Summary of SI2510 Statistical Mechanics

Yashar Honarmandi
yasharh@kth.se

September 18, 2020

Abstract

This is a summary of SI2510 Statistical Mechanics.

Contents

1 Basic Concepts

Phase Transitions Landau introduced the concept that phase transitions are defined by spontaneous symmetry breaking.

Order Parameters An order parameter is a quantity that describes spontaneous symmetry breaking. It is zero in one phase and non-zero in another. An order parameter need not have a physical interpretation.

First- and Second-Order Phase Transitions Second-order phase transitions have continuous free energy and order parameter, whereas first-order phase transitions have discontinous order parameter.

Critical Exponents Many phenomena exhibit a behaviour of the form $|T - T_\mathrm{c}|^{-c}$ close to phase transitions. The exponent c is termed the critical exponent.

Some specific critical exponents that may appear are γ for χ, α for C, $-\beta$ for the order parameter, $-\delta\beta$ for the conjugate field h, γ for the corresponding susceptibility χ and ν for the correlation length.

Density Matrices The probability distribution is of the form

$$p_n = \frac{1}{\sum\limits_m P_m} P_n = \frac{1}{Z} P_n,$$

where the summation is performed over some set of states. We now introduce the density matrix

$$\rho = \frac{1}{Z} \sum_n P_n \left|n\middle\rangle\middle\langle n\right|,$$

yielding

$$\langle O \rangle = \sum_n p_n O_{nn} = \sum_n \frac{1}{Z} P_n \langle n|O|n\rangle = \frac{1}{Z} \sum_n \sum_m P_n \langle n|m\rangle \langle m|O|n\rangle = \frac{1}{Z} \sum_n \sum_m \langle n|\rho|m\rangle \langle m|O|n\rangle = \mathrm{tr}(O\rho).$$

As a side note, ρ takes the form

$$\rho = \frac{1}{Z} e^{-\beta K},$$

where K is the Hamiltonian in the canonical ensemble and $H - \mu N$ in the grand canonical ensemble. In these cases, the partition function Z may be computed according to

$$Z = \sum_m P_m = \sum_m e^{-\beta K_m} = \sum_m \langle m|e^{-\beta K}|m\rangle = \mathrm{tr}\left(e^{-\beta K}\right).$$

The Ising Model The Ising model is a simple model of magnets. In this model, a magnet is a collection of interacting spins on a lattice under the influence of an external field. Its generalized coordinates are σ_i, which may take the values ± 1, signifying a particular spin pointing up or down. The Hamiltonian is

$$\mathcal{H} = -J \sum_i \sum_{j = \mathrm{nn}(i)} \sigma_i \sigma_j - h \sum_i \sigma_i.$$

The inner summation is carried out over the nearest neighbours of site i in the Ising model, but is generally a sum over the whole lattice. The order parameter defining its phase transition for the case $J > 0$ is $m = \langle \sigma_i \rangle$.

This model will be used to demonstrate core concepts in the course.

Scaling Laws While we will use the Ising model as a reference in this summary, it has been found that very different systems display similar critical exponents. Such exponents are also found to obey so-called scaling laws. As an example, it may be shown that

$$\chi(C_H - C_M) = T \left(\left(\frac{\partial M}{\partial T} \right)^2 \right)_H.$$

This can only be satisfied if

$$\chi C_H > T\left(\left(\frac{\partial M}{\partial T}\right)^2\right)_H.$$

Introducing the critical exponent of the involved quantities, we must therefore have

$$-\gamma-\alpha \leq 2(\beta-1),\ \ \gamma+\alpha+2\beta \geq 2.$$

The intriguing part is that each side is in fact almost equal, meaning that we may treat them as such. Three other scaling laws are:

$$\beta(\delta-1)=\gamma,\ \ \nu d = 2-\alpha,\ \ \gamma=\nu(2-\eta).$$

We have also included the exponent $\Gamma(r, T_{\mathrm{c}}) \propto r^{-(d-2+\eta)}$.

Exact Solution in One Dimension To solve the Ising model in one dimension we will impose periodic boundary conditions $\sigma_N = \sigma_0$. The Hamiltonian may be written as

$$\mathcal{H} = -J\sum_{i=0}^{N-1}\sigma_i\sigma_{i+1} - \frac{1}{2}h\sum_{i=0}^{N-1}\sigma_i+\sigma_{i+1}.$$

The partition function is thus

$$Z=\sum_{\sigma_0=\pm 1}\cdots\sum_{\sigma_{N-1}=\pm 1} e^{\beta\left(J\sum\limits_{i=0}^{N-1}\sigma_i\sigma_{i+1}+\frac{1}{2}h\sum\limits_{i=0}^{N-1}\sigma_i+\sigma_{i+1}\right)}.$$

Introducing the transfer matrix with elements $t_{\sigma\sigma'} = e^{\beta J\sigma\sigma' + \frac{1}{2}h(\sigma+\sigma')}$ we have

$$Z=\sum_{\sigma_0=\pm 1}\cdots\sum_{\sigma_{N-1}=\pm 1}\prod_{i=0}^{N-1} t_{\sigma_i\sigma_{i+1}}.$$

Now consider some particular spin and perform the summation over this one first. We obtain

$$\sum_{\sigma_j=\pm 1} t_{\sigma_{j-1}\sigma_j}t_{\sigma_j\sigma_{j+1}} = t^2_{\sigma_{j-1}\sigma_{j+1}}.$$

This process is repeated until you obtain

$$Z=\mathrm{tr}\left(t^N\right).$$

The matrix representation of the transfer matrix is

$$t=\begin{bmatrix} e^{\beta(J+h)} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta(J-h)}\end{bmatrix}.$$

Its eigenvalues are the solutions to

$$\left(\lambda - e^{\beta(J+h)}\right)\left(\lambda - e^{\beta(J-h)}\right) - e^{-2\beta J} = 0,$$

and are given by

$$\begin{aligned}\lambda^2 - 2e^{\beta J}\cosh(\beta h)\lambda + 2\sinh(2\beta J) = 0,\\ \lambda_\pm &= e^{\beta J}\cosh(\beta h) \pm \sqrt{e^{2\beta J}\cosh^2(\beta h) - 2\sinh(2\beta J)}\\ &= e^{\beta J}\cosh(\beta h) \pm \sqrt{e^{2\beta J}\sinh^2(\beta h) + e^{-2\beta J}}.\end{aligned}$$

Now that we have the eigenvalues, we identify the partition function as

$$Z=\lambda_+^N + \lambda_-^N.$$

This can be further simplified to

$$Z = \lambda_+^N \left(1 + \left(\frac{\lambda_-}{\lambda_+}\right)^N\right) \approx \lambda_+^N.$$

Next, the free energy is given by

$$G = -k_\mathrm{B}T\left(N\ln(\lambda_+) + \ln\left(1 + \left(\frac{\lambda_-}{\lambda_+}\right)^N\right)\right).$$

The magnetization is given by

$$\begin{aligned}
m &= -\frac{1}{N}\left(\frac{\partial G}{\partial \beta h}\right)_T \\
&\approx \frac{e^{\beta J}\sinh(\beta h) + \frac{e^{2\beta J}\sinh(\beta h)\cosh(\beta h)}{\sqrt{e^{2\beta J}\sinh^2(\beta h)+e^{-2\beta J}}}}{e^{\beta J}\cosh(\beta h) + \sqrt{e^{2\beta J}\sinh^2(\beta h) + e^{-2\beta J}}} \\
&= \sinh(\beta h)\frac{1 + \frac{\cosh(\beta h)}{\sqrt{\sinh^2(\beta h)+e^{-4\beta J}}}}{\cosh(\beta h) + \sqrt{\sinh^2(\beta h) + e^{-4\beta J}}} \\
&= \frac{\sinh(\beta h)}{\sqrt{\sinh^2(\beta h) + e^{-4\beta J}}}.
\end{aligned}$$

If $h = 0$ there is no spontaneous magnetization. However, at low temperatures a very small field will produce saturation magnetization. This corresponds to a phase transition at $T = 0$.

Next consider the pair distribution function

$$g(j) = \langle \sigma_0 \sigma_j \rangle.$$

The error introduced by assuming uncorrelated spins, as will be done later, is

$$\Gamma(j) = \langle \sigma_i \sigma_{i+j} \rangle - \langle \sigma_i \rangle \langle \sigma_{i+j} \rangle.$$

In a general case with different couplings between spins and without an external field we have

$$\begin{aligned}
\langle \sigma_i \sigma_{i+j} \rangle &= \frac{1}{Z}\sum \sigma_i \sigma_{i+j} e^{\beta \sum\limits_{i=0}^{N-1} J_i \sigma_i \sigma_{i+1}} \\
&= \frac{1}{Z}\sum \sigma_i \sigma_{i+1}\sigma_{i+1}\sigma_{i+2}\ldots\sigma_{i+j-1}\sigma_{i+j} e^{\beta \sum\limits_{i=0}^{N-1} J_i \sigma_i \sigma_{i+1}} \\
&= \frac{1}{Z}\frac{\partial}{\partial \beta J_i}\cdots\frac{\partial}{\partial \beta J_{i+j}} Z.
\end{aligned}$$

Using the fact that

$$Z = 2\prod_{i=1}^{N} 2\cosh(\beta J_i)$$

we obtain

$$\langle \sigma_i \sigma_{i+j} \rangle = \prod_{k=i}^{i+j} \tanh(\beta J_k).$$

In one dimension there is no magnetization. In the case where all couplings are the same we obtain

$$\Gamma(j) = \tanh^j(\beta J) = e^{-\frac{j}{\xi}},$$

where we have introduced the correlation length

$$\xi = -\frac{1}{\ln(\tanh(\beta J))}.$$

Kadanoff Block Spins Consider the Ising model on a d-dimensional hypercubic lattice with lattice constant a. Divide the lattice into blocks with l spins in each direction, meaning that each block contains l^d lattice sites. These blocks then also form a lattice with lattice constant la. If we assume that $la \ll \xi$, most of the spins in a block are correlated, allowing us to introduce a scaling hypothesis. Denoting the total spin of some block as $S_{I,\text{tot}}$ we introduce the new variables

$$S_I = l^{-x_S} S_{I,\text{tot}}.$$

Assume now that the Hamiltonian looks the same when expressed in terms of the block spins. The free energy should then be unchanged by the choice of new variables. Introducing the variable

$$t = \frac{T - T_\text{c}}{T_\text{c}},$$

we expect both this parameter and h to scale with the transformation of variables according to

$$t \to tl^{x_t}, \ h \to hl^{x_h}.$$

We now expect the free energy to scale according to

$$G = Ng(t, h) = N_l g(t_l, h_l) = \frac{N}{l^d} g(t_l, h_l).$$

The correlation length is also expected to scale as

$$\xi \to \frac{\xi}{l}.$$

This implies that in our new variables, the reduced temperature is increased. Similarly the field is given by

$$h = h \sum S_i = h \sum S_{I,\text{tot}} = hl^{x_S} \sum S_I,$$

and has thus increased. The scaling hypothesis is now

$$g(t, h) = l^{-d} g(tl^{x_t}, hl^{x_h}).$$

In other words, the free energy per particle is a homogenous function of t and h.

Near $t = 0$ the correlation length is the only characteristic length scale. As it diverges in this limit, the system is invariant under scale transformations. This implies that all thermodynamic functions are homogenous, somehow.

Let us now derive some critical exponents in this way. For instance, we compute the order parameter as

$$m(t, h) = \frac{\partial}{\partial h}\left(l^{-d} g(tl^{x_t}, hl^{x_h})\right) = l^{x_h - d} m(tl^{x_t}, hl^{x_h}).$$

This is true for any scaling factor according to the hypothesis, meaning that we may choose $l = |t|^{-\frac{1}{x_t}}$ (and $h = 0$, I believe), which one can squint to see as

$$\beta = -\frac{y_h - d}{y_t}.$$

To determine δ, we choose $t = 0, \ l = |h|^{-\frac{1}{y_h}}$ to obtain

$$\delta = -\frac{y_h}{y_h - d}.$$

To determine γ we instead use

$$\chi(t, h) = \frac{\partial}{\partial h}\left(l^{x_h - d} m(tl^{x_t}, hl^{x_h})\right) = l^{2x_h - d} \chi(tl^{x_t}, hl^{x_h}).$$

Choosing the scale factor $l = |t|$ nets

$$\gamma = \frac{2y_h - d}{y_t}.$$

To determine α, we use

$$C(t,h) = -t\frac{\partial^2 g}{\partial t^2} = l^{2y_h - d} C(tl^{x_t}, hl^{x_h}).$$

This yields

$$\alpha = \frac{2y_t - d}{y_t}.$$

The scaling assumptions must be verified experimentally. For one method, consider the order parameter. Choosing $l = |t|^{-\frac{1}{y_t}}$ we obtain

$$|t|^{-\beta} m(t,h) = m(\pm 1, h|t|^{-\Delta}),$$

where Δ is the gap exponent, given by $\frac{y_h}{y_t}$. In other words, plotting $|t|^{-\beta} m(t,h)$ against $h|t|^{-\Delta}$, one should see different curves depending on the sign of t.

We now turn to the correlation length. We expect

$$\xi(t,h) = l\xi(tl^{x_t}, hl^{x_h}).$$

Setting $l = |t|^{-\frac{1}{x_t}}$ we find

$$\nu = \frac{1}{x_t}.$$

This yields

$$\nu d = 2 - \alpha.$$

Next, the correlation function. We have

$$\begin{aligned}
\Gamma(r_l, t_l) &= \langle S_I S_J \rangle - \langle S_I \rangle \langle S_J \rangle \\
&= l^{-2x_h} \sum_{i \in I, j \in j} \langle S_i S_j \rangle - \langle S_i \rangle \langle S_j \rangle \\
&= l^{d-2x_h} (\langle S_i S_j \rangle - \langle S_i \rangle \langle S_j \rangle) \\
&= l^{d-2x_h} \Gamma(r,t),
\end{aligned}$$

implying

$$\Gamma(r,t) = l^{2x_h - d} \Gamma\left(\frac{r}{l}, l^{x_t}\right).$$

Setting $l = r$ and $t = 0$ we obtain

$$\Gamma(r,0) = r^{2x_h - d} \Gamma(1,0).$$

By the order-of-magnitude estimate $\Gamma \propto r^{d-2+\nu}$ for the correlation function we thus have

$$2 - \eta = 2x_h - d = \frac{\gamma}{\nu}$$

and finally

$$\gamma = \nu(2 - \eta).$$

2 Mean-Field Theory

The Idea The idea between mean-field theory is to separate the Hamiltonian by for any component of the system replace the ones with which it is interacting by an expectation value. Alternatively one may also replace the interactions themselves by mean values. This way a partition function may be computed more easily.

One usually requires self-consistency, meaning that expectation values should be computable from this partition function. This implies that the partition function implicitly gives the expectation values which were used to separate the Hamiltonian.

Mean-Field Theory of the Ising Model To construct a mean-field theory for the Ising model, we write the Hamiltonian as

$$\mathcal{H} = -\sum_i \sigma_i \left(J \sum_{j=\text{nn}(i)} \sigma_j + h \right).$$

We note that the replacement $\sigma_j \to m$ will make the Hamiltonian separable, hence we propose the mean-field Hamiltonian

$$\mathcal{H} = -\sum_i \sigma_i \left(J \sum_{j=\text{nn}(i)} m + h \right) = -\sum_i \sigma_i \left(qJm + h \right) = -h' \sum_i \sigma_i$$

where we have introduced the number q of nearest neighbours to any one site and the effective field

$$h' = h + qJm.$$

Using this Hamiltonian, the partition function is now given by

$$\begin{aligned}
Z &= \operatorname{tr}\left(e^{-\beta\mathcal{H}}\right) = \sum_{\{\sigma\}} e^{\beta h' \sum_i \sigma_i} = \left(\sum_{\sigma=\pm 1} e^{\beta h' \sigma} \right)^N = 2^N \cosh^N\left(\beta h'\right), \\
m &= \frac{1}{N} \left\langle \sum_i \sigma_i \right\rangle = \frac{1}{NZ} \sum \sum_i \sigma_i e^{\beta h' \sum_j \sigma_j} = \frac{1}{N} \frac{\mathrm{d}}{\mathrm{d}\beta h'} \ln\left(2^N \cosh^N\left(\beta h'\right)\right) = \tanh\left(\beta h'\right).
\end{aligned}$$

Note that this implies that all spins are expected to point in the same direction. Baked into the process there is a specific idea of the structure of the solution, and it is therefore important to make a good guess about this. We proceed with the ferromagnetic case, where the guess is good, and obtain

$$m = \tanh(\beta\left(h + qJm\right)).$$

This equation can be solved graphically to yield the magnetization, but we will discuss it qualitatively here. The expression above is anti-symmetric in m, meaning we may consider $m \geq 0$. Depending on the parameters, the number of solutions is between one and three. In the case where $h = 0$, one solution is $m = 0$, and two other solutions may be found at $m = \pm m_0$. At low temperatures the right-hand side approaches ± 1, yielding $m_0 = 1$. For $h = 0$ solutions are only found where the right-hand side grows faster than the left-hand side at zero. This is satisfied if

$$\beta q J > 1.$$

This defines the critical temperature

$$T_\text{c} = \frac{qJ}{k_\text{B}},$$

above which no non-zero solutions are found.

As the temperature approaches the critical temperature, above which no spontaneous magnetism is found, m_0 is small (but non-zero) and we obtain

$$\begin{aligned}
m_0 &\approx \beta q J m_0 - \frac{1}{3}\left(\beta q J m_0\right)^3, \\
\left(\frac{T_\text{c}}{T}\right)^3 m_0^2 &= 3\left(\frac{T_\text{c}}{T} - 1\right), \\
m_0 &= \left(\frac{T}{T_\text{c}}\right)^{\frac{3}{2}} \sqrt{3\left(\frac{T_\text{c}}{T} - 1\right)} \\
&= \sqrt{\frac{3}{T_\text{c}}\left(\frac{T}{T_\text{c}}\right)^2 \left(T_\text{c} - T\right)}.
\end{aligned}$$

This identifies the critical exponent $\beta = 1$. One of the uses of mean-field theory is determining asymptotic behaviour close to phase transitions.

One more thing should be mentioned, namely the assertion that there actually is spontaneous magnetization. After all, if three solutions are possible, who is to say that one of the non-zero ones are found? To do this, we consider the entropy of an ideal paramagnet, which can be shown to be

$$S = -Nk_{\text{B}}\left(\frac{1-m}{2}\ln\left(\frac{1-m}{2}\right) + \frac{1+m}{2}\ln\left(\frac{1+m}{2}\right)\right).$$

We see that the non-zero solutions maximize entropy - a relief.

Critical Behaviour Using the mean-field result, we may now study other quantities close to the phase transition. The susceptibility is given by

$$\chi = \left(\frac{\partial m}{\partial h}\right)_T.$$

The implicit equation for the magnetization yields

$$\begin{aligned}
\chi &= \frac{\beta}{\cosh^2(\beta(qJm+h))}(qJ\chi+1),\\
\chi &= \frac{\beta}{\cosh^2(\beta(qJm+h)) - \beta qJ} = \frac{1}{k_{\text{B}}\left(T\cosh^2(\beta(qJm+h)) - \frac{qJ}{k_{\text{B}}}\right)}.
\end{aligned}$$

Introducing the critical temperature, we write this as

$$\chi = \frac{1}{k_{\text{B}}\left(T\cosh^2(\beta(qJm+h)) - T_{\text{c}}\right)}.$$

In particular, for $h = 0$ and temperatures above T_{c}, where there is no magnetization, we obtain

$$\chi = \frac{1}{k_{\text{B}}\left(T - T_{\text{c}}\right)}.$$

When approaching the phase transition from below for $h = 0$, we use the asymptotic expression for the magnetization to obtain

$$\begin{aligned}
\chi &= \frac{1}{k_{\text{B}}\left(T\cosh^2\left(\beta qJ\sqrt{\frac{3}{T_{\text{c}}}\left(\frac{T}{T_{\text{c}}}\right)^2(T_{\text{c}}-T)}\right) - T_{\text{c}}\right)}\\
&= \frac{1}{k_{\text{B}}\left(T\cosh^2\left(\sqrt{3\left(1-\frac{T}{T_{\text{c}}}\right)}\right) - T_{\text{c}}\right)}\\
&= \frac{1}{k_{\text{B}}T_{\text{c}}\left(\frac{T}{T_{\text{c}}}\left(1+3\left(1-\frac{T}{T_{\text{c}}}\right)\right) - 1\right)}\\
&= \frac{1}{k_{\text{B}}T_{\text{c}}\left(4\frac{T}{T_{\text{c}}} - 3\left(\frac{T}{T_{\text{c}}}\right)^2 - 1\right)}\\
&= \frac{1}{k_{\text{B}}T_{\text{c}}\left(3\frac{T}{T_{\text{c}}}\left(1-\frac{T}{T_{\text{c}}}\right) + \frac{T}{T_{\text{c}}} - 1\right)}\\
&\approx \frac{1}{k_{\text{B}}T_{\text{c}}\left(3\left(1-\frac{T}{T_{\text{c}}}\right) + \frac{T}{T_{\text{c}}} - 1\right)}\\
&= \frac{1}{2k_{\text{B}}\left(T_{\text{c}} - T\right)}.
\end{aligned}$$

The Bragg-Williams Approximation The Bragg-Williams approximation to mean-field theory starts with constructing the availability in terms of the order parameter. In the case of the Ising model, we introduce the numbers $N_\pm$ of spins with values ± 1. Furthermore, we introduce the numbers $N_{\pm\pm}$ of spin pairs of any kind. The Hamiltonian is thus

$$\mathcal{H} = -J(N_{++} + N_{--} - N_{+-}) - h(N_+ - N_-).$$

Treating the spins as independent allows us to write

$$S = -k_{\mathrm{B}} \left(N_+ \ln(N_+) + N_- \ln(N_-)\right).$$

The number of pairs is given by

$$N_{\pm\pm} = \frac{qN_\pm^2}{2N},\ N_{+-} = \frac{qN_+N_-}{N}.$$

To proceed, we re-express the spin numbers in terms of the order parameter by using $N = N_+ + N_-$ and $\sigma = N_+ - N_-$ to obtain

$$N_+ = \frac{1}{2}N(1+m),\ N_- = \frac{1}{2}N(1-m).$$

The Hamiltonian is now given by

$$\begin{aligned}
\mathcal{H} &= -\frac{qJ}{2N}(N_+^2 + N_-^2 - 2N_+N_-) - Nhm \\
&= -\frac{qJN}{8}((1+m)^2 + (1-m)^2 - 2(1+m)(1-m)) - Nhm \\
&= -\frac{qJN}{2}m^2 - Nhm,
\end{aligned}$$

and the free energy is somehow

$$\begin{aligned}
G(h,T) &= \mathcal{H} - TS \\
&= -\frac{qJN}{2}m^2 - Nhm + \frac{1}{2}Nk_{\mathrm{B}}T\left((1+m)\ln\left(\frac{1}{2}N(1+m)\right) + (1-m)\ln\left(\frac{1}{2}N(1-m)\right)\right).
\end{aligned}$$

Minimizing it with respect to the order parameter yields

$$\begin{aligned}
&- qJNm - Nh + \frac{1}{2}Nk_{\mathrm{B}}T\left(\ln\left(\frac{1}{2}N(1+m)\right) + 1 - \ln\left(\frac{1}{2}N(1-m)\right) - 1\right) = 0, \\
&- qJm - h + \frac{1}{2}k_{\mathrm{B}}T\ln\left(\frac{1+m}{1-m}\right) = 0.
\end{aligned}$$

Its solution is

$$\begin{aligned}
\frac{1+m}{1-m} &= e^{2\beta(qJm+h)}, \\
1+m &= (1-m)e^{2\beta(qJm+h)}, \\
m(1+e^{2\beta(qJm+h)}) &= e^{2\beta(qJm+h)} - 1, \\
m &= \tanh(\beta(qJm+h)),
\end{aligned}$$

as expected.

Inaccuracies of Mean-Field Theories The mean-field arguments predict the existence of a phase transition, but this cannot be the case in one dimension. To see this, consider a chain in its ground state and the set of excitations that flips all spins to the right of some spin k. The change in energy is $2J$, and the number of possible states corresponding to this energy is $N-1$, hence the free energy changes by $2J - k_{\mathrm{B}}\ln(N-1)$. For large N such states are thus always preferable. Their removal of translation invariance implies that there is no magnetization, in contradiction of the mean-field results.

A slightly better result is obtained for a $N \times N$ lattice in two dimensions. The set of excitations now consists of excitations that split the system in two distinct magnetic domains. Any particular excitation is described by a chain running through the bonds. Each segment crosses one bond, and the energy change due to the excitation is $2LJ$, where L is the number of segments. The typical length is $2N$. The next segment may always be placed in at least two sites, neglecting the boundaries. Including the N possible starting points, the multiplicity of the chain is $N2^L$, and the free energy change of the excitation is

$$\Delta G = 4NJ - k_{\mathrm{B}}T\ln\left(2^{2N}N\right).$$

The phase transition occurs when this energy change is negative, i.e. when

$$k_{\mathrm{B}}T\ln\left(2^{2N}N\right) > 4NJ,\ k_{\mathrm{B}}T(2N\ln(2) + \ln(N)) > 4NJ,\ T < T_{\mathrm{c}} \approx \frac{2J}{k_{\mathrm{B}}\ln(2)},$$

which is decently close to the analytically obtained results.

Antiferromagnetism The case of $J < 0$ is another interesting case, and gives rise to some interesting phases. One phenomenon which may occur is frustration, where the lattice structure is such that for any group of spins, the state of lowest energy is degenerate and not such that all interactions are beneficial. Assuming this not to be the case, we find that the ground state corresponds to an ordered phase such that two lattice translations combined leave the system invariant. This symmetry is broken by the phase transition, hence we introduce the order parameter

$$m = \frac{1}{N}\sum(-1)^j\sigma_j,$$

which may be written as

$$m = \frac{1}{2}(m_A - m_B),$$

where the lattice has been divided into two sublattices.

The mean-field Hamiltonian is

$$\mathcal{H} = \mathcal{H}_A + \mathcal{H}_B = -Jq\left(m_B\sum_A\sigma_i + m_A\sum_B\sigma_i\right),$$

and by the same methods as previously, the mean-field partition function is

$$Z = 2^N\cosh^{\frac{1}{2}N}(\beta Jqm_B)\cosh^{\frac{1}{2}N}(\beta Jqm_A).$$

From this we obtain

$$m_A = \tanh(\beta Jqm_B).$$

By construction we have $m_A = -m_B$, which yields

$$m_A = -\tanh(\beta Jqm_A).$$

This is the same as we obtained for the ferromagnetic case, and we may immediately identify the critical temperature

$$T_{\mathrm{c}} = -\frac{qJ}{k_{\mathrm{B}}}.$$

3 Landau Theory

The Idea Landau's theory is a general theory of phase transitions. The core idea is to series expand the free energy in terms of the order parameter close to the phase transition. This isn't really valid, but this method works nevertheless.

The general form of the series expansion is

$$G(m,T) = a_0(T) + \sum_i\frac{1}{i}a_i(T)m^i,$$

where certain terms may be zero depending on the symmetry of the system. We assume the order parameter to be finite at equilibrium, meaning that the free energy must be bounded from below. One generally truncates this sum to make it possible to handle, and the free energy being bounded is guaranteed by the highest-order term to be even in m and positive.

Landau Theory of the Ising Model For the Ising model we expect the system to be invariant with respect to flipping all spins, hence we have the series expansion

$$G(m,T) = a_0(T) + \sum_i \frac{1}{2i} a_{2i}(T) m^{2i}.$$

At $T = 0$ you will have $|\mathbf{m}| = 1$.

Suppose now that as the temperature is lowered, a_2 is the first coefficient to change sign (at least one coefficient must do this in order for a minimum to exist). Close to the temperature T_c at which it changes sign, it may be linearized as

$$a_2(T) = a_{2,0}(T - T_\mathrm{c}).$$

At equilibrium we have

$$\frac{\partial G}{\partial m} = \sum_i a_{2i}(T) m^{2i-1} = 0.$$

Assuming that other coefficients are approximately constant close to the transition temperature and approaching T_c from below, where m is small but non-zero, we have

$$a_{2,0}(T - T_\mathrm{c}) + \sum_{i>2} a_{2i}(T_\mathrm{c}) m^{2i-2} = 0.$$

We truncate this sum at order 2 in m to obtain

$$m = \sqrt{\frac{a_{2,0}}{a_4(T_\mathrm{c})}(T_\mathrm{c} - T)},$$

which reproduces the correct critical exponent.

We see that the series expansion and our assumption about which coefficients change sign produce a theory with a second-order phase transition. This is profound, and both the nature of the phase transition and the critical exponent are general. This is part of the power of Landau theory.

Next we study the heat capacity

$$C = T\left(\frac{\partial S}{\partial T}\right)_h.$$

We have

$$S = -\frac{\partial G}{\partial T} = -\frac{\mathrm{d}a_0}{\mathrm{d}T} - \sum_i \frac{1}{2i}\left(\frac{\mathrm{d}a_{2i}}{\mathrm{d}T} m^{2i} + a_{2i}\frac{\mathrm{d}m^{2i}}{\mathrm{d}T}\right).$$

Close to and below the critical temperature we have

$$\begin{aligned}
C &= T\left(-\frac{\mathrm{d}^2 a_0}{\mathrm{d}T^2} - \sum_i \frac{1}{2i}\left(\frac{\mathrm{d}^2 a_{2i}}{\mathrm{d}T^2} m^{2i} + \frac{\mathrm{d}a_{2i}}{\mathrm{d}T}\frac{\mathrm{d}m^{2i}}{\mathrm{d}T} + \frac{\mathrm{d}a_{2i}}{\mathrm{d}T}\frac{\mathrm{d}m^{2i}}{\mathrm{d}T} + a_{2i}\frac{\mathrm{d}^2 m^{2i}}{\mathrm{d}T^2}\right)\right) \\
&= T\left(-\frac{\mathrm{d}^2 a_0}{\mathrm{d}T^2} - \sum_i \frac{1}{2i}\left(\frac{\mathrm{d}^2 a_{2i}}{\mathrm{d}T^2} m^{2i} + 2\frac{\mathrm{d}a_{2i}}{\mathrm{d}T}\frac{\mathrm{d}m^{2i}}{\mathrm{d}T} + a_{2i}\frac{\mathrm{d}^2 m^{2i}}{\mathrm{d}T^2}\right)\right).
\end{aligned}$$

We have

$$\begin{aligned}
\frac{\mathrm{d}m^{2i}}{\mathrm{d}T} &= i m^{2(i-1)}\frac{\mathrm{d}m^2}{\mathrm{d}T} = -i m^{2(i-1)}\frac{a_{2,0}}{a_4(T_\mathrm{c})}, \\
\frac{\mathrm{d}^2 m^{2i}}{\mathrm{d}T^2} &= i(i-1) m^{2(i-2)}\frac{a_{2,0}^2}{a_4^2(T_\mathrm{c})},
\end{aligned}$$

and thus

$$
\begin{aligned}
C =& T\left(-\frac{\mathrm{d}^2 a_0}{\mathrm{d}T^2} - \sum_{i=1}^{\infty}\frac{1}{2i}\left(\frac{\mathrm{d}^2 a_{2i}}{\mathrm{d}T^2}m^{2i} - 2i\frac{\mathrm{d}a_{2i}}{\mathrm{d}T}\frac{a_{2,0}}{a_4(T_\mathrm{c})}m^{2(i-1)} + a_{2i}i(i-1)m^{2(i-2)}\frac{a_{2,0}^2}{a_4^2(T_\mathrm{c})}\right)\right)\\
=& -T\frac{\mathrm{d}^2 a_0}{\mathrm{d}T^2} + T\frac{a_{2,0}^2}{a_4(T_\mathrm{c})} - T\sum_{i=2}^{\infty}\frac{1}{2i}\left(\frac{\mathrm{d}^2 a_{2i}}{\mathrm{d}T^2}m^{2i} - 2i\frac{\mathrm{d}a_{2i}}{\mathrm{d}T}\frac{a_{2,0}}{a_4(T_\mathrm{c})}m^{2(i-1)} + a_{2i}i(i-1)m^{2(i-2)}\frac{a_{2,0}^2}{a_4^2(T_\mathrm{c})}\right)\\
=& -T\frac{\mathrm{d}^2 a_0}{\mathrm{d}T^2} + T\frac{a_{2,0}^2}{a_4(T_\mathrm{c})} - \frac{1}{4}T\left(\frac{\mathrm{d}^2 a_4}{\mathrm{d}T^2}m^4 - 4\frac{\mathrm{d}a_4}{\mathrm{d}T}\frac{a_{2,0}}{a_4(T_\mathrm{c})}m^2 + 2a_4\frac{a_{2,0}^2}{a_4^2(T_\mathrm{c})}\right)\\
&- T\sum_{i=3}^{\infty}\frac{1}{2i}\left(\frac{\mathrm{d}^2 a_{2i}}{\mathrm{d}T^2}m^{2i} - 2i\frac{\mathrm{d}a_{2i}}{\mathrm{d}T}\frac{a_{2,0}}{a_4(T_\mathrm{c})}m^{2(i-1)} + a_{2i}i(i-1)m^{2(i-2)}\frac{a_{2,0}^2}{a_4^2(T_\mathrm{c})}\right)\\
\approx& -T\frac{\mathrm{d}^2 a_0}{\mathrm{d}T^2} + T\frac{a_{2,0}^2}{2a_4(T_\mathrm{c})},
\end{aligned}
$$

where we have ignored terms containing the magnetization. Above the critical temperature the magnetization is instead identically zero, netting

$$C = -T\frac{\mathrm{d}^2 a_0}{\mathrm{d}T^2}.$$

Suppose instead that a_4 is the first coefficient to change sign. In this case a discontinuous step in the order parameter might occur. To show that such a step exists, we need to show that $a_2(T_\mathrm{c}) > 0$. We investigate this by comparing $G(m_0, T_\mathrm{c})$ to $G(0, T_\mathrm{c})$, where m_0 is the magnetization at the minimum. The phase transition occurs when the two are equal, i.e. when

$$G(m_0, T_\mathrm{c}) - G(0, T_\mathrm{c}) = \sum_i \frac{1}{2i}a_{2i}(T_\mathrm{c})m_0^{2i} = 0.$$

The magnetization corresponds to a minimum of G, and thus satisfies

$$\sum_i a_{2i}(T_\mathrm{c})m_0^{2i-1} = 0.$$

Ignoring terms above order 6 we have

$$a_2(T_\mathrm{c})m_0 + a_4(T_\mathrm{c})m_0^3 + a_6(T_\mathrm{c})m_0^5 = 0,\ \frac{1}{2}a_2(T_\mathrm{c})m_0^2 + \frac{1}{4}a_4(T_\mathrm{c})m_0^4 + \frac{1}{6}a_4(T_\mathrm{c})m_0^6 = 0.$$

The non-trivial value satisfies

$$a_2(T_\mathrm{c}) + a_4(T_\mathrm{c})m_0^2 + a_6(T_\mathrm{c})m_0^4 = 0,\ \frac{1}{2}a_2(T_\mathrm{c}) + \frac{1}{4}a_4(T_\mathrm{c})m_0^2 + \frac{1}{6}a_4(T_\mathrm{c})m_0^4 = 0.$$

Combining the equation nets

$$
\begin{aligned}
\frac{1}{2}a_4(T_\mathrm{c})m_0^2 + \frac{2}{3}a_6(T_\mathrm{c})m_0^4 &= 0,\\
\frac{1}{2}a_4(T_\mathrm{c}) + \frac{2}{3}a_6(T_\mathrm{c})m_0^2 &= 0,\\
m_0^2 &= -\frac{3a_4(T_\mathrm{c})}{4a_6(T_\mathrm{c})}.
\end{aligned}
$$

For this to work, we must have $a_6(T_\mathrm{c}) > 0$ to keep the global minimum at finite magnetization and $a_4(T_\mathrm{c}) < 0$ per our assumption of the existence of a local minimum, making the magnetization real. Inserting this into a previous expression yields

$$a_2(T_\mathrm{c}) - a_4(T_\mathrm{c})\frac{3a_4(T_\mathrm{c})}{4a_6(T_\mathrm{c})} + a_6(T_\mathrm{c})\frac{9a_4^2(T_\mathrm{c})}{16a_6^2(T_\mathrm{c})} = a_2(T_\mathrm{c}) - \frac{3}{4}\frac{a_4^2(T_\mathrm{c})}{a_6(T_\mathrm{c})} + \frac{9}{16}\frac{a_4^2(T_\mathrm{c})}{a_6(T_\mathrm{c})} = a_2(T_\mathrm{c}) - \frac{3}{16}\frac{a_4^2(T_\mathrm{c})}{a_6(T_\mathrm{c})} = 0,$$

and thus

$$a_2(T_\mathrm{c}) = \frac{3}{16}\frac{a_4^2(T_\mathrm{c})}{a_6(T_\mathrm{c})} > 0,$$

as we wanted to show.

Non-Symmetric Cases Suppose we have some case where this symmetry does not hold. Then we would instead use the series expansion

$$G(m,T) = a_0(T) + \sum_i \frac{1}{i} a_i(T) m^i.$$

It might be of interest to remove linear terms. This can be done by introducing a new order parameter $\tilde{m} = m + \Delta$ (the tilde will be omitted from now) where Δ is chosen appropriately so that

$$G(m,T) = a_0(T) + \sum_{i=2} \frac{1}{i} a_i(T) m^i.$$

The coefficients have implicitly been modified as well. Truncating the sum at a_4, we have

$$a_2(T_\mathrm{c}) m_0 + a_3(T_\mathrm{c}) m_0^2 + a_4(T_\mathrm{c}) m_0^3 = 0.$$

In addition, at the transition point we have

$$\frac{1}{2} a_2(T_\mathrm{c}) m_0^2 + \frac{1}{3} a_3(T_\mathrm{c}) m_0^3 + \frac{1}{4} a_4(T_\mathrm{c}) m_0^4 = 0.$$

The non-zero solution satisfies

$$\frac{1}{3} a_3(T_\mathrm{c}) + \frac{1}{2} a_4(T_\mathrm{c}) m_0 = 0,\ m_0 = -\frac{2}{3} \frac{a_3(T_\mathrm{c})}{a_4(T_\mathrm{c})}$$

and

$$a_2(T_\mathrm{c}) - \frac{2}{3} \frac{a_3^2(T_\mathrm{c})}{a_4(T_\mathrm{c})} + \frac{4}{9} \frac{a_3^2(T_\mathrm{c})}{a_4(T_\mathrm{c})} = 0,\ a_2(T_\mathrm{c}) = \frac{2}{9} \frac{a_3^2(T_\mathrm{c})}{a_4(T_\mathrm{c})}$$

Ginzburg-Landau Theory Landau theory characterizes a system in terms of a single order parameter. Ginzburg-Landau theory instead characterizes the system in terms of a field $m(\mathbf{r})$. This field could be thought of as at any particular point describing the order parameter when calculated based only on the vicinity of that point. In other words, it is a high-resolution version of Landau theory.

The order parameter extremizes the free energy, which in this theory is given by

$$F = \int \mathrm{d}^d\mathbf{x}\, a_0(T) + \sum_i \frac{1}{2i} a_{2i}(T) m^{2i} + \frac{1}{2} f (\vec{\boldsymbol{\nabla}} m)^2.$$

The series expansion generalize Landau theory, whereas the last term is a simple extra term that gives non-trivial behaviour of m. We assume $f > 0$ as fluctuations should add to the free energy.

The corresponding extensive variable (the external field in the Ising model) is in this theory given by

$$h = \frac{\delta F}{\delta m}.$$

We have

$$\delta F = \int \mathrm{d}^d\mathbf{x}\, \delta m \sum_{2i} a_{2i}(T) m^{2i-1} + f \vec{\boldsymbol{\nabla}}(\delta m) \cdot \vec{\boldsymbol{\nabla}} m.$$

Fixing boundary conditions and integrating by parts yields

$$\delta F = \int \mathrm{d}^d\mathbf{x}\, \delta m \left(\sum_i a_{2i}(T) m^{2i-1} - f \nabla^2 m \right),$$

and finally

$$h = \sum_i a_{2i}(T) m^{2i-1} - f \nabla^2 m.$$

Truncating the sum yields the result

$$m_0^2 = -\frac{a_2(T)}{a_4(T)}$$

below the critical temperature for a second-order transition.

Suppose we add some perturbation $h\delta(\mathbf{x})$ from $h = 0$, which changes the field to $m_0 + \phi$. T, where m_0 is a constant. Truncating the sum at $i = 2$ we obtain

$$h_0\delta(\mathbf{x}) = a_2(T)(m_0 + \phi) + a_4(T)(m_0 + \phi)^3 - f\nabla^2 m_0 - f\nabla^2\phi.$$

Neglecting higher-order terms in ϕ we obtain

$$\nabla^2\phi - \frac{a_2(T)}{f}\phi - \frac{a_2(T)}{f}m_0 - \frac{3a_4(T)m_0^2}{f}\phi - \frac{a_4(T)}{f}m_0^3 = -\frac{h_0}{f}\delta(\mathbf{x}).$$

This simplifies to

$$\nabla^2\phi + \frac{2a_2(T)}{f}\phi = -\frac{h_0}{f}\delta(\mathbf{x})$$

below the critical temperature and

$$\nabla^2\phi - \frac{a_2(T)}{f}\phi = -\frac{h_0}{f}\delta(\mathbf{x})$$

above the critical temperature. The solution to these equations is

$$\phi = \frac{h_0}{4\pi f}\frac{e^{-\frac{r}{\xi}}}{r},$$

where

$$\xi = \begin{cases} \sqrt{\frac{f}{a_2(T)}}, & T > T_\mathrm{c}, \\ \sqrt{-\frac{f}{2a_2(T)}}, & T < T_\mathrm{c}. \end{cases}$$

ξ is the correlation length, and according to the linearization $a_2 = a_{2,0}(T - T_\mathrm{c})$ it diverges when approaching the phase transition.

Next, if we add a term

$$-\int \mathrm{d}^3\mathbf{x}\, mh$$

to the Hamiltonian, we obtain

$$\langle m \rangle = \frac{\operatorname{tr}\left(me^{-\beta\left(\mathcal{H} - \int \mathrm{d}^3\mathbf{x} mh\right)}\right)}{\operatorname{tr}\left(e^{-\beta\left(\mathcal{H} - \int \mathrm{d}^3\mathbf{x} mh\right)}\right)}.$$

This implies

$$\begin{aligned}\frac{\delta\langle m\rangle}{\delta h(0)} &= \frac{\beta\operatorname{tr}\left(m(\mathbf{0})me^{-\beta\left(\mathcal{H}-\int \mathrm{d}^3\mathbf{x}mh\right)}\right)\operatorname{tr}\left(e^{-\beta\left(\mathcal{H}-\int \mathrm{d}^3\mathbf{x}mh\right)}\right) - \beta\operatorname{tr}\left(me^{-\beta\left(\mathcal{H}-\int \mathrm{d}^3\mathbf{x}mh\right)}\right)\operatorname{tr}\left(m(\mathbf{0})e^{-\beta\left(\mathcal{H}-\int \mathrm{d}^3\mathbf{x}mh\right)}\right)}{\left(\operatorname{tr}\left(e^{-\beta\left(\mathcal{H}-\int \mathrm{d}^3\mathbf{x}mh\right)}\right)\right)^2} \\ &= \beta(\langle m(\mathbf{0})m\rangle - \langle m(\mathbf{0})\rangle\langle m\rangle) \\ &= \beta\Gamma(\mathbf{r}).\end{aligned}$$

Somehow this is supposed to be equal to $\frac{\phi}{h_0}$, hence ϕ is an order parameter correlation function. The susceptibility is given by

$$\chi = \int \mathrm{d}^3\mathbf{x}\,\beta\Gamma(\mathbf{r}),$$

and this implies that the mean-field result $\chi \propto |T_\mathrm{c} - T|^{-1}$ is obtained.

The Ginzburg Criterion The Ginzburg criterion is a self-consistency criterion for mean-field or Landau theories.

To obtain it, we generalize to d dimensions. In such cases the solution above does not hold, but we may still use the order-of-magnitude approximation

$$\phi = \frac{e^{-\frac{r}{\xi}}}{r^{d-2}}.$$

We would like to crudely approximate the correlation function at large distances. This is expected to be valid if

$$\frac{\int\limits_{\Omega(\xi)} \mathrm{d}^d\mathbf{x}\,\langle m(\mathbf{0})m\rangle - \langle m(\mathbf{0})\rangle\,\langle m\rangle}{\int\limits_{\Omega(\xi)} \mathrm{d}^d\mathbf{x}\,m_0^2} \ll 1,$$

where $\Omega(\xi)$ is the d-dimensional hypersphere of radius ξ. I believe this is the Ginzburg criterion

We will now use the Ginzburg criterion to try to estimate the dimensionality for which Landau theory correctly predicts the critical behaviour. Close to the critical point we should have $m_0^2 \approx |T_\mathrm{c} - T|^{2\beta}$ for some β. Using ϕ as an estimate of the correlation function we have

$$\frac{\int\limits_{\Omega(\xi)} \mathrm{d}^d\mathbf{x}\,\frac{e^{-\frac{r}{\xi}}}{r^{d-2}}}{\int\limits_{\Omega(\xi)} \mathrm{d}^d\mathbf{x}\,|T_\mathrm{c} - T|^{2\beta}} \ll 1.$$

Computing this in spherical coordinates yields

$$\frac{d\int\limits_0^\xi \mathrm{d}r\,re^{-\frac{r}{\xi}}}{\xi^d|T_\mathrm{c} - T|^{2\beta}} = \frac{d\xi^2\int\limits_0^1 \mathrm{d}u\,ue^{-u}}{\xi^d|T_\mathrm{c} - T|^{2\beta}} = \xi^{2-d}|T_\mathrm{c} - T|^{-2\beta}d\int\limits_0^1 \mathrm{d}u\,ue^{-u}.$$

Introducing the critical exponent ν for the correlation length, the left-hand side is proportional to

$$|T_\mathrm{c} - T|^{2\beta+(d-2)\nu}.$$

The inequality is thus satisfied if and only if

$$d > 2 + \frac{2\beta}{\nu}.$$

4 Renormalization

Decimation of the Ising Chain Before proceeding, we define the dimensionless operator

$$H = -\beta\mathcal{H} = K\sum_i \sum_{j=\mathrm{nn}(i)} \sigma_i\sigma_j + \frac{1}{2}h\sum_i \sigma_i + \sigma_{i+1}.$$

When computing the partition function, we may do this by first summing over spins with odd indices. For any one of these, we obtain the (partial) sum

$$2\cosh(K(\sigma_{i-1} + \sigma_{i+1}) + h).$$

Next we may write

$$2e^{\frac{1}{2}h(\sigma_{i-1}+\sigma_{i+1})}\cosh(K(\sigma_{i-1}+\sigma_{i+1})+h) = e^{2g+K'\sigma_{i-1}\sigma_{i+1}+\frac{1}{2}h'(\sigma_{i-1}+\sigma_{i+1})}$$

for quantities

$$\begin{aligned}
K' &= \frac{1}{4}\ln\left(\frac{\cosh(2K+h)\cosh(2K-h)}{\cosh^2(h)}\right),\\
h' &= h + \frac{1}{2}\ln\left(\frac{\cosh(2K+h)}{\cosh(2K-h)}\right),\\
g &= \frac{1}{8}\ln\big(16\cosh(2K+h)\cosh(2K-h)\cosh^2(h)\big).
\end{aligned}$$

Thus the trace over odd-numbered spins is

$$e^{Ng+K'\sum\limits_i \sigma_{2i}\sigma_{2i+2}+h'\sum\limits_i \sigma_{2i}}.$$

When computing the trace over the even-numbered spins, we notice that it has the same structure as for the total chain, apart from the modified constants. We thus obtain

$$Z(N,K,h) = e^{Ng(K,h)}Z\left(\frac{1}{2}N, K', h'\right).$$

We may repeat this procedure indefinitely. Hence we obtain

$$-\frac{\beta G}{N} = \sum_{j=0}^{\infty}\frac{1}{2^j}g(K_j,h_j).$$

A More General Discussion The more general approach is to consider a model on a d-dimensional lattice with a set of degrees of freedom σ_i and coupling constants K_α, and suppose that a transform of the kind we have discussed, now termed a renormalization transform, preserves the form of the Hamiltonian. We describe the system with the dimensionless quantity

$$H = -\beta\mathcal{H} = \sum_{\alpha=1}^{n} K_\alpha\psi_\alpha(\sigma)$$

where the functions $\psi_\alpha(\sigma)$ describe one particular kind of interaction. As an example, the Ising model with coupling all over the lattice may have one function ψ_1 which contains all nearest-neighbour interactions, one function ψ_2 which describes next-to-nearest neighbours and so on. The renormalization transforms the Hamiltonian to

$$H' = Ng(K) + \sum_{\alpha=1}^{n} K'_\alpha\psi_\alpha(\sigma'),$$

where σ' is the set of reduced degrees of freedom, which must have the same algebraic property as the non-reduced ones.