Notes for Lecture 18

Monte Carlo simulation, Exact solutions

So far, our discussions on the theory of phase transition have been more or less concentrated on the mean field approximation, while in the last lecture many very general statements have been made. Here, in this lecture, we now survey methods that can give exact results – numerical methods and analytical exact solutions.

18.1 Monte Carlo method

The Monte Carlo method is a primary numerical method of choice in statistical physics and one of the several important numerical methods in many body physics. Here, we will cover **the classical Monte Carlo method**. Here, the word "classical" needs some explanation, since it is used in a loose sense. So far, we have been using this word strictly in the sense that the underlying dynamics is Newtonian. However, here we use "classical" in a somewhat mathematical sense. We will use it not only for problems involving Newtonian dynamics as we have been doing so far, but also for quantum problems that are described fully by one particle operators that *commute* with one another.

In this sense, the Ising problem that we have been considering so far *is* a classical problem, since dynamical variables involved are σ_z matrices for different spins: all of them commute with one another. Thus, the terminology of "classical Ising model" is used despite the fact that the half integer spin is one of the most fundamental quantum phenomena. However, note the following: if one puts the Ising system in a transversal field ($\mathcal{H} = -J\sum_{\langle i,j\rangle}\sigma_{z,i}\sigma_{z,j} - h\mu_B\sum_i \sigma_{x,i}$), then the problem becomes a quantum problem.

It turns out that all quantum problems are reducible to classical problems as far as the equilibrium statistics is concerned[1]. In the so-called "quantum Monte Carlo" method, the quantum problem in d dimensions is mapped to a classical problem in $d+1$ dimensions, where the classical Monte Carlo method is applied. So, the quantum Monte Carlo method is equivalent to the classical Monte Carlo method in this sense. However, not all is well in this mapping: e.g., the infamous "sign problem" plagues the quantum Monte Carlo method, limiting the use of the method to a small system only when the sign problem occurs. This sign problem is a serious issue for problems involving fermions or frustrated spin interactions.

18.1.1 Trajectory

The Monte Carlo method is concerned with calculating the ensemble average of an observable O

$$\overline{\langle O \rangle} = \frac{\mathrm{tr}\,\{O \exp(-\beta\mathcal{H})\}}{Z} \tag{18.1}$$

where the partition function is given by

$$Z = \mathrm{tr}\,\{\exp(-\beta\mathcal{H})\}\,. \tag{18.2}$$

h

In the classical Monte Carlo method, which we will assume from this point on, we can take the basis as the simultaneous eigenstates of the commuting one particle operators, per our definition above.

To be concrete, we shall stick to the (longitudinal) Ising model, for which the Hamiltonian is

$$\mathcal{H} = -J \sum_{\langle i,j \rangle} \sigma_i \sigma_j - H\mu_B \sum_i \sigma_i \tag{18.3}$$

where, as in previous lectures, the subscript z is dropped but implied for all Pauli matrices. With the basis states taken as simultaneous eigenstates for all σ_i *operators*, we can regard σ_i's appearing in $\mathcal{H}$ above, as mere *numbers*, ±1, again as we have been doing in previous lectures.

---

[1] Sondhi, Girvin, Carini, and Shahar, Rev. Mod. Phys. 69, 315 (1997).

Then, the partition function is

$$Z = \sum_{\{\sigma_i\}} e^{-\beta \mathcal{H}(\{\sigma_i\})}. \tag{18.4}$$

Similarly for any observable O, we have

$$\overline{\langle O \rangle} = \frac{1}{Z} \sum_{\{\sigma_i\}} O(\{\sigma_i\}) \exp\Big[-\beta \mathcal{H}(\{\sigma_i\}) \Big]. \tag{18.5}$$

Thus, in a classical Monte Carlo method, we have at our hands a "mere" summation problem. The problem here is that the summation involves exponentially many terms (2^N terms in the current case). For a truly classical problem, we would have an integral over the phase space, whose dimension is $6N$. If each axis is divided into n bins, then we would have a sum over n^{6N} terms. In either case what we have is a sum over a truly enormous number of terms. This is where the random sampling method comes in real handy – so the Monte Carlo method.

Just as a well-conducted opinion poll sampling a tiny segment of the population can produce a result that can be informative with a small error bar, it is possible to do a huge sum like the above one by sampling only a very small subset of all possible configurations $\{\sigma_i\}$. However, the Monte Carlo method would be *very inefficient* if the sampling is done in a completely random manner, i.e. if any configuration is picked with equal chance. This is because in a sum like the one above, we already know, by the rule of large numbers, that the summand is peaked for those configurations that minimize the free energy. We will describe shortly the actual method of sampling those configurations efficiently. For the moment, let us assume that we have an efficient sampling method.

Let us denote the configuration $\{\sigma_i\}$ as ν. So, ν represents a list of N numbers, each 1 or -1. Let us consider a time series of ν, with ν_t representing ν at time values $t = 1, 2, ...$, where the unit of time corresponds to one loop of computation. As we assumed above, the time series is governed by our judicious sampling algorithm. For a quantity Q that is a function of ν, the average over a trajectory of T steps is given by

$$\langle Q \rangle_T = \frac{1}{T} \sum_{t=1}^{T} Q(\nu_t). \tag{18.6}$$

The ensemble average is obtained by

$$\overline{\langle Q \rangle} = \lim_{T \to \infty} \langle Q \rangle_T \,, \tag{18.7}$$

by assuming that the trajectory eventually explores all configurations relevant for the equilibrium state. It goes without saying that in reality the trajectory is terminated at a finite value of T and it is thus important to estimate the error. It would be guaranteed that, *if* subsequent steps were independent or weakly correlated, then the error would scale as $1/\sqrt{T}$, by the central limit theorem. However, subsequent steps are in fact highly correlated, and our analysis below will show that the error diminishes more slowly than this naive expectation.

18.1.2 Importance sampling

In general, the sampling of important configurations according to the probability distribution of the random variable is termed "**importance sampling**." The concept is simple. If a random variable x obeys the probability distribution $p(x)$ then it is not advised to sample x uniformly. Rather, it is much more efficient to sample y uniformly (on average), where $dy = p(x)dx$: thus, the sampling density ($\propto 1/dx$) in the original x space must be proportional to $p(x)$.

More specifically, let us consider the strictly semi-classical statistical mechanics case, where by x we mean all $6N$ phase space variables. The ensemble average of any physical quantity Q is given by

$$\langle Q \rangle = \frac{\int dx\, p(x) Q(x)}{\int dx\, p(x)} \tag{18.8}$$

where $p(x)$ is the Boltzmann factor $\exp(-\beta E(x))$ and $\int dx\, p(x)$ is the partition function up to a multiplicative factor. The *importance sampling* as described above means that we can approximate the integral as

$$\langle Q \rangle \approx \frac{\Delta y \sum_{t=1}^{T} Q(x_t)}{\Delta y \sum_{t=1}^{T}} = \frac{\sum_{t=1}^{T} Q(x_t)}{T} \tag{18.9}$$

where x_t's are chosen so that $dy = p(x)dx$ is uniformly spaced on average and Δy is the total volume in the y space divided by T. In contrast, a naive approach would give $\langle Q \rangle \approx \frac{\sum_{t=1}^{T} p(x_t) Q(x_t)}{\sum_{t=1}^{T} p(x_t)}$ where x_t's are sampled uniformly in the x space: this would be quite wasteful since most $p(x_t)$ values will be quite small!

Our current example of the Ising model can be treated just the same, if we replace x by ν, the integral $\int dx$ by sum $\sum_\nu$, and $\langle Q \rangle$ by $\overline{\langle Q \rangle}$ (recall that we have been using this somewhat elaborate symbol for the quantum ensemble average to avoid confusion with the quantum expectation value).

So, our goal is now clear. We must sample in the ν space in such a manner that the density of points sampled is proportional to the Boltzmann factor $\exp(-\beta E(\nu))$. Namely, we must consider the Boltzmann weight *as we select* points in ν space, *not after* we have randomly selected them.

It is of importance to note that this procedure that we are employing is exactly the kind that ensures that the free energy is minimized (page 4 of LN 7). This can be seen as follows. No matter which scheme of sampling we use, it is clear that the energy value that will contribute greatly to the sum is the energy value for which there is a large degeneracy (i.e., many ν points with the same energy). In other words, energy value with large entropy will contribute to the sum more importantly than the sum with low entropy. Both the naive sampling and the importance sampling already take this into account for the trivial reason that the large entropy means a better chance to be sampled. By importance sampling, we are *additionally* requiring that the low energy states be treated with preference. So, overall, we are requiring that the entropy be high and the energy be low. These two balancing requirements determine the minimum free energy, which is how the equilibrium is brought about when the system is modeled as a canonical ensemble, i.e. when the system in contact with a constant temperature reservoir.

Note that in this discussion, we did not really specify what "free energy" is. It could be the Helmholtz free energy, if the reservoir is only a heat reservoir. It could be the Gibbs free energy, if the reservoir is both a heat reservoir and a work reservoir. In any case, we consider the Hamiltonian to represent[2] $\mathcal{E} - J_R \chi$, not only the intrinsic energy of the system, $\mathcal{E}$, but also the coupling to a force J_R provided by the reservoir. You can note that, indeed, this is what happens, by convention, quite frequently (Sections 8.4.2, and 16.1.2; Homework 3.1, etc.).

18.1.3 Metropolis sampling

So, now we know what we want: we must implement a mechanism in the evolution of the trajectory in such a way that the frequency in which a certain point ν appears is proportional to its Boltzmann factor.

There is much flexibility in accomplishing this goal: there is no unique solution to this goal. One of the most widely used sampling methods is the Metropolis sampling method, named after its main progenitor.

With the Metropolis sampling (steps 4 and 5), here is the full Monte Carlo algorithm.

1. Say we have a configuration at step t, ν_t. Compute its energy E.

---

[2] Here, we use a slightly different symbol χ for the random variable version of the displacement x. Note that in the case of chemical work $x = N$. Indeed, often one finds that $-\mu N$ is included in a many body Hamiltonian!

2. Choose one spin randomly (or sequentially, or some other scheme[3]) and flip it.
3. For the new configuration ν', compute E'.
4. If $E' \le E$, then $\nu_{t+1} = \nu'$.
5. If $E' > E$, then $\nu_{t+1} = \nu'$ or ν. Pick a random number from 0 to 1. If that random number is less than $\exp(-\beta(E'-E))$, then accept ν', otherwise keep ν.

Let us analyze why this sampling method is consistent with the importance sampling requirement for the equilibrium distribution. In general the rate at which the probability for a configuration ν changes is given by the master equation

$$\dot{p}_\nu = \sum_{\nu'} \left[-p_\nu w_{\nu,\nu'} + p_{\nu'} w_{\nu',\nu} \right] \tag{18.10}$$

where $w_{\nu,\nu'}$ is the transition rate from configuration ν to ν'. An equilibrium condition would be met ($\dot{p}_\nu = 0$) if **the detailed balance** condition is satisfied:

$$-p_\nu w_{\nu,\nu'} + p_{\nu'} w_{\nu',\nu} = 0. \tag{18.11}$$

Since $p_{\nu'}/p_\nu = \exp(-\beta(E'-E))$ in equilibrium, the detailed balance condition can be written as

$$\frac{w_{\nu,\nu'}}{w_{\nu',\nu}} = \exp(-\beta(E'-E)). \tag{18.12}$$

Note the meaning of this. It means not only that this equation is valid in equilibrium, but also that $w_{\nu,\nu'}$ satisfying this detailed balance condition will be able to cause the equilibrium to be established through the master equation.

Now, we can see that the above Metropolis sampling ensures the detailed balance condition. Indeed, the Metropolis sampling method is equivalent to

$$w_{\nu,\nu'} = 1 \qquad \text{if } E' \le E, \tag{18.13}$$

$$w_{\nu,\nu'} = \exp(-\beta(E'-E)). \qquad \text{if } E' \ge E \tag{18.14}$$

By simply swapping unprimed and primed symbols, we also get

$$w_{\nu',\nu} = 1 \qquad \text{if } E' \ge E, \tag{18.15}$$

$$w_{\nu',\nu} = \exp(-\beta(E-E')). \qquad \text{if } E \le E' \tag{18.16}$$

Therefore, we get

$$\frac{w_{\nu,\nu'}}{w_{\nu',\nu}} = \exp(-\beta(E'-E)) \tag{18.17}$$

when $E' \le E$ as well as $E' \ge E$, proving the consistency of the Metropolis sampling with the Boltzmann factor.

---

[3] Here there is room for flexibility as well. Any method of updating configuration can be used, as long as it achieves the goal of accessing the important configurations of the equilibrium quickly.

18.1.4 Error analysis

Work employing a numerical method such as the Monte Carlos method is kind of an experiment. Therefore, it is extremely important to know the error bar of the result.

Error and auto-correlation

Consider $\langle Q\rangle_T$ as considered in Eq. 18.6. For the notational convenience within this section, we call it Q_m, the mean value of Q from the Monte Carlo experiment,

$$Q_m \equiv \langle Q\rangle_T = \frac{1}{T}\sum_{t=1}^{T} Q_t. \qquad\qquad Q_t \equiv Q(\nu_t) \tag{18.18}$$

We like to know what error bar we must put for Q_m. Each Q_t is taken from a microstate of the ensemble, and so it has its own standard deviation, which we denote as σ_Q.

$$\sigma_Q^2 \equiv \overline{\langle Q_t^2\rangle} - \overline{\langle Q_t\rangle}^2 = \overline{\left\langle \left(Q_t - \overline{\langle Q_t\rangle}\right)^2\right\rangle} \tag{18.19}$$

where, by assuming time invariance (no explicit time dependence of $\mathcal{H}$), we write σ_Q instead of σ_{Q_t}. How do we get this value? The answer is from the Monte Carlo experiment itself! All you need to do is to measure Q^2 in addition to Q. Now an important point. *If*, within our Monte Carlo experiment, successive configurations (ν_t's) are uncorrelated, then we can take Q_t's in Eq. 18.18 as independent, and then we would get $\sigma_{Q_m} = \sigma_Q/\sqrt{T}$ by the central limit theorem. However, if one ponders about the Monte Carlo algorithm as written down in the last section, one can be sure that each neighboring ν_t's are *strongly correlated* – they differ by only one spin value! Therefore, we have some more work to do, as follows.

$$\begin{aligned}
\sigma_{Q_m}^2 &= \frac{1}{T^2}\sum_{t,u=1}^{T}\left\{\overline{\langle Q_t Q_u\rangle} - \overline{\langle Q_t\rangle}\,\overline{\langle Q_u\rangle}\right\} && \text{from Eq. 18.18} \quad (18.20)\\
&= \frac{\sigma_Q^2}{T} + \frac{2}{T^2}\sum_{u>t}\left\{\overline{\langle Q_t Q_u\rangle} - \overline{\langle Q_t\rangle}\,\overline{\langle Q_u\rangle}\right\} && \text{split } u=t \text{ and } u\neq t\text{, use Eq. 18.19} \quad (18.21)\\
&= \frac{\sigma_Q^2}{T} + \frac{2}{T^2}\sum_{t=1}^{T}\sum_{v=1}^{T-t}\left\{\overline{\langle Q_t Q_{t+v}\rangle} - \overline{\langle Q_t\rangle}\,\overline{\langle Q_{t+v}\rangle}\right\} && (18.22)\\
&= \frac{\sigma_Q^2}{T} + \frac{2}{T^2}\sum_{t=1}^{T}\sum_{v=1}^{T-t}\left\{\overline{\langle Q_1 Q_{1+v}\rangle} - \overline{\langle Q_1\rangle}^2\right\} && \text{time invariance} \quad (18.23)\\
&= \frac{\sigma_Q^2}{T} + \frac{2}{T^2}\sum_{v=1}^{T-1}\left\{\overline{\langle Q_1 Q_{1+v}\rangle} - \overline{\langle Q_1\rangle}^2\right\}(T-v). && \text{count same terms} \quad (18.24)
\end{aligned}$$

This result can be summarized as

$$\boxed{\sigma_{Q_m} = \frac{\sigma_Q}{\sqrt{T}}\sqrt{2\tau_{Q,int}} \geq \frac{\sigma_Q}{\sqrt{T}}. \qquad \text{error bar for } Q_m \qquad (18.25)}$$

where

$$\tau_{Q,int} \equiv \frac{1}{2} + \sum_{t=1}^{T-1} A(t)\left(1 - \frac{t}{T}\right), \qquad \text{integrated auto-correlation time} \qquad (18.26)$$

$$A(t) \equiv \frac{\overline{\langle Q_1 Q_{1+t}\rangle} - \overline{\langle Q_1\rangle}\;\overline{\langle Q_{1+t}\rangle}}{\sigma_Q^2}. \qquad \text{(normalized) auto-correlation} \qquad (18.27)$$

Asymptotically, $A(t)$ tends to behave as

$$A(t) \propto \exp\left(-\frac{t}{\tau_{Q,exp}}\right) \qquad (18.28)$$

where $\tau_{Q,exp}$ and $\tau_{Q,int}$ are expected to be of the same order (as $\sum A(t) \sim \int dt A(t) \sim \tau_{Q,exp}$), but not exactly the same. Assuming that $T \gg \tau_{Q,\mathrm{exp}} \gg 1/2$, then, we can ignore the factor $1 - t/T$ in the above expression for $\tau_{Q,int}$

$$\tau_{Q,int} \approx \frac{1}{2} + \sum_{t=1}^{T-1} A(t) \approx \sum_{t=1}^{T-1} A(t). \qquad (18.29)$$

Note that for the system settle down to equilibrium from a non-equilibrium state we need to wait at least $\tau_{Q,exp}$ or $\tau_{Q,int}$. We expect that the equilibration time scale is on the order of these time scales.

Near the critical point, these time scales diverge, due to a **critical slowdown**.

$$\tau_{Q,int} \propto \xi^z \qquad T \to T_c \qquad (18.30)$$

where ξ is, as introduced in the last lecture, the correlation length of the order parameter. The reason for the critical slowdown, where **the relaxation time** diverges as the critical point is reached, is the flat free energy at T_c (graph like in page 10 of LN 16 becomes very flat at $m = 0$ as $T \to T_c$). In the case of a finite system, as in a simulation,

$$\tau_{Q,int} \propto L^z \qquad T \to T_c \qquad (18.31)$$

where L is the linear dimension of the system. This is bad news for numerical simulation. As the critical region is reached, more and more time is necessary to get correct information. For a local update method such as the one outlined in the previous section, the exponent $z \approx 2$. However, some clever update methods such as "cluster algorithm" (Swendsen, Wang, Niedermayer, Wolf) have been invented to reduce the exponent z to 0.25. Here, "update method" refers to step 2 of the algorithm presented in Section 18.1.3.

Binning analysis

Consider T points in the trajectory binned to N_B bins with k points in each bin. One can calculate the mean value for the n-th bin

$$Q_{B,n} = \frac{1}{k}\sum_{t=1}^{k} Q_{(n-1)k+t}. \qquad n = 1, 2, ..., N_B \tag{18.32}$$

The idea is that, assuming $k \gg \tau_{Q,int}$, each bin is uncorrelated. Then,

$$\sigma^2_{Q_m} = \frac{\sigma^2_B}{N_B}. \tag{18.33}$$

By Eq. 18.25, this must be equal to $2\sigma^2_Q \tau_{Q,int}/T$. Therefore, we can express $\tau_{Q,int}$ as follows.

$$\boxed{\tau_{Q,int} = \frac{k}{2}\frac{\sigma^2_B}{\sigma^2_Q}.} \tag{18.34}$$

In practice, this is how $\tau_{Q,int}$ is found: $\frac{k}{2}\frac{\sigma^2_B}{\sigma^2_Q}$ is plotted as a function of k, and it will be seen to converge to a value as k becomes large enough – that value gives $\tau_{Q,int}$. Of course, the self-consistency must be checked: $k \gg \tau_{Q,int}$ on convergence.

18.2 Exact solutions for the Ising model

It is a pleasure for us to have some analytically exact solutions in interacting problems. This is an exception rather than a rule, by far, but it does not diminish the impact of such an exact solution to shaping and guiding our thinking.

18.2.1 Ising model in one dimension

Often times, models in one dimension (1D) are easily solvable due to the kinematic simplicity. The Ising model in 1D is an elementary case in point. In the following treatment, we do not need to assume what the sign of J is. That is, the following treatment applies both to the ferromagnetic case and the anti-ferromagnetic case.

Given the Hamiltonians of Eq. 15.12 and 15.13, it is rather simple to solve them. Let us first attack Eq. 15.12 and write its form in 1D,

$$\mathcal{H} = -J \sum_{i=1}^{N-1} \sigma_i \sigma_{i+1}. \tag{18.35}$$

Here, we choose to work with an open boundary condition[4]. The virtue of the open boundary condition becomes obvious when one considers the Gibbs partition function

$$X_N = \sum_{\sigma_1=\pm1\ \dots\ \sigma_N=\pm1} \exp\left[\beta J \sum_{i=1}^{N-1} \sigma_i \sigma_{i+1}\right] \tag{18.36}$$

Consider the sum

$$\sum_{\sigma_N=\pm1} \exp\left[\beta J \sigma_{N-1} \sigma_N\right] = 2\cosh(\beta J) \tag{18.37}$$

which multiplies every term in the sum for X_N. This result is independent of the value of $\sigma_{N-1} = \pm1$. Thus, we have a recursion relation

$$X_N = X_{N-1} \cdot 2\cosh(\beta J), \tag{18.38}$$

$$X_1 = 1. \tag{18.39}$$

Thus, we get

$$X_N = \left[2\cosh(\beta J)\right]^{N-1}. \tag{18.40}$$

Therefore, in the thermodynamic limit, the Gibbs free energy is given by

$$G = -N k_B T \log\left[2\cosh(\beta J)\right]. \tag{18.41}$$

Note that this Gibbs free energy is completely analytic. So, for example, we do not expect the heat capacity to show any anomaly. On the other hand, note that the mean field solution (Section 16.1.4) shows a discontinuity in the heat capacity. So, we can infer from the above result that a phase transition never occurs at any finite temperature for a one dimensional Ising model.

---

[4] The physics of the bulk must be independent of the boundary condition that we employ.

Now, let us solve this problem with a field, to make this point even clearer.

$$\mathcal{H} = -J\sum_{i=1}^{N}\sigma_i\sigma_{i+1} - H\mu_B\sum_{i=1}^{N}\sigma_i. \tag{18.42}$$

Here, we use a different boundary condition, the periodic boundary condition. The second term can be symmetrized by shifting the index by 1 and averaging over the two identical sums

$$\mathcal{H} = -J\sum_{i=1}^{N}\sigma_i\sigma_{i+1} - \frac{H}{2}\mu_B\sum_{i=1}^{N}\left(\sigma_i + \sigma_{i+1}\right). \tag{18.43}$$

The benefit of this symmetrization will become apparent shortly. Now, the partition function is given by

$$X = \sum_{\{\sigma_i=\pm1\}}\prod_{i=1}^{N}\exp\left[\beta J\sigma_i\sigma_{i+1} + \beta\frac{H}{2}\mu_B\left(\sigma_i + \sigma_{i+1}\right)\right]. \tag{18.44}$$

This motivates a **transfer matrix**, which is defined as

$$\overleftrightarrow{P}_{\sigma_i,\sigma_{i+1}} \equiv \begin{bmatrix} P_{1,1} & P_{1,-1} \\ P_{-1,1} & P_{-1,-1}\end{bmatrix} \tag{18.45}$$

where $P_{1,1} = \exp(\beta J + \beta H\mu_B)$, $P_{-1,-1} = \exp(\beta J - \beta H\mu_B)$, and $P_{1,-1} = P_{-1,1} = \exp(-\beta J)$. Now the virtue of the above symmetrization on the H part is clear – the transfer matrix is a real symmetric matrix. In terms of the transfer matrix, the partition function is now given by

$$X = \mathrm{tr}\left\{\overleftrightarrow{P}^N\right\}. \tag{18.46}$$

The program to follow from here is straightforward. Diagonalize $\overleftrightarrow{P}$, and thus diagonalize $\overleftrightarrow{P}^N$. The trace of the diagonalized $\overleftrightarrow{P}^N$ gives the answer, since the trace is a scalar[5]. By diagonalizing $\overleftrightarrow{P}$, we get two eigenvalues

$$\lambda_{1,2} = e^{\beta J}\cosh(\beta\mu_B H) \pm \sqrt{e^{2\beta J}\sinh^2(\beta\mu_B H) + e^{-2\beta J}}. \tag{18.47}$$

Note that λ_1 and λ_2 are both positive, if J is positive[6]. In any case, $|\lambda_2| < \lambda_1$. Therefore,

$$X = \lambda_1^N + \lambda_2^N = \lambda_1^N\left(1 + \left(\frac{\lambda_2}{\lambda_1}\right)^N\right) \approx \lambda_1^N, \tag{18.48}$$

---

[5] Of course, physically, X *must* be a scalar.

[6] $\overleftrightarrow{P}$ is positive definite only when $J > 0$.

where the last step comes from the thermodynamic limit. Thus, we get

$$X = \left[e^{\beta J}\cosh(\beta\mu_B H) + \sqrt{e^{2\beta J}\sinh^2(\beta\mu_B H) + e^{-2\beta J}}\right]^N, \tag{18.49}$$

$$G = -Nk_BT\log\left[e^{\beta J}\cosh(\beta\mu_B H) + \sqrt{e^{2\beta J}\sinh^2(\beta\mu_B H) + e^{-2\beta J}}\right]. \tag{18.50}$$

It can be checked that this free energy becomes Eq. 18.41 if $H = 0$. This is left for your exercise. The magnetization, the order parameter for a ferromagnetic Ising model, is found as

$$m = -\frac{1}{\mu_B N}\left(\frac{\partial G}{\partial H}\right)_T = \frac{\sinh(\beta\mu_B H)}{\sqrt{\sinh^2(\beta\mu_B H) + e^{-4\beta J}}}. \tag{18.51}$$

Thus, we see that, as expected from the discussion after Eq. 18.41, $m(H = 0) = 0$ at any finite temperature. This is in stark contrast to the mean field theory result that predicts a finite m for T below $T_c = Jz/k_B$ no matter what the spatial dimension is. The reason is that in one dimension thermal fluctuations completely dominate in determining the equilibrium state at any finite temperature. (The following argument had been given already in page 17 of LN 17.) Consider the fact that, without field, the Gibbs free energy is given by $G = E - TS$. So, the minimization of G requires balancing the entropy maximization (the second term) and the energy minimization (the first term). If one considers a domain formation for a one dimensional crystal, it becomes obvious that at the cost of a slight energy increase, a large entropy can be gained. For instance, if one considers all the ways that a crystal can be divided into two equal parts with opposite spin alignments, the entropy is $k_B\log\frac{N}{2}$ while the energy cost to create such a domain structure is merely $4J$ (which is negligible compared to $\log N$ in the thermodynamic limit). Thus, at a small but finite value of T, it becomes completely advantageous to maximize the entropy, which is why there cannot be any finite temperature phase transition for the Ising model.

18.2.2 Ising model in two dimensions

The *tour de force* solution by Onsager for the Ising model in two dimensions is a celebrated example. It is beyond the scope of this course to cover it, however. It is covered in books such as Huang.

Some main results are worth mentioning. The partition function is given by

$$X(T, H = 0) = \left[2\cosh(\beta J)e^{I(T)}\right]^N, \tag{18.52}$$

$$I = \frac{1}{2\pi}\int_0^\pi d\phi\log\left\{\frac{1}{2}\left[1 + (1 - \kappa^2\sin^2\phi)^{1/2}\right]\right\}, \tag{18.53}$$

$$\kappa = 2\sinh(2\beta J)/\cosh^2(2\beta J). \tag{18.54}$$

This expression is very complicated compared to Eq. 18.41, and importantly is non-analytic at $\kappa = 1$, defining the critical temperature, which is given by

$$\begin{aligned} \sinh \frac{2J}{k_B T_c} &= 1 && (18.55) \\ T_c &= \frac{2J}{k_B} \frac{1}{\log\left(1+\sqrt{2}\right)} \approx \frac{2.269J}{k_B} && (18.56) \end{aligned}$$

The critical exponents are given by

$$\alpha = 0 \left(C \propto \log \frac{1}{|T - T_c|}\right),\ \beta = \frac{1}{8},\ \gamma = \frac{7}{4},\ \delta = 15,\ \eta = \frac{1}{4},\ \nu = 1 \tag{18.57}$$

where they are defined by

$$C \propto |t|^{-\alpha},\ m \propto |t|^{\beta},\ \chi \propto |t|^{-\gamma},\ m \propto H^{1/\delta} \tag{18.58}$$

and ν and η as defined in Eqs. 17.57 and 17.58.