Notes for Lecture 17

Correlation and ξ

We have considered the ferromagnetic Ising model from the mean field point of view and arrived at the Landau theory of phase transition. Here, we shall discuss an essential characteristic of the problem, the correlation and the correlation length (ξ), and look towards the renormalization group theory. Note that in this lecture note, our theory is general, definitely going beyond the mean field theory that we used in the previous lecture, while examples that we use will be limited, for now, to the mean field theory examples.

17.1 χ, the fluctuation-dissipation theorem

Within the mean field theory of the ferromagnetic Ising model, we have seen, in the previous lecture, that the magnetic susceptibility χ diverges as $T \to T_c$. Combining this fact with what we know for the general theory of statistical mechanics (Lecture 7) we can learn some important facts about the **correlation**, a key concept in the theory of phase transition and beyond.

From the general theory of various thermodynamic ensembles and potentials (Lecture 7), we get

$$M = \frac{1}{\beta}\left(\frac{\partial \log X(T,H)}{\partial H}\right)_T. \tag{17.1}$$

The susceptibility that we defined per spin, in unit of μ_B per field (Eqs. 16.21 and 16.2), $\chi \equiv \frac{\partial m}{\partial H} = \frac{1}{N\mu_B}\left(\frac{\partial M}{\partial H}\right)_T$, is then given by

$$\chi = \frac{1}{\beta N \mu_B}\left(\frac{\partial^2 \log X(T,H)}{\partial H^2}\right)_T. \tag{17.2}$$

On the other hand, from Lecture 7, we recognize that this must be related to the 2nd order cumulant of the magnetization random variable (cf. Eq. 16.1)

$$\mathcal{M} \equiv \mu_B \sum_{i=1}^{N} \sigma_i \tag{17.3}$$

where $i = 1, ..., N$ is used to count spins on the lattice[1]. Since the coupling of H and $\mathcal{M}$ enters in X, through the exponential term $\exp(\beta H \mathcal{M})$ in summation over microstates (Eq. 7.5), we see (from Eq. 7.7) that[2]

$$\frac{1}{\beta^2}\left(\frac{\partial^2 \log X}{\partial H^2}\right)_T = \left\langle \mathcal{M}^2 \right\rangle_c ,$$

which means that

$$\chi = \frac{\beta}{N\mu_B}\left\langle \mathcal{M}^2 \right\rangle_c \qquad \textbf{fluctuation-dissipation theorem } (\chi). \tag{17.4}$$

As indicated, this is an example of the so-called **fluctuation-dissipation theorem**, which states that how susceptible the system is to an external driving "force" is proportional to how large the fluctuation of the system's "displacement" is *without* the driving force. When the magnetic field H is small, χ is independent of the H field, as $M \propto H$. Therefore, it follows that the right hand side can be evaluated with $H = 0$. A kind of laymen's way of explaining the fluctuation dissipation theorem is this: a system's ability to make something happen when a chance arises is proportional to the amount of the fluctuation of that something *inherent* in the system[3]. We have actually seen this theorem in action, before, Eq. 6.39:

$$C_x = k_B \beta^2 \left\langle \mathcal{E}^2 \right\rangle_c \qquad \textbf{fluctuation-dissipation theorem } (C_x). \tag{6.39}$$

In this case, the system's ability to raise energy when an external "force" (dT) is applied, is proportional to the inherent energy fluctuations of the system. As χ is the susceptibility of the system in the sense that $dM = \chi dH$, C_x can be understood as the "susceptibility" in the sense that $dE = C_x dT$. You might have guessed by now that the fluctuation dissipation theorem must apply to non-equilibrium states as well.

[1] Note that we have been using the one dimensional index i to count spins, no matter what the actual spatial dimension of the lattice is. Any multi-dimensional integer tuple indices $(i_1, i_2, i_3, ...)$ can be mapped to one dimensional indices i.

[2] Note that in Lecture 7, we have been using χ for the random variable for the equilibrium "displacement" variable x. *That* χ is identified with $\mathcal{M}$ in the current magnetic problem (and J is identified with H). That freed up the symbol χ, which is the standard symbol for the magnetic susceptibility, just in time for us. Please be ware that we have been using χ for the magnetic susceptibility whenever we considered a magnetic problem (cf. footnote 12 in page 12 of LN 8)!

[3] By changing system to person, I think this theorem agrees with the notion why the totally "useless" things that people (= mathematicians, physicists, or students in a liberal art college) do sometimes turn out to be very useful!

Indeed, the theorem is better known for these states. For instance, let us consider an electrical current flowing in the x direction. the electrical conductivity, σ is given by $j_q = \sigma E$ (Joule's law), where $j_q = qnv_d$ (n is the particle number density $= N/V$, q is the charge per carrier, and v_d is the "drift velocity" $\vec{v}_d \equiv \langle \vec{v} \rangle$) is the electrical current density and E is the electric field: σ is proportional to the *equilibrium* fluctuation of j. Likewise, the diffusion constant, D, is a "susceptibility" in the sense that $j_n = -D\nabla n$ (Fick's law), where n is the particle number density as before, and $j_n = nv_d$ is the particle number current density[4].

17.2 χ, the spin correlation

Now, let us re-focus our attention on χ. Expanding in terms of σ_i's we get

$$\begin{aligned}
\chi &= \frac{\beta}{N\mu_B}\left(\overline{\langle \mathcal{M}^2 \rangle} - \overline{\langle \mathcal{M} \rangle}^2\right) \\
&= \frac{\beta\mu_B}{N}\left(\sum_{i,j}\overline{\langle \sigma_i\sigma_j \rangle} - \sum_{i,j}\overline{\langle \sigma_i \rangle}\,\overline{\langle \sigma_j \rangle}\right) \\
&= \frac{\beta\mu_B}{N}\sum_{i,j=1}^{N}\langle \sigma_i, \sigma_j \rangle_c \qquad \text{cross-cumulant, } \textbf{correlation} \\
&= \beta\mu_B\sum_{j=1}^{N}\langle \sigma_1, \sigma_j \rangle_c \qquad \text{assuming crystal translational invariance} \\
&= \beta\mu_B\sum_{j=1}^{N}\left(\overline{\langle \sigma_1\sigma_j \rangle} - \overline{\langle \sigma_1 \rangle}\,\overline{\langle \sigma_j \rangle}\right) \qquad (17.5)
\end{aligned}$$

Note that this is a very general result, valid independent of any approximation employed (or not) to solve the problem. If we define

$$\Gamma_j \equiv \overline{\langle \sigma_1\sigma_j \rangle} - \overline{\langle \sigma_1 \rangle}\,\overline{\langle \sigma_j \rangle}, \qquad \textbf{correlation} \qquad (17.6)$$

[4] The proper discussion of these topics falls into the realm of "**Kubo formula**," which involves time and position dependent current-current correlation functions. You are encouraged to research into this topic, if you are interested in the quantum many body formalism of transport phenomena (and susceptibilities as well; see Kubo, Rep. Prog. Phys. vol. 29, 255 (1966), e.g.). In the simpler Boltzmann theory, one can obtain, by taking the relaxation time approximation, the following results for DC transport along the x axis and clearly demonstrate the fluctuation-dissipation theorem nature of σ and D: $\sigma = \frac{\tau\beta}{n}\left\langle {j_{q,x}}^2 \right\rangle_c$ and $D = \frac{\tau}{n^2}\left\langle {j_{n,x}}^2 \right\rangle_c$, where τ is the relaxation time. Using $j_{q,x} = qnv_x$, and $j_{n,x} = nv_x$, we get $\sigma = q^2\tau\beta n\left\langle {v_x}^2 \right\rangle_c$ and $D = \tau\left\langle {v_x}^2 \right\rangle_c$. Dividing these two equations, we get $\sigma/D = q^2\beta n$. Defining the *carrier mobility* μ as $\mu \equiv \sigma/(nq)$, we get $\mu/D = q\beta$. This result is the so-called **Einstein relation**: $D/(k_BT) = \mu/q$.

then our result can be put in a more compact form as

$$\chi = \beta\mu_B \sum_j \Gamma_j. \tag{17.7}$$

Note that[5] $\overline{\langle\sigma_1\rangle} = \overline{\langle\sigma_j\rangle}$, and so Γ_j can be 1 at the most. And, it cannot be less than -2. Thus, for χ to become a very large positive value, it is necessary that $\Gamma_j > 0$ for many j values. Clearly for χ to become infinitely large, as it does at the critical point (Section 16.1.3)[6], Γ_j must be non-zero for *infinite* numbers of j. This means that **there must exist a long range correlation of spins at the critical point**[7].

We shall investigate the near critical point behavior in the subsequent sections. How about far from the critical point? It is clear that, if $T \gg T_c$, spins will be all random, really, and there would be no correlation at all at a long length scale: $\Gamma_j = 0$ for any large j. This is also true for $T \ll T_c$. At first sight, one might be misled by the expression "correlation" that there is also an infinite range correlation at low temperature, since all spins are aligned along the same direction. However, the correlation must be understood as **the correlation of fluctuation**, since it can be re-written as

$$\Gamma_j = \overline{\left\langle \left(\sigma_1 - \overline{\langle\sigma_1\rangle}\right)\left(\sigma_j - \overline{\langle\sigma_j\rangle}\right)\right\rangle}. \tag{17.8}$$

Suppose one focuses on one spin, and "holds" that spin so that its value is significantly different from its thermal equilibrium value. The correlation measures how this deviation from equilibrium is remembered by all other spins. Away from T_c, this memory is only short-ranged. The following argument explains why this is plausible. If $T \ll T_c$, then the stability of the ordered phase requires that the system be stable under an "occasional flip of the spin." If holding a spin creates a long range correlation, then the system is not stable. If $T \gg T_c$, then the system is in a completely

[5] This is because we are considering a ferromagnetic Ising model, here. The mathematics of what we have carried out so far in this section, however, is still valid for an anti-ferromagnetic Ising model. It is just that an anti-ferromagnetic Ising model ($J < 0$) does not show a divergent magnetic susceptibility at the transition, which is understandable since the correlation Γ_j will alternate in sign for an anti-ferromagnetic fluctuation. Instead, the so-called *staggered* magnetic susceptibility diverges at the anti-ferromagnetic phase transition. The staggered magnetic susceptibility requires a magnetic field that flips sign at nearest neighbor lattice sites, and thus it remains a theoretical susceptibility only.

[6] More correct solutions show that the mean field theory prediction that $\chi \to \infty$ at T_c *is* a correct prediction. However, the exact manner in which χ diverges is not correctly predicted by the mean field theory in dimensions lower than the "upper critical dimension," mentioned in class.

[7] The discussion here makes it clear that the mean field theory *does* predict some long range correlation even if it starts out by assuming no correlation between spins. Otherwise, we would never obtain a divergent susceptibility! It would thus be *incorrect* to assume that $\Gamma_j = 0$ for $j \neq 1$ in Eqs. 17.6,17.7; although such an assumption would be consistent with the *initial* assumption of the mean field theory, it is not consistent with the *result* of the mean field theory. One may note that this state of affairs is typical of an approximate method such as the perturbation theory.

disordered state, and by definition an occasionally occurring fixed spin (e.g., due to a magnetic impurity) in the lattice cannot bring any order. **On other hand, at T_c, a truly long range memory of such a fluctuation is formed!**

17.3 Correlation, Bayesian view

Let us try to quantify what we just discussed a little more. By explicitly writing down the ensemble average, it is straightforward to re-write Eq. 17.6 as

$$\Gamma_j \;=\; p_+\overline{\langle\sigma_j\rangle}_+ - p_-\overline{\langle\sigma_j\rangle}_- - (p_+ - p_-)\overline{\langle\sigma_j\rangle}.$$

Here, **the conditional ensemble averages** $\overline{\langle\ \rangle}_\pm$ are defined for any random variable $x(\sigma_1, \sigma_2, ..., \sigma_N)$ as[8]

$$\overline{\langle x\rangle}_\pm \;\equiv\; \frac{\sum_{\sigma_2,\sigma_3,...,\sigma_N} x(\sigma_1 = \pm 1, \sigma_2, \sigma_3, ..., \sigma_N)\, e^{-\beta\mathcal{H}(\sigma_1=\pm1,\sigma_2,\sigma_3,...,\sigma_N)}}{\sum_{\sigma_2,\sigma_3,...,\sigma_N} e^{-\beta\mathcal{H}(\sigma_1=\pm1,\sigma_2,\sigma_3,...,\sigma_N)}}. \tag{17.9}$$

Using this definition, it is straightforward to show that

$$\overline{\langle x\rangle} \;=\; p_+\,\overline{\langle x\rangle}_+ + p_-\,\overline{\langle x\rangle}_- \tag{17.10}$$

where p_+ and p_- are unconditional probabilities (Eq. 4.3) for σ_1,

$$p_\pm \;=\; \frac{\sum_{\sigma_2,\sigma_3,...,\sigma_N} e^{-\beta\mathcal{H}(\sigma_1=\pm1,\sigma_2,\sigma_3,...,\sigma_N)}}{\sum_{\sigma_1,\sigma_2,\sigma_3,...,\sigma_N} e^{-\beta\mathcal{H}(\sigma_1,\sigma_2,\sigma_3,...,\sigma_N)}}. \tag{17.11}$$

As simple examples,

$$\overline{\langle\sigma_j\rangle} \;=\; p_+\overline{\langle\sigma_j\rangle}_+ + p_-\overline{\langle\sigma_j\rangle}_-, \tag{17.12}$$

$$\overline{\langle\sigma_1\sigma_j\rangle} \;=\; p_+\overline{\langle\sigma_j\rangle}_+ - p_-\overline{\langle\sigma_j\rangle}_-, \tag{17.13}$$

which have been used for the derivation of the above expression for Γ_j, the first equation with $j = 1$ and the second equation as is. For $j = 1$, note that these two equations become

$$\overline{\langle\sigma_1\rangle} \;=\; p_+ - p_-, \tag{17.14}$$

$$\overline{\langle\sigma_1^2\rangle} \;=\; p_+ + p_- \;=\; 1, \tag{17.15}$$

[8] This equation is written in the common "two equations in one convention," where the choice of the upper symbol or the lower symbol from the sign $\pm$ (or $\mp$, if used) is to be made uniformly throughout the equation. This convention is used whenever $\pm$ or $\mp$ appears in our equation.

as expected. As we discussed at the end of Section 17.2, the correlation Γ_j is concerned with the memory of the fluctuation

$$\delta\sigma_j \;=\; \sigma_j - \overline{\langle\sigma_j\rangle} \tag{17.16}$$

in terms of which the above expression for Γ_j becomes

$$\Gamma_j \;=\; p_+\overline{\langle\delta\sigma_j\rangle}_+ - p_-\overline{\langle\delta\sigma_j\rangle}_-. \tag{17.17}$$

Note that

$$0 \;=\; \overline{\langle\delta\sigma_j\rangle} \;=\; p_+\overline{\langle\delta\sigma_j\rangle}_+ + p_-\overline{\langle\delta\sigma_j\rangle}_-. \tag{17.18}$$

Therefore, we get a simple and quite general result for Γ_j.

$$\Gamma_j \;=\; 2p_+\,\overline{\langle\delta\sigma_j\rangle}_+ \;=\; -2p_-\,\overline{\langle\delta\sigma_j\rangle}_-. \tag{17.19}$$

This result shows clearly the essence of the discussion given at the end of Section 17.2. This result is valid in all settings, since we did not really assume anything, other than the basic principles of statistical mechanics. It is valid at any temperatures. It is valid for any correct solution of the problem, regardless of any approximation taken or not: mean-field solution or exact solution. It is also valid with or without an external H field. Finally, it is valid for any j values, including $j = 1$ (auto-correlation).

Our primary interest is a zero field correlation. So, assume $H = 0$. It is now easy to see why $\Gamma_j \approx 0$ for any j if $T \ll T_c$: either p_+ or $p_- \approx 0$. For $T \gg T_c$, we get $\Gamma_{j=1} \approx 1$ since $p_+ \approx p_- \approx 1/2$ and $\delta\sigma_j \approx \sigma_j$. However, for large j, we expect that σ_j fluctuate randomly, independent of σ_1, and thus, $\Gamma_j \approx \overline{\langle\sigma_j\rangle}_+ \approx 0$. Indeed, measurements[9] show that, for $T \gg T_c$, the spin correlation ceases to exist beyond the nearest neighbor. Only near T_c, Γ_j remains finite for a macroscopic number of j values. Note that for any $T \approx T_c$ or $T > T_c$, $\Gamma_{j=1} \approx 1$. As we shall see later, this correlation diminishes as a power law at T_c, while it diminishes exponentially if $T \neq T_c$.

17.4 ξ

The range of correlation that we just discussed is quantified by the parameter

$$\xi \;=\; \textbf{correlation length}$$

[9] The neutron scattering is the primary technique by which the spin correlation function can be measured directly.

which is arguably **the most important number in the theory of phase transition**. Here, we will use the mean field theory to introduce this concept, and then discuss it in more general terms.

Our focus here is on temperature values near the critical temperature, $T \approx T_c$. More precisely speaking, what we will discuss in this section is applicable when the zero-field order parameter is very small, i.e., $T \approx T_c$ or $T > T_c$.

For our discussion, it helps to be more specific about the crystal structure. In a three dimensional cubic lattice with lattice constant a, the lattice vectors are given by

$$\vec{r}_i \equiv a\left(i_1\,\hat{x} + i_2\,\hat{y} + i_3\,\hat{z}\right) \tag{17.20}$$

where $i_1, i_2, i_3 =$ integers, a is the lattice constant, and $i = 1, ..., N$ is the serial index that is mapped from the group index (i_1, i_2, i_3). From a large scale point of view, $\vec{r}_i$ becomes a continuous variable, and we can write it as $\vec{r}$. In the following, we shall leave the dimensionality of the space as general (d), and so we will consider a cubic crystal in d dimensions.

$$\vec{r}_i \equiv a\sum_{\alpha=1}^{d} i_\alpha \hat{e}_\alpha \tag{17.21}$$

where $\hat{e}_\alpha$ is the α-th unit vector in the d dimensional Cartesian coordinate system. Following are some definitions ($\equiv$), and facts ($=$).

$$\int d\vec{r} \equiv \int d^d\vec{r}, \quad \int d\vec{k} \equiv \int d^d\vec{k}, \qquad \text{(hyper-)volume integral} \tag{17.22}$$

$$v_0 \equiv a^d, \qquad \text{unit cell (hyper-)volume} \tag{17.23}$$

$$\tilde{f}(\vec{k}) \equiv \int d\vec{r}\; e^{-i\vec{k}\cdot\vec{r}} f(\vec{r}), \qquad \text{Fourier transform} \tag{17.24}$$

$$f(\vec{r}) = \frac{1}{(2\pi)^d}\int d\vec{k}\; e^{i\vec{k}\cdot\vec{r}}\tilde{f}(\vec{k}), \qquad \text{Inverse Fourier transform} \tag{17.25}$$

$$(2\pi)^d\delta(\vec{k}=0) = V, \qquad \because \quad \int d\vec{r}\; e^{-i\vec{k}\cdot\vec{r}} = (2\pi)^d\delta(\vec{k}) \tag{17.26}$$

$$\delta(\vec{r}=0) = \frac{1}{v_0}. \qquad \because \quad \sum_i = \frac{1}{v_0}\int d\vec{r} \tag{17.27}$$

The total magnetization random variable is then given by

$$\mathcal{M} = \mu_B\sum_i \sigma_i = \frac{\mu_B}{v_0}\int d\vec{r}\;\sigma(\vec{r}), \tag{17.28}$$

$$\sigma(\vec{r}) \equiv \sigma(\vec{r}_i) = \sigma_i. \tag{17.29}$$

The ensemble average of $\sigma(\vec{r})$ is the per spin dimensionless magnetization

$$\overline{\langle\sigma(\vec{r})\rangle} \equiv m = \frac{M}{N\mu_B}. \qquad \text{(Eq. 16.2)} \tag{17.30}$$

Let us consider the mean field theory as we did in Lecture 15:

$$\overline{\langle\sigma_i\rangle} \;=\; \tanh\left(\beta H_i \mu_B\right), \qquad \text{solution for } \overline{\langle\sigma\rangle} \tag{15.25}$$

$$H_i \;=\; H + \frac{J}{\mu_B} \sum_{j=\mathrm{nn}(i)} \overline{\langle\sigma_j\rangle}. \qquad \text{Eq. 15.23 plus external field } H \tag{17.31}$$

Using the notation Eq. 17.29, and using the multi-variable Taylor expansion,

$$\sum_{j=\mathrm{nn}(i)} \sigma_j \;=\; z\sigma_i + \left(\sum_j \vec{\delta}_j\right) \cdot \vec{\nabla}\sigma_i + \frac{1}{2}\left(\sum_j \left(\vec{\delta}_j \cdot \vec{\nabla}\right)^2\right)\sigma_i + \ldots, \qquad \sigma_i = \sigma(\vec{r}).$$

where z is the "coordination number," i.e. the number of the nearest neighbors, as before (LN 15, page 10). For our cubic crystal, $z = 2d$, but we leave this as symbol z, anticipating the application of our result here to other sufficiently symmetric non-cubic crystal structure as well (see footnote 10). Here, $\vec{\delta}_j(\vec{r})$ vectors are nearest neighbor displacement vectors: $\vec{r} + \vec{\delta}_j(\vec{r})$ corresponds to a nearest neighbor lattice point for $\vec{r}$ at index i. For any inversion symmetric crystal, we get $\sum_j \vec{\delta}_j(\vec{r}) = 0$, and so the second term vanishes. In addition, for our cubic crystal[10] it is easy to see that $\frac{1}{2}\sum_j \left(\vec{\delta}_j \cdot \vec{\nabla}\right)^2 = a^2\vec{\nabla}^2$, where a is the lattice constant. Thus, we can write

$$H_i \;\approx\; H + \frac{Jz}{\mu_B}\,\overline{\langle\sigma_i\rangle} + \frac{Ja^2}{\mu_B}\,\vec{\nabla}^2\overline{\langle\sigma_i\rangle}.$$

Now, we plug this into Eq. 15.25. Assuming that we are near the critical point, that the H field is small, and that the spatial variation of $\overline{\langle\sigma_i\rangle}$ is small, we get, using $\tanh\delta \approx \delta - \frac{1}{3}\delta^3$,

$$\overline{\langle\sigma_i\rangle} \;\approx\; \beta\mu_B H + \beta J z\,\overline{\langle\sigma_i\rangle} + \beta J a^2\,\vec{\nabla}^2\overline{\langle\sigma_i\rangle} - \frac{1}{3}(\beta J z)^3\,\overline{\langle\sigma_i\rangle}^3$$

where only the leading order terms for H and $\vec{\nabla}\overline{\langle\sigma_i\rangle}$ are kept and terms up to $\overline{\langle\sigma_i\rangle}^3$ are kept. Using dimensionless variables t and h (Eqs. 16.14, 16.15), and noting our assumption $t \approx 0$, this equation is approximated as

$$t\,\overline{\langle\sigma_i\rangle} - \frac{a^2}{z}\vec{\nabla}^2\overline{\langle\sigma_i\rangle} + \frac{1}{3}\overline{\langle\sigma_i\rangle}^3 \;=\; h. \tag{17.32}$$

For a spatially uniform field that we assumed so far, the equilibrium magnetization $\overline{\langle\sigma_i\rangle}$ would also be uniform, and thus the second term vanishes, and we get the same equation, $tm + \frac{1}{3}m^3 = h$, with $m \equiv \overline{\langle\sigma_i\rangle}$, as in the last lecture. Then, we repeat the

[10] However, a similar conclusion holds for a sufficiently symmetric crystal structure, such as a two dimensional hexagonal crystal structure: $\frac{1}{2}\sum_j \left(\vec{\delta}_j \cdot \vec{\nabla}\right)^2 = Aa^2\vec{\nabla}^2$, where A is a number of order 1, and a is the lattice constant.

same result that we obtained there: $tm + \frac{1}{3}m^3 = h$ is identical with Eq. 16.25, which is in fact applicable for *any* $t \approx 0$; for $t > 0$ the m^3 term need be dropped, since it is negligible. We certainly do not want to just repeat what we already know. What is really new in this new mean field calculation?

The first thing to note is that nothing that led to the above equation so far prevents us from assuming a *position dependent field* $H(\vec{r})$! In this case, we expect that the equilibrium magnetization will be position dependent also. Thus, with

$$m(\vec{r}) \equiv \overline{\langle \sigma_i \rangle} \tag{17.33}$$

the above equation becomes

$$\left(\vec{\nabla}^2 - \frac{zt}{a^2}\right) m(\vec{r}) - \frac{z}{3a^2} m^3(\vec{r}) = -\frac{z h(\vec{r})}{a^2}. \tag{17.34}$$

The second thing to note is that, based on the discussions in the previous two sections, a good way to "probe the system" is to turn on the h field only at the origin ($i = 1$). With such a delta function field, we would be able to polarize the spin at origin, more so, if h is stronger. For the time being, let us suppose that we have *completely* polarized σ_1 so that $\overline{\langle \sigma_1 \rangle} = +1$ with a suitable "saturation field" h_s. Then, the magnetization distribution $m(\vec{r})$ that results from this field will be precisely $\overline{\langle \sigma_j \rangle}_+$. By measuring the difference between this $m(\vec{r})_{h_s}$ and any zero field magnetization $m(\vec{r})_{h=0}$, we will be measuring precisely what we referred to as $\overline{\langle \delta\sigma_j \rangle}_+$ in Eq. 17.19 for zero field. To summarize, we get

$$m(\vec{r})_{h_s} - m(\vec{r})_{h=0} = \overline{\langle \delta\sigma(\vec{r}) \rangle}_{+,h=0} \quad \begin{cases} \approx \Gamma(\vec{r})_{h=0} & T \lesssim T_c \quad (p_+ \approx p_- \approx 1/2) \\ = \Gamma(\vec{r})_{h=0} & T \geq T_c \quad (p_+ = p_- = 1/2) \end{cases}$$

where in the $T \lesssim T_c$ case we used Eq. 17.19, and noted that

$$p_+ = 1 - p_- = \frac{1}{2} + \frac{1}{2}\sqrt{-3t} \approx \frac{1}{2}, \qquad h = 0, \quad \overline{\langle \sigma_1 \rangle} \approx \sqrt{-3t} \tag{17.35}$$

the result of Eq. 16.26, with $h = 0$ with the + sign chosen by our convention for the saturation direction. However, there is a problem with this argument, and some careful discussion of the problem is necessary.

First, the problem. We note that in order to (nearly) saturate the spin, we need $h_s \sim O(1)$! To see this, we must go all the way back to Eq. 15.25, with $i = 1$ in mind (and H applied only to the $i = 1$ spin): $\overline{\langle \sigma_1 \rangle} = \tanh\left(\beta\mu_B H + \beta J \sum_j \overline{\langle \sigma_{j=nn} \rangle}\right)$: since $\tanh 2 = 0.964$ and $\tanh 5 = 0.99991$, we see that to get $\overline{\langle \sigma_1 \rangle} \to 1$, we need $h_s \approx \beta\mu_B H \gtrsim 4$.

Second, the solution is rather simple. We do not need to fully polarize the spin 1 at all. All we need is to apply an infinitesimal field and polarize the field just a little bit. Suppose that we have done so. Let h_b be such an infinitesimal polarizing field. What would be the induced spin moment at spin 1? It is an equally small amount. If the field was uniform, then the susceptibility will diverge as $t \to 0$, but this is not expected for the local field h_b. This can be seen, again, from the equation: $\overline{\langle\sigma_1\rangle} = \tanh\left(\beta\mu_B H + \beta J \sum_j \overline{\langle\sigma_{j=nn}\rangle}\right)$. Assuming that $t > 0$, we get $\overline{\langle\sigma_1\rangle} = h + \frac{T_c}{T}\overline{\langle\sigma_{nn}\rangle}$, where $\overline{\langle\sigma_{nn}\rangle}$ is the average induced moment at nearest neighbor sites. For a localized biasing field, $\overline{\langle\sigma_{nn}\rangle}$ can *not* be identical with $\overline{\langle\sigma_1\rangle}$, but it would be somewhat reduced. Writing $\overline{\langle\sigma_{nn}\rangle} = \alpha m_1$ where $m_1 = \overline{\langle\sigma_1\rangle}$ and $0 < \alpha < 1$, we then get $m_1\left(1 - \frac{\alpha}{1+t}\right) = h$. So, the susceptibility for $t > 0$ is always finite. In particular, for any small $|t|$, $m_1 \approx \frac{h}{1-\alpha}$. A similar argument shows that the local susceptibility remains finite and independent of t for $t < 0$ as well.

Therefore, for **an infinitesimal local biasing field** h_b and small $|t|$,

$$\overline{\langle\sigma_1\rangle}_{h_b} - \overline{\langle\sigma_1\rangle}_{h=0} = \Lambda h_b. \qquad \Lambda = \text{constant; dimensionless local susceptibility} \tag{17.36}$$

This means that, with h_b,

$$p_+ = 1 - p_- = \frac{1 + \overline{\langle\sigma_1\rangle}_{h=0} + Ch_b}{2}. \tag{17.37}$$

Now, consider

$$\delta_b m(\vec{r}) \equiv m(\vec{r})_{h_b} - m(\vec{r})_{h=0} \tag{17.38}$$

for any $\vec{r} \neq 0$. Since

$$\begin{aligned} m(\vec{r})_{h_b} &= \frac{1 + \overline{\langle\sigma_1\rangle}_{h=0} + Ch_b}{2}\,\overline{\langle\sigma(\vec{r})\rangle}_{+,h=0} + \frac{1 - \overline{\langle\sigma_1\rangle}_{h=0} - Ch_b}{2}\,\overline{\langle\sigma(\vec{r})\rangle}_{-,h=0} \\ m(\vec{r})_{h=0} &= \frac{1 + \overline{\langle\sigma_1\rangle}_{h=0} + Ch_b}{2}\,m(\vec{r})_{h=0} + \frac{1 - \overline{\langle\sigma_1\rangle}_{h=0} - Ch_b}{2}\,m(\vec{r})_{h=0} \end{aligned}$$

where $h = 0$ condition in the first line is due to $\vec{r} \neq 0$, we get

$$\delta_b m(\vec{r}) = \frac{1 + \overline{\langle\sigma_1\rangle}_{h=0} + Ch_b}{2}\,\overline{\langle\delta\sigma(\vec{r})\rangle}_{+,h=0} + \frac{1 - \overline{\langle\sigma_1\rangle}_{h=0} - Ch_b}{2}\,\overline{\langle\delta\sigma(\vec{r})\rangle}_{-,h=0}. \tag{17.39}$$

Now, from Eq. 17.19, we get

$$\overline{\langle\delta\sigma(\vec{r})\rangle}_{\pm,h=0} = \frac{\pm 1}{2p_{\pm,h=0}}\Gamma(\vec{r})_{h=0} = \frac{\pm 1}{1 \pm \overline{\langle\sigma_1\rangle}_{h=0}}\Gamma(\vec{r})_{h=0} \approx \pm\left(1 \mp \overline{\langle\sigma_1\rangle}_{h=0}\right)\Gamma(\vec{r})_{h=0}.$$

Using this in the previous equation, we get, after keeping leading order terms only,

$$\delta_b m(\vec{r}) \;=\; \Lambda h_b \Gamma(\vec{r})_{h=0}. \tag{17.40}$$

This plausible result makes it possible to measure the correlation function by applying a small local field.

To summarize, the change of the magnetization pattern caused by an infinitesimal local spin-biasing field, h_b,

$$\delta_b m(\vec{r}) \;=\; m(\vec{r})_{h_b} - m(\vec{r})_{h=0} \;=\; \Lambda h_b\, \Gamma(\vec{r})_{h=0} \tag{17.41}$$

gives the full information about **the zero-field correlation function**.

So, let us investigate $\delta_b m(\vec{r})$. The applied field is given by

$$h(\vec{r}) \;=\; h_b \delta_{i,1} \;=\; h_b v_0 \delta(\vec{r}). \qquad \text{Eq. 17.27} \tag{17.42}$$

Plugging this into Eq. 17.34, we get the following differential equation to solve:

$$\left(\vec{\nabla}^2 - b(T)\right) m(\vec{r}) - c m^3(\vec{r}) \;=\; -\lambda \delta(\vec{r}), \qquad T \approx T_c \tag{17.43}$$

$$\text{where} \quad b(T) \;=\; \frac{zt}{a^2},\; c \;=\; \frac{z}{3a^2},\; \lambda \;=\; \frac{z h_b v_0}{a^2}. \tag{17.44}$$

Our goal is to turn the above differential equation to the one for $\delta_b m(\vec{r}) = \Lambda h_b \Gamma(\vec{r})$ and then solve for $\Gamma(\vec{r})$ (we omit the $h = 0$ subscript, for brevity). First, we consider the $T \geq T_c$ case, where the cubic term is negligible, and $\Lambda h_b\, \Gamma(\vec{r}) = \delta_b m(\vec{r}) = m(\vec{r})$, since $m(\vec{r})_{h=0} = 0$. Thus, we get the following results:

$$\left(\vec{\nabla}^2 - b(T)\right) \Gamma(\vec{r}) \;=\; -\lambda' \delta(\vec{r}), \qquad T \geq T_c \tag{17.45}$$

$$\lambda' \;=\; \frac{\lambda}{\Lambda h_b} \;=\; \frac{z v_0}{a^2 \Lambda}, \tag{17.46}$$

$$\left(k^2 + b(T)\right) \tilde{\Gamma}(\vec{k}) \;=\; \lambda', \qquad T \geq T_c \tag{17.47}$$

$$\tilde{\Gamma}(\vec{k}) \;=\; \frac{\lambda'}{k^2 + b(T)}. \qquad T \geq T_c \tag{17.48}$$

This form of the correlation function is referred to as **the Ornstein-Zernike form**. Note that $\sqrt{1/b(T)}$ defines a length scale of the problem. This is what we are mainly after, here! Before we write down what this length scale is, let us consider the $T \lesssim T_c$ case, also. In this case, we must explicitly consider $\delta_b m(\vec{r}) \equiv m(\vec{r})_{h_b} - m(\vec{r})_{h=0}$, since $m(\vec{r})_{h=0} = \pm\sqrt{\frac{-b(T)}{c}} = \pm\sqrt{-3t}$ (as we find again here, or as we already found in Eq. 16.18) is not zero. For the same reason, the cubic term ($\propto m^3$) in the differential

equation cannot be ignored. Writing $m(\vec{r})_{h_b} = m(\vec{r})_{h=0} + \delta_b m(\vec{r})$, and keeping terms linear in $\Lambda h_b \, \Gamma(\vec{r}) = \delta_b m(\vec{r})$ only, we get

$$\left(\vec{\nabla}^2 + 2b(T)\right)\Gamma(\vec{r}) \;=\; -\lambda'\delta(\vec{r}), \qquad T \lesssim T_c \tag{17.49}$$

$$\left(k^2 - 2b(T)\right)\tilde{\Gamma}(\vec{k}) \;=\; \lambda', \qquad T \lesssim T_c \tag{17.50}$$

$$\tilde{\Gamma}(\vec{k}) \;=\; \frac{\lambda'}{k^2 - 2b(T)}. \qquad T \lesssim T_c \tag{17.51}$$

So, we get this very important result.

Correlation length in the mean field approximation

$$\xi(T) \;=\; \begin{cases} \dfrac{1}{\sqrt{b(T)}} \;=\; \dfrac{a}{\sqrt{zt}}, & T \geq T_c \\ \dfrac{1}{\sqrt{-2b(T)}} \;=\; \dfrac{a}{\sqrt{-2zt}}. & T \lesssim T_c \end{cases} \tag{17.52}$$

The essential physics that we can learn from this mean field exercise is that **the $\xi(T)$ diverges as** $t \to 0$, i.e., as the critical point is approached.

At the critical point, the correlation length of the order parameter is infinite, while the system is neither ordered nor disordered.

Here in the last part "not ordered" means that the system does not display a net magnetic moment in any finite fraction of the system, on average. And, "not disordered" means that the correlation length is infinite.

The infinite correlation length at the critical point is, as expected from our discussion in Section 17.2, robust, remaining valid beyond the mean field theory. However, the manner in which this divergence occurs turns out to be incorrectly predicted by the mean field theory, as you may have guessed.

The above Ornstein-Zernike correlation function can be summarized as

$$\tilde{\Gamma}(\vec{k}) \;=\; \frac{\lambda'}{k^2 + \frac{1}{\xi^2}}. \qquad\qquad \lambda' \;=\; \frac{z v_0}{a^2 \Lambda} \tag{17.53}$$

It is possible to find the real space correlation function, by inverse-Fourier transform, Eq. 17.25. But, first of all, by writing down the integral, doing the dimensional analysis, and studying the integral in terms of the dimensionless variable $\vec{k}\cdot\vec{r}$, we get the general form

$$\Gamma(\vec{r}) \;=\; z\left(\frac{a}{r}\right)^{d-2} f\left(\frac{r}{\xi}\right), \tag{17.54}$$

where f is a function. You might find this functional form particularly familiar in the case of $d = 3$. In this case, the function that we encounter here has the "screened Coulomb potential" form or the "Yukawa potential" form.

$$\Gamma(\vec{r}) \;=\; \frac{\lambda'}{4\pi r}\exp\left(-\frac{r}{\xi}\right). \qquad\qquad d = 3, \quad \text{any } r \neq 0 \tag{17.55}$$

which gives a precise form that defines ξ as the finite correlation length scale of the problem for any non-zero t. In general, it can be shown[11] that

$$\Gamma(\vec{r}) \propto \begin{cases} z\cdot\dfrac{a^{d-2}}{\xi^{\frac{d-3}{2}} r^{\frac{d-1}{2}}}\exp\left(-\dfrac{r}{\xi}\right), & t \neq 0,\ r \gg \xi \\ z\cdot\left(\dfrac{a}{r}\right)^{d-2}. & t = 0,\ \xi = \infty,\ r \to \infty,\ r \ll \xi,\ d > 2 \end{cases} \tag{17.56}$$

Despite the appearances here, note that the value of the correlation function in real space cannot be greater than 1 (Section 17.2) and it is intuitively clear that it must be a decreasing function of r. As we saw in 17.3, $\Gamma(\vec{r} = 0) \approx 1$, indeed, for $T \approx T_c$ (or any $T > T_c$). What we calculated above excludes $\vec{r} = 0$ (page 10), and so the minimum value for r is a. In any case, $\Gamma(\vec{r})$ has a strictly finite range: $-2 \leq \Gamma(\vec{r}) \leq 1$ (Section 17.2)[12].

The power law behavior of $\Gamma(\vec{r})$ at $t = 0$ corresponds to **the long range correlation**, where $\Gamma(\vec{r})$ remains finite at an infinite number of lattice sites, as discussed qualitatively in Section 17.2. Thus, this long-range correlation is the reason why the susceptibility diverges at $t = 0$.

In general, the correlation length and the correlation function are the two most important ingredients of the theory of the phase transition.

[11] Huang. Pathria and Beale.

[12] It would appear that this property is broken in some cases. For instance, if $d = 2$ and $t = 0$ ($\xi = \infty$), then the solution at large r is given by $\Gamma(\vec{r}) \propto \log\frac{\xi}{r}$, and if $d = 1$ and $t = 0$, $\Gamma(r) \propto |x|/a$. Such solutions should be taken with a grain of salt. First, if $|\Gamma(\vec{r})|$ becomes large, then the Taylor series expansion (Eq. 17.32) breaks down. Second, if $|\Gamma(\vec{r})|$ becomes large, then it means that the flucatuation becomes very large and so the mean field theory that we employed here breaks down! Indeed, if $d < 4$, the mean field theory is not valid, and the effect of fluctuation becomes stronger as the dimension is lowered. As we shall see shortly, in one dimension, the fluctuation is so large that the finite temperature phase transition disappears.

Critical exponents ν and η

$$\xi \propto |t|^{-\nu} \qquad |t| \ll 1 \tag{17.57}$$

$$\Gamma(r) \propto r^{2-d-\eta} \qquad t = 0,\ r \gg a,\ d = \text{spatial dimension} \tag{17.58}$$

In the theory of the critical phenomena, the correlation length and the correlation function play the central role. **These two exponents ν and η completely determine other exponents**, such as α ($C \propto |t|^{-\alpha}$; Section 16.1.4), β ($m \propto |t|^{\beta}$; Section 16.1.2), γ ($\chi \propto |t|^{-\gamma}$; Section 16.1.3), and δ (the equation of state $h \propto m^{\delta}$ at $t = 0$; Section 16.4). Within the mean field theory that we have used so far, we see that $\nu = 1/2$. The correct value for ν is 1 for $d = 2$, 0.6301 for $d = 3$, and 1/2 for $d \geq 4$, justifying our mean field solution above only for $d \geq 4$. The η value is 0, within the mean field theory. The correct values are 1/4 ($d = 2$) and 0.0364 ($d = 3$). The interrelationship between these exponents is the subject of the "scaling hypothesis," to be discussed later with the renormalization group theory.

17.5 χ, revisited

Eq. 17.7 can be re-written as $\chi = \frac{\beta\mu_B}{v_0} \int d\vec{r}\, \Gamma(\vec{r})$, which is equal to, by Eq. 17.24,

$$\chi = \frac{\beta\mu_B}{v_0}\, \tilde{\Gamma}(\vec{k} = 0). \tag{17.59}$$

Within the mean field theory, we know that $\chi \sim \frac{\mu_B}{Jz|t|}$ (Section 16.1.3) near T_c. Thus, it must be that $\tilde{\Gamma}(\vec{k} = 0) \sim \frac{v_0 k_B T_c}{Jz|t|} = \frac{v_0}{|t|}$. Since $\lambda' = \frac{zv_0}{a^2\Lambda}$, and $\xi \sim \frac{a}{\sqrt{z|t|}}$, we get, from Eq. 17.53, $\tilde{\Gamma}(\vec{k} = 0) = \lambda'\xi^2 \sim \frac{zv_0}{a^2\Lambda}\frac{a^2}{z|t|} = \frac{v_0}{|t|\Lambda}$, which is indeed the expected behavior. In addition, note that the constant Λ can be determined by equating the two χ's calculated, one calculated in the last lecture and the other calculated using $\tilde{\Gamma}(\vec{k} = 0)$. Such exercise will confirm that Λ is a smallish finite constant, as discussed prior to Eq. 17.36.

17.6 Landau theory, generalized

Considering the Hamiltonian

$$\mathcal{H} = -\mu_B H \sum_i \sigma_i - J \sum_{\langle i,j \rangle} \sigma_i \sigma_j, \tag{15.13}$$

in the continuum limit, as we did in the previous section, we get

$$\mathcal{H} = -\frac{\mu_B H}{v_0} \int d\vec{r} \sigma(\vec{r}) - \frac{Jz}{2v_0} \int d\vec{r}\, \sigma(\vec{r})^2 - \frac{Ja^2}{2v_0} \int d\vec{r}\, \sigma(\vec{r})\, \vec{\nabla}^2 \sigma(\vec{r}) + ...$$

Integrating the last term by parts[13], Thus, it is straightforward to write

$$\mathcal{H} = \int d\vec{r}\, \hbar(\vec{r}), \tag{17.60}$$

$$\beta v_0 \hbar(\vec{r}) = -h\sigma(\vec{r}) - \frac{Jz\beta}{2}\sigma(\vec{r})^2 - \frac{Ja^2\beta}{2}\left(\vec{\nabla}\sigma(\vec{r})\right)^2 + ... \tag{17.61}$$

Now, consider the Gibbs partition function

$$X = \sum_{\{\sigma(\vec{r})\}} e^{-\beta\mathcal{H}}, \tag{17.62}$$

where the summation is over all possible functions[14] of $\sigma(\vec{r})$. At this point, let us pause and think about what we have done above. We assumed that the order parameter is slowly varying. This is why we could do the Taylor expansion such as Eq. 17.32 in the first place. The reason that we did so is not because there are not short range fluctuations, but because those long range fluctuations are singularly responsible for the phase transition physics. The proper formalism in which one averages over the short range fluctuations and bring out the long range fluctuation physics to the "effective Hamiltonian" is the renormalization group theory, which we will discuss in a future lecture. At this point, it suffices to say that, assuming that such a procedure has been carried out, we can consider $\vec{r}$ as **a coarse grained position variable**, where many atomic sites are considered a single point, and similarly $\sigma(\vec{r})$ must be taken as **a coarse grained spin variable**, the so-called block spin, the sum or average over many atomic spins. Then, one notes that any block spin variable $\sigma(\vec{r})$ will carry an entropy as well. With this in mind, we change the variable $\sigma(\vec{r}) \to m(\vec{r})$.

As in the last lecture, we write m instead of $\mathfrak{m}$, even if what we mean is the random variable, not the thermal average, for the purpose of discussing the Landau free energy.

[13] The "surface integral" vanishes, if one applies the periodic boundary condition or the fixed boundary condition, procedures well-justified for the study of bulk physics.

[14] So, this summation would represent a functional integral, if the sum can be converted to an integral.

Now, one must write X as

$$X = \sum_{\{m(\vec{r})\}} e^{-\beta(\mathcal{H}-T\mathcal{S})} = \sum_{\{m(\vec{r})\}} e^{-\beta G[m(\vec{r})]} \tag{17.63}$$

where the Gibbs free energy *random variable* G is a *functional* of $m(\vec{r})$. To be specific, it is an integral over a function of m:

$$G[m(\vec{r})] = \int d\vec{r}\, g(m(\vec{r})). \tag{17.64}$$

where $g(m(\vec{r}))$ is the Gibb free energy density random variable. As before, note that $G = \mathcal{H}_0 - H\mathcal{M} - T\mathcal{S}$, where $\mathcal{H}_0$ is the internal energy of the system, corresponding to E in thermodynamics. We write G instead of $\mathcal{G}$, and g instead of $\mathcal{g}$, for the purpose of discussing the Landau free energy, just as write m instead of .

In the Landau theory, $g(m)$ is expressed as an analytic function of $m(\vec{r})$

$$g(m(\vec{r})) = -h(\vec{r})m(\vec{r}) + a(T) + \frac{1}{2}|\vec{\nabla}m(\vec{r})|^2 + \frac{1}{2}b(T)m(\vec{r})^2 + \frac{1}{4}c(T)m(\vec{r})^4 + ... \tag{17.65}$$

In this expression, we have further scaled $m(\vec{r})$ so that it is not dimensionless any more. Rather, its dimension is determined from the requirement that g or $|\vec{\nabla}m(\vec{r})|^2$ has the dimension of energy per volume. Likewise, the h field is no longer dimensionless, and thus is *not* equal to the h field that we have been discussing so far[15].

Note that the above form of $g(m)$ is well motivated by our above elementary examination of the Hamiltonian density $\mathcal{h}$ (Eq. 17.61) and the expected Taylor expansion of the local Gibbs free energy (cf. page 4 of LN 16). It is different significantly from Eq. 16.54, the mean field Landau free energy for a uniform h field, only in that we have now included the gradient square term of the magnetization.

It is now important to realize that **as a theory examining $g(m(\vec{r}))$ the Landau theory is no longer a mean field theory**. Rather, it is a theory that examines the long range fluctuations of the order parameter.

For a given set of thermodynamic variables, T, h, the system seeks to minimize the Landau free energy $G[m] = \int d\vec{r}\; g(m)$, as it approaches the equilibrium. This minimization fixes the thermodynamic variables of the system. However, there is no reason why this this minimization occurs with a single functional form of $m(\vec{r})$ at all.

This is easy to understand if one considers a one dimensional ferromagnetic Ising chain. Considers a finite temperature, and no field. The free energy, $\mathcal{E} - T\mathcal{S}$, must

[15] The same comment is applicable to the h field included in Eq. 16.34. It is *not* equal to other h fields used in that lecture note, while m in Eq. 16.34 can be taken to be the same as other m's used in that lecture note.

be minimized. Within the periodic boundary condition, one can consider the Ising chain as a ring of spins. Naively, one might expect that all spins lining up in the same direction is the best thing to do, assuming that T is low enough, but finite. However, this is not so. Consider bisecting the ring in two. For one half, all spins point up, and for the other half, all spins point down. This leads to an increase in $\mathcal{E}$, by $2J$. However, it is to be noted that there are $N/2$ ways to bisect the ring, contributing $-T\Delta\mathcal{S} = -k_BT\log(N/2)$. So, the free energy for such bisected configurations is lower by $2J - k_BT\log(N/2)$, which is a very large negative number in the thermodynamic limit. This is the reason why in a one dimensional Ising chain, the average magnetic moment vanishes at any finite temperature. And it is not just bisected ring configurations. One can divide the ring into three parts, or four parts, and so on, while keeping the total magnetization fixed at zero.

This one dimensional Ising chain makes it clear that (exponentially) many different functions $m(\vec{r})$ contribute to the equilibrium state[16]. Also, from this discussion, it is clear why the mean field theory prediction of a finite temperature phase transition is completely broken in the one dimensional case. These two facts must be related to each other.

Indeed, if one looks back on lecture 16, one realizes that there we treated each spin as acting exactly the same way as all other spins. This is the assumption of the mean field theory. Such assumption automatically implies only one functional form for $m(\vec{r})$ – a uniform function in the case of the ferromagnetic case. Even when the external field $h(\vec{r})$ was taken to be position dependent (Section 17.4), we could use the mean field theory and assumed the following. In the case of a delta function field, we assumed that all spins following the average behavior, spin maximized at the center of the delta function site and decreasing as it moves away from the site, *is sufficient to describe* the thermodynamic state.

Making these observations, the meaning of the mean field approximation can be put in a very useful form, within the Landau theory, as follows. In the expression for the partition function, Eq. 17.63, one cannot[17] use the rule of large numbers ("the saddle point approximation"; Section 4.3) with respect to the functional sum

[16] For one thing, we have no rule for large numbers for minimizing $G[m]$, since the integral $G = \int d\vec{r}\, g(m)$ does *not* involve any exponential function.

[17] One cannot over-emphasize the difference that would arise if one expressed the sum as the sum over total magnetization variable: $X = \sum_{\mathcal{M}} e^{-\beta G(\mathcal{M},T,h)}$. In this case, the rule of large numbers must hold, *of course*, since $\mathcal{M} = O(N)$, and we get $X = e^{-\beta G_{min}}$, which is the *reason why* $G = G_{min} = -k_BT\log X$ in equilibrium. This was the very basis of the principle of statistical mechanics that we studied with care in lectures 6 and 7. In those lectures, note that we never used the rule of large numbers when the summation was over μ (microstates, which corresponds to $m(\vec{r})$, if not coarse-grained, of the current problem); we used the rule of large numbers only when the sum is over macroscopic thermodynamic variables such as energy, volume, or magnetization (see page 7 of LN 6, and also examine Eqs. 7.2, 7.5, 7.6, 7.11, 7.12, etc. carefully).

over $m(\vec{r})$; the number of possible functions is exponentially large, invalidating the assumption for Section 4.3.1. **If one insists on using the rule or large numbers, i.e., the saddle point approximation, for the sum expressed in Eq. 17.63, nevertheless, then the resulting theory is the mean field approximation.**

Clearly, this is a crude approximation. However, within the Landau theory, a phenomenological theory at its heart, the situation is not as bad as one might think, since (1) the formalism already includes some short scale physics in the correct way by the coarse graining procedure assumed for $m(\vec{r})$, and (2) the free energy density, Eq. 17.65, can be used to go beyond the mean field theory.

Therefore, the mean field theory consists of taking the $m(\vec{r})$ that satisfies

$$\delta G = \int d\vec{r} \frac{\delta g}{\delta m} \delta m = 0. \tag{17.66}$$

as determining the equilibrium value of $G[m]$ completely[18]. The equation $\delta g/\delta m = 0$ for Eq. 17.65, with unwritten higher order terms ignored, gives

$$-h(\vec{r}) - \vec{\nabla}^2 m(\vec{r}) + b(T) m(\vec{r}) + c(T) m(\vec{r})^3 = 0. \tag{17.67}$$

Assuming

$$b(T) = b_0 t, \qquad b_0 > 0 \tag{17.68}$$

$$c(T) = c > 0 \tag{17.69}$$

similarly as in Eqs. 16.32 and 16.33, we get

$$\left(\vec{\nabla}^2 - b_0 t\right) m(\vec{r}) - c\, m(\vec{r})^3 = -h(\vec{r}), \tag{17.70}$$

which indeed has precisely the same structure and the same t dependence as Eq. 17.34.

17.7 Ginzburg-Landau criterion

Having applied the mean field theory, and knowing that it could be wrong, as discussed in the previous section, one might ask how can one judge whether or not the theory is incorrect? Generally, this would be the topic of the field theory, in which the Landau free energy introduced in the previous section can be seen as quite similar to the ϕ^4 field theory for mesons.

[18] Of course, one should, in general, make sure that the solution $\delta g/\delta m = 0$ gives the minimum, not the maximum, of G.

However, using what we already know, we can make some simple physical argument, known as the Ginzburg-Landau criterion. Simply, one would require that the fluctuation of the order parameter be much smaller than its mean value. The typical magnitude of the fluctuation squared can be taken as the value of $\Gamma(\vec{r})$ at $r = \xi$. So, we require that

$$\frac{\Gamma(r=\xi)}{m^2} \ll 1. \tag{17.71}$$

Using $m^2 = -3t$ (Eq. 16.18), $\Gamma(\xi) \sim z\left(\frac{a}{\xi}\right)^{d-2}$ (Eq. 17.54), and $\xi \sim a/\sqrt{z|t|}$ (Eq. 17.52), we then get $z(z|t|)^{\frac{d}{2}-1} \ll |t|$, which can be re-arranged to give

$$|t|^{4-d} \gg z^d. \tag{17.72}$$

Note that this condition will be satisfied as $t \to 0$, only if $d > 4$, which is the reason why $d = 4$ is the upper critical dimension of the Ising model.

However, it is very interesting to consider the $d < 4$ case as well! All the more so, since $d < 4$ is the more relevant case for the real world. In this case, the above condition is equivalent to

$$|t| \gg z^{\frac{d}{4-d}}. \qquad d < 4 \tag{17.73}$$

This is an interesting piece of information, if we combine it with what we already know about our mean field calculation so far: the mean field example calculation carried out in Section 17.4 is valid for *any* $T \geq T_c$, not just around T_c. Thus, we see that for $T \gg T_c$ and $d < 4$, the mean field theory is *quite valid* (cf. the importance of the Curie-Weiss law in that temperature regime, discussed in Section 16.1.3) for describing physical properties. As T is lowered, and gets close to T_c, predicted by the mean field theory, however, we see that the condition $|t| \gg z^{d/(4-d)}$ will begin to break down. This temperature regime where the critical fluctuations destroy the mean field picture is referred to as **the critical region**. It is a region of size $|t| \sim z^{\frac{d}{4-d}}$ up to a numerical factor. As the result of this kicking in of the critical fluctuations as a function of temperature, the mean field transition *is not realized* in real systems for $d < 4$. Instead, the transition temperature must be lowered further for the transition to occur. In some cases ($d = 1$), the transition is pushed all the way down to 0 K – the transition disappears! Spectroscopic tools such as neutron scattering and X-ray scattering tend to detect strong increase in short range fluctuations at the mean field transition temperature. Only when the temperature reaches the true transition temperature, long range fluctuations are detected.

Note that the size of the critical region vanishes as $z \to \infty$. So, even for a low dimensional material, if the coordination number becomes infinity, then the mean field solution is reliable. This property is exploited in theories such as the "dynamical mean field theory."

17.8 χ, re-revisited

Before closing this lecture, let us pay attention to a seemingly less important but still interesting topic. We have seen that the magnetic susceptibility diverges as $T \to T_c$. However, this behavior was obtained under the assumption that, as far as the condensed phase is concerned, we do not change the direction of the magnetization. When we discussed the critical exponent in Eq. 16.24, we were considering *smooth* fluctuation around the actual value of the magnetization, avoiding the discontinuity at $h = 0$ (see figure in page 7, LN 16). Of course, we *can* reverse the magnetization direction *in an abrupt manner*, if we apply a small field in the opposite direction to the magnetization[19], within the mean field theory. In this sense, it is also correct to say that the magnetic susceptibility[20] is infinite for all $T \leq T_c$, within the mean field theory. *This* susceptibility for $T < T_c$ can be calculated as follows. First, assume that $\overline{\langle \sigma_i \rangle} = 0$ for any i (mean field approximation). Second, let us calculate $\overline{\langle \sigma_1 \sigma_j \rangle}$. We get

$$\overline{\langle \sigma_1 \sigma_j \rangle} \;=\; m(T) > 0 \tag{17.74}$$

since both $\sigma_1 = 1$ and $\sigma_1 = -1$ act as an effective magnetic field to break the symmetry into m and $-m$, respectively, where we take m to be positive. Thus,

$$\Gamma_j \;=\; m(T) > 0. \qquad \text{for any } j\text{; from Eq. 17.6 and } \overline{\langle \sigma_i \rangle} = 0 \text{ for any } i \text{ (assumption)} \tag{17.75}$$

And, therefore, from Eq. 17.7,

$$\chi \;=\; \beta \mu_B N m. \tag{17.76}$$

Remember that this susceptibility is per-spin susceptibility. Thus, the fact that it is linear in N means that it is divergent. It follows that the magnetic field of $1/(\beta \mu_B N)$ is needed to reach the thermodynamic equilibrium magnetization m. This is the symmetry breaking field required to force the choice of the magnetization, and it is extremely small, $O(1/N)$.

Two comments. (1) The above argument can be generalized as long as $|\overline{\langle \sigma \rangle}_i| < m(T)$: for any partially magnetized state, we have $\chi = O(N)$. (2) Interesting as it is, *this* susceptibility *does not* diverge for $T < T_c$, if fluctuations are included. **However, $\chi(T \to T_c) \to \infty$ no matter what – within or beyond the mean field theory – as we saw in the rest of this lecture!**

[19] In reality, this process involves domain dynamics and hysteresis in reality, and also in theory, if we include fluctuations, but, here, we stick to the simple mean field picture.

[20] Note that the susceptibility has been defined as $\frac{\partial M}{\partial H}$ in this course. One may define the susceptibility as M/H, also. In the second definition, the susceptibility would be infinite for any magnetized ground state, whether or not the mean field theory is valid, or whether or not there is a hysteresis.