

Lecture 2: Meanfield Approximation, Variational Principle and Landau Expansion

Naoki KAWASHIMA

ISSP, U. Tokyo

April 15, 2024

In this lecture, we see ...

- Molecular-field theory yields the equation of the state, not the free energy.
- Gibbs-Bogoliubov-Feynman inequality give us a systematic and flexible framework for constructiong the mean-field-type approximations.

Molecular field theory revisited

- In the molecular-field theory, the effect of **environment** is replaced by an additional term. For the Ising ferromagnet, the many-body problem is reduced to a 1-body problem about a spin, say S_0 ,

$$\mathcal{H} = -J \sum_{(ij)} S_i S_j - H \sum_i S_i \rightarrow \mathcal{H}_{\text{MF}} = -H_{\text{MF}} S_0 - H S_0$$

where (ij) is the nearest neighbor pair of sites.

- It is also argued that the right choice of the effective field is

$$H_{\text{MF}} = J \sum_j \langle S_j \rangle$$

- The uniformity condition $m = \langle S_i \rangle$ (independent of i) yields,

$$m = \tanh(\beta(H + zJm)) \quad (z = (\text{number of nearest-neighbors}))$$

Why should we complain?

- It is not generally obvious how we should choose the mean-field. (The order parameter may not have a simple form in terms of the model degrees of freedom.)
- In principle, we have multiple solutions of the self-consistent equation.
- The molecular-field theory does not tell us which solution we should choose.

Is there any general framework for the mean-field approximation?

Gibbs-Bogoliubov-Feynman (GBF) inequality

Theorem 1 (GBF inequality)

For two Hamiltonians $\mathcal{H}(S)$ and $\mathcal{H}_0(S)$ defined on the same space $S \in \Omega$,

$$F_v \equiv F_0 + \langle \mathcal{H} - \mathcal{H}_0 \rangle_0 \geq F, \quad (1)$$

where F and F_0 are the free-energies of \mathcal{H} and \mathcal{H}_0 respectively and $\langle \cdots \rangle_0$ is the thermal average with respect to \mathcal{H}_0 .

Variational calculation

When $\mathcal{H}(S)$ is the Hamiltonian of the system that we want to study but is not solvable, by taking $\mathcal{H}_0(\Lambda, S)$ that depends on a set of parameters Λ and is solvable for any Λ , the resulting $F_v(\Lambda)$ provides us with the computable upper bound of the correct free energy. To obtain the best upper-bound, we minimize it w.r.t. Λ .

GBF inequality from information-scientific view-point

Theorem 2 (Relation to Kullback-Leibler divergence)

The “error” in the variational free-energy is proportional to the Kullback-Leibler divergence of the equilibrium distribution of \mathcal{H}_0 relative to that of \mathcal{H} .

More precisely,

$$F_v - F = k_B T D_{\text{KL}}(\rho_0 | \rho) \quad (2)$$

where

$$\rho_0 \equiv e^{-\beta \mathcal{H}_0} / Z_0, \quad \text{and} \quad \rho \equiv e^{-\beta \mathcal{H}} / Z \quad (3)$$

and

$$D_{\text{KL}}(P|Q) \equiv \sum_S P(S) \log \frac{P(S)}{Q(S)} \quad (4)$$

(Z and Z_0 are partition functions of \mathcal{H} and \mathcal{H}_0 respectively.)

Proof of Theorem 2

$$\begin{aligned} D_{\text{KL}} \left(Z_0^{-1} e^{-\beta \mathcal{H}_0} \mid Z^{-1} e^{-\beta \mathcal{H}} \right) &= \sum_S \left(\frac{e^{-\beta \mathcal{H}_0(S)}}{Z_0} \log \left(\frac{e^{-\beta \mathcal{H}_0(S)}}{Z_0} \right) - \frac{e^{-\beta \mathcal{H}_0(S)}}{Z_0} \log \left(\frac{e^{-\beta \mathcal{H}(S)}}{Z} \right) \right) \\ &= -\log Z_0 + \langle -\beta \mathcal{H}_0 \rangle_0 + \log Z - \langle -\beta \mathcal{H} \rangle_0 \\ &= \beta F_0 - \beta \langle \mathcal{H}_0 \rangle_0 - \beta F + \beta \langle \mathcal{H} \rangle_0 \\ &= \beta (F_v - F) \end{aligned}$$

Kullback-Leibler information measure is positive

$$\begin{aligned} D_{\text{KL}}(P|Q) &= \sum_S P(S) \log \frac{P(S)}{Q(S)} = - \sum_S P(S) \log \frac{Q(S)}{P(S)} \\ &\geq - \sum_S P(S) \left(\frac{Q(S)}{P(S)} - 1 \right) \quad (\text{because } \log(x) \leq x - 1) \\ &= \sum_S (Q(S) - P(S)) = 1 - 1 = 0 \end{aligned}$$

Remark

This inequality together with Theorem 2 proves Theorem 1.

Quantum extension

- The theorems have been proved for classical systems. The corresponding quantum version of them can be also proved.
- For the extension, the KL divergence must be generalized to

$$D_{\text{KL}}(P|Q) \equiv \text{Tr } P(\log P - \log Q) \quad (5)$$

where P and Q are now density operators. (If a positive semi-definite operator X satisfies the normalization $\text{Tr } X = 1$, we call it a density operator.)

- The only non-trivial part in the proof of the quantum extension is the positivity of the KL information. The rest is straight-forward simply by replacing \sum_S by Tr .

Positivity of quantum version of KL divergence (1/2)

Theorem 3 (Positivity of quantum KL divergence)

For any density operators P and Q ,

$$D_{\text{KL}}(P|Q) \equiv \text{Tr } P \log P - P \log Q \geq 0 \quad (6)$$

Proof:

- Let us take the basis set in which P is diagonal, i.e., $P_{ij} = p_i \delta_{ij}$.
- Let U be the unitary operator U diagonalizing Q , i.e., $\Delta \equiv U^\dagger Q U$ is a diagonal operator. Then, $Q_{ij} = u_{ik} q_k u_{jk}^*$.
- With this u_{ij} ,

$$D_{\text{KL}}(P|Q) = \sum_i p_i \log p_i - \sum_{ij} p_i a_{ij} \log q_j \quad (a_{ij} \equiv |u_{ij}|^2) \quad (7)$$

- Notice that $a_{ij} \geq 0$, $\sum_i a_{ij} = \sum_j a_{ij} = 1$.

Positivity of quantum version of KL divergence (2/2)

- By defining $p'_{ij} = p_i a_{ij}$, $q'_{ij} = q_j a_{ij}$, we can regard them as the “classical” distribution function in the squared Hilbert space $H \times H \equiv \{(ij) \mid i, j \in H\}$ because

$$p'_{ij} \geq 0, \quad q'_{ij} \geq 0, \quad \sum_{ij} p'_{ij} = \sum_{ij} q'_{ij} = 1$$

- The two terms in (7) can be expressed as, respectively,

$$\begin{aligned} \sum_{ij} p'_{ij} \log p'_{ij} &= \sum_i p_i \log p_i + \sum_{ij} p_i a_{ij} \log a_{ij} \\ \sum_{ij} p'_{ij} \log q'_{ij} &= \sum_{ij} p_i a_{ij} \log q_j + \sum_{ij} p_i a_{ij} \log a_{ij}, \end{aligned}$$

- By subtracting the second from the first, we obtain $D_{KL}(P|Q) = \sum_{ij} p'_{ij} \log(p'_{ij}/q'_{ij})$, and this must be non-negative because the RHS is the classical KL divergence. (QED)

Some examples

- Ising model:** $\Lambda = (\Gamma)$

$$\mathcal{H} \equiv -J \sum_{(ij)} S_i S_j - H \sum_i S_i, \quad \mathcal{H}_0 \equiv -\Gamma \sum_i S_i$$

($\Gamma \neq 0$ represents ferromagnetism.)

- Hubbard model:** $\Lambda = (\{\Lambda_{ij\mu\nu}\})$

$$\mathcal{H} \equiv -t \sum_{(ij)\sigma=\uparrow\downarrow} c_{i\sigma}^\dagger c_{j\sigma} + \frac{1}{2} \sum_{i,j} V_{ij} n_i n_j - \mu \sum_i n_i$$

$$\mathcal{H}_0 \equiv \sum_{ij\mu\nu} \Lambda_{i\mu,j\nu} a_{i\mu} a_{j\nu} \quad (\Lambda_{i\mu,j\nu} = -\Lambda_{j\nu,i\mu})$$

where $n_i \equiv \sum_\sigma c_{i\sigma}^\dagger c_{i\sigma}$, and $(a_{i1}, a_{i2}, a_{i3}, a_{i4}) = (c_{i\uparrow}, c_{i\downarrow}, c_{i\uparrow}^\dagger, c_{i\downarrow}^\dagger)$.
($\Lambda_{i1,j2} \neq \Lambda_{i4,j3}$ represents superconductivity.)

Variational approximation to the Ising model (1/2)

- For the target Hamiltonian $\mathcal{H} \equiv -J \sum_{\langle ij \rangle} S_i S_j - H \sum_i S_i$, let us take the “trial” Hamiltonian $\mathcal{H}_0 \equiv -\Lambda \sum_i S_i$. where Λ is a variational parameter.
- Then, our variational free energy is

$$\begin{aligned} F_v &= F_0 + \langle \mathcal{H} - \mathcal{H}_0 \rangle_0 \\ &= \langle \mathcal{H} \rangle_0 - S_0 T \end{aligned}$$

where $S_0 \equiv T^{-1}(\langle \mathcal{H}_0 \rangle_0 - F_0)$ is the entropy of the \mathcal{H}_0 system

- By introducing $m \equiv \langle S_i \rangle_0 = \tanh \beta \Lambda$,

$$\langle \mathcal{H} \rangle_0 = -\frac{z}{2} N J m^2 - H N m \quad S_0 = N \sigma(m) \quad (8)$$

$$\sigma(m) \equiv -k_B \left(\frac{1+m}{2} \log \frac{1+m}{2} + \frac{1-m}{2} \log \frac{1-m}{2} \right) \quad (9)$$

Variational approximation to the Ising model (2/2)

Variational free-energy density

$$f_v \equiv \frac{F_v}{N} = -\frac{z J m^2}{2} - H m - T \sigma(m)$$

- Moreover, the GBF inequality tells us that we should take the solution with the smallest f_v with respect to λ . Since f_v depends on λ only through m , the stationary condition is $\partial f_v / \partial m = 0$.
- From this, we obtain the same as the molecular-field approx.:

$$\frac{\partial f_v}{\partial m} = 0 \quad \Rightarrow \quad m = \tanh \beta H_{\text{MF}} \quad (H_{\text{MF}} \equiv z J m + H). \quad (10)$$

Mean-field free energy — Landau expansion

- We have also obtained the explicit expression for the free energy.
- Since its behavior near $m \approx 0$ is most important for the critical phenomena, let us expand f_v with respect to m .

$$f_v = -\frac{zJ}{2}m^2 - Hm - k_B T \left(\log 2 - \frac{m^2}{2} - \frac{m^4}{12} + \dots \right) \quad (11)$$

- From the condition (coefficient of m^2) = 0, we obtain $k_B T_c = zJ$.
- Near $T \approx T_c$, we can get a generic form of Landau expansion:

$$\Delta f_v \approx -Hm + tm^2 + um^4 \quad (12)$$

where $\Delta f_v \equiv f_v - f_v^0$, $f_v^0 \equiv -zJ \log 2$, $t \equiv (k_B T - zJ)/2$ and $u \equiv zJ/12$.

- It is also clear we should choose non-zero m solution below T_c .

Exercise 2.1: Verify (8) and (9).

$$\begin{aligned} \langle \mathcal{H} \rangle_0 &= -J \sum_{(ij)} \langle S_i S_j \rangle_0 - H \sum_i \langle S_i \rangle_0 = -J \sum_{(ij)} m^2 - H \sum_i m \\ &= N \times \left(-\frac{2}{z} J m^2 - H m \right) \end{aligned}$$

$$S_0 = \frac{1}{T} (\langle \mathcal{H}_0 \rangle_0 - F_0)_0 = \frac{1}{T} (-\Lambda m + NT \log (e^{\beta \Lambda} + e^{-\beta \Lambda}))$$

$$m \equiv \tanh \beta \Lambda \Rightarrow \beta \Lambda = \frac{1}{2} \log \frac{1+m}{1-m}$$

$$\begin{aligned}
S_0/N &= -\frac{m}{2} \log \frac{1+m}{1-m} + \frac{1}{2} \log (e^{2\beta\Lambda} + 2 + e^{-2\beta\Lambda}) \\
&= -\frac{m}{2} \log \frac{1+m}{1-m} + \frac{1}{2} \log \left(\frac{1+m}{1-m} + 2 + \frac{1-m}{1+m} \right) \\
&= -\frac{1+m}{2} \log \frac{1+m}{2} - \frac{1-m}{2} \log \frac{1-m}{2}
\end{aligned}$$

Exercise 2.2: At the point where we have arrived at (10), m is just the magnetization of the variational Hamiltonian. How can we interpret it as the magnetization of the target Hamiltonian within the mean-field approximation?

Within the mean-field approximation, the variational free energy F_v , not F_0 , with the optimal value of Λ is our approximation to the target free energy. Therefore, for the variational minimum $F_v^* \equiv \min_{\Lambda} F_v(\Lambda)$, if we can prove that $Nm = -\partial F_v^*/\partial H$, then it would establish the interpretation. We can do this as follows:

$$\frac{\partial f^*}{\partial H} = \left(\frac{\partial f^*}{\partial H} \right)_m + \left(\frac{\partial m}{\partial H} \right) \left(\frac{\partial f^*}{\partial m} \right)_H = -m$$

because $(\partial f^*/\partial m)_H = 0$.

Exercise 2.3: Based on the variational principle, propose an improvement on the simplest mean-field approximation of the Ising model discussed in the lecture, and obtain its equation of state. Specifically, propose a variational Hamiltonian that would be closer to \mathcal{H} than \mathcal{H}_0 discussed in the lecture but still solvable.

For example, we may consider the Hamiltonian of non-interacting pairs instead of the Hamiltonian of non-interacting single spins considered in the lecture.

Specifically, $\mathcal{H}_0 = \sum_{(ij)} \mathcal{H}_{ij}$ with $\mathcal{H}_{ij} \equiv -\Lambda_1(S_i + S_j) - \Lambda_2 S_i S_j$ where $\sum_{(ij)}$ is the summation over all mutually non-overlapping nearest-neighbor pairs of spins that covers the whole system. Similarly, we can improve the approximation systematically by considering bigger local clusters: $\mathcal{H}_0 = \sum_c \mathcal{H}_c$ where \sum_c is the summation over all mutually non-overlapping clusters of spins that covers the whole system, and \mathcal{H}_c is some parameterized function of spins in the cluster. (The required computational effort would increase very fast, though, as we increase the cluster size.)

Exercise 2.4: Obtain the Landau expansion of the 3-state Potts model by taking the trial Hamiltonian $\mathcal{H}_0 \equiv -\Lambda \sum_i \delta_{S_i,1}$ with $S_i = 1, 2, 3$. What is the symmetry-breaking order-parameter and what is the essential difference from the Ising case?

By defining $m \equiv \langle \delta_{S_i,1} \rangle - \frac{1}{3}$, we obtain

$$\begin{aligned} \langle \mathcal{H} \rangle_0 &= -\frac{z}{2} N J \left(\frac{1}{3} + \frac{3}{2} m^2 \right) \\ S_0 &= -p_1 \log p_1 - p_2 \log p_2 - p_3 \log p_3 \\ &= -\left(\frac{1}{3} + m \right) \log \left(\frac{1}{3} + m \right) - 2 \left(\frac{1}{3} - \frac{m}{2} \right) \log \left(\frac{1}{3} - \frac{m}{2} \right) \end{aligned}$$

From this we obtain the Landau free energy similar to that of the Ising model discussed in the lecture. However, the important difference lies in the existence of the 3rd order term $-vm^3$ with a positive constant v . This makes the transition the 1st-order one, in contrast to the Ising model.

Exercise 2.5: Consider the $S=1$ Ising model:

$$\mathcal{H} \equiv -J \sum_{\langle ij \rangle} S_i S_j - D \sum_i (S_i)^2. \quad (S_i = -1, 0, 1)$$

Propose a trial Hamiltonian for this model, and, by the variational approximation with it, obtain the $T - D$ phase diagram.

Let the variational Hamiltonian be

$$\mathcal{H}_0 = - \sum_i (\Gamma S_i + \Lambda S_i^2).$$

Or, equivalently, let us denote the probability of having $S_i = 1, 0, -1$ under this trial Hamiltonian as p, r, q respectively. Then,

$$m \equiv \langle S_i \rangle_0 = p - q,$$

$$n \equiv \langle S_i^2 \rangle_0 = p + q.$$

It follows that

$$p = \frac{n + m}{2}, \quad q = \frac{n - m}{2}, \quad r = 1 - n.$$

Obviously, we can regard (m, n) or (p, q) as our variational parameters, instead of (Γ, Λ) . Now, the expectation value of the target Hamiltonian becomes

$$\langle \mathcal{H} \rangle_0 / N = -\frac{z}{2} J m^2 - D n.$$

As for the entropy, using Shannon's entropy $S = - \sum_{\alpha} p_{\alpha} \log p_{\alpha}$,

$$\begin{aligned} S_0 &= -p \log p - q \log q - r \log r \\ &= -\frac{n + m}{2} \log \frac{n + m}{2} - \frac{n - m}{2} \log \frac{n - m}{2} - (1 - n) \log(1 - n) \end{aligned}$$

Therefore, the variational free energy per spin is, in the unit of zJ ,

$$\begin{aligned}\hat{f} &= (\langle \mathcal{H}_0 \rangle_0 - TS_0)/(NzJ) \\ &= -\frac{1}{2}m^2 - gn \\ &\quad + t \times \left(\frac{n+m}{2} \log \frac{n+m}{2} + \frac{n-m}{2} \log \frac{n-m}{2} + (1-n) \log(1-n) \right)\end{aligned}$$

where $g \equiv D/(zJ)$ and $t \equiv T/(zJ)$. The stationary conditions are

$$0 = \frac{\partial \hat{f}}{\partial m} = -m + \frac{t}{2} \log \frac{n+m}{n-m} \quad (13)$$

$$0 = \frac{\partial \hat{f}}{\partial n} = -g + \frac{t}{2} \log \frac{(n+m)(n-m)}{4(1-n)^2} \quad (14)$$

For the second order phase transition, we may expand these conditions with respect to m , which yields

$$\begin{aligned}m &= \frac{t}{2} \log \frac{n+m}{n-m} = \frac{t}{n}m + \mathcal{O}(m^3) \\ \frac{g}{t} &= \log \frac{n}{2(1-n)} + \mathcal{O}(m^2)\end{aligned}$$

From the first equation, we obtain the transition temperature as a function of n ,

$$t_c = n$$

By replacing n by t in the second equation, we obtain the equation that determines the critical temperature t_c as a function of g :

$$g = t \log \frac{t}{2(1-t)}. \quad (15)$$

Solution of this equation is plotted in the figure.

So far, we have been assuming that the transition is of the second order, which means that the order parameter is continuous function of the state parameter, i.e., $\lim_{t \rightarrow t_c+0} m(t) = \lim_{t \rightarrow t_c-0} m(t)$. However, this is not the case with the first order transition. As a result, some part of the phase boundary defined by (15) is taken over by the first order transition and not realized. The first-order transition line is determined by the condition that the $m = 0$ solution and the $m \neq 0$ solution to (13) and (14) have the same free energy.

The detailed analysis is given in papers;
 Capel, Physica 32 966 (1966);
 Blume, Phys. Rev. 141 517 (1966).

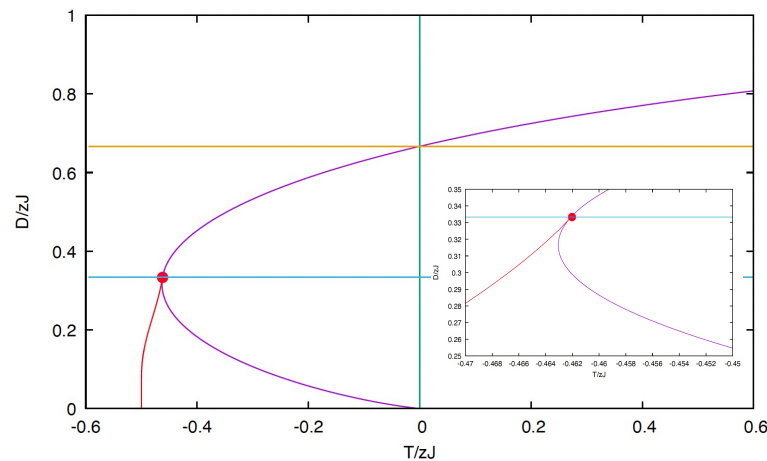


Figure: The transition line is separated by the point located at $(T/zJ, D/zJ) = (\frac{1}{3} \log \frac{1}{4}, \frac{1}{3}) \approx (-0.462028 \dots, 0.333333 \dots)$, which is marked by filled red circle. Above this point, the purple curve represents the critical (the 2nd order transition) line whereas the red curve below is the first-order transition line. The inset is the close-up view.

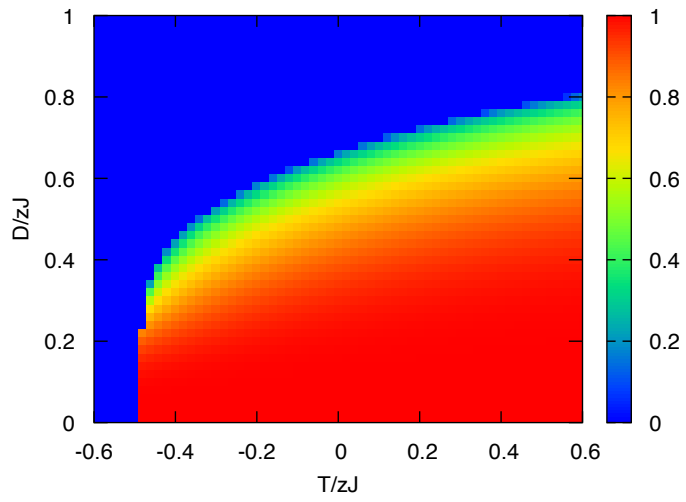


Figure: The magnetization.