Notes for Lecture 18

Monte Carlo, Renormalization group

We have covered the mean field theory and some exact solutions as part of our efforts to try to understand the phenomena of phase transition. Here, we introduce briefly the Monte Carlo method that can be used to study the phase transition numerically. Then, we proceed to the renormalization group method, arguably one of the most important conceptual progresses made in physics in the second half of the last century. It should also be noted that the application of these two important methods goes well beyond the theory of phase transition.

18.1 Monte Carlo method

The Monte Carlo method is a primary numerical method of choice in statistical physics and one of the few 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 *commuting* one particle operators.

In this sense, the Ising problem that we have been considering so far is a classical problem, since the variables involved are only 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 ($H = -J\sum_{\langle i,j\rangle}\sigma_{z,i}\sigma_{z,j} - h\mu_B\sum_i \sigma_{x,i}$), 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] So, 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 "fermion sign problem" plagues the quantum Monte Carlo method, limiting the use of the method to a small system only for fermion problems or frustrated spin problems.

18.1.1 Trajectory

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

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

where the partition function is given by

$$Z = \mathrm{tr}\left\{\exp(-\beta H)\right\} \tag{18.2}$$

In 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

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

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 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

$$Z = \sum_{\{\sigma_i\}} \exp\left[-\beta H(\{\sigma_i\})\right] \tag{18.4}$$

Similarly for any observable O, we have

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

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 maximize the free energy (Sections 7.2.1 and 7.2.2). 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 ν as the configuration $\{\sigma_i\}$. ν is a list of N numbers, each 1 or -1. Let us consider a time series of ν, with ν_t is ν at time value $t = 1, 2, ...$, where the unit of time corresponds to one loop in computation. The time series represents our judicious sampling. 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 is expected that, *if* subsequent steps are independent or weakly correlated, then the error scales as $1/\sqrt{T}$, from 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 $dy = p(x)dx$. This simply means that the sampling density ($\propto 1/dx$) in the original x space must be proportional to $p(x)$.

More specifically, let us consider the purely classical 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 (Eqs. 7.17,7.18,7.19). 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. This is the essential physics of Eqs. 7.17,7.18,7.19, and you cannot forget it!

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, as there is no unique solution to this. 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. Choose one spin randomly (or sequentially, or some other scheme[2]) and flip it.
3. For the new configuration ν', compute E'.
4. If $E' \le E$, then $\nu_{t+1} = \nu'$.

[2] 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 and quickly.

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 $p_{\nu'}/p_\nu = \exp(-\beta(E' - E))$ at 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 at 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.

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,

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

We like to know what error bar we must put for Q_m. Each Q_t is part 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 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 v_t's are *strongly correlated* – they differ by only one spin value! Therefore, we have some more work to do, as follows.

$$\sigma_{Q_m}^2 = \frac{1}{T^2}\sum_{t,u=1}^{T}\left\{\overline{\langle Q_t Q_u\rangle} - \overline{\langle Q_t\rangle}\,\overline{\langle Q_u\rangle}\right\} \qquad \text{from Eq. 18.18} \tag{18.20}$$

$$= \frac{1}{T}\sigma_Q^2 + \frac{2}{T^2}\sum_{u>t}\left\{\overline{\langle Q_t Q_u\rangle} - \overline{\langle Q_t\rangle}\,\overline{\langle Q_u\rangle}\right\} \qquad \text{split } u=t \text{ and } u\neq t\text{, use Eq. 18.19} \tag{18.21}$$

$$= \frac{\sigma_Q^2}{T} + \frac{2}{T^2}\sum_{t=1}^{T}\sum_{v=1}^{T-t}\left\{\overline{\langle Q_t Q_{t+v}\rangle} - \overline{\langle Q_t\rangle}\,\overline{\langle Q_{t+v}\rangle}\right\} \tag{18.22}$$

$$= \frac{\sigma_Q^2}{T} + \frac{2}{T^2}\sum_{t=1}^{T}\sum_{v=1}^{T-t}\left\{\overline{\langle Q_1 Q_{1+v}\rangle} - \overline{\langle Q_1\rangle}^2\right\} \qquad \text{time invariance} \tag{18.23}$$

$$= \frac{\sigma_Q^2}{T} + \frac{2}{T^2}\sum_{v=1}^{T-1}\left\{\overline{\langle Q_1 Q_{1+v}\rangle} - \overline{\langle Q_1\rangle}^2\right\}(T-v) \qquad \text{count same terms} \tag{18.24}$$

This result can be summarized as

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

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

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

Asymptotically, $A(t)$ behaves as

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

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

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

Note that for the system settle down to equilibrium from a non-equilibrium state we need to wait at least $\tau_{Q,int}$ or $\tau_{Q,exp}$. We might 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.*

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

where ξ is, as introduced in the last lecture (Section 17.2), 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 8 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,

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

where L is the linear dimension of the system. So, this is something of bad news for numerical simulation. As the critical region is reached, more and more time need be spent 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

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

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

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

In practice, $\tau_{Q,int}$ can be plotted as a function of k, and it will be seen to converge to a value as k becomes large enough.

18.2 Renormalization group

The "group" in the expression renormalization group refers to the mathematical group. And, it is a bit of a misnomer, mathematically speaking. However, the motivation is clear enough. The group theory forms a very important backbone of physics. This is because the group theory is the mathematical language of the symmetry and conservation principles of physics. Here, we are also talking about a sort of symmetry, indeed. We have already noted that a critical state is a scale-invariant state (Section 17.2). It is the "scaling operation" that is embodied as a "symmetry operation" in the renormalization group theory. In reality, this scaling operation is a one-way operation – a "zooming-out" operation or a "coarse-graining" operation. Zooming out operations, without their inverses, cannot form a mathematical group. Thus, a bit of a misnomer. Also, the renormalization group operations are not unique. This is different from other more clearcut symmetry operations such as translation, reflection, rotation, charge conjugation etc. These are well-defined operations. However, the renormalization group operations are, as we shall see, somewhat loosely defined. For a given system of spins, researcher A might define "zooming out" as deleting every other spin and computing the effective Hamiltonian for the rest of the spins, while researcher B might work in k-space and integrate out high k components of the Hamiltonian to define the effective Hamiltonian. In either case, it is important to note that the core process of integrating out short range physics is common. Thus, the renormalization group seeks to examine the behavior at large length scales, or, equivalently, at small energy scales. As long as this process of integrating out is correctly done, the correct physics will be captured, and the details of the actual "symmetry operation" can be defined somewhat idiosyncratically.

The "renormalization" in the expression renormalization group refers to the fact that the parameters of a given physical model change from their initial values as the renormalization group operation is applied. The way the parameters get renormalized tell us what the essential character of the low energy physics is. In particular, at a critical point, we expect that these parameters not be renormalized, as the system is scale-invariant. Thus, we expect to find a fixed point for a critical point. However, the converse is not true. Just because we have a fixed point does not mean that we have a critical point. It can be a trivial fixed point, as we shall see shortly.

Many of these points can be demonstrated by a simple model: the one dimensional Ising model.

18.2.1 1D Ising model

Here, we will use a very simple renormalization group theory on a one dimensional Ising model.

The partition function is given by

$$Z(K,N) = \sum_{\{\sigma_i\}} \exp\left[K(\sigma_1\sigma_2 + \sigma_2\sigma_3 + ... + \sigma_N\sigma_1)\right] \qquad K \equiv J\beta \tag{18.35}$$

where the periodic boundary condition is used for N spins. We can coarse-grain simply by integrating out every other spin, e.g. spin at even indices. The result is

$$\begin{aligned} Z(K,N) = \sum_{\{\sigma_1,\sigma_3,...\}} & \{\exp[K(\sigma_1+\sigma_3)] + \exp[-K(\sigma_1+\sigma_3)]\} \\ & \{\exp[K(\sigma_3+\sigma_5)] + \exp[-K(\sigma_3+\sigma_5)]\} \; ... \end{aligned} \tag{18.36}$$

Now, this looks quite different from the original partition function, at first sight. However, we are looking for some possible scale invariant form. So, we might hope to put this new partition function in a form that looks like $\sum_{\{\sigma_1,\sigma_3,...\}} \exp\left[K'(\sigma_1\sigma_3 + \sigma_3\sigma_5 + ...)\right]$. It is possible to do something like it if we require that

$$\exp[K(\sigma_i+\sigma_{i+2})] + \exp[-K(\sigma_i+\sigma_{i+2})] = f(K)\exp[K'\sigma_i\sigma_{i+2}] \tag{18.37}$$

where f and K' are two quantities to be determined. That these two quantities can be determined follows from the fact that our requirements consist of just two equations

$$2\cosh(2K) = f(K)\exp(K') \qquad \text{if } \sigma_i\sigma_{i+2} > 0 \tag{18.38}$$

$$2 = f(K)\exp(-K') \qquad \text{if } \sigma_i\sigma_{i+2} < 0 \tag{18.39}$$

These have solutions

$$\exp(2K') = \cosh(2K) \tag{18.40}$$

$$f(K) = 2\sqrt{\cosh(2K)} \tag{18.41}$$

And the partition function can be written as

$$Z(K,N) = 2^{N/2}\cosh(2K)^{N/4}\, Z(K', N/2) \tag{18.42}$$

These last three equations constitute the "renormalization group (RG) equations." If we take *log* of the last equation and define

$$\log Z(K,N) \equiv N\, g(K) \qquad -k_BT\, g \text{ is free energy per spin} \tag{18.43}$$

we get

$$g(K') = 2g(K) - \log\left(2\sqrt{\cosh(2K)}\right) \quad (18.44)$$

Note that the first RG equation can be rewritten as

$$\exp(2K') + \exp(2K') = \exp(2K) + \exp(-2K)$$

We will focus on the ferromagnetic case $K \geq 0$, for the moment, commenting on the anti-ferromagnetic case at the end. Note that $K' < K$ if $K > 0$, and $K' = K$ if $K = 0$. Thus, $K = 0$ is a fixed point. Another fixed point is $K = \infty$. K' is renormalized with respect to K by application of the RG equation. In order to examine the low energy behavior of a given system, the above RG equations must be applied repeatedly, after putting $K = K'$ from one step. Thus, in the RG equation, K is a *running coupling constant*. As the RG equation is repeatedly applied from a finite value of K, K goes to 0. In the RG language, we say that the RG *flows* to the $K = 0$ fixed point.

What is the nature of the $K = 0$ fixed point? It is where the effect of the interaction is zero, either due to J being zero or T being very high. It is the disordered phase. While such a state is also scale invariant, it is not scale invariant for an interesting reason. It is a trivial fixed point.

The other fixed point is $K = \infty$, which represents the $T = 0$ phase. This is a completely ordered phase. At $T = 0$, the system is fully magnetized. It is clear why such a state is scale invariant. However, this scale invariance is not as interesting as the scale invariance at a finite critical temperature – the scale invariance without any order. Nevertheless, note that, if K is near this $K = \infty$ fixed point, the correlation length becomes very large.[3]

The RG flow shows what happens when the system is viewed in longer and longer length scales. Without a long range order, the system eventually looks disordered at sufficiently large length scale. The fact that in the current example the RG flows to the $K = 0$ fixed point, starting from any finite value of K, means that there is no long range order at any finite temperature. This is consistent with the exact solution (Section 17.3.1).

[3] $\xi = [\log\coth K]^{-1}$. So, the correlation length diverges exponentially, not as a power law, as $T \to 0$. The derivation of the formula for ξ is left for your work. Here are some guidelines. First, the correlation function (Eq. 17.8) in the disordered phase is given by $\Gamma_j = \overline{\langle\sigma_1\sigma_j\rangle} = \overline{\langle\sigma_1\sigma_2\sigma_2\sigma_3...\sigma_{j-1}\sigma_j\rangle}$, since $\sigma_i^2 = 1$. Then, write $-\beta H = \sum K_i\sigma_i\sigma_{i+1}$, where K_i's are taken to be independent variables that will be put to K at the end of the calculation. The partition function for this Hamiltonian can be obtained as easily as that for uniform K (steps leading to Eq. 17.24). And, $\Gamma_j = \frac{1}{Z}\frac{\partial^{j-1}Z}{\partial K_1\partial K_2...\partial K_{j-1}}$.

What would have happened if there *is* a finite temperature critical point, T_c? We would have found that the corresponding *finite* value of $K_c = J/(k_B T_c)$ is a fixed point. We will see later that this happens in higher dimensions.

Finally, what if $J < 0$? The RG transformation set up above is not appropriate to study the anti-ferromagnetism. One may note that, starting from a negative value of K, the first application of the RG equation above immediately gives a positive K, turning the model into a ferromagnetic one. This does *not* mean that an anti-ferromagnetic Ising model is mapped to a ferromagnetic Ising model. It means that the RG transformation set up above is not good enough to study both the ferromagnetic Ising model and the anti-ferromagnetic Ising model. This is clear from the fact that $K = -\infty$ is *not* a fixed point if we use the above RG equation.