_{\alpha,\alpha'} = q$ for some $q \in [0, 1]$.*

(We will not prove this result. See [2].)

Corollary 4.3 *Recalling the definition $d\mu(z) = e^{-z^2/2}dz/\sqrt{2\pi}$, one has the formula*

$$\phi(n) = \frac{\beta^2}{4}(1 - 2q - (n-1)q^2) + \log(2) + \frac{1}{n}\log\int d\mu(z)\,\cosh^n[\beta h + \beta\sqrt{q}\,z] \tag{34}$$

where q solves the implicit equation $q = \int d\alpha_{n,q}(z)\tanh^2(\beta[h + \sqrt{q}\,z])$, in which

$$d\alpha_{n,q}(z) = \frac{d\mu(z)\,\cosh^n(\beta[h+\sqrt{q}\,z])}{\int d\mu(z')\,\cosh^n(\beta[h+\sqrt{q}\,z'])}\,. \tag{35}$$

Sherrington and Kirkpatrick also arrived at the conclusion of the corollary, albeit with several unjustified assumptions. The "replica trick" which comes next is an attempt to extrapolate the value of $p(\beta, h)$ from the values of $\phi(\beta, h)$. In the corollary above, n plays the role of a real parameter. Sherrington and Kirkpatrick, simply took $n \to 0$, formally to arrive at their "result".

Claim 4.1 *(Sherrington and Kirkpatrick : not universally valid)*

$$p(\beta, h) = \frac{\beta^2}{4}(1-q)^2 + \log(2) + \int d\mu(z)\,\log\cosh(\beta[h+\sqrt{q}\,z]) \tag{36}$$

where $q = q(\beta, h)$ solves the implicit equation

$$q = \int d\mu(z)\,\tanh^2(\beta[h+\sqrt{q}\,z])\,. \tag{37}$$

This would be quite an attractive result, if it were true. In fact, there are ranges of (β, h) where it is true, although this was proved by completely different methods, and is generally a hard result. However it is not true for all (β, h). Sherrington and Kirkpatrick, themselves, later pointed out that it gives a negative entropy at zero temperature, which is impossible in a discrete spin system, since the entropy equals the log of the number of ground state configurations, and there is always at least one. In the next lecture we will learn Parisi's rather creative "solution" to this dilemma. It is commonly believed that the result of Parisi's ansatz is correct, although the current attempts at rigorizing the result use entirely different methods.

4.1 Proofs

Proof (of Corollary 2.3). An important identity is the following Hubbard-Stratonovich formula

$$\forall\lambda \in \mathbb{C}\,,\quad e^{\lambda^2/2} = \int_{\mathbb{R}} d\mu(z)e^{\lambda z}\,. \tag{38}$$

To prove this, note that $\int d\mu(z) e^{\lambda z - \frac{1}{2}\lambda^2} = \int d\mu(z-\lambda)$. If $\lambda \in \mathbb{R}$, this integral is clearly 1. Using Cauchy's integral formula, one determines that $\int_{\mathbb{R}} d\mu(z+it) = \int_{\mathbb{R}} d\mu(z)$ for any $t \in \mathbb{R}$, because $e^{-(z+is)^2} \to 0$ uniformly for $s \in [-t,t]$, which proves the identity.

We observe that

$$\mathcal{F}_n = -\frac{\beta^2}{2}\binom{n}{2} q^2 + \log\left(\sum_{S \in \Omega_n} e^{\beta h \sum_{\alpha=1}^n S_\alpha + (\beta^2 q/2)\left(\sum_{\alpha=1}^n S_\alpha\right)^2}\right) - \frac{n\beta^2 q}{2} \tag{39}$$

when $Q_{\alpha,\alpha'} \equiv q$. By the Hubbard-Stratonovich formula,

$$\begin{aligned}\sum_{S \in \Omega_n} e^{\beta h \sum_{\alpha=1}^n S_\alpha + (\beta^2 q/2)\left(\sum_{\alpha=1}^n S_\alpha\right)^2} &= \int d\mu(z) \sum_{S \in \Omega_n} e^{\beta[h+\sqrt{q}\,z]\sum_{\alpha=1}^n S_\alpha} \\ &= \int d\mu(z)\, \cosh^n(\beta[h+\sqrt{q}\,z])\,.\end{aligned} \tag{40}$$

This leads to (34).

To obtain the implicit formula for the optimal q, observe that any $\boldsymbol{Q}$ which optimizes $\mathcal{F}_n$ solves

$$Q_{\alpha,\alpha'} = \langle S_\alpha S_{\alpha'}\rangle_{\beta h, \beta^2 \boldsymbol{Q}} \tag{41}$$

where $\langle \cdots \rangle_{\beta h, \beta^2 \boldsymbol{Q}}$ is the Gibbs state at inverse-temperature $k_B/T = 1$ for the Hamiltonian

$$\widetilde{H}_n(S_1, \dots, S_n) = -\beta h \sum_{\alpha=1}^n S_\alpha - \beta^2 \sum_{\alpha < \alpha'} Q_{\alpha,\alpha'} S_\alpha S_{\alpha'}\,. \tag{42}$$

Using Theorem 2, we know that there is an optimizer $\boldsymbol{Q}$ such that $Q_{\alpha,\alpha'} = q$ for all $\alpha < \alpha'$. Therefore, this q must satisfy the same equation. Using the Hubbard-Stratonovich formula allows one to write

$$\langle S_\alpha S_{\alpha'}\rangle_{\beta h, \beta^2 \boldsymbol{Q}} = \frac{\int d\alpha_{n,q}(z) \tanh^2(\beta[h+\sqrt{q}\,z])}{\int d\alpha_{n,q}(z)} \tag{43}$$

for this choice of $\boldsymbol{Q}$. □

Proof (of Theorem 4.1). Observe that

$$Z_N(\beta,h)^n = \sum_{\sigma^1,\dots,\sigma^n \in \Omega_N} e^{-[\beta H_N(\sigma^1) + \cdots + \beta H_N(\sigma^n)]}\,. \tag{44}$$

Therefore,

$$\mathbb{E}\{Z_N(\beta,h)^n\} = \sum_{\sigma^1,\dots,\sigma^n \in \Omega_N} e^{\beta h \sum_{i=1}^N \sum_{\alpha=1}^n \sigma_i^\alpha} \mathbb{E}\{\exp(\frac{\beta}{\sqrt{2N}} \sum_{i,j=1}^N J_{i,j} \sum_{\alpha=1}^n \sigma_i^\alpha \sigma_j^\alpha)\} \tag{45}$$

Now $K_N^{(n)}(\sigma^1,\dots,\sigma^n)$, defined by

$$K_N^{(n)}(\sigma^1,\dots,\sigma^n) = \frac{\beta}{\sqrt{2N}} \sum_{i,j=1}^N J_{i,j} \sum_{\alpha=1}^n \sigma_i^\alpha \sigma_j^\alpha) \tag{46}$$

is a Gaussian random variable, with variance equal to

$$\mathbb{E}\{K_N^{(n)}(\sigma^1,\dots,\sigma^n)^2\} = \frac{\beta^2 N}{2}(n + 2\sum_{\alpha<\alpha'} q_{\alpha,\alpha'}^2) \tag{47}$$

where $q_{\alpha,\alpha'}$ is our abbreviation for the overlap between σ^α and $\sigma^{\alpha'}$. Since we know how to take the expectation of an exponential of a Gaussian random variable, we have

$$\mathbb{E}\{Z_N(\beta,h)^n\} = e^{\beta^2 N n/4} \sum_{\sigma^1,\dots,\sigma^n \in \Omega_N} e^{\beta h \sum_{i=1}^N \sum_{\alpha=1}^n \sigma_i^\alpha} e^{(\beta^2 N/2)\sum_{\alpha<\alpha'} q_{\alpha,\alpha'}^2}\,. \tag{48}$$

One uses the Hubbard-Stratonovich transform, $\binom{n}{2}$ times, to obtain

$$\begin{aligned} e^{(\beta^2 N/2)\sum_{\alpha<\alpha'} q^2_{\alpha,\alpha'}} &= \prod_{\alpha<\alpha'} \int \frac{dQ_{\alpha,\alpha'} e^{-Q^2_{\alpha,\alpha'}/2}}{\sqrt{2\pi}} \, e^{\beta\sqrt{N} Q_{\alpha,\alpha'} q_{\alpha,\alpha'}} \\ &= \prod_{\alpha<\alpha'} \int dQ_{\alpha,\alpha'} \sqrt{\frac{N\beta^2}{2\pi}} \exp(\beta^2 N(-\frac{1}{2}Q^2_{\alpha,\alpha'} + Q_{\alpha,\alpha'} q_{\alpha,\alpha'}))\,. \end{aligned} \tag{49}$$

Plugging this back into the formula for $\mathbb{E}\{Z_N(\beta,h)^n\}$ yields

$$\begin{aligned} \mathbb{E}\{Z_N(\beta,h)^n\} = e^{\beta^2 Nn/4} \int \cdots \int \left[\prod_{\alpha<\alpha'} dQ_{\alpha,\alpha'}\right] \left(\frac{\beta N}{2\pi}\right)^{n/2} \exp(-\frac{\beta^2 N}{2} \sum_{\alpha<\alpha'} Q^2_{\alpha,\alpha'}) \times \\ \times \sum_{\sigma^1,\ldots,\sigma^n \in \Omega_N} \exp\left(\beta h \sum_{i=1}^{N}\sum_{\alpha=1}^{n} \sigma_i^\alpha + \beta^2 \sum_{i=1}^{N}\sum_{\alpha<\alpha'} Q_{\alpha,\alpha'} \sigma_i^\alpha \sigma_i^{\alpha'}\right) \end{aligned} \tag{50}$$

by virtue of the definition of $q^{\alpha,\alpha'}$.

Now observe that

$$\begin{aligned} &\sum_{\sigma^1,\ldots,\sigma^n \in \mathbb{Z}_2^N} \exp\left(\beta h \sum_{i=1}^{N}\sum_{\alpha=1}^{n} \sigma_i^\alpha + \beta^2 \sum_{i=1}^{N}\sum_{\alpha<\alpha'} Q_{\alpha,\alpha'} \sigma_i^\alpha \sigma_i^{\alpha'}\right) \\ &\quad = \prod_{i=1}^{N} \sum_{\sigma_i^1,\ldots,\sigma_i^n \in \mathbb{Z}_2} \exp\left(\beta h \sum_{\alpha=1}^{n} \sigma_i^\alpha + \beta^2 \sum_{\alpha<\alpha'} Q_{\alpha,\alpha'} \sigma_i^\alpha \sigma_i^{\alpha'}\right) \\ &\quad = \sum_{(S_1,\ldots,S_n)\in\Omega_n} \exp\left(N\beta h \sum_{\alpha=1}^{n} S_\alpha + N\beta^2 \sum_{\alpha<\alpha'} Q_{\alpha,\alpha'} S_\alpha S_{\alpha'}\right)\,, \end{aligned} \tag{51}$$

where we have introduced the notation $S = (S_1, \ldots, S_n) \in \Omega_n$ to stand for a typical n-fold replica of a spin at a site $i \in \{1, \ldots, N\}$. Therefore, one obtains

$$\begin{aligned} \mathbb{E}\{Z_N(\beta,h)^n\} = e^{\beta^2 Nn/4} \int \cdots \int \left[\prod_{\alpha<\alpha'} dQ_{\alpha,\alpha'}\right] \left(\frac{\beta N}{2\pi}\right)^{n/2} \\ \times \sum_{S\in\Omega_n} \exp\left(N\left[-\beta^2 \sum_{\alpha<\alpha'} \frac{1}{2} Q^2_{\alpha,\alpha'} + \beta^2 \sum_{\alpha<\alpha'} Q_{\alpha,\alpha'} S_\alpha S_{\alpha'} + \beta h \sum_{\alpha=1}^{n} S_\alpha\right]\right)\,. \end{aligned} \tag{52}$$

If so desired, one can make the role of N even more explicit by defining

$$\mathcal{F}_n(\boldsymbol{Q}) = -\frac{\beta^2}{2} \sum_{\alpha<\alpha'} Q^2_{\alpha,\alpha'} + \log\left(\sum_{S\in\Omega_n} e^{\beta h \sum_{\alpha=1}^{n} S_\alpha + \beta^2 \sum_{\alpha<\alpha'} Q_{\alpha,\alpha'} S_\alpha S_{\alpha'}}\right)\,. \tag{53}$$

This allows one to rewrite

$$\mathbb{E}\{Z_N(\beta,h)^n\} = e^{\beta^2 Nn/4} \int \cdots \int \left[\prod_{\alpha<\alpha'} dQ_{\alpha,\alpha'}\right] \left(\frac{\beta N}{2\pi}\right)^{n/2} e^{N\mathcal{F}_n(\boldsymbol{Q})}\,. \tag{54}$$

From this one concludes

$$\phi_N(\beta,h;n) = \frac{\beta^2}{4} + \frac{1}{2N}\log(\beta N/2\pi) + \frac{1}{Nn}\log\int\left[\prod_{\alpha<\alpha'} dQ_{\alpha,\alpha'}\right] e^{N\mathcal{F}_n(\boldsymbol{Q})}\,. \tag{55}$$

Now, as $\mathcal{F}^{(n)}(\boldsymbol{Q})$ is a smooth function on $\mathbb{R}^{n(n-1)/2}$, such that as $\boldsymbol{Q}^2$ goes to infinity $\mathcal{F}^{(n)}(\boldsymbol{Q}) \sim -\frac{\beta^2}{2}\boldsymbol{Q}^2$, the method of steepest descents applies (c.f., [5]). Therefore, one determines that $\phi(\beta, h; n)$ exists (in a different way than Homework 4), and that

$$\phi(\beta, h; n) = \frac{\beta^2}{3} + \frac{1}{n} \max_{\boldsymbol{Q} \in \mathbb{R}^{n(n-1)/2}} \mathcal{F}^{(n)}(\boldsymbol{Q}) \,, \tag{56}$$

which was to be shown. □

References

[1] F. Guerra and F. Toninelli. The thermodynamic limit in mean field spin glass models. *Comm. Math. Phys.*, 230(1):71–79, 2002.

[2] J.L. van Hemmen and R.G. Palmer The replica method and a solvable spin glass model. *J. Phys. A* 12(4):563–580, 1979.

[3] Kiyosi Itô and Henry P. McKean. Diffusion Processes and their Sample Paths. Berlin, 1974.

[4] Kumar Joag-dev, Michael D. Perlman and Loren D. Pitt. Association of Normal Random Variables and Slepian's Inequality *Ann. Probab.* 11(2):451–455, 1983.

[5] M.J. Hoffman and J.E. Marsden, Basic complex analysis. Second edition. W. H. Freeman and Company, New York, 1987.

[6] J.-P. Kahane Une inégalité du type Slepian et Gordon sur les processus gaussiens. *Israel J. Math.*, **55**, 109–110, 1986.

[7] E.H. Lieb and M. Loss Analysis. 2 ed. AMS : GSM 14. Providence, RI, 2001.

[8] L.A. Pastur and M.V. Scherbina Absence of self-averaging of the order parameter in the Sherrington-Kirkpatrick model. *J. Stat. Phys.*, **62**, 1–19, 1991.

[9] The Classical Moment Problem as a Self-Adjoint Finite Difference Operator *Adv.in Math.*, **137**, 82–203, 1988.