Summary of SI2510 Statistical Mechanics

Yashar Honarmandi
yasharh@kth.se

October 23, 2020

Abstract

This is a summary of SI2510 Statistical Mechanics.

Someone should probably give a brief statement about what this course is about. My attempt is the following: In the first course we only studied ideal systems in order to develop the machinery. Ideal systems generally have no phase transitions (bosonic gases being the exception, but quantum systems are a whole other can of worms). Having developed the machinery for describing ideal systems, we now wish to extend our studies to systems with interactions in order to describe phase transitions. This can generally not be done analytically as we have previously done, hence this course is dedicated to developing both a better language to understand phase transitions and approximate analytical methods.

I am sorry about these notes. They are unfinished and a mess in general. I blame poor teaching.

On that note, everyone seems very eager to tell you that statistical mechanics is so hard to understand, a proposition that I would counter with the following: No one seems interested in making an effort at teaching statistical mechanics. As an analogy, consider an infant swimming. If you look this up, you will see that babies have a remarkable affinity for swimming and being in water. However, a parent throwing their child in a pool and leaving it can hardly attribute the subsequent and unavoidable drowning to the intricacies of swimming. Likewise, if your teacher attributes their difficulties in teaching to the complexity of the subject matter, you can be sure that they have thrown you into the deep end with no regard for your learning. You are being gaslit. But do not let this discourage you. I have made an honest effort within my abilities to give comprehensive explanations both of what I am doing and why. If you are reading this, I hope that you will take the time to improve upon these notes, both for your own sake and that of future students.

Contents

1 Basic Concepts

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

Spontaneous Symmetry Breaking Consider a Hamiltonian with some particular symmetry and suppose that a system with this Hamiltonian solves the equation of motion in such a way that the aforementioned symmetry does not leave the system invariant. This is called spontaneous symmetry breaking.

A classical example is a particle on a sombrero. The Hamiltonian of such a system is invariant with respect to rotations about the hat's axis of symmetry. This symmetry also leaves the system invariant if the ball is on top of the sombrero. However, if the ball rolls down from the top, the system loses its rotational symmetry (the Hamiltonian is still invariant, but the system has changed). This process happens spontaneously, hence this system spontaneously breaks rotational symmetry.

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, $-\delta$ 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

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

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

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

Scaling Laws Suppose that close to a phase transition, the free energy is dominated by a term G which scales according to

$$G(t, h) = \lambda^{-d} G(\lambda^{y_t} t, \lambda^{y_h} h).$$

The idea of scaling laws is to use such relations to compute critical exponents. Namely, we obtain

$$\begin{aligned} m(t, h) &= \lambda^{-d+y_h} m(\lambda^{y_t} t, \lambda^{y_h} h), \\ \chi(t, h) &= \lambda^{-d+2y_h} \chi(\lambda^{y_t} t, \lambda^{y_h} h), \\ C_h(t, h) &= \lambda^{-d+2y_t} C_h(\lambda^{y_t} t, \lambda^{y_h} h). \end{aligned}$$

To identify the critical exponents, we approach the critical point from $h = 0$. The scaling laws apply for any scale factor, hence we may choose $\lambda^{y_t} = \frac{1}{|t|}$, yielding

$$\begin{aligned} m(t,0) &= |t|^{\frac{d-y_h}{y_t}} m(\pm 1, 0), \\ \chi(t,0) &= |t|^{\frac{d-2y_h}{y_t}} \chi(\pm 1, 0), \\ C_h(t,0) &= |t|^{\frac{d}{y_t}-2} C_h(\pm 1, 0). \end{aligned}$$

On the right-hand side we have the different quantities evaluated at points far from criticality, meaning they are finite. This also implies that if critical exponents exist on either side of the critical point, they are identical (this is for instance not the case for the order parameter, as it is zero on the other side of the transition). We now identify

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

By instead approaching from $t = 0$ and choosing $\lambda^{y_h} = \frac{1}{|h|}$ we obtain

$$m(t,h) = |h|^{\frac{d}{y_h}-1} m(0, \pm 1),$$

and we identify

$$\delta = \frac{1}{\frac{d}{y_h} - 1} = \frac{y_h}{d - y_h}.$$

Based on this, we may identify certain general scaling laws. Two of these are

$$\alpha + 2\beta + \gamma = 2, \ \beta(\delta - 1) = \gamma.$$

These are the Rushbrook and Widom scaling laws. As can be seen, they do not uniquely specify the critical exponents. This can also be seen from the scaling laws themselves.

We started the derivation with the assumption that the free energy satisfies some generalized homogeneity relation. The same must be true for the correlation length and correlation function:

$$\xi(t,h) = \lambda^{-b} \xi(\lambda^{y_t} t, \lambda^{y_h} h), \ \Gamma(\mathbf{k}, t, h) = \lambda^{-c} \Gamma(\lambda^{y_\mathbf{k}} \mathbf{k}, \lambda^{y_t} t, \lambda^{y_h} h).$$

For the correlation length we also added its dependence on the wavevector. As will be shown, the correlation function is proportional to the susceptibility for $\mathbf{k} = \mathbf{0}$, hence

$$\Gamma(\mathbf{k}, t, h) = \lambda^{-d+2y_h} \Gamma(\lambda^{y_\mathbf{k}} \mathbf{k}, \lambda^{y_t} t, \lambda^{y_h} h).$$

At $t = h = 0$ we have

$$\Gamma(\mathbf{k}) \propto k^{-2+\eta},$$

which defines the exponent η. This result is obtained by choosing $\lambda = k^{-\frac{1}{y_\mathbf{k}}}$, and we thus have

$$\frac{2y_h - d}{y_\mathbf{k}} = \frac{\gamma y_t}{y_\mathbf{k}} = 2 - \eta.$$

The correlation length is given by

$$\xi^2 \propto \frac{1}{C(0)} \left. \frac{\partial C}{\partial k^2} \right|_{k=0},$$

which scales as $\lambda^{-y_\mathbf{k}}$, hence $b = y_\mathbf{k}$. By a scaling law, we further obtain

$$\nu = \frac{y_\mathbf{k}}{y_t}$$

and Fisher's law

$$\frac{\gamma}{\nu} = 2 - \eta.$$

The laws we have derived thus far have considered the order parameter and susceptibility or the correlation function and correlation length by themselves. To derive the final scaling law, the hyperscaling law, we must connect the two. A scaling of the correlation length by a factor l should scale the energy density by l^{-d}. A rescaling of the correlation length by λ^{-b} is possible either by direct rescaling or by rescaling the field and temperature. If the previous claim is true, then we must have

$$\lambda^d G(t,h) = \lambda^{bd} G(t,h).$$

This can only be true if

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

The Universality Hypothesis The universality hypothesis is an experimentally based hypothesis. It states that all phase transitions can be sorted into a set of universality classes characterized by the dimensions of the system and order parameter, and that the critical exponents are the same for all phase transitions in a given universality class.

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 for Magnetic Systems 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

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

This is the Rushbrook scaling law back in action.

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

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

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

$$Z = \sum_{\sigma_0=\pm1} \cdots \sum_{\sigma_{N-1}=\pm1} \prod_{i=0}^{N-1} t_{\sigma_i\sigma_{i+1}}.$$

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

$$\sum_{\sigma_j=\pm1} t_{\sigma_{j-1}\sigma_j} t_{\sigma_j\sigma_{j+1}} = t^2_{\sigma_{j-1}\sigma_{j+1}}.$$

This process is repeated until you obtain

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

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

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

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

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

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

$$t \to t l^{x_t}, \; h \to h l^{x_h}.$$

We now expect the free energy to scale according to

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

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

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

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

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

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 general idea of mean-field theory is to study a system by averaging over certain degrees of freedom or interactions in order to more easily compute a free energy or partition function.

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 The idea in this approach is to separate the Hamiltonian by for any component of the system replace the ones with which it is interacting by an expectation value.

To construct a mean-field theory for the Ising model, we write the Hamiltonian as

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

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

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

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

$$Z = \mathrm{tr}\Big(e^{-\beta\mathcal{H}}\Big) = \sum_{\{\sigma\}} e^{\beta h' \sum_i \sigma_i} = \left(\sum_{\sigma=\pm 1} e^{\beta h'\sigma}\right)^N = 2^N \cosh^N\big(\beta h'\big),$$

$$m = \frac{1}{N}\left\langle \sum_i \sigma_i \right\rangle = \frac{1}{NZ}\sum\sum_i \sigma_i e^{\beta h' \sum_j \sigma_j} = \frac{1}{N}\frac{\mathrm{d}}{\mathrm{d}\beta h'} \ln\big(2^N \cosh^N\big(\beta h'\big)\big) = \tanh\big(\beta h'\big).$$

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

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

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

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

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

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

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

$$S = -N k_\mathrm{B}\left(\frac{1-m}{2}\ln\left(\frac{1-m}{2}\right) + \frac{1+m}{2}\ln\left(\frac{1+m}{2}\right)\right).$$

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

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

Introducing the critical temperature, we write this as

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

In particular, for $h = 0$ and temperatures above T_c, where there is no magnetization, we obtain

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

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

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

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

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

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 phase transition 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. This is also expected as the truncated free energy series is expected to transition between having a single minimum and two minima.

The heat capacity is given by

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

In the case of a second-order phase transition 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

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

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

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

Hence there is a step in the phase transition of

$$-\frac{a_{2,0}^2 T_\mathrm{c}}{2a_4(T_\mathrm{c})}$$

when approaching the phase transition from below.

Suppose instead that a_4 is the first coefficient to change sign. In this case the free energy might have several local minima, meaning that 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

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

Truncating at 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 of the order parameter 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, the equilibrium magnetization satisfies

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

In addition, at the transition point the free energy is equal at the post-transition equilibrium and zero, yielding

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

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

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

3 Ginzburg-Landau Theory

Formulation of 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 more resolved 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

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

Close to a second-order phase transition we may re-obtain the result from Landau theory, namely

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

by setting $h = 0$.

Ginzburg-Landau theory allow us to study fluctuations, which is what we will do next. Suppose we add some small perturbation $h_0\delta(\mathbf{x})$ from $h = 0$, which changes the field to $m_0 + \phi$, where m_0 depends on temperature and ϕ on position. 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.

The perturbation ϕ may be related to a correlation function. To see this, we add a term

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

to the Hamiltonian, yielding

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

This implies

$$\frac{\delta\langle m\rangle}{\delta h_0} = \frac{\beta\operatorname{tr}\left(m(\mathbf{0})me^{-\beta\left(\mathcal{H}-\int \mathrm{d}^3\mathbf{x} mh\right)}\right)\operatorname{tr}\left(e^{-\beta\left(\mathcal{H}-\int \mathrm{d}^3\mathbf{x} mh\right)}\right) - \beta\operatorname{tr}\left(me^{-\beta\left(\mathcal{H}-\int \mathrm{d}^3\mathbf{x} mh\right)}\right)\operatorname{tr}\left(m(\mathbf{0})e^{-\beta\left(\mathcal{H}-\int \mathrm{d}^3\mathbf{x} mh\right)}\right)}{\left(\operatorname{tr}\left(e^{-\beta\left(\mathcal{H}-\int \mathrm{d}^3\mathbf{x} mh\right)}\right)\right)^2}$$

$$= \beta(\langle m(\mathbf{0})m\rangle - \langle m(\mathbf{0})\rangle\langle m\rangle)$$
$$= \beta\Gamma(\mathbf{r}).$$

As ϕ contains the entire effect of the perturbation, this is 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 ϕ will generally have a different form, 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 the correlation function is small compared to the overall order parameter, which is satisfied 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. Using our estimate of the correlation function and introducing the critical exponent for the order parameter, we have

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

when close to criticality. Computing this in spherical coordinates yields

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

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 The Renormalization Group

The Idea The concept which we will study considers models on a d-dimensional lattice with a set of degrees of freedom σ_i and coupling constants K_α. We describe the system with the dimensionless Hamiltonian

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

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. We suppose that there exists a transformation $R_b(\sigma)$ which rescales the system and introduces degrees of freedom σ' that are a factor of b^d fewer than the original degrees of freedom. We also suppose that it works in such a way that the Hamiltonian expressed in the new degrees of freedom has the same form as that describing the original degrees of freedom, up to the addition of constants, i.e.

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

where σ' is the set of reduced degrees of freedom, which must have the same algebraic property as the non-reduced ones, and we have new coupling coefficients which are functions of the old. Such a transform is called a renormalization transform.

For a renormalization transform we obtain a new free energy

$$-\frac{\beta}{N}G\left(N,K_\alpha\right)=f\left(K_\alpha\right)=g(K_\alpha)+b^{-d}f\left(K'_\alpha\right).$$

Repeated application of such a transformation will yield

$$-\frac{\beta}{N}G\left(N,K_\alpha\right)=\sum_{j=0}b^{-jd}g(K_{\alpha,j}).$$

In other words, by determining the sequence of coupling constants and the function g, the free energy may be written as a series of this form. A more relevant use, however, comes from the fact that singularities in the free energy must be captured in the function f and not g as they also occur in the non-transformed system. Hence we may study the critical behaviour of the system by studying f. As we will see later, we will be doing this using scaling theory.

A characteristic feature of renormalization transforms is that $R_b^2(\sigma)=R_{b^2}(\sigma)$. This was a motivation to dub the set of all renormalization transforms the renormalization group. This is a misnomer, however, as there is no inverse renormalization transform - the price we pay for doing this is losing information about microscopic details of the system.

Fixed Point Analysis Consider a transformation as described above. There exist values of the coupling constants such that $K'_\alpha(K_\alpha)=K_\alpha$. These are called fixed points. The renormalization transform causes the coupling constants to flow in coupling constant space. Depending on the direction of this flow, the fixed point is termed either stable or unstable.

To study the critical behaviour of the system, we will now study the critical points of such model. By assuming that we are close to criticality and linearizing, we hope that we obtain scaling laws for the system.

We introduce this approach by considering a model described by two coupling constants. The set of critical points for this model will form a curve in coupling constant space. Changing the temperature for a fixed system corresponds to moving along a straight line in this space. We will try to relate the flow under renormalization transformations to that of changing temperature.

Under a renormalization transform, the coupling constant flow cannot approach the critical line, as the correlation length there is infinite and generally decreased by a renormalization transform (when measured in transformed length units). The long-range order may not be changed by such transforms either, hence the flow cannot make points cross the critical line. Certain points on the critical line may be fixed points, but generally all of them are not. Hence the flow around a fixed point is parallel to the critical line and away from the fixed point, or towards the fixed point for points starting on the critical line.

To proceed, consider some particular fixed point $K_1^\star,K_2^\star$ and a flow close to this point. Denoting the functions R_α, which perform the flow, we introduce displacements ΔK from the critical point, and obtain to first order

$$K_1'=R_1(K_1^\star+\Delta K_1,K_2^\star+\Delta K_2)=K_1^\star+\Delta K_1\frac{\partial R_1}{\partial K_1}+\Delta K_2\frac{\partial R_1}{\partial K_2},$$

with a similar expression for K_2'. Introducing the matrix

$$M_{ij}=\frac{\partial R_i}{\partial K_j}$$

and the displacements

$$\Delta K_1'=K_1'-K_1^\star$$

we have

$$\Delta K_1'=M_{11}\Delta K_1+M_{12}\Delta K_2,\ \Delta K_2'=M_{21}\Delta K_1+M_{22}\Delta K_2.$$

To continue, we change variables as determined by the eigenvectors of M. Naming these $\boldsymbol{\phi}_1$ and $\boldsymbol{\phi}_2$, the new variables, called scaling fields, are

$$U_i = \phi_{1,i}\Delta K_1 + \phi_{2,i}\Delta K_2.$$

The property of renormalization transforms imply that the corresponding eigenvalues are $\lambda_i = b^{y_i}$, hence

$$U_i' = b^{y_i}U_i.$$

At this point we take a break to obtain an explicit expression for the exponents. We have

$$y_i = \frac{\ln(\lambda_i)}{\ln(b)}.$$

In other words, they are given by the structure of the fixed points.

Due to our assumptions about the flow close to the fixed point, one of the y_i must be positive while the other is not. The other exponent corresponds to an eigenvector which is tangential to the critical line at the fixed point. The possibility of the non-positive exponent being zero corresponds to a line of fixed points.

Returning to the energy function, we expect g to be analytic as it only contains information about short-range fluctuations, hence the singular part of f satisfies the recursion relation

$$f\left(K_\alpha\right) = b^{-d}f\left(K_\alpha'\right).$$

Close to the critical point, this may be written as

$$f(U_1', U_2') = b^{-d}f\left(b^{y_1}U_1', b^{y_2}U_2'\right).$$

Suppose now that a change in temperature affects only U_1. At any point on the critical line close to the one around which we linearized, we must have $U_1 = 0$. The converse is true elsewhere in the surrounding. Hence we may identify $U_1 = t$ and obtain.

$$f(t, U_2') = b^{-d}f\left(b^{y_1}t, b^{y_2}U_2'\right).$$

Similarly, U_2 may be identified with some other variable. By combining this with the previously obtained expressions for the exponents, we may obtain scaling exponents and thus critical exponents from the structure of the fixed points.

As renormalization transforms are expected to either enhance or weaken long-range interactions, we must have $y_1 > 0$. Hence by approaching criticality at fixed U_2, we obtain that the flow is in fact independent of U_2 (you can verify this by using a scaling argument analogous to the ones previously performed). Hence, all systems whose Hamiltonians flow to the same critical fixed points have the same critical exponents, which is reminiscent of the previously studied universality arguments.

The previous analysis generalizes to higher-dimensional coupling constant spaces.

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

$$H = -\beta\mathcal{H} = K\sum_i \sigma_i\sigma_{i+1} + \frac{1}{2}h\sum_i \sigma_i + \sigma_{i+1}.$$

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

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

The partition function is thus

$$\begin{aligned} Z &= tr(e^H) \\ &= \sum_{\sigma_1}\sum_{\sigma_3}\cdots\sum_{\sigma_2}\sum_{\sigma_4}\cdots e^{K\sum_i \sigma_i\sigma_{i+1}+\frac{1}{2}h\sum_i \sigma_i+\sigma_{i+1}} \\ &= \sum_{\sigma_1}\sum_{\sigma_3}\cdots\sum_{\sigma_{N-1}}\prod_{i=1}^{\frac{N}{2}} 2e^{\frac{1}{2}h(\sigma_{2i-1}+\sigma_{2i+1})}\cosh(K(\sigma_{2i-1}+\sigma_{2i+1})+h). \end{aligned}$$

Reindexing according to $2i-1 \to i,\ 2i+1 \to i+1$ and introducing the new number of degrees of freedom $N' = \frac{1}{2}N$, we obtain

$$Z = \sum_{\{\sigma_1\}} \prod_{i=1}^{N'} 2e^{\frac{1}{2}h(\sigma_i+\sigma_{i+1})} \cosh(K(\sigma_i+\sigma_{i+1})+h).$$

This is a Hamiltonian of an Ising chain if we can identify quantities g, K', h' such that

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

This must be true for any value of the involved functions $\sigma_i+\sigma_{i+1},\ \sigma_i\sigma_{i+1}$. There are three possible combinations of these values, yielding the four equations

$$\begin{aligned} 2e^{h}\cosh(2K+h) &= e^{2g+K'+h'}, \\ 2e^{-h}\cosh(-2K+h) &= e^{2g+K'-h'}, \\ 2\cosh(h) &= e^{2g-K'}. \end{aligned}$$

Multiplying the former yields

$$4\cosh(2K+h)\cosh(2K-h) = e^{4g+2K'}.$$

Multiplying or dividing by the square of the latter yields

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

We thus have

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

Inserting this into one of the former yields

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

The partition function may thus be written

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

Repeating this procedure generates a sequence of constants and a prefactor given by the same function g. Doing this an infinite number of times yields

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

Renormalization And Universality Consider an Ising model with coupling between nearest and next-to-nearest neighbours and suppose that we are able to identify a renormalization transform of the system. We would thus have

$$f\left(t, h, \frac{J_2}{J_1}\right) = l^{-d} f\left(t l^{y_t}, h l^{y_h}, \frac{J_2}{J_1} l^{y}\right).$$

The quotient between the coupling constants is also necessary to specify the free energy in this case. We expect that $y < 0$ for such a transform as there are no interactions beyond next-to-nearest neighbours. Choosing $l = |t|^{-\frac{1}{y_t}}$ will yield

$$f\left(t, h, \frac{J_2}{J_1}\right) = |t|^{\frac{d}{y_t}} l^{-d} f\left(\pm 1, h|t|^{-\frac{y_h}{y_t}}, \frac{J_2}{J_1}|t|^{-\frac{y}{y_t}}\right).$$

Close to criticality the coupling constant argument on the right-hand side vanishes, hence coupling may be neglected close to criticality. This is an example of universality emerging from renormalization transforms.

The Cumulant Method For more examples, we study the Ising model on a two-dimensional triangular lattice. The dimensionless Hamiltonian is

$$H = \frac{1}{2} \sum_{i,j} K_{ij} \sigma_i \sigma_j + h \sum_i \sigma_i,$$

where we now allow interactions throughout the lattice. Note that the wedge brackets signify that each bond is to be counted once.

The triangular lattice may be divided into triangular blocks of three spins. The lattice of such blocks is a triangular lattice with a lattice constant a factor $\sqrt{3}$ greater than that of the original lattice. We will perform a renormalization procedure on these blocks defined by the projection operator $P(\mu_I, \{\sigma_I\})$, which transforms the spins into the block into new variables μ_I. For our transformation to work, we must have

$$e^{Ng(K,h)+H'(\{\mu\},K',h')} = \mathrm{tr}_{\{\sigma\}}\left(\left(\prod_I P(\mu_I, \{\sigma_I\})\right) e^{H(\{\sigma\},K,h)}\right).$$

In other words, the trace over spin configurations is carried out in such a way that the configuration of new variables is pre-determined. If we require the projection operators to have trace 1, we may trace out the new variables and reorder the summation to obtain

$$\mathrm{tr}_{\{\mu\}}(e^{Ng+H'}) = \mathrm{tr}_{\{\sigma\}}(e^{H}),$$

and the free energy is preserved.

We wish to treat this system using the cumulant method, and do this by dividing the Hamiltonian into two terms

$$H(\{\sigma\}, K, h) = H_0(\{\sigma\}, K, h) + V(\{\sigma\}, K, h).$$

H_0 contains all terms that operate within the blocks and V contains all interactions between blocks. We introduce K_n as the coupling constant n-th nearest neighbours and thus have

$$H_0 = \sum_I K_1(\sigma_{I,1}\sigma_{I,2} + \sigma_{I,1}\sigma_{I,3} + \sigma_{I,2}\sigma_{I,3}) + h(\sigma_{I,1} + \sigma_{I,2} + \sigma_{I,3}),\ V = \frac{1}{2} \sum_{I,J,n} K_n \sum_{\alpha,\beta} \sigma_{I,\alpha}\sigma_{J,\beta}.$$

This yields

$$\mathrm{tr}_{\{\sigma\}}\left(\left(\prod_I P(\mu_I, \{\sigma_I\})\right) e^{H(\{\sigma\},K,h)}\right) = Z_0 \left\langle e^V \right\rangle$$

where

$$Z_0 = \mathrm{tr}_{\{\sigma\}}\left(\left(\prod_I P(\mu_I, \{\sigma_I\})\right) e^{H_0(\{\sigma\},K,h)}\right)$$

and the expectation value is defined in terms this partition function, i.e. considering only interactions within the block. Note that it is done for some particular configuration of the new variables. This is not yet an approximation.

The approximation comes now, when we introduce a so-called truncated cumulant expansion

$$\left\langle e^V \right\rangle = e^{\langle V\rangle + \frac{1}{2!}\left\langle (V-\langle V\rangle)^2\right\rangle + \frac{1}{3!}\left\langle (V-\langle V\rangle)^3\right\rangle + \dots}.$$

We name the terms in the exponent C_i. Where does this approximation come from? When computing the free energy we have

$$\begin{aligned}
F &= F_0 - k_\mathrm{B}T \ln\left(\left\langle e^V \right\rangle\right) \\
&= F_0 - k_\mathrm{B}T \ln\left(\left\langle \sum_{i=0}^{\infty} \frac{(-1)^i}{i!} V^i \right\rangle\right) \\
&= F_0 - k_\mathrm{B}T \ln\left(\sum_{i=0}^{\infty} \frac{(-1)^i}{i!} \left\langle V^i \right\rangle\right) \\
&= F_0 - k_\mathrm{B}T \ln\left(1 + \sum_{i=1}^{\infty} \frac{(-1)^i}{i!} \left\langle V^i \right\rangle\right) \\
&= F_0 + k_\mathrm{B}T \left(\sum_{j=1}^{\infty} \frac{(-1)^j}{j} \left(\sum_{i=1}^{\infty} \frac{(-1)^i}{i!} \left\langle V^i \right\rangle\right)^j\right),
\end{aligned}$$

and working out some terms and computing the exponential should yield the correct result. It might be nice to show this.

Let us now include only nearest-neighbour interactions. We thus write the partition function as

$$Z_0 = \mathrm{tr}_{\{\sigma\}} \left(\left(\prod_I P(\mu_I, \{\sigma_I\}) \right) e^{\sum_I K_1(\sigma_{I,1}\sigma_{I,2} + \sigma_{I,1}\sigma_{I,3} + \sigma_{I,2}\sigma_{I,3}) + h(\sigma_{I,1}+\sigma_{I,2}+\sigma_{I,3})} \right).$$

For the renormalization transform to work, Z_0 must represent the interaction-free parts of the renormalized Hamiltonians. We must therefore have

$$Z_0 = \prod_I e^{A + B\mu_i}.$$

Asssuming the new variables to be designated according to a majority rule, we fix these. By tracing out the individual blocks we obtain

$$e^{A+B} = e^{3K+3h} + 3e^{-K+h}, \; e^{A-B} = e^{3K-3h} + 3e^{-K-h},$$

with solutions

$$A = \frac{1}{2}\ln\Big((e^{3K+3h} + 3e^{-K+h})(e^{3K-3h} + 3e^{-K-h})\Big), \; B = \frac{1}{2}\ln\left(\frac{e^{3K+3h} + 3e^{-K+h}}{e^{3K-3h} + 3e^{-K-h}}\right).$$

As the blocks do not interact in the base Hamiltonian, we obtain

$$\langle V\rangle = \frac{1}{2}\sum_{I,J,n} K_n \sum_{\alpha,\beta} \langle \sigma_{I,\alpha}\sigma_{J,\beta}\rangle = \frac{1}{2}\sum_{I,J,n} K_n \sum_{\alpha,\beta} \langle \sigma_{I,\alpha}\rangle \langle \sigma_{J,\beta}\rangle .$$

In particular, as there are only nearest-neighbour interactions we have

$$\langle V\rangle = K \sum_{\langle I,J\rangle} \sum_{\langle \alpha,\beta\rangle} \langle \sigma_{I,\alpha}\rangle \langle \sigma_{J,\beta}\rangle ,$$

where the wedge bracket is taken to mean that each interaction be included only once. Translational invariance yields that all these expectation values must be equal. To reproduce a Hamiltonian of the same form in the new variables, we must have

$$\langle \sigma_{I,\alpha}\rangle = C + D\mu_I.$$

These expectation values can be calculated directly from Z_0 according to

$$\langle\sigma\rangle = \frac{1}{e^{A+B\mu_I}} \sum_{\alpha} \sigma_1 e^{K(\sigma_1\sigma_2+\sigma_1\sigma_3+\sigma_2\sigma_3)+h(\sigma_1+\sigma_2+\sigma_3)},$$

where we chose some arbitrary spin with respect to which we compute the expectation values. The two choices of the new variables yield

$$\begin{aligned} C+D &= \frac{1}{e^{A+B}}(e^{3K+3h}+2e^{-K+h}-e^{-K+h}) = \frac{e^{3K+3h}+e^{-K+h}}{e^{3K+3h}+3e^{-K+h}}, \\ C-D &= \frac{1}{e^{A-B}}(-e^{-3K-3h}-2e^{-K-h}+e^{-K-h}) = -\frac{e^{-3K-3h}+e^{-K-h}}{e^{3K-3h}+3e^{-K-h}}, \end{aligned}$$

with solution

$$\begin{aligned} C &= \frac{1}{2}\left(\frac{e^{3K+3h}+e^{-K+h}}{e^{3K+3h}+3e^{-K+h}} - \frac{e^{-3K-3h}+e^{-K-h}}{e^{3K-3h}+3e^{-K-h}}\right), \\ D &= \frac{1}{2}\left(\frac{e^{3K+3h}+e^{-K+h}}{e^{3K+3h}+3e^{-K+h}} + \frac{e^{-3K-3h}+e^{-K-h}}{e^{3K-3h}+3e^{-K-h}}\right). \end{aligned}$$

This is combined with the renormalization anzats to yield

$$\begin{aligned} Ng(K,h)+H'(\{\mu\},K',h') &= \frac{1}{3}NA(K,h)+B\sum_I \mu_I + K\sum_{\langle I,J\rangle}\sum_{\langle\alpha,\beta\rangle}(C+D\mu_I)(C+D\mu_J) \\ &= \frac{1}{3}NA(K,h)+B\sum_I \mu_I + K\sum_{\langle I,J\rangle}\sum_{\langle\alpha,\beta\rangle} C^2+CD(\mu_I+\mu_J)+D^2\mu_I\mu_J. \end{aligned}$$

Let us now count terms. The inner summation gives two contributions for any fixed pair I, J, hence

$$Ng(K,h)+H'(\{\mu\},K',h') = \frac{1}{3}NA(K,h)+B\sum_I \mu_I + 2K\sum_{\langle I,J\rangle} C^2+CD(\mu_I+\mu_J)+D^2\mu_I\mu_J.$$

The last term is now fine. The first term may be multiplied by a factor $\frac{1}{3}N$ due to the number of Is, 6 due to the number of Js for each I and $\frac{1}{2}$ due to double counting. Next, relabelling J to I in the middle term adds a factor 2. We also add a factor 6 from summing over J and divide by 2 due to double counting. The final result is

$$Ng(K,h)+H'(\{\mu\},K',h') = \frac{1}{3}NA(K,h)+2KNC^2+(B+12KCD)\sum_I \mu_I + 2KD^2\sum_{\mathrm{nn}(I,J)}\mu_I\mu_J.$$

This yields the recursion relations

$$\begin{aligned} g(K,h) &= \frac{1}{3}A(K,h)+2KC^2(K,h), \\ K' &= 2KD^2(K,h), \\ h' &= B(K,h)+12KC(K,h)D(K,h), \ . \end{aligned}$$

For $h=0$, we obtain $B=C=0$, and thus

$$K' = 2K\left(\frac{e^{3K}+e^{-K}}{e^{3K}+3e^{-K}}\right)^2, \ h'=0.$$

For small K the recursion becomes $K'=\frac{1}{2}K$, and for large K it is $K'=2K$, meaning that the flow reverses at some point. The final fixed point is $K=0.3356$. Using the fact that $b=\sqrt{3}$ one obtains $y_t=0.882,\ y_h=2.034$. These results do not agree particularly well with the exact results, but nevertheless demonstrate the power of the framework.

Next we consider the second-order approximation.

ϵ-Expansion Consider some system of $N = n^d$ vectorized degrees of freedom on a d-dimensional simple cubic lattice with lattice parameter a. The Hamiltonian is given by

$$\mathcal{H} = -\sum_{\mathbf{r}} \sum_{\mathbf{r}'} J_{\mathbf{r}-\mathbf{r}'} \mathbf{S}_{\mathbf{r}} \cdot \mathbf{S}_{\mathbf{r}'} - h \sum_{\mathbf{r}} S_{\mathbf{r},\alpha},$$

where all sums are performed over all lattice vectors, meaning $r_i = ma,\ m = 0, \dots, n-1$, and α is some particular direction in which the field points. We introduce a weighting function W to allow the magnitudes of the vectors $\mathbf{S}$ to vary. The partition function is thus

$$Z = \left(\prod_{\mathbf{r}} \int \mathrm{d}^d \mathbf{S}_{\mathbf{r}}\, W(S_{\mathbf{r}}) \right) e^{-\beta \mathcal{H}}.$$

To proceed, we impose periodic boundary conditions on the spins, namely $\mathbf{S}_{\mathbf{r}+na\mathbf{e}_i} = \mathbf{S}_{\mathbf{r}}$ and Fourier transform the spins according to

$$\mathbf{S}_{\mathbf{r}} = \frac{1}{\sqrt{N}} \sum_{\mathbf{q}} \mathbf{S}_{\mathbf{q}} e^{i\mathbf{q}\cdot\mathbf{r}},\ q_i = \frac{2\pi}{na} m,\ m = 0, \pm 1, \dots, \pm(n-1).$$

The Fourier components must satisfy $\mathbf{S}_{\mathbf{q}}^{\star} = \mathbf{S}_{-\mathbf{q}}$. We thus have

$$\sum_{\mathbf{r}} S_{\mathbf{r},\alpha} = \frac{1}{\sqrt{N}} \sum_{\mathbf{r}} \sum_{\mathbf{q}} S_{\mathbf{q},\alpha} e^{i\mathbf{q}\cdot\mathbf{r}}.$$

The term such that $\mathbf{q} = \mathbf{0}$ produces a factor N. I wave my hands to remove the rest for now. Next we have

$$\begin{aligned}
\sum_{\mathbf{r}} \sum_{\mathbf{r}'} J_{\mathbf{r}-\mathbf{r}'} \mathbf{S}_{\mathbf{r}} \cdot \mathbf{S}_{\mathbf{r}'} &= \frac{1}{N} \sum_{\mathbf{r}} \sum_{\mathbf{r}'} J_{\mathbf{r}-\mathbf{r}'} \sum_{\mathbf{q}} \sum_{\mathbf{k}} \mathbf{S}_{\mathbf{q}}^{\star} \cdot \mathbf{S}_{\mathbf{k}} e^{i(\mathbf{q}\cdot\mathbf{r} - \mathbf{k}\cdot\mathbf{r}')} \\
&= \frac{1}{N} \sum_{\mathbf{q}} \sum_{\mathbf{k}} \mathbf{S}_{-\mathbf{q}} \cdot \mathbf{S}_{\mathbf{k}} \sum_{\mathbf{r}} \sum_{\mathbf{r}'} J_{\mathbf{r}-\mathbf{r}'} e^{i(\mathbf{q}\cdot\mathbf{r} - \mathbf{k}\cdot\mathbf{r}')}
\end{aligned}$$

We also choose a weighting function

$$W(S) = e^{-\frac{1}{2} b S^2 - c S^4}$$

to simplify calculations. If $b < 0$ and $c = \frac{b}{4}$ then the exponent has a maximum at $S = 1$. By choosing b to be large this function may be arbitrarily sharply peaked around the maximum value. We now introduce the effective Hamiltonian

$$\begin{aligned}
H &= -\beta \mathcal{H} + \sum_{\mathbf{r}} \ln(W(S_{\mathbf{r}})) \\
&= -\sum_{\mathbf{q}} \left(\frac{1}{2} b - K(\mathbf{q}) \right) \mathbf{S}_{\mathbf{q}} \cdot \mathbf{S}_{-\mathbf{q}} - \sum_{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3} (\mathbf{S}_{\mathbf{q}_1} \cdot \mathbf{S}_{\mathbf{q}_2})(\mathbf{S}_{\mathbf{q}_3} \cdot \mathbf{S}_{-\mathbf{q}_1 - \mathbf{q}_2 - \mathbf{q}_3}) + \sqrt{N} h_0 S_{\mathbf{0}}^{\alpha}
\end{aligned}$$

where

$$K(\mathbf{q}) = \beta \sum_{\mathbf{a}} J_{\mathbf{a}} e^{-i\mathbf{q}\cdot\mathbf{a}},\ h_0 = \beta h.$$

By introducing $\beta J_{\mathbf{a}} = \frac{1}{2} K_0$ only for nearest-neighbours and restricting ourselves to a simple cubic lattice we have in the long-wavelength limit

$$K(\mathbf{q}) = K_0 \sum_{j=1}^{d} d K_0 - \frac{1}{2} K_0 a^2 q^2.$$

We may somehow rescale things to obtain

$$H = -\frac{1}{2} \sum_{\mathbf{q}} \left(r + q^2 \right) \mathbf{S}_{\mathbf{q}} \cdot \mathbf{S}_{-\mathbf{q}} - \frac{u}{N} \sum_{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3} (\mathbf{S}_{\mathbf{q}_1} \cdot \mathbf{S}_{\mathbf{q}_2})(\mathbf{S}_{\mathbf{q}_3} \cdot \mathbf{S}_{-\mathbf{q}_1 - \mathbf{q}_2 - \mathbf{q}_3}) + \sqrt{N} h_0 S_{\mathbf{0}}^{\alpha}$$

for

$$r = \frac{b - 2dK_0}{K_0 a^2},\ h = \frac{h_0}{\sqrt{K_0 a^2}},\ u = \frac{c}{K_0^2 a^4}.$$

We now proceed with the case $u = 0$, known as the Gaussian model. Its partition function is given by

$$Z = \left(\prod_{\mathbf{q}} \int \mathrm{d}^d \mathbf{S}_{\mathbf{q}} \right) e^{-\frac{1}{2}\sum_{\mathbf{q}} (r+q^2)\mathbf{S}_{\mathbf{q}} \cdot \mathbf{S}_{-\mathbf{q}} + \sqrt{N} h_0 S_{\mathbf{0}}^{\alpha}}.$$

Rather than computing this exactly, we will approach it with the renormalization group. To do that, we first carry out the functional integration for all $\mathbf{S}_{\mathbf{q}}$ such that $q > q_l = \frac{\pi}{l}$ for some parameter $l > 1$. This corresponds to dividing the spins into blocks. The resulting partition function is

$$Z = A \left(\prod_{q < q_l} \int \mathrm{d}^d \mathbf{S}_{q<l} \right) e^{-\frac{1}{2}\sum_{\mathbf{q}} (r+q^2)\mathbf{S}_{\mathbf{q}} \cdot \mathbf{S}_{-\mathbf{q}} + \sqrt{N} h_0 S_{\mathbf{0}}^{\alpha}}$$

for some constant A. Next we introduce $\mathbf{q}' = l\mathbf{q}$ - in other words, we rescale lengths - and multiply the summation by a factor l^{-d} to compensate for the extra degrees of freedom. The rescaled dimensionless Hamiltonian is thus

$$H' = -\frac{1}{2} \sum_{\mathbf{q}'} \left(r + \left(\frac{q'}{l} \right)^2 \right) \mathbf{S}_{\frac{1}{l}\mathbf{q}'} \cdot \mathbf{S}_{-\frac{1}{l}\mathbf{q}'} + \sqrt{N} h_0 S_{\mathbf{0}}^{\alpha}.$$

For this to have the same functional form as before, we must rescale the degrees of freedom according to

$$\mathbf{S}_{\frac{1}{l}\mathbf{q}'} = l^{1+\frac{d}{2}} \mathbf{S}_{\mathbf{q}'}.$$

This will yield

$$H' = -\frac{1}{2} \sum_{\mathbf{q}'} \left(r' + (q')^2 \right) \mathbf{S}_{\mathbf{q}'} \cdot \mathbf{S}_{-\mathbf{q}'} + \sqrt{N} h' S_{\mathbf{0}}^{\alpha},\ r' = r l^2,\ h' = h l^{1+\frac{d}{2}}.$$

Assuming r to be temperature-like, we have the scaling relation

$$g(t, h) = l^{-d} g(t l^2, h l^{1+\frac{d}{2}}).$$

Some critical exponents are $\alpha = 2 - \frac{d}{2},\ \beta = \frac{d}{4} - \frac{1}{2},\ \gamma = 1$.

The S^4 Model Consider an effective Hamiltonian

$$\begin{aligned} H = &-\frac{1}{2} \sum_{\mathbf{q}} (r + q^2) \mathbf{S}_{\mathbf{q}} \cdot \mathbf{S}_{-\mathbf{q}} - \frac{u}{N} \sum_{\mathbf{q}_1, \ldots, \mathbf{q}_3} (\mathbf{S}_{\mathbf{q}_1} \cdot \mathbf{S}_{\mathbf{q}_2})(\mathbf{S}_{\mathbf{q}_3} \cdot \mathbf{S}_{-\mathbf{q}_1 - \mathbf{q}_2 - \mathbf{q}_3}) \\ &- \frac{w}{N^2} \sum_{\mathbf{q}_1, \ldots, \mathbf{q}_3} (\mathbf{S}_{\mathbf{q}_1} \cdot \mathbf{S}_{\mathbf{q}_2})(\mathbf{S}_{\mathbf{q}_3} \cdot \mathbf{S}_{\mathbf{q}_4})(\mathbf{S}_{\mathbf{q}_5} \cdot \mathbf{S}_{-\mathbf{q}_1 - \mathbf{q}_2 - \mathbf{q}_3 - \mathbf{q}_4 - \mathbf{q}_5}) + \cdots + h\sqrt{N} S_{\mathbf{0}}^{\alpha}. \end{aligned}$$

We will handle this using a renormalization transform, but as traces over the higher-order terms are difficult to compute, we will use a cumulant expansion.

5 Percolation

The Idea In many contexts the existence of certain behaviour depends on the structure of microscopic clusters. The study of such clusters is called percolation theory.

Percolation models come in two forms: site percolation, in which lattice sites are occupied with some probability p, and bond percolation, in which nearest-neighbour bonds are formed with some probability p.

Definition of Quantities We now define the following quantities for such models:

- the occupation probability p.
- the site percolation probability p_c, which is the lowest probability such that infinite clusters form in the thermodynamic limit.
- $n_s(p)$, which is the number of clusters of size s per lattice site given the occupation probability.
- $P(p)$, which is the fraction of occupied sites belonging to a spanning cluster (a cluster which goes from one end of the lattice to the other).
- $S(p)$, which is the mean size of finite clusters.
- the pair connectedness $C(p, r)$, which is the probability that two occupied sites separated by a distance r belong to the same cluster.

Given these definitions, we may infer the following: First, the product $sn_s(p)$ is equal to the number of sites in clusters of size s per lattice site, i.e. the probability that a particular site is occupied and belongs to a cluster of size s. Next, we must have $P(1) = 1$ and $P(p) = 0$ for $p < p_c$ in the thermodynamic limit. In addition, all occupied sites either belong to a spanning cluster or a cluster of finite size, implying

$$pP(p) + \sum_s sn_s(p) = p,$$

where the summation is over clusters of finite size. Finally, we have

$$S(p) = \frac{\sum_s s^2 n_s(p)}{\sum_s sn_s(p)}.$$

Scaling Theory For doing scaling theory, we postulate that the quantity

$$G(p) = \sum_s n_s(p)$$

is analogous to the free energy per site in such models. According to this analogy, $P(p)$ plays the role of the order parameter, $S(p)$ the role of the susceptibility and $C(p, r)$ that of the correlation function. Based on this, we define critical exponents as previously. The interesting part is that if universality holds, the critical exponents will not depend on most of the details of our work.

To this we add an extra piece of conjecture, namely that there exists a typical cluster size s_ξ close to $p = p_c$ which dominates the critical behaviour of G, P and S. We define its critical exponent according to

$$s_\xi \propto |p - p_c|^{-\frac{1}{\sigma}}.$$

We also assume that

$$n_s(p) = n_s(p_c) f\left(\frac{s}{s_\xi}\right).$$

f must, by construction, approach zero as its argument increases to infinity and approach 1 as its argument approaches zero. For large s, it has been observed that $n_s(p_c)$ behaves as a power law in s with power $-\tau$, which is dependent on dimensionality. Thus we have

$$n_s(p) = s^{-\tau}\phi\left(s|p - p_c|^{\frac{1}{\sigma}}\right).$$

We may now write G as

$$G = \sum_s s^{-\tau}\phi\left(s|p - p_c|^{\frac{1}{\sigma}}\right) \approx \int_0^\infty \mathrm{d}s\, s^{-\tau}\phi(s|p - p_c|^{\frac{1}{\sigma}}) = |p - p_c|^{\frac{\tau-1}{\sigma}} \int_0^\infty \mathrm{d}x\, x^{-\tau}\phi(x).$$

Assuming the integral to converge and treating p as a temperature-like parameter we may use scaling theory to obtain

$$\alpha = 2 - \frac{\tau - 1}{\sigma},\ \gamma = \frac{3 - \tau}{\sigma},\ \beta = \frac{2 - \tau}{\sigma}.$$

Next, we expect that the integral of C should be equal to S. This will yield

$$S \propto \int \mathrm{d}r\, r^{d-1} \frac{e^{-\frac{r}{\xi}}}{r^{d-2+\eta}} = \xi^{d-1-(d-2+\eta)} \int \mathrm{d}u\, u^d \frac{e^{-u}}{u^{d-2+\eta}} \propto |p - p_\mathrm{c}|^{-\nu(2-\eta)},$$

yielding

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

We also assume the dominance of clusters of size s_ξ around the critical point. Their concentration is proportional to $|p - p_\mathrm{c}|^{\frac{\tau}{\sigma}}$. This concentration should be inversely proportional to ξ^d as the clusters have size ξ, yielding

$$\frac{\tau}{\sigma} = d\nu,$$

which is (definitely not, but anyway) equivalent to $d\nu = 2 - \alpha$.