

PHYS 353 (Advanced Statistical Mechanics) Notes

Rio Weil

This document was typeset on December 31, 2024

Introduction:

This is a set of lecture notes taken from UChicago's PHYS 353 (Advanced Statistical Mechanics), taught by Peter Littlewood. Topics covered include the discrete to continuous transition, the ferromagnetic and antiferromagnetic Heisenberg chain, Ginzburg-Landau Theory, Fluctuations, The Scaling Hypothesis, The Renormalization Group, Perturbative RG, Continuous Symmetries, the non-linear σ model, the 2-D XY model, Disorder and random system, Random fields, Spin glasses, Replica symmetry breaking, Neural networks and Boltzmann machines, and Non-equilibrium dynamics.

Contents

1 Introduction	4
1.1 Probabilities in Statistical Mechanics	4
1.2 Spin Models	4
1.3 From discrete models to fields	4
1.4 Phase Transitions	5
2 From Particles to Fields	7
2.1 Phonons	7
2.2 Ferromagnetic Heisenberg Chain	10
3 Antiferromagnetic Heisenberg Chain,	12
3.1 Review of the Heisenberg Chain	12
3.2 Antiferromagnetic Case	12
3.3 Phase Transitions in the Classical Magnet	14
3.4 Ginzburg-Landau Theory	16
4 Ginzburg-Landau Theory Continued	16
4.1 Review from last class + Symmetries	16
4.2 Minimal Model	17
4.3 Mean Field Theory from Saddle Point Approximation	17
4.4 Fluctuations	19
5 Fluctuations, Continued	20
5.1 Phase and Amplitude Fluctuations	20
5.2 Technical Aside: Gaussian Integrals	22
5.3 Back to Physics - Transverse Fluctuations and Coloumb Emergence	23
6 Scaling Theory	24
6.1 Recap	24
6.2 Massive Modes and Upper Critical Dimension	25
6.3 Scaling Theory	27

7	Scaling Hypothesis Part II and Renormalization Group	29
7.1	Recap	29
7.2	Renormalization Group: Conceptual View	30
7.3	Getting Exponents out of RG	31
7.4	RG Fixed Points and Flows	32
8	Renormalization Group Part II	33
8.1	Exact Solution of Gaussian Model	34
8.2	RG approach to Gaussian model	35
8.3	Scaling dimension	37
9	Perturbative Renormalization Group	38
9.1	Perturbative RG for ϕ^4 theory	38
10	Continuous Symmetries - Beyond Spin Waves	42
10.1	Non-linear σ models	43
10.2	RG for Non-linear σ model	44
10.3	Vortices	45
11	2-D XY model	46
11.1	Single Vortices	46
11.2	Pairs of Vortices	47
11.3	Transitions	48
11.4	Polarizability + Looking ahead	48
12	Disorder and Random Systems	49
12.1	Introduction and Motivation	49
12.2	Disorder Type I - Impurities/Vacancies	50
12.3	Computing quenched averages - Replica Trick	51
12.4	Transition Temperatures and Concentration	51
12.5	Computing with Replicas	52
13	Random Fields and Spin Glasses	53
13.1	Setting up the Random Field Problem	53
13.2	Solving the model	53
13.3	Ergodicity	55
13.4	Order parameters	57
14	Spin Glass	58
14.1	Order parameters	58
14.2	Sherrington-Kirepatrick model	59
14.3	p -spin spherical model	59
15	Replica Symmetry Breaking, Neural Nets, Boltzmann Machines	62
15.1	Replica Symmetry Breaking, Concluded	62
15.2	Neural Nets	63
15.3	Boltzmann Machines	64

16 Non-equilibrium Dynamics	65
16.1 Diffusion Equation ($V = 0$)	65
16.2 Fokker-Planck Equation ($V \neq 0$)	65
16.3 Generalizing to a field - Model A	66
16.4 Model B	67
16.5 Noise and Optimal Trajectories	67
16.6 Where have we come?	69

1 Introduction

I was 34 minutes late to lecture due to misunderstanding the start time - thanks to Kalpak for providing his notes!

1.1 Probabilities in Statistical Mechanics

In stat mech, probabilities are given by:

$$P(\mu) = \frac{e^{-\beta H(\mu)}}{Z} \quad (1.1)$$

for a microstate μ . Here, Z is the partition function:

$$Z = \sum_{\mu} e^{-\beta H(\mu)} \quad (1.2)$$

and $\beta = \frac{1}{k_B T}$ is the inverse temperature. The free energy can be obtained as:

$$F = -\frac{1}{\beta} \ln Z \quad (1.3)$$

What is meant by a state μ ? In some sense, there is a tension, as one slices up the continuous state space of classical physics. This is resolved by fields.

1.2 Spin Models

A rich setting to explore statistical mechanics is spin models. For example, the celebrated *Ising model* is defined by the Hamiltonian:

$$H_{\text{Ising}} = \sum_{ij} J_{ij} S_i S_j \quad (1.4)$$

with J_{ij} the coupling between classical spins $S_i \in \{1, -1\}$. We can also generalize this for arbitrary spin directions by promoting the spins to vectors (giving rise to a continuous degree of freedom), yielding the Heisenberg model:

$$H_{\text{Heisenberg}} = \sum_{ij} J_{ij} \mathbf{S}_i \cdot \mathbf{S}_j. \quad (1.5)$$

As is clear from the sums appearing in the above expressions, we are considering discrete spins (perhaps on some lattice). We can convert to momentum space via a Fourier transform, and then via a "coarse-graining procedure" (formalized by the renormalization group) we obtain a roughly continuous spectrum, which we may treat as a field.

1.3 From discrete models to fields

For example, we can consider the linear dispersion relation:

$$\omega(k) = c|k| \quad (1.6)$$

which gives rise to the density of states:

$$d^3 n = u(k) d^3 k = u(\omega) d\omega \quad (1.7)$$

Which by then considering a sphere in k -space of surface area $4\pi k^2$ this becomes:

$$qu(\omega) = u(k) 4\pi k^2 \frac{dk}{d\omega} \quad (1.8)$$

The energy density $u(k)$ we take to be constant, i.e. just $V/(2\pi)^d = L^3/(2\pi)^d$ (for a box of length L) and $\frac{dk}{d\omega}$ is obtained from the dispersion relation, yielding:

$$d^3n = \frac{L^3}{2\pi^2} \frac{\omega^2}{c^3} \quad (1.9)$$

This is the Rayleigh argument for black-body radiation. Of course this is incorrect, as the density of energy diverges as $\omega \rightarrow \infty$. This is fixed by quantum mechanics, where the Einstein blackbody argument yields:

$$u(\nu) = \frac{8\pi h\nu^3}{c^3} \frac{1}{e^{\frac{h\nu}{k_B T}} - 1}. \quad (1.10)$$

Let's now return to the spin model, and consider a field construction of the system. We define the magnetization:

$$m = \sum_{i=1}^N S_i \quad (1.11)$$

and we now coarse grain, taking $m = m(x)$ to be a function of x . We look at a window/block of spins and describe it with a theory in the continuum limit. For a given window/block i , we can consider:

$$m(x_i) = \sum_{i,j < N} S_{i+j} \quad (1.12)$$

Phenomenologically, we can guess the theory with local terms:

$$H = \int dx \left[am^2(x) + bm^3(x) + cm^4(x) + k(\nabla m)^2 + k'(\nabla^2 m)^2 + \dots \right] \quad (1.13)$$

and use this to probe the explore of the system. Note that the Hamiltonian cannot have terms of order m or Δm as this breaks the \mathbb{Z}_2 inversion symmetry. However, such terms can enter in the forms of external fields (yielding terms like hm) which break the symmetry.

1.4 Phase Transitions

We can also think about phases of matter and phase transitions between them. E.g. phase transition between solid/liquid/gas phases are discontinuous. Can have discontinuous phase transitions between different phases; find free energy minimums of two different phases, and they cross. There's a subtlety at the critical point, however, which we return to later (it turns out that this point is scale invariant, and described by a conformal field theory). Also note that there are V for which no apparent symmetries broken in the liquid-to-gas transition, i.e. the phase transition is continuous (but it does turn out that an emergent symmetry becomes broken).

In the magnet case, we can have magnetic and non-magnetic phase. In magnetic phase, has to choose a direction for the magnetization to point in. This breaks the \mathbb{Z}_2 symmetry. Looking at the phase diagram for a magnet, we have the parameters T temperature and h the external magnetic field. For $h > 0$ we have magnetization up and for $h < 0$ we have magnetization down. At zero temperature the transition is abrupt. There is a critical temperature T_c above which the transition is smooth, and below T_c we have varying degrees of an abrupt transition. Interestingly, the critical point between liquid and gas "looks like" an Ising transition.

We can obtain scaling laws for the magnetization:

$$m(T, h = 0) = \begin{cases} 0 & T > T_c \\ |t|^\beta & T < T_c \end{cases} \quad (1.14)$$

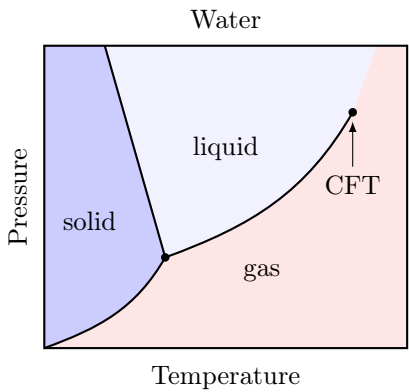


Figure 1.1: Phase diagram of water. There is a lot of rich phenomenology in the phase transitions, for example at the critical point where the system is scale invariant and described by a CFT.

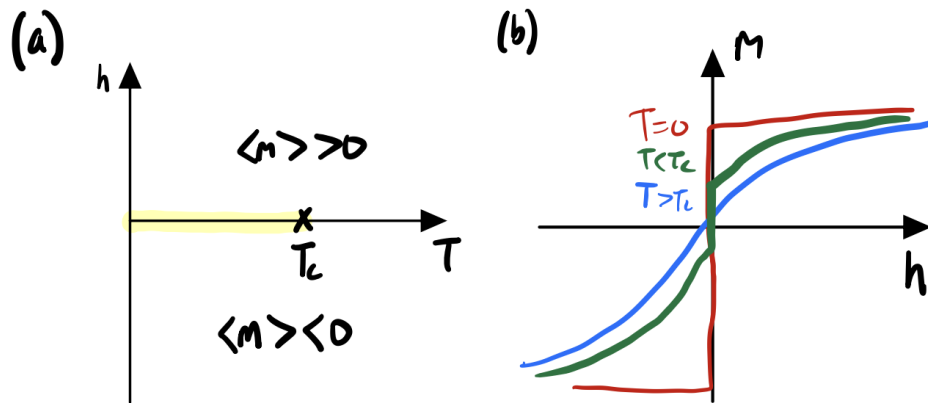


Figure 1.2: (a) h (external field) vs. T (temperature) phase diagram for the Ising model and (b) behaviour of m (magnetization) as a function of h for different T . For $h > 0$ we have $\langle m \rangle > 0$ and vice versa for $h < 0$. For $T < T_c$, there is an abrupt transition in m at $h = 0$ (yellow line). This can be seen by studying the behaviour of m for different temperatures; for $T = 0$ the transition is abrupt (as we would expect; since the external field is the only contribution to the energy, any finite magnetic field fully polarizes the spins). Below T_c the transition stays abrupt, though with a smaller “jump” in the magnetization, and above T_c the transition is smooth.

$$m(T = T_c, h) \sim h^{1/\delta} \quad (1.15)$$

There are other quantities we may compute; for example we could measure the magnetic susceptibility $\chi = \frac{\partial m}{\partial h}(t)$. Because the system becomes very polarizable near the phase transition, we may expect that χ has a peak/divergence at $T = T_c$, which is characterized by a further exponent. We could also compute the specific heat of the system, which also has its associated exponent and so on.

There's a whole zoo of critical exponents, and it turns out that there are (universality) classes where the exponents are the same - i.e. near these phase transitions very different systems are governed by the same theories.

Note that the theory described by Eq. (1.13) is known as Ginzburg-Landau theory (or the *mean-field* theory), and was written down to describe phenomenology. We have a power series expansion in the magnetization, with the coefficients (e.g. a) as smooth functions governed by T, h , and so on.

We can consider a ϕ^4 theory:

$$F = a\phi^2 + b\phi^4 + h\phi \quad (1.16)$$

with a phenomenological guess of $a = a_0(T - T_c)$.

$$\frac{dF}{d\phi} = 0, \quad \phi^2 = -\frac{a}{b} \text{ or } \phi = 0 \quad (1.17)$$

We then have $\phi \sim (T_c - T)^{1/2}$, i.e. $\beta = 1/2$.

General procedure: Start with mean field theory. Then, aspects of field theory will turn out. One aspect will be thermal fluctuations about the mean field theory states. Then we think about terms like $k(\nabla m)^2$, which tells us that there are energy costs to (e.g.) mis-aligned spins. This is known still as the Gaussian model as the Hamiltonian with this term is still quadratic in the fields. The moment I have a gradient term, I get modes in the system that can propagate. If I have a lot of modes, I don't get a phase transition.

Scaling Hypothesis - I know there are a set of critical exponents. How are they related? Via scaling. So, suppose I look at the system on different scales. As I average over different scales, if the system has to look self-similar. RG gives us a way to obtain critical exponents.

Other systems - we will also look at systems in low-dimensions, which can be interesting. There are also systems with inherent disorder/randomness, e.g. glasses which are frozen but not ordered systems. There is at least one model that realizes this, which Parisi got a Nobel for. We will also look at dynamical systems; what happens to open/driven systems? There are ways to formally write them in a very similar language. E.g. friction, forced flow, growing things at an interface, biology. Finally, we may study information and data. The straightforward introduction for physicists is spin glasses, which can (in theory) be used to encode data via learning a set of interactions J_{ij} .

2 From Particles to Fields

In this lecture, we will look at a chain of atoms, and a chain of ferromagnetic spins.

2.1 Phonons

We consider atoms of mass m connected by springs of spring constant k_s , with equilibrium a . We can notate the displacement of the n th atom from its equilibrium position as ϕ_n , and x_n the distance of the n th atom from the end of the string.

The Lagrangian of the system is then:

$$L = T - V = \sum_{i=1}^N \left[\frac{1}{2} m \dot{x}_i^2 - \frac{k_s}{2} (x_{i+1} - x_i - a)^2 \right] = \sum_{n=1}^N \left[\frac{1}{2} m \dot{\phi}_n^2 - \frac{k_s}{2} (\phi_{n+1} - \phi_n)^2 \right] \quad (2.1)$$

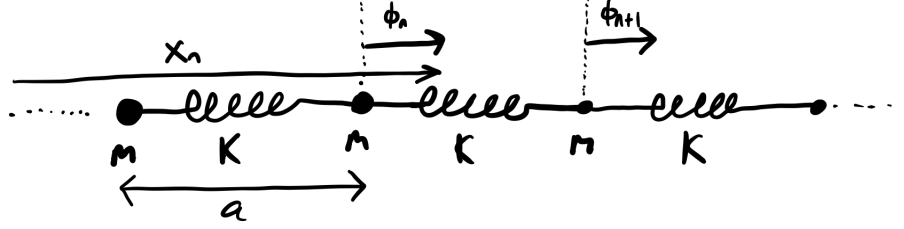


Figure 2.1: Diagram of a chain of atoms connected by springs.

notice there is a potential energy piece and a kinetic energy piece, both quadratic. We will want to solve for the eigenmodes of these problem. There are N degrees of freedom and as such we expect to find N eigenmodes.

This problem is solvable! We can write down the equation of motion of the springs; this is just nearest neighbour couplings so we can write down an analytic solution.

We promote:

$$\phi_n \rightarrow a^{1/2}\phi(x) \quad (2.2)$$

with the $a^{1/2}$ introduced for dimensional purposes, namely in such a way that we measure position in units of $\frac{x}{a}$, i.e. making it dimensionless. Then:

$$(\phi_{n+1} - \phi_n) \rightarrow a^{3/2}\partial_x\phi(x) \quad (2.3)$$

Finally, we convert from a discrete sum to an integral:

$$\sum \rightarrow \frac{1}{a} \int_0^{L=Na} dx \quad (2.4)$$

Thus the Lagrangian becomes a *functional* (i.e. a function of a function) of a field:

$$L[\phi] = \int_0^x \mathcal{L}(\phi, \partial_x\phi, \dot{\phi}) \quad (2.5)$$

where \mathcal{L} is the Lagrangian density:

$$\mathcal{L} = \frac{1}{2}m\dot{\phi}^2 - \frac{k_s a^2}{2}(\partial_x\phi)^2. \quad (2.6)$$

We then write down the classical action:

$$S[\phi] = \int dt L[\phi] \quad (2.7)$$

Currently the field ϕ is not determined; it is a function of x, t . Our next task is to look for solutions that minimize the action.

To this end, we have the Euler-Lagrange equation:

$$\frac{\partial \mathcal{L}}{\partial \phi} - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{\phi}} - \frac{d}{dx} \frac{\partial \mathcal{L}}{\partial (\partial_x \phi)} = 0 \quad (2.8)$$

which is derived from the principle of least action. Working out the Euler Lagrange equation, we get the expected result:

$$\boxed{[m\partial_t^2 - (k_s a^2)\partial_x^2]\phi(x, t) = 0} \quad (2.9)$$

this is a wave equation! What we can see is it describes sound waves. Notice that the above equation is translation invariant, and as such the general solution is the sum of a right and left handed wave:

$$\phi_+(x + vt) - \phi_-(x - vt) \quad (2.10)$$

with velocity:

$$v = a\sqrt{\frac{k_s}{m}}. \quad (2.11)$$

Often however, it is common to look at this equation and say that it is simplest to study it in terms of Fourier modes; consider solutions:

$$\phi(x, t) = \phi_0 e^{i(\omega t - kx)} \quad (2.12)$$

which plugging into the wave equation gives us the dispersion relation:

$$\omega = vk \quad (2.13)$$

The waves oscillate in time and in space.

Going back to the discrete case, we have the eigenmodes:

$$\phi_n(t) = \sum_k \frac{1}{\sqrt{N}} e^{i(\omega_k t - kna)} \quad (2.14)$$

with dispersion:

$$\omega(k) = 2\sqrt{\frac{k_s}{m}} \left| \sin\left(\frac{ka}{2}\right) \right| \quad (2.15)$$

This solution knows about the discreteness of the lattice. Going to the continuum, we lost something, but we have a much simpler theory. In particular, the theory is good for small k /large wavelengths. The coarse graining made us lose information about the short-wavelength modes; we should not use the field theory for microscopic wavelengths. But, say for populating the chain with low- T bosons (i.e. low energy bosons) this captures the phenomenology well.

The dispersion curves looks like:

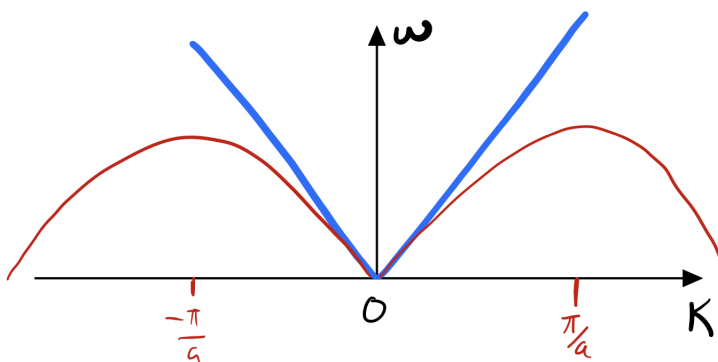


Figure 2.2: Continuum and Discrete Dispersion. The continuum dispersion is linear in k , while the discrete dispersion is $\sim |\sin(k)|$. The two dispersion relations agree for small k /large wavelengths/small energies.

We can now use this to do thermodynamics! We can write down the internal energy:

$$E(T) = E_0 + \sum_k \hbar\omega(k) \left[\langle n_k \rangle + \frac{1}{2} \right] \quad (2.16)$$

where:

$$\langle n_k \rangle = \frac{1}{e^{\frac{\hbar\omega}{k_B T}} - 1} \quad (2.17)$$

In the limit where $k_B T \ll \hbar\omega$, $\hbar\omega(k) \rightarrow \hbar vk$ and the energy goes as $\sim T^2$. This is the well known thermodynamic dependence of the spring chain.

2.2 Ferromagnetic Heisenberg Chain

We consider the quantum Hamiltonian:

$$\hat{H} = -J \sum_{m=1}^{N-1} \hat{\mathbf{S}}_m \cdot \hat{\mathbf{S}}_{m+1} \quad (2.18)$$

where we take the coupling $J > 0$ such that the spins aligning result in lowered energy (i.e. we are describing a ferromagnet). At each site we have spin operators:

$$[\hat{S}_m^i, \hat{S}_n^j] = i\delta_{mn}\epsilon^{ijk}\hat{S}_n^k \quad (2.19)$$

with δ_{mn} the Kronecker delta (which tells us that the spin operators only have nontrivial commutation on the same site) and ϵ^{ijk} is the totally anti-symmetric tensor. We denote the total spin as S , a number (can be $1/2, 1, 3/2$, etc.). We can consider the eigenstates of \hat{S}^z :

$$\hat{S}_m^z |s_m\rangle = s |s_m\rangle. \quad (2.20)$$

where $|s_m\rangle$ form a ladder of states from $s = -S, -S+1, \dots, S-1, S$. We guess the ground state:

$$|\Omega\rangle = \bigotimes_{m=1}^N |S_m\rangle \quad (2.21)$$

i.e. all the spins are spin up. This clearly minimizes the energy, as for each coupling term $\hat{\mathbf{S}}_m \cdot \hat{\mathbf{S}}_{m+1}$ we get energy $-JS$ and so the total energy of $|\Omega\rangle$ is $-JSN$. Note however in writing this ground state down we broke some symmetry, as indeed the spins could be all aligned in any direction, and we would get the same ground state energy.

Recall that on a single site, we can have the raising/lowering operators:

$$\hat{S}_m^\pm = \hat{S}_m^x \pm i\hat{S}_m^y \quad (2.22)$$

which takes us up/down the ladder of $|s_m\rangle$ eigenstates. We can now rewrite our Hamiltonian to be of the form:

$$\hat{H} = -J \sum_m \left(\hat{S}_m^z \hat{S}_{m+1}^z + \frac{1}{2} \left(\hat{S}_m^+ \hat{S}_{m+1}^- + \hat{S}_m^- \hat{S}_{m+1}^+ \right) \right) \quad (2.23)$$

We now play a trick, namely the Holstein-Primakoff transformation. We define new operators (sometimes known as magnon variables):

$$\hat{S}_m^z = S - \hat{a}_m^\dagger \hat{a}_m \quad (2.24)$$

$$\hat{S}_m^+ = \left(2S - \hat{a}_m^\dagger \hat{a}_m \right)^{1/2} \hat{a}_m \quad (2.25)$$

$$\hat{S}_m^- = \hat{a}_m^\dagger \left(2S - \hat{a}_m^\dagger \hat{a}_m \right)^{1/2} \quad (2.26)$$

where:

$$[\hat{a}_m, \hat{a}_n^\dagger] = \delta_{mn} \quad (2.27)$$

It can be verified that this transformation is *exact* (though it looks strange!) as it preserves the spin commutation relations. Why does this work? Intuitively, we have introduced the bosons through the operators \hat{a}, \hat{a}^\dagger ; the ladder of spin states look like the ladder of spin states (just that one is an infinite, and one is a finite ladder). But the square roots actually make sure that the dimensionality of our Hilbert space is preserved, because when I hit the walls of $\pm S$, the state vanishes. However it is instructive to consider the case where S is large:

Let us Taylor expand \hat{S}_m^\pm :

$$\hat{S}_m^+ \approx \sqrt{2S}\hat{a}_m + \frac{O(\hat{a}_m^\dagger\hat{a}_m\hat{a}_m)}{\sqrt{2S}} \quad (2.28)$$

if $S \rightarrow \infty$, then we can throw away all but the first order term. Interestingly, it has a habit of working even when $S = \frac{1}{2}$. Let us write the Hamiltonian to second order in the operators:

$$\hat{H} = -JNS^2 + JS \sum_m \left[\hat{a}_m^\dagger \hat{a}_m + \hat{a}_{m+1}^\dagger \hat{a}_{m+1} - \hat{a}_m^\dagger \hat{a}_{m+1} - \hat{a}_{m+1}^\dagger \hat{a}_m \right] + O(S^0) \quad (2.29)$$

The number operator terms come from the \hat{S}^z , while the exchange terms come from the \hat{S}^\pm cross terms. As such, we have now rewrote the Hamiltonian in terms of bosonic modes. Because this is periodic, the natural thing to do is to go back to the phonon problem in the discrete sense. There, we went into momentum space, and we shall do that again here:

$$\tilde{a}_k = \frac{1}{\sqrt{N}} \sum_m e^{ikm} \hat{a}_m \quad (2.30)$$

where the \tilde{a}_k corresponds to a momentum annihilation operator and \hat{a}_m corresponds to a position annihilation operator. Note that the momentum operators obey the same algebra:

$$[\tilde{a}_k, \tilde{a}_{k'}^\dagger] = \delta_{kk'} \quad (2.31)$$

which transforms the Hamiltonian into:

$$\hat{H} = -JNS^2 + \sum_k \omega_k \tilde{a}_k^\dagger \tilde{a}_k \quad (2.32)$$

where:

$$\omega_k = 4J \sin^2\left(\frac{k}{2}\right) \quad (2.33)$$

This is a quadratic (at small k) dispersion relation. Note that we still have a zero mode at $k = 0$ where the symmetry is preserved. Note also that in the HW, you will solve the classical version of this problem, and find the same dispersion.

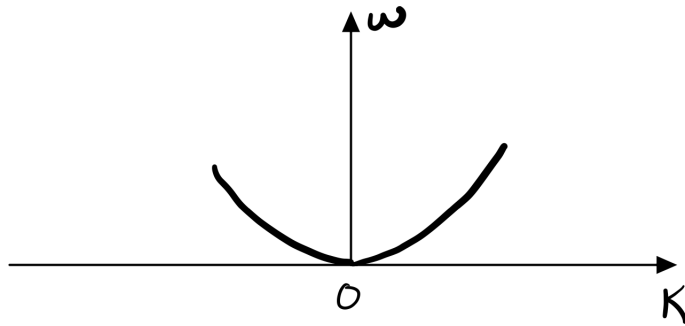


Figure 2.3: Small k dispersion relation for the ferromagnetic Heisenberg spin chain. For small k , the dispersion is quadratic.

We call these objects “spin waves”. This is clearly an approximation, as we have ignored the $O(S^0)$ terms in the Hamiltonian which include terms such as $\hat{a}^\dagger a \hat{a}^\dagger a$ which correspond to interactions between spin waves. So long as we stay near the zero mode, there is “nowhere to go” so the interactions can be neglected (if we look at large k modes on the other hand, the interactions become relevant).

A naive field theory for this would produce $\omega_k \sim k$, but it turns out spin conservation + rotational invariance gives us $\omega_k \sim k^2$. On Friday, we will look at the antiferromagnetic version of this problem, where we will see we obtain $\omega_k \sim k$ (due to vacuum fluctuations).

3 Antiferromagnetic Heisenberg Chain,

3.1 Review of the Heisenberg Chain

Recall the Heisenberg spin chain Hamiltonian:

$$\hat{H} = -J \sum_m \left(\hat{S}_m^z \hat{S}_{m+1}^z + \frac{1}{2} \left(\hat{S}_m^+ \hat{S}_{m+1}^- + \hat{S}_m^- \hat{S}_{m+1}^+ \right) \right) \quad (3.1)$$

and after the Holstein-Primakoff transformation, this became:

$$\hat{H} = -JNS^2 + JS \sum_m \left(\hat{a}_m^\dagger \hat{a}_m + \hat{a}_{m+1}^\dagger \hat{a}_{m+1} - \hat{a}_m^\dagger \hat{a}_{m+1} - \hat{a}_{m+1}^\dagger \hat{a}_m \right) + O(S^0) \quad (3.2)$$

Note to get here we first did an exact transformation into bosonic operators, then expanded this transformation in a Taylor series to obtain a Hamiltonian quadratic in \hat{a}/\hat{a}^\dagger . This can be turned into a field theory; loosely, we can replace the a s by field operators:

$$\hat{a}_m \rightarrow \hat{\phi}(x), \quad \hat{a}_m^\dagger \rightarrow \hat{\phi}^\dagger(x) \quad (3.3)$$

Further, we can consider:

$$\hat{a}_{m+1} \approx \hat{a}_m + \frac{\partial}{\partial x} \hat{a}_m \quad (3.4)$$

which can give us $\nabla \hat{\phi}$ terms. We also have the promotion of the commutators to the continuum:

$$[\hat{a}_m, \hat{a}_{m'}^\dagger] = \delta_{mm'} \rightarrow [\hat{\phi}(x), \hat{\phi}^\dagger(x')] = \delta(x - x')a \quad (3.5)$$

Note that this would give us a free field theory (as the bosonic operator terms are all quadratic).

3.2 Antiferromagnetic Case

Formally, we change the sign of J in Eq. (3.1) so that it is now $J < 0$. Now, we guess our ground state to be alternating between spin up/down, with spins on the A sublattice being up and those on the B sublattice being down:

So, let us consider applying a π rotation on the B sublattice of spins, i.e. taking:

$$S_B^z \rightarrow -S_B^z, \quad S_B^y \rightarrow -S_B^y, \quad S_B^x \rightarrow S_B^x \quad (3.6)$$

So the Hamiltonian becomes:

$$\hat{H} = -|J| \sum_m \left(\hat{S}_m^z \hat{S}_{m+1}^z - \frac{1}{2} \left(\hat{S}_m^+ \hat{S}_{m+1}^+ + \hat{S}_m^- \hat{S}_{m+1}^- \right) \right) \quad (3.7)$$

or with the H.P. transformation:

$$\hat{H} = -|J|NS^2 + |J|S \sum_m \left(\hat{a}_m^\dagger \hat{a}_m + \hat{a}_{m+1}^\dagger \hat{a}_{m+1} - \hat{a}_m \hat{a}_{m+1} - \hat{a}_{m+1}^\dagger \hat{a}_m^\dagger \right) + O(S^0) \quad (3.8)$$

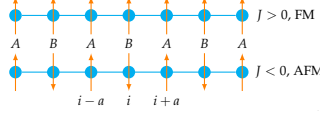


Figure 3.1: Cartoon picture of FM and AFM Heisenberg chain. $J > 0$ (ferromagnet) encourages the spins to align, while $J < 0$ (antiferromagnet) encourages the spins to anti-align. Our guess for the ground state in this case is to bipartition the chain into two sublattices, A, B for spins pointing up and down. We do note that this is an ansatz/guess for the ground state, and is not an exact eigenstate of the quantum Hamiltonian.

If we diagonalize this (which we can always do - it is a quadratic form!) we obtain:

$$\hat{H} = -N|J|S(S+1) + |J|S \sum_k \begin{pmatrix} \hat{a}_k^\dagger & \hat{a}_k \end{pmatrix} \begin{pmatrix} 1 & \cos k \\ \cos k & 1 \end{pmatrix} \begin{pmatrix} \hat{a}_{-k} \\ \hat{a}_{-k}^\dagger \end{pmatrix} \quad (3.9)$$

where:

$$\hat{a}_k = \sum_m e^{ikm} \hat{a}_m \quad (3.10)$$

Now doing a Bogoulibov transformation:

$$\begin{pmatrix} \hat{a}_k \\ \hat{a}_{-k}^\dagger \end{pmatrix} = \begin{pmatrix} \cosh \theta_k & -\sinh \theta_k \\ -\sinh \theta_k & \cosh \theta_k \end{pmatrix} \begin{pmatrix} \hat{a}_k \\ \hat{a}_{-k}^\dagger \end{pmatrix} \quad (3.11)$$

with $\tanh 2\theta_k = \cos k$. With this rotation, \hat{H} becomes:

$$\hat{H} = -N|J|S^2 + 2|J|S \sum_k (\sin k) \hat{a}_k^\dagger \hat{a}_k \quad (3.12)$$

Which gives us the spectrum - it is linear! We can think of these eigenstates as linear superpositions of spin waves in the original basis.

Unfortunately, we aren't quite done. These operators are not conserved. Because the ground state is defined by $\hat{a}_k|0\rangle = 0$, but \hat{a}_k is a combination of a^\dagger/a s and so can create excitations.

Let us look at the magnetization of the ground state:

$$M = \langle 0 | \frac{1}{N} \sum_i \hat{S}_i^z | 0 \rangle = S - \frac{1}{N} \sum_i \langle 0 | a^\dagger a | 0 \rangle = S - \frac{1}{N} \sum_k \sinh^2 \theta_k \quad (3.13)$$

I.e. we see it is nonzero! These are quantum fluctuations on top of the classical ground state. Looking at the integral form for this, we have:

$$M = S - \int_0^{1/a} dk \frac{1}{k} k^{d-1} \quad (3.14)$$

which diverges at $k = 0$ for $d = 1$. An important note to make; fluctuations, even quantum ones, can destroy phase transitions! E.g. the 1D classical Heisenberg model orders, the quantum one does not. But of course when we look at thermal fluctuations thing disorder even more.

The fact that M diverges in 1D tells us that the approximation (i.e doing the sublattice divisions...) we made at the beginning was incorrect. This sublattice division is incorrect as there is no local $SU(2)$ symmetry (only a global one). This might have been okay if I was just asking about long range order. If M turned out to be finite, we would still have a magnet... but here that is not the case.

3.3 Phase Transitions in the Classical Magnet

We consider a classical magnet (i.e. a lattice of classical spins), subject to two parameters, an external magnetic field h and a temperature T . We consider the magnetization $m(T)$, which is something we may measure in a lab:

$$m(T) = \frac{1}{V} \lim_{h \rightarrow 0} M(h, T) \quad (3.15)$$

where:

$$M(h, T) = \frac{1}{N} \sum_i S_i \quad (3.16)$$

is a quantity that depends on the external field and temperature.

We have the scaling of the magnetization:

$$m(T, h = 0) = \begin{cases} 0 & t > 0 \\ |t|^\beta & t < 0 \end{cases} \quad (3.17)$$

with $t = \frac{T_c - T}{T}$. We can also consider the scaling of m w.r.t. h at the critical temperature:

$$m(T = T_c, h) \propto h^{1/\delta}. \quad (3.18)$$

The β, δ appearing above are the critical exponents that quantify the phase transition.

We may also consider the measurement of a response function, i.e. the response of the system to some perturbation that we put in. An example is the magnetic susceptibility χ :

$$\chi(T, h = 0) = \frac{\partial m}{\partial h} \quad (3.19)$$

or the specific heat:

$$C = \left. \frac{dE}{dT} \right|_T \propto c |t|^{-\alpha \pm} \quad (3.20)$$

How do we calculate these quantities? We start with some partition function:

$$Z = \text{Tr}(e^{-\beta(H_0 + m.h.)}) \quad (3.21)$$

with $\beta = (k_B T)^{-1}$, $H_0 = -J \sum_{ij} S_i S_j$ and $m.h. = -h \sum_i S_i$ (the magnetization in field h). The trace here is a shorthand to say we compute the exponential for every possible configuration of spins, with the terms weighted by the energy (you could do this numerically, e.g. by doing a Monte Carlo simulation).

If I have a partition function, I can compute things! In particular, if I want to compute the (average) magnetization:

$$\langle M \rangle = \frac{\partial \ln Z}{\partial (\beta h)} = \frac{1}{Z} \text{Tr}(M e^{-\beta(H_0 + m.h.)}) \quad (3.22)$$

We can then compute χ :

$$\chi = \frac{\partial M}{\partial h} = \beta \left[\frac{1}{Z} \text{Tr}(M^2 e^{-\beta H}) - \frac{1}{Z^2} \left(\text{Tr}(M e^{-\beta H}) \right)^2 \right] \quad (3.23)$$

Of we look at this carefully, we see the first term is the expectation of M^2 and the second term is the square of the expectation value of M , i.e.:

$$\chi = \beta \left(\langle M^2 \rangle - \langle M \rangle^2 \right). \quad (3.24)$$

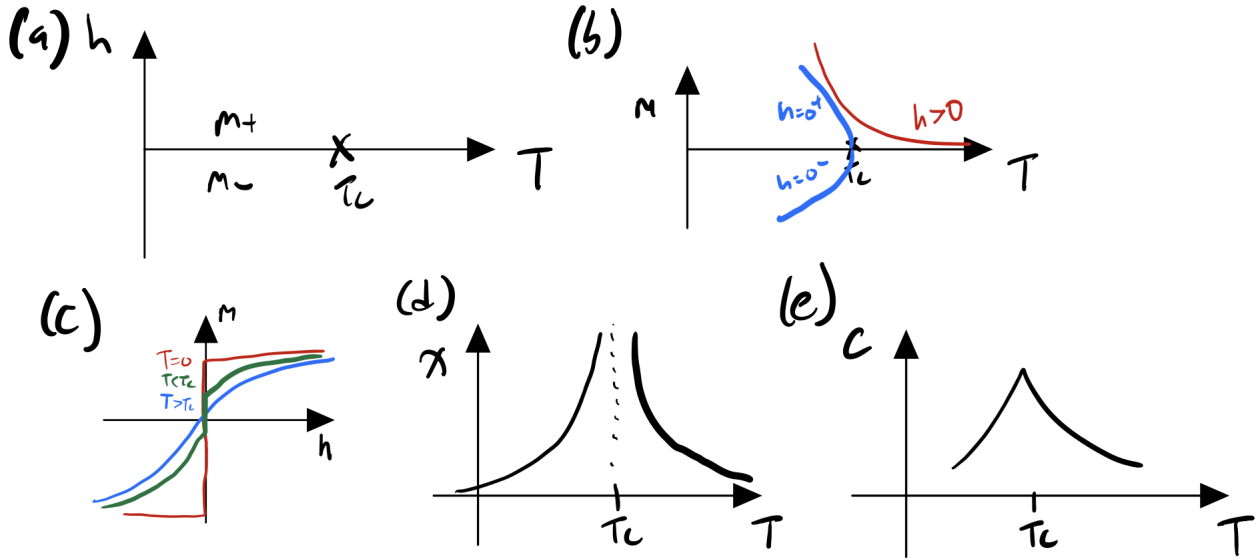


Figure 3.2: Plots of various measurable quantities and their behaviours near the critical temperature. (a) External field vs. Temperature. (b) Magnetization vs. Temperature for $h = 0$ and $h > 0$. (c) Magnetization vs. External field for various T . (d) Magnetic susceptibility vs. Temperature. (e) Heat capacity vs. Temperature.

Thus, the susceptibility tells you about fluctuations from the average.

This M we have written here is an average over all configurations and over all space. We may also be interested in an average just over space. To this end, we may consider that the macroscopic magnetization is just the integral over the course-grained position dependent magnetization $m(x)$:

$$M = \int d^3r m(x) \quad (3.25)$$

from which:

$$\chi = \beta \int d^3r d^3r' \left(\langle m(r)m(r') \rangle - \langle m(r) \rangle \langle m(r') \rangle \right). \quad (3.26)$$

Now, if r, r' are closeby, we expect the magnetization to be correlated, and if they are not, it depends on the ordering of the state. To this end, let us define here a connected average of C :

$$\langle m(r)m(r') \rangle_C = \langle (m(r) - \langle m(r) \rangle)(m(r') - \langle m(r') \rangle) \rangle = G(r - r') - m^2 \quad (3.27)$$

where G is a (Green's) function that depends on the distance $r - r'$. With this, the integral for χ becomes:

$$\chi = \frac{V}{k_B T} \int dr \langle m(r)m(0) \rangle_C. \quad (3.28)$$

So, interestingly, the susceptibility tells us something about the correlation function of spins. Generically, I expect this to decay:

$$G(r) \sim e^{-r/\xi} \quad (3.29)$$

which introduces the idea of a correlation length ξ . As I approach the phase transition, far-away degrees of freedom become more and more correlated- from the divergence of χ at the phase transition, we can also conclude that the correlation length ξ must also diverge. Thus, we hypothesize:

$$\xi \sim |t|^{-\nu}. \quad (3.30)$$

Also, notice that this happens on *both* sides of the transition.

3.4 Ginzburg-Landau Theory

Another way we can think about the partition function:

$$Z(T) = \text{Tr}(e^{-\beta H}) = \int \mathcal{D}[m(r)] W(m(r)) \quad (3.31)$$

I.e. look at all spatial configurations $\mathcal{D}[m(r)]$ of the order parameter, for each of these compute the statistical weight W , and add them up.

We can use the W to define a Hamiltonian:

$$\beta H[m(r)] = -\ln W[m(r)] \quad (3.32)$$

we do this to avoid writing down a microscopic Hamiltonian. Then, $\beta H[m(r)]$ is a functional of m , so:

$$\beta H[m(r)] = \int d^d r \phi[m(r), r] = \int d^d r \phi[m, \nabla m, \nabla^2 m, h, \dots] \quad (3.33)$$

We assume that ϕ is an analytic function of all of these variables, and stop at some point because we would consider some terms to not be relevant.

4 Ginzburg-Landau Theory Continued

4.1 Review from last class + Symmetries

We will go into the phenomenological G-L theory, showing how the mean field theory are derived from the saddle point approximation. We then consider small Gaussian fluctuations about this saddle point. Sometimes these fluctuations are not actually small, which will get us into discussion of the renormalization group.

We start with the partition function:

$$Z(T) = \text{Tr}(e^{-\beta H}) = \int \mathcal{D}[\bar{m}(\mathbf{r})] W(\bar{m}(\mathbf{r})) \quad (4.1)$$

Here, \bar{m} is a coarse-grained averaged of the order parameter, the coarse-grained version of the microscopic spins S_i s. We take an integral over all configurations, and we associate with each configuration a statistical weight W .

We then define the (effective) Hamiltonian as:

$$\beta H_{\text{eff}}[\bar{m}(\mathbf{r})] = -\ln W[\bar{m}(\mathbf{r})] \implies W \sim e^{-\beta H_{\text{eff}}} \quad (4.2)$$

We remark that RG will eventually allow us to derive the effective Hamiltonian from the microscopic one. Then, $\beta H[\bar{m}(\mathbf{r})]$ is a functional of m , and we assume that Φ , whose integral gives the effective Hamiltonian, is local:

$$\beta H_{\text{eff}} = \int d^d r \Phi[\bar{m}(\mathbf{r}), \mathbf{r}] \quad (4.3)$$

and then we may expand it perturbatively:

$$\Phi[\bar{m}(\mathbf{r}), \nabla \bar{m}(\mathbf{r}), \nabla^2 \bar{m}(\mathbf{r}), \dots]. \quad (4.4)$$

The next assumption that Φ is an analytic function. We now consider symmetries of the system. For example for the Heisenberg chain, we have rotational symmetry:

$$H[R\bar{m}] = H[\bar{m}] \quad (4.5)$$

This symmetry implies that the leading order term must be $\bar{m}^2 = \bar{\mathbf{m}} \cdot \bar{\mathbf{m}}$, and then \bar{m}^4 and so on (the rotational symmetry forbids odd powers). Further, inversion symmetry forbids odd powers in ∇m , so we consider:

$$(\nabla \bar{m})^2 = \sum_{i=1}^n \sum_{\alpha=1}^d (\partial_\alpha \bar{m}_i)(\partial_\alpha \bar{m}_i) \quad (4.6)$$

What if the system is anisotropic? E.g. we have a unit cell that has different x/y lengths; then we can rescale via redefinition $y \rightarrow yb, x \rightarrow x/a$ to recover isotropy. The symmetries must then be preserved.

4.2 Minimal Model

We consider a scalar field m , where:

$$\beta H = \text{const} + \int d^d r \left\{ \frac{1}{2} t m^2(\mathbf{r}) + \frac{1}{4} u m^4(\mathbf{r}) + \frac{1}{2} k (\nabla m)^2 - h m \right\} \quad (4.7)$$

Where we note the model has inversion symmetry up to the external magnetic field term. At the moment, the constants appearing in the model are all phenomenological, e.g. $t = t(T, B, P, \dots)$. For now, we don't care about the parameters precisely - we are just trying to see if there is any interesting behaviour generically interesting in theory. Are there any points in parameter space where it changes its behaviour? This is quite a different approach from trying to "solve" a specific microscopic model.

For this theory to be sensible, we must have $u > 0$; else the magnetization will grow to infinity (as $m \rightarrow \infty$ would then prove to be the lowest energy configuration), and there is nothing to stabilize the model (if $u < 0$ we require a m^6 term to stabilize it). $k > 0$ is also known from the phonons in the model; it gives the dispersion at long wavelengths; if $k < 0$, it would disperse downwards, and at finite wavevectors the model would be unstable.

4.3 Mean Field Theory from Saddle Point Approximation

The next question - a technical one - is what does $\mathcal{D}[m(\mathbf{r})]$ mean? The idea will be to discretize:

$$\int \mathcal{D}[m] F\left(m, \frac{\partial m}{\partial x}, \dots\right) = \lim_{N \rightarrow \infty} \prod_{i=1}^N \int dm_i F\left[m_i, \frac{m_{i+1} - m_i}{\delta}, \dots\right] \quad (4.8)$$

Large fluctuations have a high energy cost (due to the gradient term) and thus have low statistical weight.

The first thing we are tempted to do when minimizing h is to choose configurations where m is uniform; the zeroth order guess is to say that $m(\mathbf{r}) = m + \delta m(\mathbf{r})$, where $\delta m(\mathbf{r})$ are the fluctuations we return to in a moment. We then have:

$$Z \propto \text{Const} \int dm \exp\left(-V \left(\frac{1}{2} t m^2 + \frac{1}{4} u m^4 - h m\right)\right) \quad (4.9)$$

where V is the volume. Now, this is an easy integral; since V is large, the dominant term is the minimum of the function in the exponential. We can now compute the free energy:

$$F = -\frac{1}{\beta} \ln Z \pm \text{Const} + V \min_m [\Phi(m)] \quad (4.10)$$

An aside; there are all kinds of constants that flow around in this problem due to taking limits and large numbers of integrals. We can ignore these, because all multiplicative constants that go into the partition function end up going into an additive constant into F which we can drop (it just sets the zero energy).

We minimize Φ , so the condition on the minimal m is:

$$tm + um^3 - h = 0. \quad (4.11)$$

Sign of t	Minimizing m	Critical Exponents	Specific Heat
$t > 0$	$m = h/t$	$\alpha \sim \frac{1}{t}, \gamma = 1$	0
$t < 0$	$m = \pm \sqrt{\frac{t}{m}}$	$\beta = 1/2$	$\frac{1}{2u}$
$t = 0$	$m = \left(\frac{h}{u}\right)^{1/3}$	$\delta = 3$	

Table 1: Table of minimizing m and Critical exponents depending on t .

If $t > 0$ then we have one solution, if $t < 0$ then we have broken symmetry and we have two solutions. The specific heat is obtained from $C = -T \frac{\partial^2 F}{\partial T^2}$, where we note that:

$$\beta F = \begin{cases} 0 & t > 0 \\ -\frac{t^2}{u} & t < 0 \end{cases} \quad (4.12)$$

and we note that there is a discontinuity at $t = 0$.

Note that there are models where this is exact. Consider a system with finite range interactions, where the interactions drop off at some length scale R_0 .

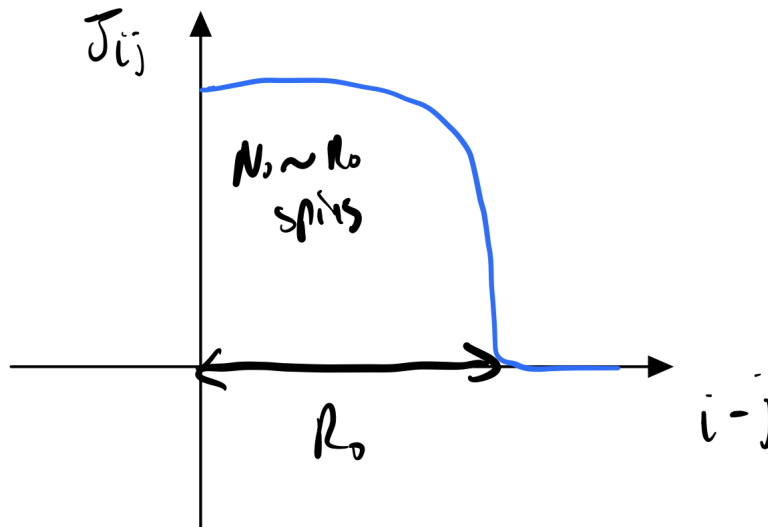


Figure 4.1: Plot of the interactions/couplings J_{ij} , where the interactions are the same up to some long length scale R_0 .

In the Ising model, this looks like:

$$H = \sum_{ij} J S_i S_j = \frac{J}{N_0} \sum_{|i-j| < R_0} S_i S_j \quad (4.13)$$

This looks like a coarse grained version:

$$H = \sum_i S_i h_i \quad (4.14)$$

where:

$$h_i = \frac{J_0}{N_0} \sum_i S_i = J_0 \langle S \rangle \quad (4.15)$$

In the $N_0 \rightarrow \infty$ limit, this theory becomes correct (i.e. long range interactions get us closer to mean field theory, which makes sense! If all spins interact with each other, then the mean field literally is the physical field felt). A variant of this was used to win the Nobel prize last week (We will discuss the Hopfield model later in the course - it is a ferromagnet).

Note that the saddle point approximation works in a high number of dimensions, but it generally fails in a small number of dimensions - despite the dominance of the saddle point, there are many ways of making fluctuations; hence entropy comes in, competing with the energetics argument we have made here.

4.4 Fluctuations

We have two classes of models we consider, characterized by their symmetries.

Continuous (Rotational) Symmetries

These are found, e.g., in X-Y models, Heisenberg models, superconductors. These generically have symmetries like $U(1), O(2), SO(2)$. The order parameters have amplitude and a phase, $\psi(x) = |\psi_0| e^{i\theta(x)}$. At the mean field level, we can write $|\nabla m|^2 = m_0^2 |\nabla \Theta|^2$. We consider Hamiltonians with fluctuations of the form:

$$\beta H = \text{Const} + \int d^d x \left\{ \frac{1}{2} k |\nabla \psi|^2 \right\} \rightarrow |\psi_0|^2 \int d^d x \frac{1}{2} k (\nabla \Theta)^2 \quad (4.16)$$

The stiffness is $\frac{1}{2} k |\psi_0|^2$, where ψ_0 is the order parameter. Because this vanishes above the transition, the modes do not have stiffness above the critical point; they do below, with the stiffness vanishing as they approach the critical temperature. We can rewrite this as:

$$\frac{1}{2} |\psi_0|^2 k \sum_a a |\Theta_a|^2 \quad (4.17)$$

If we start populating these modes (as if they are bosons), we can get a very large number of them. We saw this in the AFM spin chain. Something similar will happen here.

Why are they called soft modes? If I am far below T_c , the slope in the dispersion relation is finite. As I approach T_c in the mean field model, the slope goes to zero as the order parameters go to zero. Hence, they are "soft".

Discrete (\mathbb{Z}_2) symmetry

Next, we consider;

$$\beta H = \int d^d x \left[\phi = \frac{1}{2} t m^2 + \frac{1}{4} u m^4 + \frac{1}{2} k (\nabla m)^2 \right] \quad (4.18)$$

To which we can consider a more general E-L equation to find the extremum:

$$k \nabla^2 m = t m + u m^3 \quad (4.19)$$

We can now consider solutions that start into one well and cross into another well. We can solve this explicitly:

$$m = m_0 \tanh\left(\frac{x - x_0}{\omega}\right) \quad (4.20)$$

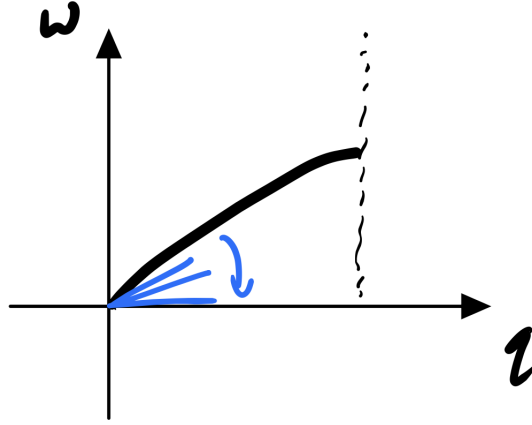


Figure 4.2: Depiction of “soft modes” in the dispersion relation, where as the temperature approaches the critical temperature below the slope in the dispersion (and the “stiffness”) approaches zero.

where the width is:

$$\omega = \sqrt{\frac{2k}{-t}} \quad (4.21)$$

This is another stationary solution; we previously found the lowest energy solution, this is a higher-lying stationary solution. As $t \rightarrow 0$, the domain wall becomes very long and the energy goes to zero. Notably, this is a particle-like excitation, with a lot of degrees of freedom as there are many points where it could be placed. Thus, we have to sum over all possible places to put 1, 2, \dots , N domain walls. Each one has finite/small energy, but there are very many of these, so the entropic contribution is likely to win.

Next class, we will discuss how to measure these fluctuations (i.e. scattering experiments).

5 Fluctuations, Continued

5.1 Phase and Amplitude Fluctuations

We consider ϕ (the effective Hamiltonian/energy):

$$\phi(\mathbf{m}) = \frac{\kappa}{2} |(\nabla \mathbf{m})|^2 + \frac{t}{2} m^2 + \frac{u}{4} m^4 \quad (5.1)$$

where \mathbf{m} is a order parameter with magnitude $m = |\mathbf{m}(\mathbf{x})|$ but also a phase (as it has a direction).

Looking at the potential, we have massless “Goldstone” modes that go between the two wells and the massive modes that oscillate inside of a given well.

We wish to find the probability distribution:

$$\mathcal{P}[\mathbf{m}(\mathbf{x})] = \exp\left(-\int d^d x \phi(\mathbf{m}, \mathbf{x})\right). \quad (5.2)$$

Let us notate $\langle m \rangle = \bar{m}$ and then write:

$$m = [\bar{m} + \phi_l(\mathbf{x})] \hat{e}_1 + \sum_{\alpha=2}^n \phi_{t,\alpha}(\mathbf{x}) \hat{e}_\alpha \quad (5.3)$$

Where \hat{e}_1 is a vector parallel to the magnetization, and \hat{e}_α for $\alpha = \{2, \dots, n\}$ are vectors perpendicular to the magnetization and mutually orthogonal. The ϕ_l, ϕ_t appearing above are the longitudinal and transverse fluctuations, respectively.

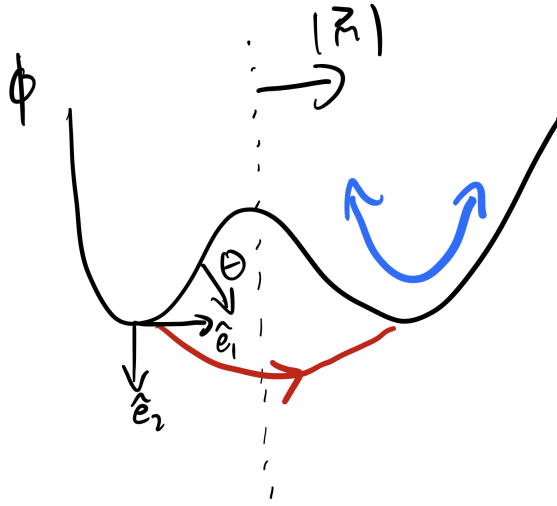


Figure 5.1: We consider a potential/energy ϕ that depends on an order parameter $\mathbf{m}(\mathbf{x})$. The order parameter has magnitude $m = |\mathbf{m}|$ and direction θ . There are massless Goldstone modes (red) that go between the wells of the potential, and massive modes that stay within a potential. For a given direction of the order parameter, we choose locally a coordinate system such that \hat{e}_1 points along the direction of magnetization and \hat{e}_2 (and other vectors) to be mutually orthogonal. This plot is for $t < 0$.

We consider the mean field solution:

$$\bar{m} = \begin{cases} 0 & t > 0 \\ \sqrt{-\frac{t}{u}} & t < 0 \end{cases} \quad (5.4)$$

Which we then obtain:

$$-\ln \mathcal{P} = V \left[\frac{t}{2} \bar{m}^2 + \frac{u}{4} \bar{m}^4 \right] + \int d^d x \left(\frac{\kappa}{2} |\nabla \phi_l|^2 + \left(\frac{t + 3u\bar{m}^2}{2} \right) \phi_l^2 \right) + \int d^d x \left(\frac{\kappa}{2} |\nabla \phi_t|^2 + \left(\frac{t + u\bar{m}^2}{2} \right) \phi_t^2 \right) + O(\phi_{l/t}^4) \quad (5.5)$$

Let us rescale:

$$\frac{\kappa}{\bar{\zeta}_l^2} = t + 3u\bar{m}^2 = \begin{cases} t & t > 0 \\ -2t & t < 0 \end{cases} \quad (5.6a)$$

$$\frac{\kappa}{\bar{\zeta}_t^2} = t + u\bar{m}^2 = \begin{cases} t & t > 0 \\ 0 & t < 0 \end{cases} \quad (5.6b)$$

Where the $\bar{\zeta}$ are called correlation lengths. Above the mean-field phase transition they are the same. Below, they are different. The 0 is the Goldstone mode (gapless, linear) and the $-2t$ is the massive mode (gapped, looks like a harmonic oscillator at long wavelengths).

We now have translation invariance, so a natural change of basis is to go into Fourier/momentum space:

$$\phi(x) = \sum_{\mathbf{q}} \phi_{\mathbf{q}} \frac{e^{i\mathbf{q} \cdot \mathbf{x}}}{\sqrt{V}} \quad (5.7)$$

We then find:

$$\mathcal{P}[\phi_{l\mathbf{q}}, \phi_{t\mathbf{q}}] \propto \prod_{\mathbf{q}, \alpha} \exp\left(-\frac{\kappa}{2} \left(\mathbf{q}^2 + \xi_\alpha^{-2}\right) |\phi_{\alpha\mathbf{q}}|^2\right) \quad (5.8)$$

We can now start to compute objects of interest, e.g. correlation functions (between fluctuations). We call this the structure factor:

$$S(\mathbf{q}) = \langle \phi_{\alpha\mathbf{q}} \phi_{\beta\mathbf{q}'} \rangle = \frac{\delta_{\alpha\beta} \delta_{\mathbf{q}, -\mathbf{q}'}}{\kappa(q^2 + \xi_\alpha^2)} \quad (5.9)$$

The decoupling of longitudinal/transverse gives the first delta function and the second comes from momentum conservation. The denominator comes from the computation of the Gaussian integrals. Plotted, it looks like:

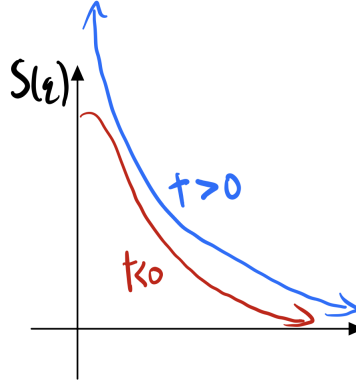


Figure 5.2: Plot of structure for transverse modes as a function of q . For $t < 0$ we have $\xi_t = 0$ and so the structure factor diverges at 0. For $t > 0$, ξ_t is finite so the structure factor at 0 does not diverge. The structure factor for the longitudinal fluctuations has the same general shape as the $t > 0$ transverse fluctuations (for $t < 0$ and $t > 0$; ξ_l is finite for both regimes, just with a different prefactor. For $t > 0$ it is identical to the transverse fluctuations).

What is the interpretation of $S(\mathbf{q})$? We have some object of interest, which we probe via a scattering experiment. We have some probe, which could be an electron, neutron, photon etc. It has an incident wavevector \mathbf{k}_i and a output wavevector \mathbf{k}_f , emerging at some angle θ to the incident. If we have elastic scattering, then the energy of the particle remains unchanged, but there is some changed momenta, i.e. $\mathbf{k}_f = \mathbf{k}_i + \mathbf{q}$. We then study the scattering amplitude:

$$A(q) \propto \langle \mathbf{k}_f | \mathbf{U} | \mathbf{k}_i \rangle \propto \sigma(\mathbf{q}) \int d^d x e^{i\mathbf{q} \cdot \mathbf{x}} \rho(\mathbf{x}) \propto \rho_{\mathbf{q}} \quad (5.10)$$

Note that the average over the scattering amplitude vanishes:

$$\langle A(q) \rangle = \langle \rho_{\mathbf{q}} \rangle = 0 \quad (5.11)$$

So to probe the structure we want to study:

$$S(q) = \langle |A(\mathbf{q})|^2 \rangle \propto \langle \rho_{\mathbf{q}} \rangle \quad (5.12)$$

5.2 Technical Aside: Gaussian Integrals

Consider the single variable integral:

$$I_1 = \int_{-\infty}^{\infty} d\phi e^{-\frac{1}{2}k\phi^2 + n\phi} \quad (5.13)$$

We solve via completing the square and then redefinition of the integration variable to $\phi' = \sqrt{k}(\phi - \frac{h}{k})$. What is left with just the familiar Gaussian integral, which we should be familiar with the result of:

$$I_1 = \int_{-\infty}^{\infty} d\phi e^{-\frac{1}{2}k(\phi - \frac{h}{k})^2 + \frac{h^2}{k}} = \sqrt{\frac{2\pi}{k}} e^{\frac{h^2}{2k}} \quad (5.14)$$

Now consider $\frac{dI_1}{dh}$, which via differentiating under the integral sign we know is equivalent to $\langle \phi \rangle$.

$$\frac{dI_1}{dh} = \int d\phi \phi e^{-(\dots)} = \langle \phi \rangle = \frac{h}{k} \quad (5.15)$$

This integral is correct up to a constant factor. We don't have to care about these constants, as eventually we take the logarithm of the partition function, and this becomes an additive constant which has no effect on the physics. We can also look at the second derivative:

$$\frac{d^2 I_1}{dh^2} = \langle \phi^2 \rangle = \frac{h}{k} + \frac{h^2}{k^2} \quad (5.16)$$

We can also look at the cumulant:

$$\langle \phi^2 \rangle_c = \langle \phi^2 \rangle - \langle \phi \rangle^2 = \frac{h}{k}. \quad (5.17)$$

In principle we could go to higher order cumulants, but you can verify for yourself that these would all vanish.

Now, let's look at the N -dimensional version:

$$I_N = \int_{-\infty}^{\infty} \prod_{i=1}^N d\phi_i \exp\left(-\frac{1}{2} \sum_{ij} \phi_i K_{ij} \phi_j + \sum_i n_i \phi_i\right) = \sqrt{\frac{(2\pi)^N}{\det K}} \exp\left(\sum_{ij} h_i \frac{K_{ij}^{-1}}{2} h_j\right) \quad (5.18)$$

How would we obtain this? K is a general, presumably symmetric matrix. We diagonalize K , and suppose it has eigenvalues $K_{ij} \hat{q}_j = K_q \hat{q}_j$. We use the \hat{q}_j as the canonical basis, and rotate into this basis by saying $\phi_i = \sum_q \tilde{\phi}_q \hat{q}_i$. Thus the integral becomes:

$$I_N = \int \prod_{q=1}^N d\tilde{\phi}_q \exp\left(-\frac{K_q}{2} \phi_q^2 + h_q \tilde{\phi}_q\right) = \prod_{q=1}^N \sqrt{\frac{2\pi}{K_q}} \exp\left(h_q \frac{K_q^{-1}}{2} h_q\right) \quad (5.19)$$

where now recognizing the determinant is just the product of the eigenvalues, and rotating back into the eigenbasis, gives the desired expression.

Remark: This relation is very useful, and it works out for QFTs as well, with some i s floating around.

5.3 Back to Physics - Transverse Fluctuations and Coloumb Emergence

Now, we make the remark that all Gaussian theories are *solvable* - they may be wrong, but they are solvable. In other words, mean field theory + fluctuations to lowest order (i.e. quadratic) are completely solvable analytically, and are known as Gaussian theories. They are simple, but already can give us interesting phenomena.

For example, for the structure factors appearing in Eq. (5.9), we were able to find this just with a Gaussian theory, and it tells us that the long wavelength/small q is the interesting regime where things are happening. We can look at the exponent η that tells us how the structure factor scales.

$$S(q, t = 0) \sim \frac{1}{q^{2-\eta}} \quad (5.20)$$

Let us focus on the transverse fluctuations, below the critical temperature, so $t < 0$ with $\eta_t = 0$. Then, we have:

$$\mathcal{P}(\Theta(q)) \propto \exp\left(-\frac{\kappa}{2} \sum_q q^2 |\Theta_q|^2\right) \quad (5.21)$$

So without computed structure factor, let us look at the probability fluctuations in the angles:

$$\langle \Theta_{\mathbf{q}} \Theta_{\mathbf{q}'} \rangle = \frac{\delta_{\mathbf{q}+\mathbf{q}'}}{\kappa q^2} \quad (5.22)$$

Going back into real space, this is the statement that:

$$\langle \Theta(\mathbf{x}) \Theta(\mathbf{x}') \rangle = \frac{1}{V} \sum_{\mathbf{q}} e^{i\mathbf{q}(\mathbf{x}-\mathbf{x}')} \kappa q^2 = \int \frac{d^d q}{(2\pi)^d} \frac{e^{i\mathbf{q} \cdot (\mathbf{x}-\mathbf{x}')}}{q^2} \quad (5.23)$$

This is just the Coloumb potential:

$$C(\mathbf{x}) = - \int \frac{d^d q}{(2\pi)^d} \frac{e^{i\mathbf{q} \cdot \mathbf{x}}}{q^2} \quad (5.24)$$

Which satisfies $\nabla^2 C = \delta^d(\mathbf{x})$, i.e. just Gauss' Law for a point charge. Now, we can ask what happens to the potential as $x \rightarrow \infty$?

$$\lim_{x \rightarrow \infty} C(x) \propto \begin{cases} \text{Const} & d > 0 \\ x^{2-d} & d < 2 \\ \ln(x) & d = 2 \end{cases} \quad (5.25)$$

This behaviour can be obtained by looking at $d^d q \sim q^{d-1} dq$ (plus various angular integrals which we don't care about).

Thus, in low dimensions the correlations fall off like a power law, so no long range order. For $d > 2$ we maybe have LRO. At $d = 2$, we have the critical/marginal dimension. We have a pesky logarithm, and we need more work (which we will study in future lectures).

Earlier on, we explained how there were models (AFM Heisenberg) that did not order at low dimensions due to the presence of soft modes. This is a souped up version of that. So, when generally looking at a theory, if I look at a theory with Goldstone modes, I should look at the phase fluctuations and see if I get LRO. Even though in principle the length scales can get very large, we can still say that at very large scales that the correlations fall off. Also, its worth noting that the LRO is temperature dependent; we dropped the β s through the calculations here, but it is relevant.

6 Scaling Theory

6.1 Recap

To recap; we are studying Gaussian theory, in particular ϕ^4 theory of a magnet. This is mean field theory + quadratic fluctuations, which is exactly solvable, but can still give us insight.

We looked at correlation length associated to longitudinal and transverse fluctuations. We studied the effect of the transverse fluctuations/Goldstone mode, and below $2-d$ the correlations die off and we have no LRO. This was the lower critical dimension. We will now study an upper critical dimension, which will turn out to be $d = 4$.

We look at the logarithm of the partition function:

$$\ln Z = \frac{t}{2} \bar{m}^2 + \frac{u}{4} \bar{m}^4 + \int d^d x \left[\frac{1}{2} \kappa (\nabla \phi_l)^2 + \bar{\zeta}_l^{-2} \phi_l^2 \right] + \int d^d x \left[\frac{1}{2} \kappa (\nabla \phi_t)^2 + \bar{\zeta}_t^{-2} \phi_t^2 \right] \quad (6.1)$$

The correlation lengths have the scaling:

$$\zeta_t = \begin{cases} \zeta_0 t^{-1/2} & t > 0 \\ (-2t)^{-1/2} & t < 0 \end{cases} \quad (6.2)$$

$$\tilde{\zeta}_t = \begin{cases} \zeta_0 t^{-1/2} & t > 0 \\ \infty & t < 0 \end{cases} \quad (6.3)$$

The important observation for the transverse modes is below the phase transition, the $\tilde{\zeta}_t \phi_t^2$ term is absent. We can do the integral by going into momentum space:

$$\int d^d x \left[\frac{1}{2} \kappa (\nabla \phi_t)^2 + \tilde{\zeta}_t^2 \phi_t^2 \right] = \int \frac{d^d q}{(2\pi)^d} \left[\frac{1}{2} \kappa (q^2 + \tilde{\zeta}_t^{-2}) \phi_t(q)^2 \right] \quad (6.4)$$

Then the partition function is:

$$Z = Z_{MF} e^{-\ln \det(\kappa(q^2 + \tilde{\zeta}_t^{-2}))} \quad (6.5)$$

Where this comes from:

$$Z = \int \mathcal{D}\phi_q e^{-\frac{1}{2} \int d^d q \kappa (q^2 + \tilde{\zeta}_t^{-2}) \phi_q^2} \quad (6.6)$$

And then we take the Gaussian integral. Then computing the correlators from this result (dropping $\tilde{\zeta}_t$ as we are $t < 0$);

$$\langle \phi_t(\mathbf{q}), \phi_t(\mathbf{q}') \rangle = \frac{\delta_{\mathbf{q}+\mathbf{q}'}}{\kappa \mathbf{q}^2} \implies \langle \phi_t(\mathbf{x}) \phi_t(\mathbf{x}') \rangle = \int \frac{d^d q}{(2\pi)^d} \frac{e^{i\mathbf{q} \cdot \mathbf{x} - \mathbf{x}'}}{\kappa \mathbf{q}^2} \quad (6.7)$$

from which we saw that:

$$\lim_{x \rightarrow \infty} C(x = |\mathbf{x} - \mathbf{x}'|) \propto \begin{cases} \text{Const} & d > 0 \\ x^{2-d} & d < 2 \\ \ln(x) & d = 2 \end{cases} \quad (6.8)$$

and found our lower critical dimension of 2.

6.2 Massive Modes and Upper Critical Dimension

We have the partition function of the longitudinal modes:

$$Z = Z_{MF} e^{-\ln \det(\kappa(q^2 + \tilde{\zeta}_t^{-2}))} \quad (6.9)$$

Let us look at the free energy, by taking the logarithm of Z:

$$F = F_{MF} + \frac{1}{2} \int \frac{d^d q}{(2\pi)^d} \ln(\kappa(\mathbf{q}^2 + t)) \quad (6.10)$$

Now looking at the heat capacity:

$$C = -T \frac{\partial^2 F}{\partial T^2} \sim -t \frac{\partial^2 F}{\partial t^2} \quad (6.11)$$

Looking at the contribution from the fluctuations:

$$C \sim \int \frac{d^d q}{(2\pi)^d} \frac{1}{(\kappa q + t)^2} \stackrel{t \rightarrow 0}{\sim} \int_0^{1/a} dq \frac{q^{d-1}}{q^4} \sim a^{4d} \quad (6.12)$$

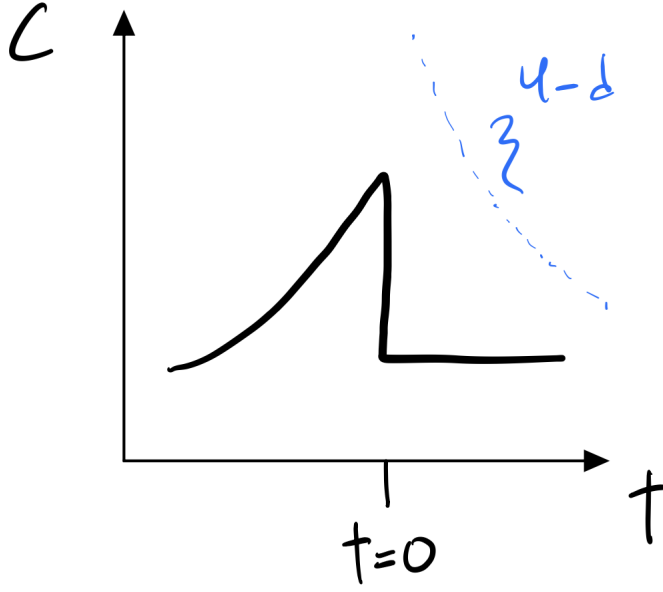


Figure 6.1: Plot of heat capacity for Gaussian theory.

we want to see how this integral behaves. For $d > 4$, we have the integral is well behaved at $q = 0$ and evaluates to $\sim a^{4d}$ as above. For $d < 4$, we need to keep the correlation length (else the integral does not converge) so:

$$\int_0^{1/a} dq \frac{q^{d-1}}{(q^2 + \xi_l^{-2})} \sim \xi^{4-d} \quad (6.13)$$

Plotting the heat capacity, we get:

So, we see that Gaussian theory fails for $d < 4$. Why? We said MF theory + small fluctuations. If the effect of small fluctuations is benign, then we are happy. If the effect of small fluctuations is large, then we aren't happy; the theory is not self-consistent. Below $d < 4$, the failure of the integral to converge (and the rapid change in behaviour of physical quantities due to the fluctuations) tells us that the theory does not apply.

So below $d = 2$ we have no LRO, and above $d = 4$ we know that MFT + Gaussian fluctuations is ok. The interesting regime left to explore is $d = 2$ to $d = 4$.

We also were studying the phase transition at $t = 0$ we can also ask about the magnitude of the fluctuations as we go away from the transitions.

In summary, we saw that the heat capacity from the mean field goes as $C_{MF} = \frac{1}{u}$ and the capacity of the fluctuations go as $C_F = \frac{\xi^{4-d}}{\kappa^2}$. The κ is a microscopic parameter that characterizes the stiffness of the bare mode, with dimensional analysis from $\kappa \nabla^2$ telling us that κ has a dimension of length square, i.e. $\kappa \sim \xi_0^2$. Thus:

$$C_F \sim \frac{\xi_0^{4-d} |t|^{-\frac{4-d}{2}}}{\xi_0^4} = \frac{1}{\xi_0^d} t^{-\frac{4-d}{2}} \quad (6.14)$$

so indeed the t piece diverges, but we have the ξ_0 piece characterizing the microscopic correlation length. So, if we have a system with a very long microscopic correlation length, we have to get very close to the transition $t = 0$ to see the divergence. This defines a temperature where the fluctuations take over and the theory fails. This temperature is typically known as the *Ginzburg temperature*.

There are systems, typically superconductors, where the microscopic correlation length is very long. So, despite the fact that the correlations diverge, we have to get very close to the transitions to see these diverging correlations.

6.3 Scaling Theory

Is it going to be the case where I need a theory that calculates a new parameter for each fluctuating quantity? Or are they all related? Our proposal is the latter. Let us make this statements precise

Scaling theory: Thermodynamic functions (e.g. the thermodynamic free energy F) are homogenous functions of scaled variables (e.g. T, h).

This hypothesis can be used to simplify the problem such that I may compute simple things. To be clear, we already had this in mean field theory. There, we looked at some (scaled) free energy which we viewed as the minimum of a power series expansion:

$$f(t, h) = \min_m \left[\frac{t}{2} m^2 + \frac{u}{4} m^4 - hm \right] \quad (6.15)$$

For example, this gave:

$$f(t, h) = \begin{cases} -\frac{t^2}{4u} & h = 0, t < 0 \\ -\frac{3}{4} \frac{h^4/3}{u^{1/3}} & h \neq 0, t = 0 \end{cases} \quad (6.16)$$

With a crossover line $h \sim |t|^{3/2}$ below which we have the first behaviour, and above which we have the second.

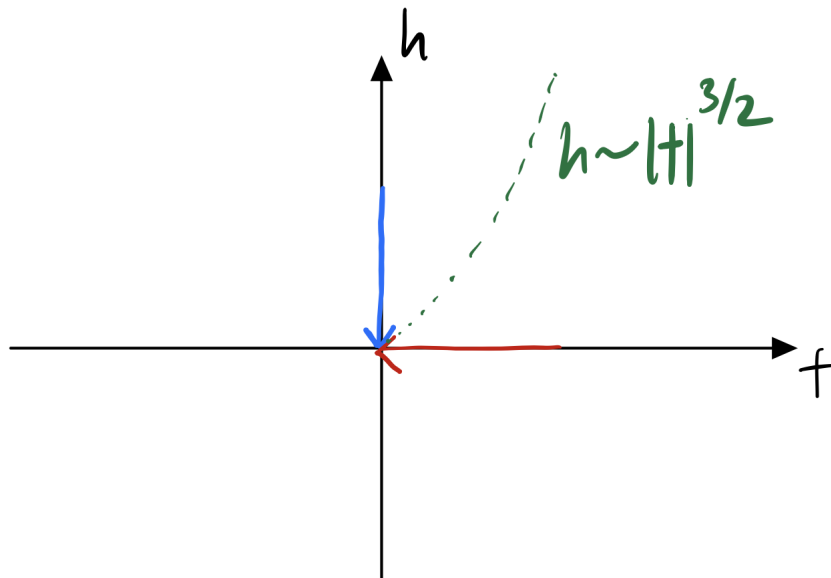


Figure 6.2: The free energy has a certain behaviour along the $h = 0$ and $t = 0$ lines. The crossover of these two behaviours occurs along $h \sim |t|^{3/2}$.

Rewriting this in another way, We have:

$$f(t, h) = |t|^2 g(h/|t|^\Delta) \quad (6.17)$$

where g is a homogenous function of $h/|t|^\Delta$. The previous calculations constrain g . E.g. at $g = 0$, we have:

$$g(0) = -\frac{1}{4u} \quad (6.18)$$

and

$$g(x \rightarrow \infty) \sim x^{4/3} \quad (6.19)$$

In the limit where h is large, we also know the answer is independent of t , therefore if we have:

$$f(t, h) = |t|^2 \left(\frac{h}{|t|^\Delta} \right)^{4/3} \quad (6.20)$$

the t -independence constrains $\Delta = 3/2$.

In MF theory, we know precisely what g is. In general with scaling analysis, we only know the asymptotic behaviour.

What we will do - we propose that there is a "better theory", where:

$$f_{\text{sing}}(t, h) = |t|^{2-\alpha} g(h/|t|^\Delta) \quad (6.21)$$

The singular component to the energy goes as:

$$\begin{aligned} E_{\text{sing}} &\sim \frac{\partial f}{\partial t} = (2 - \alpha)|t|^{1-\alpha} g(h/|t|^\Delta) - \Delta h |t|^{1-\alpha-\Delta} g'(h/|t|^\Delta) \\ &= |t|^{1-\alpha} \left((2 - \alpha)g(h/|t|^\Delta) - \Delta \left(\frac{h}{|t|^\Delta} \right) g'(h/|t|^\Delta) \right) \end{aligned} \quad (6.22)$$

where the term in brackets is another homogenous function, $g_E(h/|t|^\Delta)$. To get the heat capacity, we take the derivative again:

$$C_{\text{sing}} \sim -\frac{\partial^2 f}{\partial t^2} = |t|^{-\alpha} g_c(h/|t|^\Delta) \quad (6.23)$$

and we have the magnetization:

$$m(t, h) = \frac{\partial f}{\partial h} = |t|^{2-\alpha-\Delta} g_m(h/|t|^\Delta) \quad (6.24)$$

now, as $h \rightarrow 0$, we have $g_m \rightarrow \text{const}$, and we identify the critical exponent $\beta = 2 - \alpha - \Delta$. For large arguments, we propose $g_M(x) \sim x^P$, then:

$$m(t = 0, h) \sim |t|^{2-\alpha-\Delta} \left(\frac{h}{|t|^\Delta} \right)^P \quad (6.25)$$

so by the t -independence argument, we have:

$$2 - \alpha - \Delta = P\Delta \quad (6.26)$$

and then:

$$m(h) \sim h^{\frac{2-\alpha-\Delta}{\Delta}} \quad (6.27)$$

so then we identify the exponent:

$$\delta = \frac{\Delta}{2 - \alpha - \Delta} = \frac{\Delta}{\beta} \quad (6.28)$$

whose measurement allows us to infer Δ .

Why might this theory be true? Underlying this theory is a diverging correlation length, which gives you averaging, and is solely responsible for the singular fluctuations. The last thing we propose is that there is a correlation length ζ that is hidden behind all of this phenomenology, where:

$$\zeta \sim |t|^{-\nu} g_{\zeta} \left(\frac{h}{|t|^{\Delta}} \right) \quad (6.29)$$

where the divergence of ζ is responsible for the rest of the behaviour we have analyzed.

This is the hypothesis, and we have to work out this hypothesis in such a way that is self-consistently well behaved. The process we will go through (known as the *renormalization group*) will start with a Ginzburg-Landau theory, suggest there is a diverging length scale that is responsible for the phenomena, and then average that action on a characteristic length scale. We then scale. Then, supposing that the correlation length gets scaled, we can average again, and we want to force the averaging procedure to give us a function which is a homogenous function of the parameters. This is the process/constraint we put on scaling. We force the algebra to give us a scaling function like this. When we do this, there is a sequence of operations that will enable us to calculate these exponents. We can then check if the procedure is consistent, or not. This is what we will explore in future lectures.

7 Scaling Hypothesis Part II and Renormalization Group

7.1 Recap

We postulate a free energy:

$$f_{\text{sing}}(t, h) = |t|^{2-\alpha} g(h/|t|^{\Delta}) \quad (7.1)$$

where the exponents are different from those appearing in mean field theory, and g is a homogenous function of its argument. From this, we are able to derive for a scaling law for the magnetization m , for the specific heat C , the susceptibility χ , etc.

If this holds, we have a correlation length:

$$\zeta(t, h) = |t|^{-\nu} g_{\zeta}(h/|t|^{\Delta}). \quad (7.2)$$

In MF theory, $\nu = 1/2$, but it may be different here. We will find $\mu > 0$ so the correlation length diverges as $T \rightarrow T_c$.

We assume that $\ln Z$ is extensive, i.e. $\propto L^d$ (it is very rare that this would not be the case). More specifically, $\ln Z \propto \frac{L^d}{\zeta^d}$ plus any nonsingular terms. Then the free energy goes as:

$$f_{\text{sing}} = \frac{\ln Z}{L^d} g(h/|t|^{\Delta}) \propto |t|^{d\nu} g(\cdot) \quad (7.3)$$

which implies:

$$2 - \alpha = d\nu \quad (7.4)$$

This is known as the hyperscaling, or Josephson relation. It doesn't work in MF theory, unless $d = 4$. In one of the homework problems, we will muse on why this is.

We study the correlation function:

$$G(\gamma, 0) = \langle m(\gamma)m(0) \rangle - \langle m \rangle^2 \sim \frac{1}{|\gamma|^{d-2+\eta}} \quad (7.5)$$

We now study the renormalization group, which allows us to mathematically obtain these exponents. We start with a conceptual view.

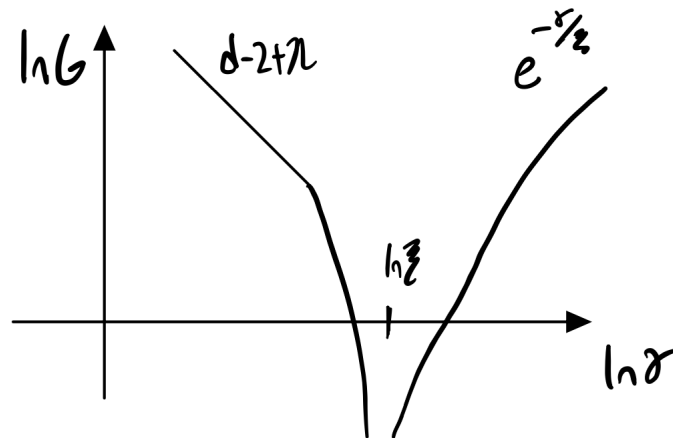


Figure 7.1: At long distances, the correlation functions die off exponentially. Near the critical point, it is a power law.

7.2 Renormalization Group: Conceptual View

Step 1: Coarse Graining

We do RG via a process of coarse graining. We take a length scale x , and convert it to bx where $b > 1$. Pictorially, we look at a $bx \times bx$ block of spins and average them to form an “averaged spin” corresponding to that whole block. I.e. the magnetization as a function of x becomes:

$$m(x) = \frac{1}{b^d} \int d^d x' m(x') \quad (7.6)$$

where we integrate over the cell center at x .

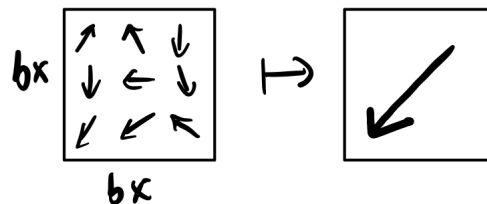


Figure 7.2: At the coarse graining step, we integrate over a block of spins and replace it with the average.

Notice that we lose information when we do this. Via course graining, we lose the microscopics.

Step 2: Rescale

We now take $x_{\text{new}} = \frac{x_{\text{old}}}{b}$, rescaling/blowing up the new cell size to be the size of the original unit cell.

Step 3: Renormalize

We renormalize the magnetization:

$$m_{\text{new}}(x_{\text{new}}) = \frac{1}{\bar{\xi} b^d} \int d^d x' m(x') \quad (7.7)$$

Where to make the theory look the same as when we started, we introduce a new renormalization parameter $\bar{\xi}$.

7.3 Getting Exponents out of RG

Having done this, we can construct a new probability distribution:

$$W[m_{\text{new}}] = e^{-\beta H_b[m_{\text{new}}]} \quad (7.8)$$

We will assert that at criticality ($t = h = 0$), the Hamiltonian is statistically similar to the one I had originally. This means of course that in this theory, we introduce a new effective temperature and field that must be a function of the old ones. I.e.:

$$\begin{aligned} t_{\text{new}} &= f(t_{\text{old}}, h_{\text{old}}) \\ h_{\text{new}} &= f'(t_{\text{old}}, h_{\text{old}}) \end{aligned} \quad (7.9)$$

but we have some constraints. $(0,0) \rightarrow (0,0)$, and $t_{\text{new}}, h_{\text{new}}$ should grow. Since we rescale the length, the correlation length should get shorter, so we grow away from the 0 point. We also hope that there is no singular behaviour in these functions. Finally, we want to preserve the symmetry; as such, we will not mix t with h . Then:

$$\begin{aligned} t_b(t, h) &= A(b)t + O(t^2) \\ h_b(t, h) &= B(b)h + O(h^2) \end{aligned} \quad (7.10)$$

We know that:

$$A(1) = B(1) = 1 \quad (7.11)$$

(This corresponds to no scaling). There is also the assumption that this could be repeated:

$$t_{b_1 b_2}(t, h) = A(b_1 b_2)t = A(b_1)A(b_2)t \quad (7.12)$$

In other words, scaling once by $b_1 b_2$ should be the same as scaling by b_2 and then b_1 . This implies that:

$$A(b) \sim b^{y_t} \quad B(b) \sim b^{y_h} \quad (7.13)$$

Notice that this is a semigroup, as we can only go one way. So, this isn't actually a group (Renormalization group is a bit of a misnomer). This is how we can see the "loss in information" of this process.

We also assume that $y_t, y_h > 0$. So, we go further from the critical point. This might cause us concern because then the higher order t^2, h^2 terms in the scaling may become relevant. But indeed we take $b \sim 1$ in order to stay within a neighbourhood of the critical point, as this is the only regime in which this analysis is actually valid. Then, looking at the the scaling of the general functions:

$$X(t, h) = b^{y_x} X(b^{y_t} t, b^{y_h} h) \quad (7.14)$$

and the new configuration becomes:

$$Z = \int \mathcal{D}m W[m] \quad (7.15)$$

but the weights of the old configuration should be the same as the old one, so:

$$Z = \int \mathcal{D}m W[m] = \int \mathcal{D}m' W[m'] = Z' \quad (7.16)$$

Which then leads us to conclude that:

$$Vf(t, h) = V'f(t', h') \quad (7.17)$$

So then looking at the scaling of the free energy:

$$f(t, h) \sim \frac{1}{b^d} f(b^{y_t} t, b^{y_h} h) \quad (7.18)$$

We now have to propose a trajectory to scale such that we eventually get to the form $|t|^{2-\alpha} g(h/|t|^\Delta)$. We take $b^{y_t} t \sim 1$ so $b \sim t^{-1/y_t}$. Then:

$$f(t, h) = \frac{1}{b^d} f(1, h/t^{y_h/y_t}) \quad (7.19)$$

So, we are taking the free energy, positing a scaling form, and we get out something that looks like Eq. (7.3) as we desired.

From this, we obtain:

$$2 - \alpha = \frac{d}{y_t} \quad (7.20)$$

$$\Delta = \frac{y_h}{y_t} \quad (7.21)$$

What we will do is find a scaling path and determine the scaling exponents y_h, y_t , and then this tells us about the critical exponents of our interest. Let's look at the others.

There will now be a new correlation length:

$$\chi' = \chi/b \quad (7.22)$$

Looking at the general form of the scaling of the correlation length:

$$\chi(t, h) = b\chi(b^{y_t} t, b^{y_h} h) = t^{-1/y_t} \zeta(1, h/t^{y_h/y_t}) \quad (7.23)$$

and thus:

$$\nu = -\frac{1}{y_t} = \frac{2 - \alpha}{d} \quad (7.24)$$

which is the Josephson scaling relation.

Let us look at the magnetization as well"

$$m(t, h) = \frac{1}{V} \frac{\partial \ln Z}{\partial h} = b^{y_h - d} m(1, h/t^{y_h/y_t}) \quad (7.25)$$

so then:

$$\beta = \frac{d}{y_t} - \frac{y_h}{y_t} = 2 - \alpha - \Delta \quad (7.26)$$

7.4 RG Fixed Points and Flows

To start, we write down the most general Hamiltonian allowed by symmetry. For example here:

$$\beta H = \int d^d x \left[\frac{1}{2} t m^2 + \frac{1}{4} u m^4 + \frac{1}{6} v m^6 + \frac{1}{2} K (\nabla m)^2 + \frac{1}{2} L (\nabla^2 m)^2 + \dots \right] \quad (7.27)$$

I now have a parameter space $S(t, u, v, K, L, \dots)$ in which the parameters define a point.

Then, we'll go through the RG process. We course grain by averaging over block size b , rescale $x \rightarrow x/b$, then normalize by ζ . We then have the magnetization:

$$m'(x') = \frac{1}{\zeta b^d} \int d^d x m(x) \quad (7.28)$$

We then construct $P[m']$ from $P[m]$ (nontrivial process!). We then get a rescaled Hamiltonian $H'(t', u', v', K', L', \dots)$. This process is the same as saying that:

$$S' = R_b S \quad (7.29)$$

i.e. we have applied a map to our parameter space.

We consider a deviation:

$$S'_\alpha + \delta S'_\alpha = S^*_\alpha + (R_b)_{\alpha\beta} \delta S'_\beta \quad (7.30)$$

where S^* is the fixed point of the RG theory (imagine we have done the calculation and found this point/the critical point), i.e.:

$$R_b S^* = S^* \quad (7.31)$$

and $\zeta(S^*) = b\zeta(R_b S^*)$ i.e. $\zeta \rightarrow \infty$. We can write the renormalization group map (the L denoting the linearized version of the map, which we assume we can do in the vicinity of the critical point) as:

$$(R_b^L)_{\alpha\beta} = \frac{\partial S'_\alpha}{\partial S_\beta} \quad (7.32)$$

i.e. the RG takes us from S_α to S'_α . Now, consider this map to have eigenvectors Θ_i and eigenvalues $\lambda(b)_i$. Then:

$$R_b^L R_b^L \Theta_i = \lambda_i(n) \lambda_i(b') = \lambda_i(b'b'). \quad (7.33)$$

thus:

$$\lambda_i(b) \sim b^{y_i} \quad (7.34)$$

Therein, the Θ_i correspond to scaling directions and the y_i correspond to anomalous dimensions (in MF theory they are all 1 - not anomalous). We now expand S around the fixed point based on these scaling directions:

$$S = S^* + \sum_i g_i \Theta_i \quad (7.35)$$

After we scale:

$$S' = S^* + \sum_i g_i b^{y_i} [\Theta] \quad (7.36)$$

Pictorially, we have the flows in Fig. 7.3.

The vectors define a basin of attraction. Relevant means the eigenvectors go away from the fixed point, so the fluctuations grow and become important. Irrelevant mean the eigenvectors go towards the fixed point, so the fluctuations vanish (hence irrelevant). Finally, we have $y_i = 0$ and then we call the flows "marginal". These directions correspond to some linear combinations to linear combinations of the parameters t, u, v, K, L, \dots of the theory. The main takeaway is that this flow analysis tells us what variables are the important ones.

Next Wednesday, we will go through this process for the Gaussian theory, and identify the Wilson-Fisher fixed point.

8 Renormalization Group Part II

Today, we will go through the RG procedure on a model which is exactly solvable.

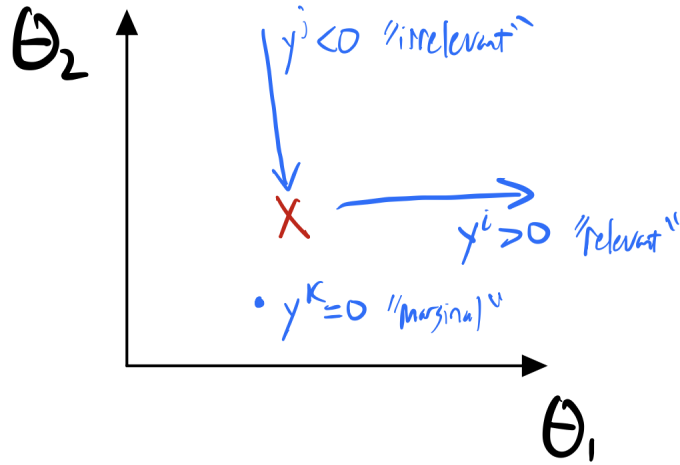


Figure 7.3: Different eigenvectors have different flow properties. Irrelevant eigenvectors go towards the fixed point, relevant eigenvectors flow away, and marginal eigenvectors are stationary.

8.1 Exact Solution of Gaussian Model

Consider the following partition function for a Gaussian model:

$$Z = \int \mathcal{D}m(x) \exp \left(- \int d^d x \left(\frac{1}{2} t m^2 + \frac{1}{2} k (\nabla m)^2 + \frac{1}{2} L (\nabla^2 m)^2 - m h \right) \right) \quad (8.1)$$

We go into momentum space:

$$m(\mathbf{q}) = \int d^d x e^{i\mathbf{q}\cdot\mathbf{x}} m(\mathbf{x}) \quad (8.2)$$

Then:

$$m(\mathbf{x}) = \sum_{\mathbf{q}} \frac{e^{-i\mathbf{q}\cdot\mathbf{x}}}{V} m(\mathbf{q}) = \int \frac{d^d q}{(2\pi)^d} m(\mathbf{q}) \quad (8.3)$$

Note the factor of volume to ensure the average magnetization does not scale with the size of the system. We then have:

$$\beta H = \sum_{\mathbf{q}} \left(\left[\frac{t + kq^2 + Lq^4}{V} \right] |m(\mathbf{q})|^2 - h m(\mathbf{q}) \delta_{\mathbf{q}0} \right) \quad (8.4)$$

Then:

$$Z = \prod_{\mathbf{q}} \left(\frac{1}{\sqrt{V}} \right)^n \int dm_{\mathbf{q}} \exp \left(- \left(\frac{t + kq^2 + Lq^4}{V} \right) |m(\mathbf{q})|^2 + h m(\mathbf{q}) \delta_{\mathbf{q}0} \right) \quad (8.5)$$

with n the number of components. Then, factoring out the $\mathbf{q} = 0$ term:

$$Z_0 = \left(\frac{2\pi}{t} \right)^{n/2} \exp \left(\frac{Vh^2}{2t} \right) \quad (8.6)$$

And the rest of the partition function is:

$$Z_{\text{rest}} = \prod_{\mathbf{q} \neq 0} \left(\frac{2\pi}{t + kq^2 + Lq^4} \right)^{n/2} \quad (8.7)$$

Z is then:

$$Z = \exp\left(\frac{Vh^2}{2t}\right) \prod_{q \neq 0} \left(\frac{2\pi}{t + kq^2 + Lq^4} \right)^{n/2} \quad (8.8)$$

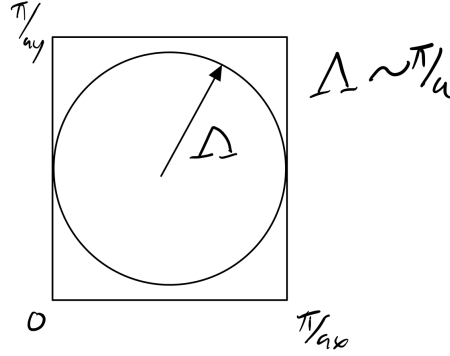
Thus:

$$F = -\frac{\ln Z}{V} = -\frac{h^2}{t} + n \int \frac{d^d q}{(2\pi)^d} \ln(t + kq^2 + Lq^4) \quad (8.9)$$

What is the range of this integral? In general, we have a lattice, and we would integrate over a Brouillin zone. We approximate this region of integration as over a hypersphere with radius $\Lambda \sim \pi/a$:

$$F = -\frac{h^2}{2t} + \frac{n\Omega_d}{2} \int_0^\Lambda dq q^{d-1} \ln(t + kq^2 + Lq^4) \Omega_d \quad (8.10)$$

where $\Omega_d = S_d \frac{1}{(2\pi)^d}$ is the solid angle of a d -dimensional hypersphere.



We rescale the integral by defining $x = q/\sqrt{t}/k$ which gives:

$$F = -\frac{h^2}{2t} + \frac{n\Omega_d}{2} \left(\frac{t}{k}\right)^{d/2} \int_0^{\Lambda/\sqrt{t}/k} dx x^{d-1} \left(\ln t + \ln\left(1 + x^2 + \frac{Ltx^4}{k^2}\right) \right) \quad (8.11)$$

Note that if $t \ll 1$ the upper limit of the integral goes to infinity. Then we study the integral, and although it looks a little scary, it does turn out to converge. Note that in this limit, L does not matter. It's a bit easier to compute the specific heat than the free energy directly. So, we can take two derivatives of F and find C :

$$C = -\frac{\partial^2 F}{\partial T^2} = -n \frac{\Omega_d}{2} \left(\frac{t}{k}\right)^{d/2} \int_0^\infty dx \frac{x^{d-1}}{t^2 \left(1 + x^2 + \frac{Lx^4 t}{k^2}\right)^2} \quad (8.12)$$

For $d > 4$ this is finite.

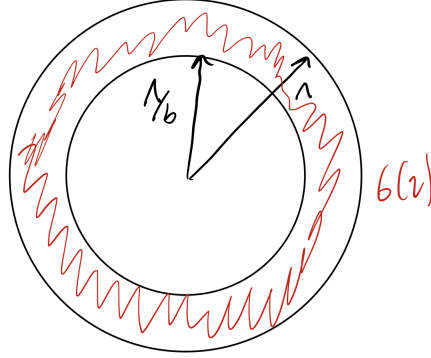
8.2 RG approach to Gaussian model

We return back to the partition function:

$$Z = \int \mathcal{D}m(\mathbf{q}) \exp\left(-\int_0^\Lambda \frac{d^d q}{(2\pi)^d} \left[\left(\frac{t + kq^2 + Lq^4}{2} \right) |m|^2 + hm(\mathbf{q}) \delta(\mathbf{q}) \right] \right) \quad (8.13)$$

Coarse Grain

We consider $a < x < ba$ and then consider splitting the momentum space region into $0 < q < \Lambda/b$ and $\Lambda/b < q < \Lambda$, this latter region we call $\sigma(\mathbf{q})$.



We then split the integral into two parts:

$$Z = \exp\left(-\frac{nV}{2} \int_{\Lambda/b}^{\Lambda} d^d q \ln(t + kq^2 + Lq^4)\right) \cdot \int \mathcal{D}m^<(\mathbf{q}) \exp\left(-\int_0^{\Lambda/b} \frac{d^d q}{(2\pi)^d} (t + kq^2 + Lq^4) |m^<|^2 + hm^<(\mathbf{q})\delta(\mathbf{q})\right) \quad (8.14)$$

where the first term corresponds to the result of doing the Gaussian integral over $\sigma(\mathbf{q})$. Note we have called the dummy variable of integration $m^<$ to remind ourselves that these are the momenta $< \Lambda/b$ we have not yet integrated over.

Rescaling

Call:

$$\delta f_b(t) = \frac{n}{2} \int_{\Lambda/b}^{\Lambda} d^d q \ln(t + kq^2 + Lq^4) \quad (8.15)$$

Now, we rescale $q' = bq$, then:

$$Z = e^{-V\delta f_b(t)} \int \mathcal{D}m^<(\mathbf{q}') \exp\left(-\int_0^{\Lambda} \frac{d^d q'}{(2\pi)^d} b^{-d} \left(\frac{t + kb^{-2}q'^2 + Lb^{-4}q'^4}{2} |m^<(\mathbf{q}')|^2 + hm^<(\mathbf{q}')\delta(\mathbf{q}')\right)\right) \quad (8.16)$$

where the appropriate dimensional quantities have picked up factors from rescaling.

Renormalization

We now renormalize:

$$m'(q) = \frac{m^<(q)}{z} \quad (8.17)$$

After we do this, the partition function now becomes:

$$Z = e^{-V\delta f_b} \int \mathcal{D}m(\mathbf{q}) \exp\left(-\int_0^{\Lambda} \frac{d^d q}{(2\pi)^d} \left[b^{-d} z^2 \left(\frac{t + kb^{-2}q^2 + Lb^{-4}q^4}{2}\right) |m|^2 + zh m(\mathbf{q})\delta(\mathbf{q}) \right]\right) \quad (8.18)$$

Now we rescale our parameters:

$$t_{\text{new}} = z^2 b^{-2} t \quad (8.19a)$$

$$h_{\text{new}} = z h \quad (8.19b)$$

$$k_{\text{new}} = z^2 b^{-d-2} k \quad (8.19c)$$

$$L_{\text{new}} = z^2 b^{-d-4} L \quad (8.19d)$$

The $t = 0, h = 0$ singular point fortunately remain the same with the new parameters. If things are to be scale invariant at the singular/critical point, then $k_{\text{new}} = k$ (since we want the Hamiltonian to transform back to itself at the critical point - we take this as the definition) which enforces:

$$z = b^{1+d/2} \quad (8.20)$$

So then:

$$L_{\text{new}} = b^{-2} L \quad (8.21)$$

so L is irrelevant. (Note: We could not enforce $L_{\text{new}} = L$ and find a consistent set of equations, as if we did, we would find that k would scale to ∞ . We can propose fixed points, but if we propose L as the leading order variable, then the other variables explode, so it is not a valid fixed point. The trivial scaling dimension of k is $d + 2$ and for L it is $d + 4$). We then find:

$$t_{\text{neq}} = b^2 t \quad (8.22)$$

$$h_{\text{neq}} = b^{1+d/2} h \quad (8.23)$$

So then:

$$y_t = 2 \quad (8.24)$$

$$t_n = 1 + d/2 \quad (8.25)$$

Which gives us the critical exponents:

$$\nu = 1/y_t = 1/2 \quad (8.26)$$

$$\Delta = y_n/y_t = \frac{1}{2} + \frac{d}{4} \quad (8.27)$$

$$\alpha = 2 - d\nu = 2 - d/2 \quad (8.28)$$

$$\gamma = 1 \quad (8.29)$$

Our fixed point Hamiltonian (which is scale invariant) is:

$$H^* = \frac{k}{2} \int d^d x (\nabla m)^2 \quad (8.30)$$

our results here agree with the exact solution obtained via solving the Gaussian theory. In the future, we will use the RG to solve problems that are not exactly solvable. But before we go to perturbative RG, a couple of technical points.

8.3 Scaling dimension

Consider:

$$u_n \int d^d x m^n \quad (8.31)$$

If we now find a fixed point and add a perturbation to the model, we can now power count after doing the RG:

$$u_n \int d^d x m^n \rightarrow b^d z^n \int d^d x' m^n \quad (8.32)$$

then, u_n scales as:

$$u'_n = u_n b^{d+n-nd/2} = u_n b^{y_n} \quad (8.33)$$

and $y_n = n - d \left(\frac{n}{2} - 1 \right)$. For different n :

$$y_1 = 1 + d/2 \quad (8.34a)$$

$$y_2 = 2 \quad (8.34b)$$

$$y_4 = 4 - d \quad (8.34c)$$

$$y_6 = 6 - 2d \quad (8.34d)$$

From y_4 , we see that the fourth order coefficient is relevant for $d < 4$. From y_6 , we see that the sixth order term is relevant for $d < 3$. This is then the framework for what we will next when we think about the Wilson-Fisher fixed point, where we include the fourth order coupling. This suggests that there might be a small parameter ϵ which is the difference in dimensions from $d = 4$; we can't do perturbations in the fields in $d = 3$ (as they grow rapidly); but if we work in dimensions in close to 4 (i.e. $d = 4 - \epsilon$), then things can be controlled. If we expand in ϵ and the radius of convergence is > 1 then this mathematical trick allows us to obtain results in $d = 3$. This is very close to dimensional regularization in QFT. Historically, a lot of these concepts were already known by Schwinger before they were applied to stat mech, and even before that known to applied mathematicians looking at badly behaved differential equations.

9 Perturbative Renormalization Group

9.1 Perturbative RG for ϕ^4 theory

We have the Hamiltonian:

$$\beta H = \beta H_0 + U \quad (9.1)$$

Where:

$$\beta H_0 = \int d^d x \left[\frac{1}{2} t |m|^2 + \frac{1}{2} k |\nabla m|^2 \right] = \frac{1}{V} \sum_q \left[\frac{t + kq^2}{2} \right] |m_q|^2 \quad (9.2)$$

and:

$$U = u \int d^d x |m|^4 = \frac{u}{(2\pi)^{4d}} \int d^d q_1 \dots d^d q_4 \delta(q_1 + q_2 + q_3 + q_4) \sum_{\alpha\beta} m_\alpha(q_1) m_\alpha(q_2) m_\beta(q_3) m_\beta(q_4) \quad (9.3)$$

We will perturb in U ; this theory is not Gaussian and therefore not exactly solvable, but we can look at perturbative corrections coming from the fourth order term.

Let's turn the RG crank and see what happens. First, we coarse grain. We have some radius which we integrate out to, $\Lambda \sim 1/a$. We then chop the magnetization up into two parts:

$$m = \begin{cases} \tilde{m}(q) & q < \Lambda/b \\ \sigma(q) & \Lambda/b < q < \Lambda \end{cases} \quad (9.4)$$

Then the partition function becomes:

$$Z = \int \mathcal{D}\tilde{m}(q) \mathcal{D}\sigma(q) \exp \left(- \int_0^\Lambda \frac{d^d q}{(2\pi)^d} \left(\frac{t + kq^2}{2} (|\tilde{m}(q)|^2 + |\sigma(q)|^2) \right) - u(\tilde{m}, \sigma) \right) \quad (9.5)$$

Since the shell that involves the $\sigma(q)$ s are sufficiently faraway from the origin, we are able to assume that for these momenta that $\sigma(q)$ is Gaussian:

$$Z_\sigma = \int \mathcal{D}\sigma(q) e^{-\beta H_\sigma(\sigma(q))} \quad (9.6)$$

So, we want to consider $\langle e^{-U(\tilde{m}, \sigma)} \rangle_\sigma$, the average over the high momenta parts of momentum space. Looking at an expectation value of an operator:

$$\langle O \rangle_\sigma = \frac{1}{Z_\sigma} \int \mathcal{D}\sigma(q) O e^{-\beta H_0(\sigma)} = \frac{1}{Z_\sigma} \int \mathcal{D}\sigma O e^{-\int_{-\Lambda/b}^{\Lambda} \frac{d^d q}{(2\pi)^d} \frac{t+kq^2}{2} \sigma^2} \quad (9.7)$$

If we then carry out this integral over U , this would leave us with:

$$Z = \int \mathcal{D}\tilde{m} \exp\left(\int_0^{\Lambda/b} \frac{d^d q}{(2\pi)^d} \frac{t+kq^2}{2} \tilde{m}^2 - \ln \langle e^{-U} \rangle_\sigma\right) \quad (9.8)$$

The integral for us to do is (expanded out in a power series):

$$\ln \langle e^{-U} \rangle = -\langle U \rangle_\sigma + \frac{1}{2} \left(\langle u^2 \rangle_\sigma - \langle u \rangle_\sigma^2 \right) + \dots \quad (9.9)$$

thus:

$$\langle U \rangle_\sigma = u \int \frac{d^d q_1 \dots d^d q_4}{(2\pi)^{4d}} \delta\left(\sum_n q_n\right) \langle [\tilde{m}(q_1) + \sigma(q_1)] \cdot [p\tilde{m}(q_2) + \sigma(q_2)] \cdot [\tilde{m}(q_3) + \sigma(q_3)] \cdot [p\tilde{m}(q_4) + \sigma(q_4)] \rangle_0 \quad (9.10)$$

The zeroth order term is:

$$\langle \sigma^0 \rangle = U(\tilde{m}) \quad (9.11)$$

the odd terms vanish, and any terms that do not depend on \tilde{m} we can ignore as this only gives a constant. Thus, we are left with:

$$\langle \sigma(q_1) \cdot \sigma(q_2) \tilde{m}(q_3) \cdot \tilde{m}(q_4) \rangle_\sigma \times 2 \quad (9.12)$$

$$\langle \sigma(q_1) \cdot \tilde{m}(q_2) \sigma(q_3) \cdot \tilde{m}(q_4) \rangle_\sigma \times 4 \quad (9.13)$$

Now, the usual Gaussian average we have seen before is:

$$\langle \sigma(q_1) \sigma(q_2) \rangle = \frac{\delta(q_1 + q_2)}{t + kq_1^2} (2\pi)^d m \quad (9.14)$$

The terms of Eq. (9.12) become:

$$\begin{aligned} 2 \langle \sigma(q_1) \sigma(q_2) \tilde{m}(q_3) \tilde{m}(q_4) \rangle_\sigma &= -2nu \int \delta(q_1 + q_2 + q_3 + q_4) \delta(q_1 + q_2) \frac{\tilde{m}(q_3) \tilde{m}(q_4)}{(t + kq_1^2)} \\ &= -2nu \int_0^{\Lambda/b} \frac{d^d q}{(2\pi)^d} |\tilde{m}(q)|^2 \int_{\Lambda/b}^{\Lambda} \frac{d^d p}{(2\pi)^d} \frac{1}{(t + kp^2)} \end{aligned} \quad (9.15)$$

where n is the number of components of the spin. Thus:

$$\beta H[\tilde{m}] = \text{const} + \int_0^{\Lambda/b} d^d q \frac{\tilde{t} + kq^2}{2} |\tilde{m}|^2 + U[\tilde{m}] \quad (9.16)$$

Where:

$$\tilde{t} = t + 4u(n+2) \int_{\Lambda/b}^{\Lambda} \frac{d^d p}{(2\pi)^d} \frac{1}{t + kp^2} \quad (9.17)$$

We have the contribution of $4un$ from the Eq. (9.12) term (with the n coming from the dot product over spin components) and the contribution of $8u$ from the (9.13) term (doesn't have an n component as we pick out a single component in the dot product between \tilde{m} and σ). So the integral over Λ/b to Λ can be viewed as a shift of the t parameter.

Now, we rescale $q = q'/b$ and renormalize $\tilde{m} = zm'$. This yields:

$$\beta H[m'] = ut + \int_0^\Lambda \frac{d^d q'}{(2\pi)^d} b^{-d} z^2 (\tilde{t} + kb^{-2} q'^2) |m(q')|^2 + uz^4 b^{-3d} \int_0^\Lambda |m|^4 \quad (9.18)$$

We can now relabel:

$$t' = b^{-d} z^2 \tilde{t} \quad (9.19)$$

$$k' = b^{-d-2} z^2 k \quad (9.20)$$

$$u' = b^{-3d} z^4 u \quad (9.21)$$

Choosing z such that it scales to keep the k constant:

$$z = b^{1+d/2} \quad (9.22)$$

As an aside (to make life a little easier):

$$b = e^{\delta l} = 1 + \delta l + O((\delta l)^2) \quad (9.23)$$

Let's make b small/differential. This allows us to carry out the momentum integrals we avoided doing earlier:

$$t + \delta l \frac{dt}{dl} = (1 + 2\delta l) \left(t + 4u(n+2) \frac{S^d}{(2\pi)^d} \frac{\Lambda^d \delta l}{(t + k\Lambda^2)} \right) \quad (9.24)$$

where:

$$\frac{dt}{dl} = 2t + 4u(n+2) \frac{S_d}{(2\pi)^d} \frac{\Lambda^d}{(t + k\Lambda^2)} \quad (9.25)$$

and:

$$\frac{du}{dl} = (4-d)u \quad (9.26)$$

Unfortunately, this is not helpful. Why?:

$$\frac{d}{dl} \begin{pmatrix} t \\ u \end{pmatrix} = \begin{pmatrix} 2 & \text{(factor)} \\ 0 & 4-d \end{pmatrix} \begin{pmatrix} t \\ u \end{pmatrix} \quad (9.27)$$

As before, we will find that $y_t = 2$ and $y_u = 4-d$. If we think about what the RG flows look like - for $d > 4$, u flows to zero (Gaussian theory, that we might expect) and for $d < 4$ unfortunately the arrows flip the other way and u flows to infinity :(.

There are also second order terms - we will not look at these as there are 36 terms and this will be painful. But we sort of know what to expect. There will be corrections of order u^2 , so:

$$\frac{dt}{dl} = 2 + u(\text{factor}) - Au^2 \quad (9.28)$$

$$\frac{du}{dl} = u(4-d) - Bu^2 \quad (9.29)$$

Now, if u starts to flow away from the origin, it will get pushed back! In that case, what happens is we get a new critical point at u^* , which is equal to:

$$u^* = \frac{4-d}{B} \quad (9.30)$$

which is where u stops flowing. There will be a corresponding critical value t^* :

$$t^* \sim -Cu^+(\text{corrections}) \quad (9.31)$$

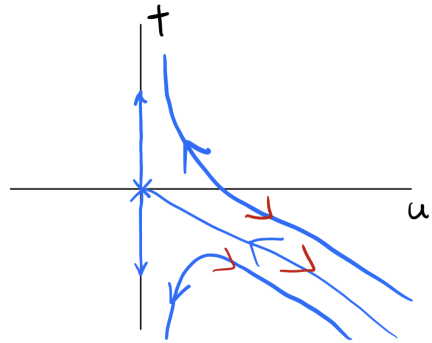


Figure 9.1: Flow for linearized RG; for $d > 4$ u flows towards 0, but for $d < 4$ u diverges to infinity.

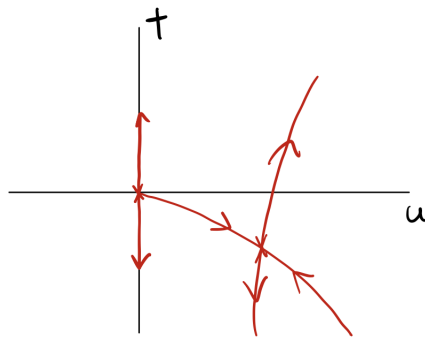


Figure 9.2: Flow for RG to quadratic order; a new fixed point emerges.

This is a rather laborious calculation to show this, but it has been done. But what is important to note is that in principle the behaviour is controlled by $4 - d$, and if we define $\epsilon = 4 - d$, then this is what is known as the ϵ -expansion, i.e. regularizing the problem by working close to a critical dimension:

$$u^* = \frac{\epsilon}{B} \quad (9.32)$$

The result of which is:

$$y_t = 2 - \frac{n+2}{n+8}\epsilon + O(\epsilon^2) \quad (9.33)$$

$$y_u = -\epsilon \quad (9.34)$$

This is a manifestation of universality (I missed why this was...)

One should also check what other kinds of coefficients get generated by this process. If we keep going on, there will be terms of ϵ^2 , and we will generate other gradient terms we did not have before. They will always appear one order higher than the previous. At this level, the new fields are generated and are irrelevant at this order.

What happens if we try $\epsilon = 1$ do get $d = 3$? The exponents turn out for Ising model are remarkably close to numerical estimates, even though there is no reason things should have been. We don't know what the radius of convergence of this series is; it is a bit of a punch in the dark.

Note that the fixed point (t^*, u^*) is the known as the Wilson-Fisher fixed point.

Reviewing exactly what we have done today:

- We take a model that has a complicated higher order term, namely the fourth order term.
- What we know is that this term can be made small by going into the vicinity of a critical dimension, here $d = 4$. (The scaling dimension of the fourth order term is $4 - d$). To the extent that this is a small number, the expansion can be controlled.
- Then, we know that the scaling law has to be augmented as we go through perturbatively through terms, with corrections u, u^2, \dots . Therefore the $\frac{dt}{dl}$ must have the power series functional form that we found.
- The signs of the term we do not necessarily know, but we can guess that we can find solutions by postulating that there exist fixed points of the form as we found.
- Then we get a scaling diagram, where in addition to the Gaussian fixed point (unstable in the u direction) we find the Wilson-Fisher fixed point, which u flows to.
- Rather than the tedious algebra, the takeaway is the structure of the theory that comes out of it, and what kind of scaling theory it gives us.

Next class, we will look at continuous symmetries, which are slightly easier to deal with. We go away from spin waves, towards objects with topology. We shall begin this with the nonlinear- σ model (spins of fixed length that can point in any direction). From there, we will go to the 2-D XY model, and start to study topology in spin space.

10 Continuous Symmetries - Beyond Spin Waves

We return back to models with Goldstone modes and transverse fluctuations, where:

$$S_t(q) \sim \frac{1}{q^2} \quad (10.1)$$

for $T < T_c$. These are models where the order parameters have a $U(1)$ symmetry - order parameters of the form:

$$\psi(r) = |\psi|e^{i\theta} \quad (10.2)$$

and correlation functions look like:

$$\langle \Theta(q)\Theta(q') \rangle = \frac{\delta(q+q')}{q^2} \quad (10.3)$$

Where:

$$\langle \Theta(r)\Theta(0) \rangle \sim \begin{cases} r^{2-d} & d < 2 \\ \ln r & d = 2 \\ C & d > 2 \end{cases} \quad (10.4)$$

The $d = 2$ case is confusing because its not clear what the $\ln r$ means. One route into thinking about this is thinking about models where we can apply RG.

10.1 Non-linear σ models

We consider the non-linear σ model (kind of a misnomer - Gell-Mann named it because he thought it would describe σ -mesons, even though it does not, but the name stuck). Here, we consider an n -component spin of length 1. We can thus write the spin degrees of freedom as:

$$\mathbf{s} = (s_1, s_2, \dots, s_n) \quad (10.5)$$

where:

$$\sum_i |s_i|^2 = 1 \quad (10.6)$$

$n = 2$ would be the XY-model, $n = 3$ would be the Heisenberg model, and we can be in arbitrary dimension. Each one of these components should have a transverse mode.

We can write down the partition function for this as follows:

$$Z = \int \mathcal{D} [\mathbf{s}(x), \delta(|\mathbf{s}|^2 - 1)] \exp\left(-\frac{k}{2} \int_{\Lambda} |\nabla \mathbf{s}|^2 dx\right) \quad (10.7)$$

Note that the stiffness k is a function of temperature, namely $k \sim \frac{1}{T}$.

We integrate over all possible parts, maintaining the constraint that the length of the spin is unity. It is easiest to re-label spins:

$$\mathbf{s} = (\pi_1, \pi_2, \dots, \pi_{n-1}, \sigma) \quad (10.8)$$

i.e. pick out one component of the spins and integrate it out. The σ component is fixed by the constraint, and so:

$$\int \mathcal{D}[\mathbf{s}(x), \delta(|\mathbf{s}|^2 - 1)] = \int \mathcal{D}\pi \mathcal{D}\sigma \delta(|\pi|^2 - \sigma^2 - 1) = \int \mathcal{D}\pi d\sigma \left[\delta(\sigma - \sqrt{1 - \pi^2}) \delta(\sigma + \sqrt{1 - \pi^2}) \right] = \int \frac{1}{2} \frac{\mathcal{D}\pi}{\sqrt{1 - \pi^2}} \quad (10.9)$$

So then we can write Z as:

$$Z = \int \frac{\mathcal{D}\pi(x)}{\sqrt{1 - \pi(x)^2}} \exp\left(-\frac{\kappa}{2} \int d^d x \left[(\nabla \pi)^2 + (\nabla(\sqrt{1 - \pi^2}))^2 \right]\right) \quad (10.10)$$

The last thing to do is to get rid of the pesky $\frac{1}{\sqrt{1 - \pi(x)^2}}$ term and exponentiate:

$$Z = \int \mathcal{D}\pi \exp\left(-\frac{\kappa}{2} \int d^d x \left[(\nabla \pi)^2 + (\nabla(\sqrt{1 - \pi^2}))^2 \right] + \frac{\rho}{2} \ln(1 - \pi^2)\right) \quad (10.11)$$

where $\rho = 1$. So far this is exact with no approximations. But of course, we will not get much further without doing an expansion.

To leading order, we find:

$$\langle \pi^2 \rangle_0 = T \quad (10.12)$$

as as $T \rightarrow 0$, the occupation of the Goldstone modes goes to zero. By expanding, we have:

$$\beta H = \beta H_0 + u_1 + u_2 + \dots \quad (10.13)$$

where:

$$\beta H_0 = i \int d^d x \frac{1}{2} \kappa (\nabla \pi)^2 \quad (10.14)$$

and:

$$u_1 = \int d^d x \left[\frac{1}{2} \kappa (\pi \nabla \pi)^2 - \rho \frac{\pi^2}{2} \right] \quad (10.15)$$

We now have an opportunity to go through RG, but in a bit of a different way; the $(\pi \nabla \pi)^2$ terms should get smaller and smaller.

Let's take this and go into Fourier space:

$$\beta H_0 = \frac{\kappa}{2} \int \frac{d^d q}{(2\pi)^d} q^2 |\pi(q)|^2 \quad (10.16)$$

$$u_1 = -\frac{\kappa}{2} \int \frac{d^d q_1 \dots d^d q_4}{(2\pi)^{4d}} \delta(\sum_i q_i) (q_1 \cdot q_3) \pi_\alpha(q_1) \pi_\alpha(q_2) \pi_\beta(q_3) \pi_\beta(q_4) - \frac{\rho}{2} \int \frac{d^d q}{(2\pi)^d} |\pi(q)|^2 \quad (10.17)$$

Now we turn the RG crank.

10.2 RG for Non-linear σ model

We divide up momentum space into a shell σ (corresponding to $\Lambda/b < q < \Lambda$) and the rest (corresponding to $\tilde{\pi}$), and integrate out that shell. Each π term appearing above can be broken up into a $\tilde{\pi}$ part and a σ part. The only terms which will survive will be the ones where σ come in pairs. For example, we will have a Gaussian average of the form:

$$\langle \sigma_\alpha(q_1) \sigma_\alpha(q_2) \rangle \pi_\beta(q_3) \pi_\beta(q_4) \quad (10.18)$$

Actually, we will only have two surviving parts:

$$\begin{aligned} (A) &= -\frac{\kappa}{2} \int d^d q_1 \dots d^d q_4 (q_1 \cdot q_3) \langle \sigma_\alpha(q_1) \sigma_\alpha(q_3) \rangle \tilde{\pi}_\beta(q_2) \tilde{\pi}_\beta(q_4) \\ &= -\frac{\kappa}{2} \int d^d q_1 \dots d^d q_4 (q_1 \cdot q_3) \delta(q_1 - q_3) \tilde{\pi}_\beta(q_2) \tilde{\pi}_\beta(q_4) \\ &= \frac{\kappa}{2} \int_{\Lambda/b}^{\Lambda} d^d k \frac{k^2}{\kappa k^2} \int_0^{\Lambda/b} \frac{d^d q}{(2\pi)^d} |\tilde{\pi}(q)|^2 \end{aligned} \quad (10.19)$$

in the last line $k = q_1 = q_3$ and $q = q_2 = q_4$. We use the known expectation value of the σ s from our previous analysis. There is also a $(n-1)$ factor coming from the components of the spin, which we don't write above. The other surviving part is:

$$\begin{aligned} (B) &= -\frac{\kappa}{2} \int d^d q_1 \dots d^d q_4 (q_1 \cdot q_2) \tilde{\pi}_\alpha(q_1) \tilde{\pi}_\beta(q_3) \langle \sigma_\alpha(q_2) \sigma_\alpha(q_4) \rangle \\ &= \frac{\kappa}{2} \int \frac{d^d q}{(2\pi)^d} q^2 |\tilde{\pi}(q)|^2 \frac{I_d}{\kappa} \end{aligned} \quad (10.20)$$

where:

$$I_d = \int_{\Lambda/b}^{\Lambda} d^d k \frac{1}{k^2} \quad (10.21)$$

We have two things that are changing here, κ and ρ , which we renormalize:

$$\tilde{\kappa} = \kappa \left[1 + \frac{I_d(b)}{\kappa} \right] \quad (10.22)$$

$$\tilde{\rho} = \rho[1 - b^{-d}] \quad (10.23)$$

where:

$$\rho = \frac{N}{V} \int_0^{\Lambda} \frac{d^d q}{(2\pi)^2} \quad (10.24)$$

Then, the two terms we had before become:

$$-\beta H' = -\tilde{\kappa} \frac{b^{d-2}}{2} z^2 \int d^d x (\nabla \pi)^2 \quad (10.25)$$

$$u'_1 = -\kappa \frac{b^{d-2} z^4}{2} \int d^d x (\pi \nabla \pi)^2 + \frac{\rho z^2}{2} \int d^d x |\pi(x)|^2 \quad (10.26)$$

Looking at the average:

$$\langle \tilde{s} \rangle_0 = \left\langle (\pi_1 + \sigma_1)(\pi_2 + \sigma_2) \dots \sqrt{1 - (\tilde{\pi} + \sigma)^2} \right\rangle = 1 - \frac{1}{2} \langle \sigma \rangle^2 + O(T)^2 = 1 - \frac{n-1}{2} \frac{I_d(b)}{\kappa} = z \quad (10.27)$$

This tells us that we have to choose the scaling exponent (that sets the size of the spin, which is fixed!) such that the above is preserved, hence we set the above to z . Then:

$$\kappa = b^{d-2} z^2 \tilde{\kappa} = b^{d-2} \left[1 - \frac{n-1}{\kappa} I_d \right]^2 \kappa \left[1 + \frac{I_d}{\kappa} \right] = b^{d-2} \kappa \left[1 - \frac{n-2}{\kappa} I_d + O\left(\frac{1}{\kappa^2}\right) \right] + \dots \quad (10.28)$$

So then $b \sim e^l$, and:

$$\frac{\partial \kappa}{\partial l} = (d-2)\kappa - (n-2)\kappa \Lambda^{d-2} \quad (10.29)$$

so temperature scales to zero or infinity here depending on dimension and the number of components.

Say, take $d < 2$. Then $\kappa \rightarrow 0$ so $T \rightarrow \infty$ and no order. But, for $d > 2$, we have $\kappa \rightarrow \infty$ so $T \rightarrow 0$ and we are ordered. The critical temperature is then:

$$T^* = \frac{d-2}{n-2} \frac{1}{\kappa_0 \Lambda^{d-2}} \quad (10.30)$$

At $n = d = 2$ we have the XY-model which is not resolved.

10.3 Vortices

What have we missed in the spin wave theory? Consider $d = 2, n = 2$. There are objects in the lattice where there are vortices; faraway it looks like the spin just slowly varies, but close to it there is an observable "vortex charge" measurable by going around it in a loop. There are charges and anti-charges, and we measure these via:

$$\oint_S \nabla \theta \cdot ds = (2\pi)m \quad (10.31)$$

for $m \in \mathbb{Z}$. Labelling these as charges is well-motivated; they behave like charges. In fact it will look exactly like EM. Indeed they actually form in pairs.

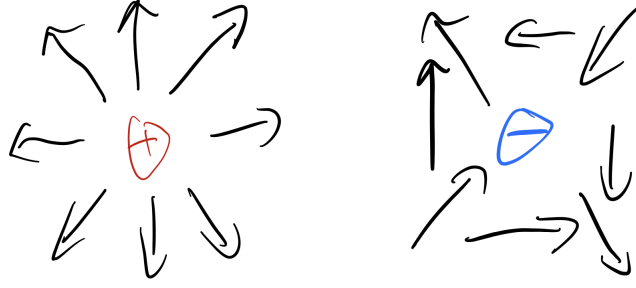


Figure 10.1: Cartoon picture of spin vortex charges

We can ask; does it make sense that as we go down to low T we have the spontaneous formation of these vortices? Normally, one would think that it takes finite energy to place them in the lattice (so they should go away at low T) but there are so many places to put them and as such there is an entropic gain as well... What happens is (even at low T) fluctuations create charges out of a vacuum, but they will be close together and be bound. This is one phase. There is also a second possible phase where the charges unbind; we go from vortex charge molecules to a plasma; they begin to overlap, and they start screening, and the system breaks itself up. We can describe this with classical EM. This transition/plasma phase has no long range order, but the confined phase has a quasi-long range order, where the effective stiffness stays finite. The transition between these two is the Berezinskii-Kosterlitz-Thouless transition.

11 2-D XY model

11.1 Single Vortices

Let us recap where we ended up last class. What has been missing thus far from our discussion of continuous systems is the presence of vortices. We were assuming that we could capture everything with slow variations of spins, but this leaves out configurations like those that appear in Fig. 10.1. We can define such vortices by looking at:

$$\oint (\nabla\theta) \cdot d\mathbf{S} = 2\pi n \quad (11.1)$$

We can make vortices that wind more than once, or those that have negative m (antivortices). Since:

$$\nabla\theta = -n\nabla \times (\hat{\mathbf{z}} \ln r) \sim \frac{1}{r} \quad (11.2)$$

Let's compute the energy of a vortex.

$$\beta\mathcal{E}_m = \beta\mathcal{E}_0(a) + \frac{1}{2}\kappa \int_a^L d^2x (\nabla\theta)^2 = \beta\mathcal{E}_0(a) + \pi\kappa n^2 \ln(L/a) \quad (11.3)$$

which diverges in the system size. This may suggest that we might not have any vortices - but this is not the case. Let's look at the partition function:

$$Z(n) \approx \left(\frac{L}{a}\right)^2 e^{-\beta\mathcal{E}_0(a)} e^{-\pi\kappa n^2 \ln(L/a)} \quad (11.4)$$

The $\mathcal{E}_0(a)$ is like the chemical potential, and we can call its exponential the fugacity. We have:

$$Z(n) \approx y_0(a) \left(\frac{L}{a}\right)^{2-\pi\kappa n^2} \quad (11.5)$$

Let us also remember that $\kappa \propto \frac{1}{T}$. So, above $T_c \sim \frac{\pi}{2} n^2$ is where the vortices proliferate (look at the term in the partition function). Although the energy contribution is infinite, the entropic gain wins out at this temperature.

11.2 Pairs of Vortices

This was a single vortex; there are also pairs of vortices, which is what we will have to pay attention to. Consider a vortex anti-vortex pair at distance d . We then find:

$$(\nabla\theta) \sim d\nabla\left(\frac{\mathbf{r} \times \mathbf{z}}{r}\right) \sim \frac{1}{r^2} \quad (11.6)$$

These have finite energy, and act like a Coloumb gas.

Consider a flow, which we treat by looking at a sum of a rotational and irrotational part.

$$\mathbf{u}(\mathbf{r}) = \nabla\theta \quad (11.7)$$

One part is curl free (this is the piece without vortices):

$$(\nabla \times u_0) = 0 \implies u_0 = \nabla\phi \quad (11.8)$$

Then:

$$\mathbf{u} = u_0 - \nabla \times (\mathbf{z}\psi) \quad (11.9)$$

where if we consider the contribution from vortices:

$$\oint \mathbf{u} \cdot d\mathbf{S} = 2\pi\hat{\mathbf{z}} \sum_i n_i \delta(r - r_i) \quad (11.10)$$

Then:

$$\nabla \times \mathbf{u} = \hat{\mathbf{z}} \nabla^2 \psi \quad (11.11)$$

We can then see:

$$\nabla^2 \psi = 2\pi \sum_i n_i \delta(r - r_i) \quad (11.12)$$

So then:

$$\psi(r) = \sum_i n_i \ln(|r - r_i|) \quad (11.13)$$

Having written the field as this, let us compute the energy:

$$\beta H = \frac{1}{2} \kappa \int d^2r \left[(\nabla\phi)^2 - 2(\nabla\phi) \cdot \nabla(\hat{\mathbf{z}}\psi) + (\nabla \times \hat{\mathbf{z}}\psi)^2 \right] \quad (11.14)$$

The middle term vanishes via integrating by parts. The first term is familiar. The third term looks like:

$$(\nabla \times \hat{\mathbf{z}}\psi)^2 = \left(-\frac{\partial\psi}{\partial x}, \frac{\partial\psi}{\partial y}, 0 \right)^2 = (\nabla\psi)^2 \quad (11.15)$$

We then have (through an integration by parts):

$$\beta H_1 = \kappa \int d^2r (\nabla\psi)^2 = -\frac{\kappa}{2} \int d^2r \psi \nabla^2 \psi = -\frac{\kappa}{2} \int d^2r \left(\sum_i n_i \ln(|r - r_i|) \right) \left(\sum_j 2\pi n_j \delta(r - r_j) \right) \quad (11.16)$$

which we can now use to define the correlation function:

$$\beta H_1 = -\frac{\kappa}{2} \int d^2r \sum_{i \neq j} n_i n_j C(r_i - r_j) \quad (11.17)$$

where:

$$C(r_i - r_j) = \frac{\ln(|r_i - r_j|)}{2\pi} \quad (11.18)$$

and we exclude a vortex interacting with itself.

11.3 Transitions

Now, we may go back and write down the full partition function:

$$Z = \int \mathcal{D}\phi e^{-\frac{\kappa}{2} \int (\nabla\phi)^2} \sum_{r_k} \int d^2r_k e^{-\sum_i \beta \mathcal{E}_0(r_i) + 4\pi^2 \kappa \sum_{ij} n_i n_j C(r_i - r_j)} \quad (11.19)$$

The first term corresponds to spin waves, and the second corresponds to vortices, so we can write:

$$Z = Z_{SW} Z_{\text{vort}} \quad (11.20)$$

We also constrain the configuration to that which where $\sum_i n_i = 0$, and restrict to $n_i = \pm 1$. To make the connection with charges more clear, we replace n_i with q_i . We then study:

$$Z_v = \sum_{N=0}^{\infty} (y_0)^N \int \prod_{k=1}^N d^2r_k \exp(4\pi^2 \kappa \sum_{ij} q_i q_j C(r_i - r_j)) \quad (11.21)$$

Suppose the fugacity is small, such as in a metal-insulator. Then, y_0 /the fugacity is small (y_0 can be thought to be controlling the density), and we should see that the vortices bind in pairs. Here, $C(x) \rightarrow C(x)/\mathcal{E}$. There is another limit where y_0 is large, and then the bound pairs of charges start to overlap highly, so much so that the charges become unbound. This is the plasma phase. and we see $C(x) \sim e^{-x/\xi}$.

We have:

$$Z \sim y_0(a) \left(\frac{L}{a}\right)^{2-\pi\kappa} \rightarrow y_0(ba) \left(\frac{L}{ba}\right)^{2-\pi\kappa} \quad (11.22)$$

where we rescale such that the partition function remains invariant. Then if we make $b \sim e^{-\delta l}$:

$$\frac{dy_0}{dl} = (2 - \pi\kappa)y_0 \quad (11.23)$$

The $x = 2 - \pi\kappa$ parameter tells us whether vortices are relevant or not.

We then assert:

$$\frac{dx}{dl} = Ay_0^2 \quad (11.24)$$

Which gives rise to the flow diagram:

So we have two regimes, one where vortices are irrelevant and another where they proliferate. There is a separatrix line which is the divisor between whether we flow to no vortices or proliferation. We can say this defines a physical transition temperature, known as the Berezinskii-Kosterlitz-Thouless temperature.

Lastly, let us think about what the correlation functions look like:

$$\langle \psi(r)\psi(0) \rangle = \frac{1}{r^{1/2\pi\kappa}} \quad (11.25)$$

Notice that this power law varies. When we get to $k = 2/\pi$, at that point the correlator goes as $\sim \frac{1}{r^{1/4}}$. But, everywhere along the line of $x = 0$ fixed points there is a (continuously varying) power law (i.e. each point along the line has a different critical exponent), up until $r^{-1/4}$, at which afterwards it becomes a short-range correlator $e^{-r/\xi}$ which exponential decay. This means, at the transition temperature, there is a jump in physical properties.

11.4 Polarizability + Looking ahead

Suppose we add an external field:

$$H = H_0 + \mathbf{E} \cdot \sum_{\mathbf{R}} \mathbf{R} \cdot n(\mathbf{R}) \quad (11.26)$$

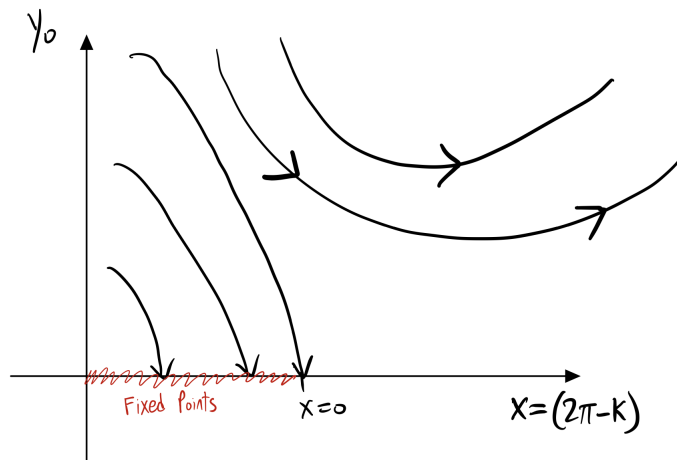


Figure 11.1: Flows of fugacity and stiffness in the 2-D XY model, corresponding to 2 different phases with differing behaviour of vortices.

Suppose now we want to compute the dielectric function:

$$\chi_{ij} = \frac{\partial f}{\partial E_i \partial E_j} \sim \sum_{\mathbf{R}} \sum_{\mathbf{R}'} R^2 \langle n(\mathbf{R}) n(\mathbf{R}') \rangle = \int dR R^3 e^{-2\pi\kappa \ln(R)} = \int_a^L dR R^{3-2\pi\kappa} L^{4-2\pi\kappa} \quad (11.27)$$

which depending on κ either diverges or vanishes. It diverges when $\kappa = 2/\pi$. This connects back up to the correlation function above.

Next day, we move to disorder and randomness. Everything we have talked about thus far assumes a perfect medium and translational invariance. We will now break this. This is both relevant in a practical sense (every physical system has disorder of some form), and also inherently interesting problems with disorder, e.g. frustration. We could consider models with sufficient randomness such that the spins have confusion about their orientation, then we can get systems that act glassy - they freeze into a disorder state. There is a sharp transition from a slowly-flowing material to something that freezes; this is called a spin glass, and we will talk about the infinite-range spin-glass, which we will be able to solve exactly (the Parisi solution - he won a Nobel prize for this)! Interestingly, these kinds of models allow you to store information/use them as a physical memory, which was John's Hopfield's idea. Rapidly coming out of this was how one "learns" via tuning parameters, the simplest of which is the Boltzmann machine. Of course, this was the Nobel prize in physics this year.

12 Disorder and Random Systems

12.1 Introduction and Motivation

Disorders and imperfections are unavoidable in real systems; do phase transitions still exist in this setting? This will be the first point which we will address, and we will be able to answer it quite sharply.

The answer will turn out to be yes, but we then have the follow-up question; how are the transitions modified? We will find that the (upper and lower) critical dimensions become modified.

Finally, we may ask if there will be new phenomena emerging from disorder; indeed there can be, and we will look at the example of a glass, where there is an abrupt transition to being frozen/carrying information.

12.2 Disorder Type I - Impurities/Vacancies

E.g. we have a ferromagnet, and randomly start removing spins. We could for example write down the Hamiltonian:

$$H(\{s\}, \{m\}) = -\frac{1}{2} \sum_{ij} J_{ij} s_i s_j m_i m_j \quad (12.1)$$

where randomly, we pick sites m_i and set them to zero such that those terms become absent from the sum. We will need to make a careful distinction about whether the system is quenched or annealed. Suppose we “freeze in” the spin configuration. Then the partition function looks like:

$$Z(\{m\}) = \text{Tr}_S e^{-\beta H(s,m)} \quad (12.2)$$

which is called “quenched”. This has the odd feature that the answer depends explicitly on the $\{m\}$. If we do the experiment again and put the impurities in different places, we get a different partition function. We could instead construct the “annealed” partition function, which is:

$$Z = \text{Tr}_m Z(\{m\}) \quad (12.3)$$

where we average over all configurations $\{m\}$ of empty sites. The annealed problem is a tough one and we won’t address it in this lecture, but we will attack the quenched partition function.

A comment - we lose translational invariance when we do this.

Let us block our system. We can consider an averaged free energy:

$$\bar{F} = \text{Tr}_m Z[\{m\}] F[\{m\}] \quad (12.4)$$

where we assume that we can compute (over a given block) $F[\{m\}]$ governed by $P[\{m\}]$ (each block will have a different value, but obeys the same distribution). The \bar{F} above represents a quenched average. Looking at the probability distribution, we have:

$$P[\{m\}] = \prod_i \left[(1-x)\delta_{m_i,1} + x\delta_{m_i,0} \right] \quad (12.5)$$

where x is the concentration of vacancies. The important thing to think about; when we do the trace over S , we average over spin configurations, but do not minimize over the free energy. It is instead, fixed here.

We distinguish the thermal average with the spatial average:

$$\langle s_i s_j \rangle \neq \overline{\langle s_i s_j \rangle} \quad (12.6)$$

where the lhs represents the thermal average, the rhs the spatial. The thermal average is dependent on the spin configuration, but the spatial average only depends on the distance between i, j i.e. $|i - j|$. Formally, we can write:

$$\overline{\langle s_i s_j \rangle} = \frac{1}{N} \sum_k \langle s_{k+i} s_{k+j} \rangle = M(|i - j|) \quad (12.7)$$

why might we want to do this? It is relevant for determining magnetic fluctuations in the system.

We can consider some types of impurities/vacancies:

- (i) Bond dilution, e.g. $J_{ij} = J + \text{small fluctuations}$.
- (ii) Random field; $H = \sum_{ij} J_{ij} s_i s_j + \sum_i h_i s_i$ where we make the h_i random. Where $\bar{h} = 0$ and $\overline{h^2} = \Delta$.
- (iii) Spin glass, where we take J_{ij} random with mean zero. In this case, since we pick couplings randomly, we can have frustrated systems. Note that is similar to bond dilution with $J = 0$, but allowing the fluctuations to be not necessarily small. We could call a system with $s_i(0)s_i(t) \neq 0$ a glass (correlations get “baked in”).

12.3 Computing quenched averages - Replica Trick

We want to compute:

$$\bar{F} = \text{Tr}_m P[\{m\}] \ln Z[\{m\}] \quad (12.8)$$

averaging logarithms is hard, but we can expand $\ln Z$ in a power series:

$$\ln Z = \lim_{n \rightarrow 0} \frac{Z^n - 1}{n} \quad (12.9)$$

Taking quenched averages of power series of Z are much easier. We interpret this as taking our Hamiltonian and *replicating* it:

$$\overline{Z^n} = \text{Tr}_m \text{Tr}_{s^a} P[\{m\}] e^{-\sum_{a=1}^n H[\{s^a\}, \{m\}]} \quad (12.10)$$

Because $P[\{m\}]$ is random, this will correspond to some SM model with randomness. But the trick is that we can now do the average over m first, so we end up with:

$$\overline{Z^n} = \text{Tr}_{s^a} e^{-\sum_{ab} H_{\text{eff}}(s^a, s^b)} \quad (12.11)$$

where by doing the average over m we get correlations between the different replicas. (Note that in principle we could have couplings between higher numbers of replicas, but since we started with two-body couplings, we have at most couplings between pairs of replicas).

12.4 Transition Temperatures and Concentration

Suppose I have some concentration of impurities $x \ll 1$, and the transition temperature T_c depends on x . Therefore, we have:

$$T_c(x) \implies \xi \sim |T - T_c(x)|^{-\nu(x)} \quad (12.12)$$

Because the system varies spatially, the concentration may be different at different locations. We thus may consider the deviation:

$$\delta T_c(\mathbf{r}) = T_c(\mathbf{r}) - T_c \quad (12.13)$$

Let us define a correlation function:

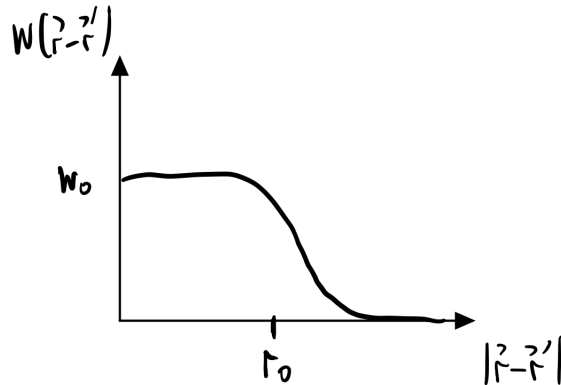
$$W(\mathbf{r} - \mathbf{r}') = \overline{T_c(\mathbf{r}) T_c(\mathbf{r}')} \quad (12.14)$$

We now ask; as we open up the block size of the system, what happens to the correlation function?

We consider an aperture/block size L such that $a \ll L \ll$ system size. Then:

$$(\Delta T_c)^2 = \int_a^L \frac{d^d \mathbf{r}}{L^d} \int_a^{r_0} \frac{d^d \mathbf{r}'}{L^d} W(\mathbf{r} - \mathbf{r}') \quad (12.15)$$

where we integrate up to a characteristic length scale of the correlation r_0 :



We then find;

$$\Delta T_c \sim W_0^{1/2} \left(\frac{r_0}{L} \right)^{d/2} \sim \zeta^{-d/2} \quad (12.16)$$

which has the interpretation; if we average over a large volume, we have the square root number of microscopic domains in the sample. The sensible L to choose here is the correlation length, because if the correlation length diverges (uniformly) this average makes sense. Thus, we should have:

$$\Delta T_c \ll |T - T_c(x)| \sim \zeta^{-\frac{1}{\nu}} \quad (12.17)$$

phrased another way, we want to get a convergent average when we average over the scale of the correlation length.

We thus obtain the criterion:

$$\boxed{\frac{d\nu}{2} > 1} \quad (12.18)$$

because only then will $\zeta^{-d/2} \ll \zeta^{-\frac{1}{\nu}}$. This is the criterion for transitions to exist, and is known as the *Harris criterion*. Note that this requires $\alpha < 0$. If this is not obeyed, the phase transition is destroyed by randomness.

12.5 Computing with Replicas

Consider:

$$H = H^* + \sum_i m_i E_i \quad (12.19)$$

where E_i is the local energy density and $m_i \in \{0, 1\}$ is the vacancy. We then have:

$$\begin{aligned} \bar{Z}^n &= \text{Tr}_m \text{Tr}_{s^a} P(\{m\}) e^{-\sum_a H_a^* - \sum_{ia} m_i E_i^a} \\ &= \exp\left(-\sum_a H^* a - \bar{m} \sum_{ai} E_i^a + \frac{1}{2} \sum_{ij} + ab \sum_{ij} (\bar{m}_i \bar{m}_j - \bar{m}^2) E_i^a E_j^b\right) \end{aligned} \quad (12.20)$$

where we have used a Gaussian distribution $P(m) \sim e^{-m^2/\Delta}$ to perform the average (if not Gaussian, there are higher order terms/cumulants).

Looking at the leading off-diagonal term, we study $\sum_{a \neq b} E_i^a E_i^b$. Look at the two point function:

$$\left\langle \sum_{a \neq b} E_i^a E_i^b \sum_{a' \neq b'} E_j^{a'} E_j^{b'} \right\rangle = 2n(n-1) \langle E_i^a E_j^a \rangle^2 = 2n(n-1) \frac{1}{|i-j|^{4x_E}} \quad (12.21)$$

Where x_E is the critical exponent from some general operator. AS a reminder:

$$\langle \phi_\alpha(0) \phi_\alpha(r) \rangle = \frac{1}{|r|^{2x_\alpha}} \quad (12.22)$$

with $x_\alpha = d - y_\alpha$ with y_α the RG eigenvalues. These come from scaling/homogeneity. This is obtained from the integral $\int d^d r' \phi(r) \phi(r+r')$. If this is associated with the energy density, then:

$$x_E = d - \frac{1}{\nu} \quad (12.23)$$

and:

$$y_E = d - 2x_E = \frac{2}{\nu} - d \quad (12.24)$$

This is the answer we get from averaging at the critical point. y_E is irrelevant if $d\nu > 2$, the same answer as we got from the Harris criterion. Of course there is the pesky n which if we take $n = 1$ (the limit of a simple replica) then everything goes away.

So, the story was; does the phase transition still exist? There was a heuristic argument based on averaging over a sufficiently large region that gave the Harris criterion. We could also answer this via the laborious RG procedure, which ended up giving the same answer.

If the variables are relevant the replicas get coupled, if the variables are irrelevant then the replicas remain uncoupled, and we get a trivial answer.

13 Random Fields and Spin Glasses

13.1 Setting up the Random Field Problem

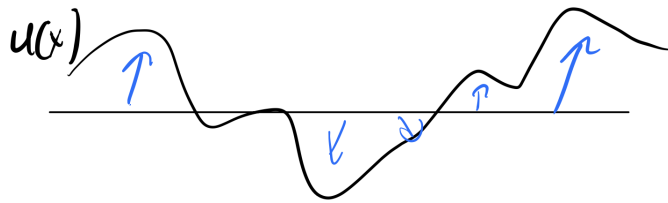
We consider the Hamiltonian:

$$H = \int d^d x \left(\frac{1}{2} \kappa (\nabla \mathbf{s})^2 + \frac{1}{2} t s^2 + \frac{1}{4} u s^4 - \mathbf{h}(x) \cdot \mathbf{s} \right) \quad (13.1)$$

where:

$$\overline{h_i(\mathbf{x})h_j(\mathbf{x}')} = \delta_{ij}\Delta\delta(\mathbf{x} - \mathbf{x}') \quad (13.2)$$

i.e. we have a random uncorrelated field. Although it is written down as a magnet, it is a kind of model that emerges very frequently in random media. For example, an elastic medium sitting on a surface which is random, causing the elastic to stretch in some ways.



Consider random fields h_i . If we RMS average over a length scale L :

$$\sqrt{\sum_{i=1}^L h_i^2} = h_0 \sqrt{cL^d} \quad (13.3)$$

where h_0 is a strength and c a concentration. This $\sim \sqrt{N}$ comes from a random walk.

If we average and get some nonzero up spin, and average in a neighbouring region and get a nonzero down spin, then we get an energy penalty. We have a gain of the form $h_0 L^{d/2}$ (from the spins within a domain aligning with the local field, from the $\mathbf{h}_i \cdot \mathbf{S}_i$ term) and a loss of the form JL^{d-1} (coming from the penalty of interacting neighbouring domains, the $\mathbf{S}_i \cdot \mathbf{S}_{i+1}$ term). When do each of the terms win out? When $\frac{d}{2} > d - 1$, i.e. when $d < 2$, the energy gain wins.

13.2 Solving the model

To attack this model, we consider the replica trick:

$$\overline{Z^n} = \int \mathcal{D}h(x) P(h(x)) \prod_{a=1}^n \int dS_a e^{-\beta H[S_a, h]} \quad (13.4)$$

We take the distribution of the fields to be Gaussian:

$$P(h) \sim e^{-\frac{h^2}{\Delta}} \quad (13.5)$$

This allows us to write:

$$\overline{Z^n} = \prod_a \int dS_a e^{-\beta H_{\text{eff}}[S_a, S_b]} \quad (13.6)$$

with:

$$H_{\text{eff}} = \int d^d x \left[\sum_{a=1}^n \frac{1}{2} \kappa (\nabla s_a)^2 + \frac{1}{2} t s_a^2 + \frac{1}{4} u s_a^4 \right] - \frac{1}{2} \beta \sum_{a,b} \Delta s_a \cdot s_b \quad (13.7)$$

(Rio note: This expression doesn't look like it makes sense with the sum over a, b s, but so be it.) Going to momentum space:

$$H_{\text{eff}} = \frac{1}{2} \int \frac{d^d k}{(2\pi)^d} \left(\sum_{a,b} (\kappa k^2 + t) \delta_{ab} - \beta \Delta J_{ab} \right) s_a(k) s_b(-k) + \frac{u}{4} s^4 \quad (13.8)$$

with J_{ab} just a matrix of 1s everywhere. If we had no fourth order term, we can view this as a Green's function:

$$\langle s_a^i(k) s_b^j(k') \rangle = G_{ab}^0 \delta(k + k') \quad (13.9)$$

with:

$$G_{ab}^0(k) = \delta_{ij} \left[\frac{T}{\kappa k^2 + t} \delta_{ab} + \frac{\Delta}{(\kappa k^2 + t)^2} J_{ab} \right] \quad (13.10)$$

If we study the scaling:

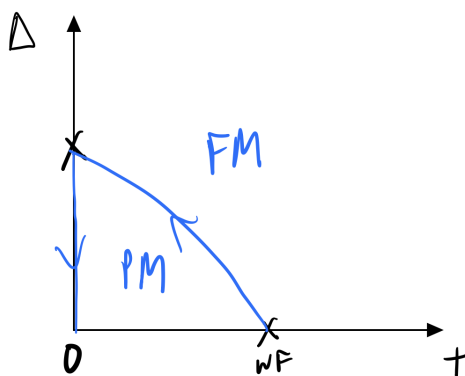
$$\kappa \rightarrow \kappa b^{d-2+2\zeta} \quad (13.11)$$

$$t \rightarrow t b^{d+2\zeta} \quad (13.12)$$

$$u \rightarrow u b^{d+4\zeta} \quad (13.13)$$

$$\frac{\Delta}{T} \rightarrow b^{d+2\zeta} \frac{\Delta}{T} = b^2 \Delta \quad (13.14)$$

with $\zeta = \frac{2-d}{2}$ where we recall that we have the RG scaling parameters from $s \rightarrow b^\zeta s$. We see that Δ has a positive eigenvalue and is therefore a relevant exponent. There is a new relevant field in the vicinity of the fixed point, and we flow away.



The fixed point Hamiltonian is:

$$H^* = \int d^d x \left[\frac{1}{2} \kappa (\nabla s)^2 \right] \quad (13.15)$$

If we ask what is $\langle \delta s^2 \rangle_{\text{thermal}}$, this is something we've seen before:

$$\langle \delta s^2 \rangle_{\text{thermal}} = \int_{1/L}^{\infty} \frac{d^d k}{(2\pi)^d} \frac{T}{\kappa k^2} \quad (13.16)$$

Now we want to ask how this scales depending on the lower momentum cutoff L . By a scaling argument, we see:

$$\langle \delta s^2 \rangle_{\text{thermal}} \sim \frac{T}{\kappa} L^{2-d} \quad (13.17)$$

If I look at the random field Hamiltonian we have the $\frac{\Delta}{(\kappa k^2 + t)^2}$ term. Then we have:

$$\langle \delta s^2 \rangle_{\text{RF}} \sim \int_{1/L}^{\infty} \frac{d^d k}{(2\pi)^d} \frac{\Delta}{\kappa^2 k^4} = \frac{\Delta}{\kappa^2} L^{4-d} \quad (13.18)$$

We then get the length scale over which one term wins:

$$\xi = \left(\frac{\kappa}{\Delta} \right)^{\frac{1}{4-d}} \quad (13.19)$$

In $D > 4$, this is minimized as $L \rightarrow \infty$, as $D < 4$ the length scale is finite.

Remark: Formal theory arguments gave $D = 3$ for a long time, in contrast with the quick argument we have here. It turns out the reason is that high-dimensional domain walls are hard. This is a reminder that back-of-the-envelope intuition is better than formal theory.

If we have an EOM:

$$\nabla^2 \psi + t \psi = h \quad (13.20)$$

with h a source, then:

$$\psi_k = \frac{h_k}{k^2 + t} \quad (13.21)$$

and so:

$$\langle \psi_k \psi_k \rangle = \frac{\langle h_k \rangle}{(k^2 + t)^2} \quad (13.22)$$

We'll stop here on random fields, but this is a rather generic problem. There are many scenarios where we get this type of equation of motion.

13.3 Ergodicity

Every system we've thought about so far has been ergodic. Namely, Z has been the average over all possible configuration, and the system is assumed to explore all of them. It turns out that the heart of exploring glasses is the breaking of this ergodicity.

Fundamentally, what this means is that if we go down to low temperature, the system gets trapped in a region of phase space, and cannot explore the rest of it. Notably, if we heat up the glass and cool it back down, it will go to a different region.

We define broken ergodicity as:

$$\langle \cdot \rangle = \sum_{\alpha} \omega_{\alpha} \langle \cdot \rangle_{\alpha} \quad (13.23)$$

with α a pure state. For example in the FM Ising model:

$$\langle \sigma \rangle = \frac{1}{2} \langle \sigma \rangle_+ + \frac{1}{2} \langle \sigma \rangle_- = 0 \quad (13.24)$$

There is a tendency for pure states to cluster; for example:

$$\langle \sigma_i \sigma_j \rangle \rightarrow \langle \sigma_i \rangle \langle \sigma_j \rangle \text{ for } |i - j| \rightarrow \infty \quad (13.25)$$

which means that the connected correlation function goes to zero.

Suppose we have some internal state α and a state β outside.



We can define a length scale r of the α droplet, and compute the free energy difference of the two terms:

$$\Delta F = \sigma r^{d-1} - (\delta f) r^d \quad (13.26)$$

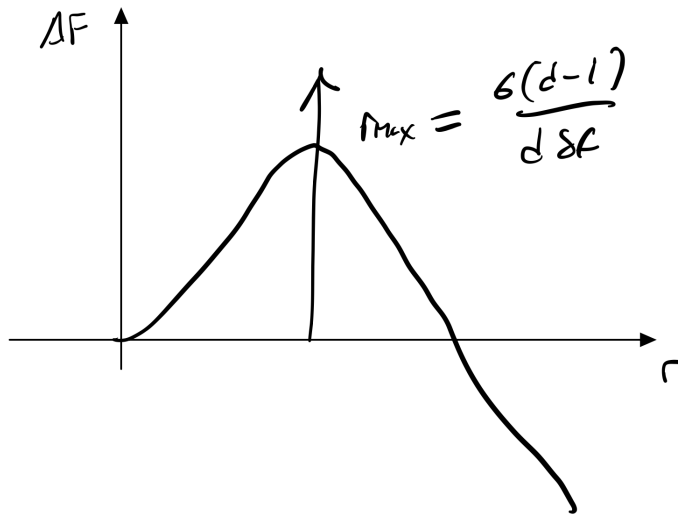
where δf is $F_\alpha - F_\beta$. This defines a r_{\max} where the free energy difference is maximized:

$$r_{\max} = \frac{\sigma(d-1)}{d\delta f} \quad (13.27)$$

Looking at the maximum difference in the free energy (the barrier):

$$\Delta F_{\max} = \sigma \left(\frac{\sigma}{\delta f} \right)^{d-1} \left(\left(\frac{d-1}{d} \right)^{d-1} - \left(\frac{d-1}{d} \right)^d \right) \quad (13.28)$$

Either $d \rightarrow \infty$ or $\delta f \rightarrow 0$ means we have large barriers between states. The latter case is found because we have two states with the same energy. $d \rightarrow \infty$ is also interesting as then the large dimension provides a big penalty/barrier.



Next day, we look at a infinite range glass model, which is exactly solvable. At infinite temperature, despite not having any order, it has many ground states whose energies are the same, related by hierarchies.

13.4 Order parameters

Previously, we looked at the magnetization:

$$m = \frac{1}{N} \sum_{i=1}^N \langle \sigma_i \rangle \quad (13.29)$$

now we are interested in states that freeze, but do not order. This gives rise to an order parameter proposed by Edwards and Anderson, which looks like:

$$q_{EA} = \frac{1}{N} \sum_{i=1}^N \langle \sigma_i \rangle^2 \quad (13.30)$$

where we look at the square of the average. The spin could in principle point in any direction, which I don't care about; I care about whether the thermal average could persist forever. This is something that one in principle could attempt to measure, and would distinguish between a paramagnetic state, and a frozen but disordered state. We can generalize this to overlaps:

$$q_{\sigma\tau} = \frac{1}{N} \sum_{i=1}^N \sigma_i \tau_i \quad (13.31)$$

which taking averages:

$$Q_{\alpha\beta} = \frac{1}{N} \sum_{i=1}^N \langle \sigma_i \rangle_{\alpha} \langle \tau_i \rangle_{\beta} \quad (13.32)$$

These overlaps allow us to probe the structure in phase space. In low temperature FM, $Q_{\alpha\alpha} = 1$ while $Q_{\alpha\beta} = 0$. Conversely, for the paramagnetic states everything is zero. More concretely, $q_{++} = q_{--} = m^2$ and $q_{+-} = q_{-+} = -m^2$.

14 Spin Glass

14.1 Order parameters

Last class we introduced the Edwards-Anderson order parameter:

$$Q_{EA} = \frac{1}{N} \sum_{i=1}^N \langle \sigma_i \rangle^2 \quad (14.1)$$

we then had the order parameter quantifying overlap:

$$q_{\sigma\tau} = \frac{1}{N} \sum_{i=1}^N \sigma_i \tau_i \quad (14.2)$$

and the average:

$$Q_{\alpha\beta} = \frac{1}{N} \sum_{i=1}^N \langle \sigma_i \rangle_{\alpha} \langle \sigma_i \rangle_{\beta} \quad (14.3)$$

where the $\langle \cdot \rangle_{\alpha}$ denotes an average over a pure state α .

In a glass, there is no spatial order. But, if we think about something that is frozen, if I look at a single spin and watch it for a very long time, it doesn't flip. So, order in a glass is really a statement about temporal order. The way to measure this would be to take the thermal average of a single spin i . If at this point I average over spins, I get zero (as there is no spatial order). But if I square it and then average, I get a finite value.

Generalizing Q_{EA} , we consider that our system may break ergodicity. It doesn't explore all of phase space; it gets stuck. There can be multiple states that survive for long times and are different. That said, they could potentially overlap. Thus there is an overlap matrix $Q_{\alpha\beta}$. If we take the configuration and break it up, there may be multiple pure states α, β and then there is some overlap between these states. We do the thermal average over a single site for two pure states, multiply and then take the sum.

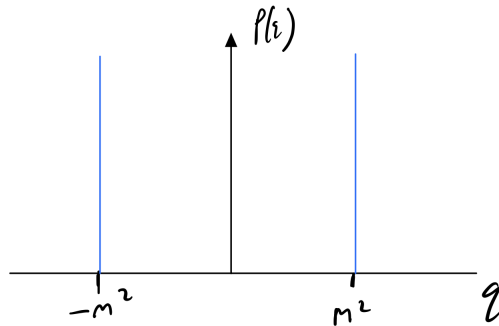
To be more precise, when we take this thermal average, we consider:

$$Q_{\alpha\beta} = \frac{1}{N} \sum_i \frac{1}{Z_{\alpha}} \int_{\sigma \in \alpha} \mathcal{D}\sigma \sigma_i e^{-H(\sigma)} \frac{1}{Z_{\beta}} \int_{\tau \in \beta} \mathcal{D}\tau \tau_i e^{-H(\tau)} = \frac{1}{Z_{\alpha} Z_{\beta}} \int_{\alpha} \mathcal{D}\sigma \int_{\beta} \mathcal{D}\tau q_{\sigma\tau} e^{-H(\sigma)} e^{-H(\tau)} \quad (14.4)$$

For a paramagnet we get $Q_{\alpha\alpha} = 0$, and for a frozen state we get $Q_{\alpha\alpha} = 1$. We now consider the overlap distribution:

$$P(q) = \frac{1}{Z^2} \int \mathcal{D}\sigma \mathcal{D}\tau e^{-H(\sigma) - H(\tau)} \delta(q - q_{\sigma\tau}) \quad (14.5)$$

For a ferromagnet, we have two equal delta peaks when $q = \pm m^2$.



But note that there is generically no reason that this distribution would just be delta functions. Generically, we have delta functions + continuum part. When we find this, we have replica symmetry breaking. The two ferromagnet states have a replica symmetry, but with a continuum no such symmetry is present.

14.2 Sherrington-Kirepatrick model

We consider the Hamiltonian:

$$H = \sum_{i,j=1}^N J_{ij} \sigma_i \sigma_j \quad (14.6)$$

where the J_{ij} are independent random variables:

$$J_{ij} = \pm \frac{J}{\sqrt{N}}. \quad (14.7)$$

This is a mostly solvable model, coming from the infinite range coupling. Note that even in a model as simple as this one, we have hidden broken symmetries and interesting phenomenology. While we are on the topic of Nobel prizes, this model was also used by Hopfield. He thought about the fact that we could use this model as a memory; if we chose $J_{ij} = \xi_i \xi_j$ with ξ a vector. Then, notice that our Hamiltonian would be minimized with this choice when $\sigma = \xi$ (just a ferromagnet that's been rotated).

14.3 p -spin spherical model

The model we will actually discuss looks more complicated, but is actually easier:

$$H = - \sum_{1 \leq i_2 < i_3 < \dots < i_p \leq N} J_{i_1, \dots, i_p} \sigma_{i_1} \sigma_{i_2} \dots \sigma_{i_p} \quad (14.8)$$

where we consider $p \geq 3$. This is a p -spin spherical model (think about the entire set of spins as a vector), with $\frac{1}{N} \sum_i \sigma_i^2 = 1$. Each site i has a real-valued spin. The J s are drawn from the Gaussian probability distribution:

$$P(J) = \exp\left(-\frac{1}{2} J^2 \frac{2N^{p-1}}{p!}\right) \quad (14.9)$$

where we note that $\sqrt{\bar{J}^2} \sim \frac{1}{N^{\frac{p-1}{2}}}$. Looking at the averaged Z (for $p = 3$):

$$\bar{Z} = \int \mathcal{D}\sigma \int \prod_{i < j < k} dJ_{ijk} \exp\left(-\left(J_{ijk}^2 \frac{N^{p-1}}{p!} + \beta J_{ijk} \sigma_i \sigma_j \sigma_k\right)\right) = \int \mathcal{D}\sigma \exp\left(\frac{\beta^2}{4N^{p-1}} p! \sum_i \sigma_i^2 \sum_j \sigma_j^2 \sum_k \sigma_k^2\right) \quad (14.10)$$

Each of these sums now just gives 1, so with the identity $p! \sum_{i > j > k} = \sum_{ijk}$, we just get the surface area of the surface:

$$\bar{Z} = \exp\left(\frac{N\beta^2}{4}\right) \Omega \quad (14.11)$$

where Ω is the solid angle. Now consider n replica version:

$$\bar{Z}^n = \int \mathcal{D}\sigma_i^\alpha \prod_{ijk} \int dJ_{ijk} \exp\left(-\frac{J^2 N^{p-1}}{p!} + J_{ijk} \beta \sum_\alpha \sigma_i^\alpha \sigma_j^\alpha \sigma_k^\alpha\right) = \int \mathcal{D}\sigma \exp\left(\frac{\beta^2}{4N^{p-1}} \sum_{\alpha, \beta=1}^n \left(\sum_i \sigma_i^\alpha \sigma_i^\beta\right)^p\right) \quad (14.12)$$

Now we have an overlap matrix! Let us call it:

$$Q_{\alpha\beta} = \frac{1}{N} \sum_i \sigma_i^\alpha \sigma_i^\beta \quad (14.13)$$

with $Q_{\alpha\alpha} = 1$. How do we deal with this integral? We've seen this a few things, and now what we will do is to introduce an auxiliary field. We introduce the trivial statement:

$$1 = \int dQ_{\alpha\beta} \delta\left(NQ_{\alpha\beta} - \sum_i \sigma_i^\alpha \sigma_i^\beta\right) = \int \mathcal{D}\lambda_{ab} e^{-\lambda(NQ_{\alpha\beta} - \sum_i \sigma_i^\alpha \sigma_i^\beta)} \quad (14.14)$$

But now doing the trick:

$$\delta(x) = \int dk e^{ikx} \quad (14.15)$$

So by exponentiating the 1, we can introduce the auxiliary field Q :

$$\bar{Z}^n = \int \mathcal{D}Q_{\alpha\beta} \mathcal{D}\lambda_{\alpha\beta} \mathcal{D}\sigma \exp\left(\frac{\beta^2 N}{4} \sum_{\alpha\beta} Q_{\alpha\beta}^p + N \sum_{\alpha\beta} \lambda_{\alpha\beta} Q_{\alpha\beta} - \sum_i \sum_{\alpha\beta} \sigma_i^\alpha \lambda_{\alpha\beta} \sigma_i^\beta\right) \quad (14.16)$$

So then:

$$\bar{Z}^n = \int \mathcal{D}Q_{\alpha\beta} \mathcal{D}\lambda_{\alpha\beta} \exp(-NS(Q, \lambda)) \quad (14.17)$$

where:

$$S(Q, \lambda) = -\frac{\beta}{4} \sum_{\alpha\beta} Q_{\alpha\beta}^p - \sum_{\alpha\beta} \lambda_{\alpha\beta} Q_{\alpha\beta} - \frac{1}{2} \log \det(2\lambda_{\alpha\beta}) \quad (14.18)$$

there are no longer fluctuations of spins, but now have two fields. We have an important factor of N ; this tells us the solution of this problem should be the saddle point of S , for everything away from the saddle point becomes exponentially suppressed. We need to consider the limits $n \rightarrow 0$ (we are interested in this because $F = \log Z = \lim_{n \rightarrow 0} \frac{Z^n - 1}{n}$), $N \rightarrow \infty$ (we want a large number of spins/system size, and want $N \rightarrow \infty$ so the saddle point is the solution). Then:

$$-\beta F = \lim_{N \rightarrow \infty} \lim_{n \rightarrow 0} \frac{1}{nN} \int \mathcal{D}Q \mathcal{D}\lambda e^{-NS} \quad (14.19)$$

The number of independent elements in $Q_{\alpha\beta}$ are $\frac{n(n-1)}{2}$, which as $n \rightarrow 0$ becomes negative. This is strange, but we press on - we will do this by interchanging the order of limits. We consider (to obtain the saddle point):

$$\frac{\partial}{\partial \lambda_{\alpha\beta}} \left[-\lambda_{\alpha\beta} Q_{\alpha\beta} + \frac{1}{2} \log \det(2\lambda) \right] = 0 \implies -Q_{\alpha\beta} + \frac{1}{2} (\lambda^{-1})_{\alpha\beta} = 0 \implies Q^{-1} = 2\lambda \quad (14.20)$$

Thus replacing the $N \rightarrow \infty$ limit integral with the saddle point value:

$$F = \lim_{n \rightarrow \infty} -\frac{1}{2\beta n} \left[\frac{\beta^2}{2} \sum_{\alpha\beta} Q_{\alpha\beta}^p + \log \det Q \right] \quad (14.21)$$

If we now minimize w.r.t Q :

$$0 = \frac{\partial F}{\partial Q} = \frac{\beta^2 p}{2} Q_{\alpha\beta}^{p-1} + Q_{\alpha\beta}^{-1} \quad (14.22)$$

What do we know? We know $Q_{\alpha\alpha} = 1$. Since there is nothing to distinguish the replicas (i.e. if the model is replica-symmetric) then we would say that all the off-diagonal entries of Q are the same, and equal to some value q_0 . Then, writing:

$$Q_{ab} = q_0 + (1 - q_0) \delta_{ab} \quad (14.23)$$

it is clear that:

$$Q_{ab}^{-1} = \frac{1}{1 - q_0} \delta_{ab} - \frac{q_0}{(1 - q_0)[1 + (n - 1)q_0]} \quad (14.24)$$

here, it's a bit interesting that the n shows up, and I can begin to compute this in the $n \rightarrow 0$ limit:

$$\lim_{n \rightarrow 0} \frac{\partial F}{\partial Q} = \frac{\beta^2 p}{2} q_0^{p-1} - \frac{q_0}{(1 - q_0)^2} = 0 \quad (14.25)$$

This has two solutions. $q_0 = 0$, which is the paramagnetic, and the pair of solutions:

$$q_0^{p-2}(1 - q_0)^2 = \frac{2}{p\beta^2} \quad (14.26)$$

This is a replica-symmetric solution. It has the unfortunate problem of not being stable. While it is a solution, if we look perturbatively about this solution, e.g. break the system into many pieces, we see that the energy is decreased. Another way to see the instability; it has to not just be a saddle point/extremum, but also a minimum. This requires the eigenvalues to be positive definite.

What does this mean? We still know that everything is dominated by a saddle point, and just need to find what it is. Let's think about the average/first moment:

$$q^{(1)} = \bar{Q}_{\alpha\beta} = \lim_{n \rightarrow 0} \int \mathcal{D}Q_{\alpha\beta} Q_{\alpha\beta} e^{-NS(Q_{\alpha\beta})} \quad (14.27)$$

what happens when we average over saddle points? When we look at the α, β s, we have to think about which ones we take. Eventually, this turns into:

$$\bar{Q}_{\alpha\beta} = \lim_{n \rightarrow 0} \frac{2}{n(n-1)} \sum_{\alpha > \beta} Q_{\alpha\beta} \quad (14.28)$$

This average is the first moment, and of course we can generate the k th moment, which gives:

$$q^{(k)} = \lim_{n \rightarrow 0} \frac{2}{n(n-1)} \sum_{\alpha\beta} \overline{Q_{\alpha\beta}^k} \quad (14.29)$$

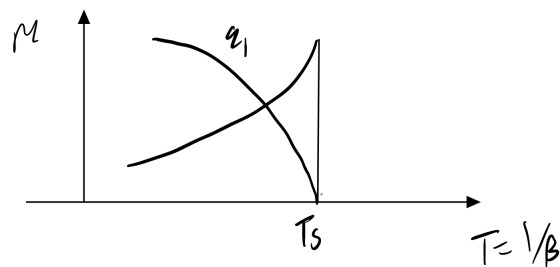
the higher moments are more and more sensitive to the broken symmetry. We then get the distribution:

$$\bar{P}(q) = \lim_{n \rightarrow 0} \frac{2}{n(n-1)} \sum_{a>b} \delta(q - Q_{ab}) \quad (14.30)$$

The distribution is governed by the fact that the likely overlap is given by a value of q . Working with this probability distribution, we can rewrite it as:

$$P(q) = (1 - m)\delta(q - q_1) - m\delta(q - q_0) \quad (14.31)$$

with $0 < m < 1$. We then have the distribution:



it is recommended to look at the homework problem before Friday - we'll come back to this last part, which we rushed a bit. What is the takehome message? The reason to writing down the odd model is that it is a mean field theory, and in MF models we are able to compute things easily from the saddle point. We are used to saddle points being very simple, but actually there is more complex structure to the saddle points than we are used to. But, if we look at finite n , we can visualize these solutions. We had a replica-symmetric solution, but the replica symmetry gets broken, leading to a fractal-like structure in the Q matrix.

15 Replica Symmetry Breaking, Neural Nets, Boltzmann Machines

I was absent for this lecture, and these lecture notes are obtained from Kalpak Duddella's notes - many thanks to him!

In this lecture, we wrap up the discussion of spin glasses and replica symmetry breaking, before moving onto the last topic of discussion for this course - Neural networks and Boltzmann machines.

15.1 Replica Symmetry Breaking, Concluded

We consider the replica matrix ansatz Q_{ab} , as you will study in HW3. Its structure “naturally clusters” due to off-diagonal (correlation) elements, forming a block structure about the diagonal. We obtain this by guessing the block form as the solution to the non-linear EoM and minimizing the free energy. But, we also have to check the stability of the solutions.

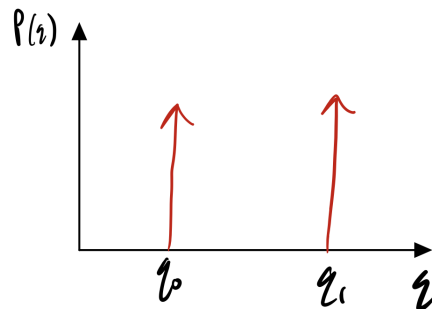
In more detail, we have 1s on the diagonal, q_1 in $m \times m$ blocks around the diagonal, and q_0 elsewhere, with $1 > q_1 > q_0$. We have the probability distribution:

$$P(q) = \frac{m-1}{n-1} \delta(q - q_1) + \frac{n-m}{n-1} \delta(q - q_0) \quad (15.1)$$

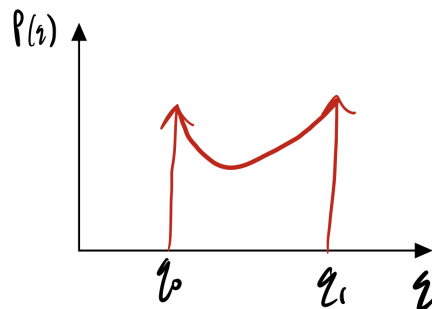
which as $n \rightarrow 0$ becomes the two delta peaks:

$$p(q) = (1-m)\delta(q - q_1) + m\delta(q - q_0) \quad (15.2)$$

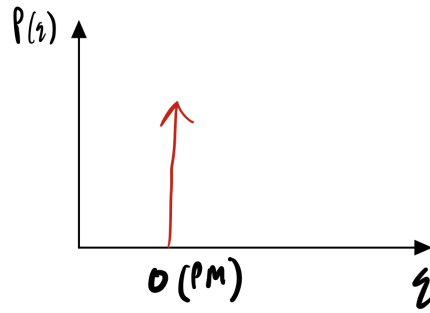
Graphically, we have the distribution as sketched last class. In this limit, we have 1-step symmetry breaking in the exact solution of the p -spin spherical model.



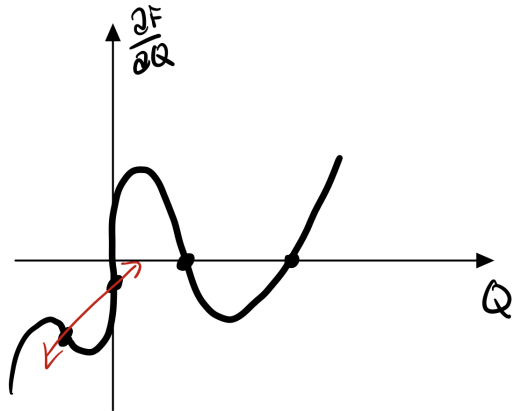
For $T < T_s$, we have a continuous solution where $P(a)$ varies smoothly between the two peaks at q_0, q_1 , with peaks getting closer together as $T \rightarrow T_s$.



For $T > T_c$, we have “ultrametricity”; full replica symmetry breaking.



This is related to lines connecting points on the $\frac{\partial F}{\partial Q}$ curve.



15.2 Neural Nets

John Hopfield considered the following model for associative memory - consider the vector χ . We then store this memory by flipping spins in an (Ising) array, with system Hamiltonian:

$$H = - \sum_{\langle ij \rangle} J_{ij} S_i S_j \quad (15.3)$$

where the couplings are:

$$J_{ij} = J \xi_i \xi_j \quad (15.4)$$

Thus:

$$H = -J \sum_i \xi_i S_i \sum_j \xi_j S_j \quad (15.5)$$

If we now generalize to store a set of bit strings $\{\chi^a\}_a$. Then for an associative memory, we take:

$$J_{ij} = \sum_a \chi_i^a \chi_j^a \quad (15.6)$$

with a indexing over different bitstrings and i indexing the position in a given bitstring. This is a "one-layer neural network" (we take $1 = \uparrow$, $0 = \downarrow$). Hopfield won a Nobel for a (insert choice word) ferromagnet!

For a $N \times N$ neural network, we can store about $\sim 0.15N$ memories.

15.3 Boltzmann Machines

Two step phases for using Boltzmann machines:

1. We first run a training phase to obtain the ξ_i .
2. We then run a readout phase to allow the system to evolve on a fixed array of J s constructed from the learned ξ s.

We consider the probability weights for a given spin configuration s

$$P(s) = \exp(-E(s))/Z \tag{15.7}$$

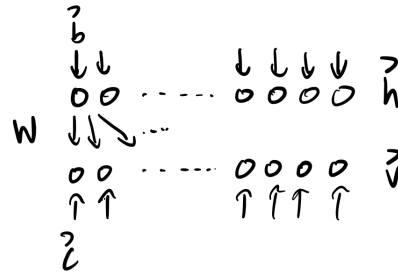
where the energy function is:

$$E(s) = \sum_{ij} U_{ij}s_i s_j - \sum_i b_i s_i \tag{15.8}$$

with U_{ij} the weights and b_i the biases. We need to *learn* the U s and b s to have the model $P(s)$ fit the data, i.e.:

$$P(s)|_{\text{model}} \approx P(s)|_{\text{data}}. \tag{15.9}$$

Comparing neural networks with spin models, a layered neural network corresponds to an auxiliary field model.



For simplicity, we consider a 2-layer network, composed of a hidden layer of nodes (vector of spins) \mathbf{h} and a visible layers of nodes \mathbf{v} . We have a bias \mathbf{b} on the hidden nodes and \mathbf{c} on the visible nodes. We also have a weight-matrix W between the two layers which computes “scores” as a linear transformation.

The energy of a neuron configuration specified by \mathbf{v}, \mathbf{h} is:

$$E(v, h) = -\mathbf{b} \cdot \mathbf{v} - \mathbf{c} \cdot \mathbf{h} - v_i W_{ij} h_j \tag{15.10}$$

We then have the conditional probabilities:

$$P(\mathbf{h}|\mathbf{v}) = \prod_i P(h_i|\mathbf{v}) \tag{15.11}$$

$$P(\mathbf{v}|\mathbf{h}) = \prod_i P(v_i|\mathbf{h}) \tag{15.12}$$

Which we can also write as:

$$P(\mathbf{h}|\mathbf{v}) = \frac{P(\mathbf{h}, \mathbf{v})}{P(\mathbf{v})} = \frac{1}{P(\mathbf{v})Z} e^{\mathbf{b} \cdot \mathbf{v} + \mathbf{c} \cdot \mathbf{h} + v_i W_{ij} h_j} = \frac{1}{Z'} e^{\mathbf{c} \cdot \mathbf{h} + v_i W_{ij} h_j} \tag{15.13}$$

With $\frac{1}{Z'} = \frac{1}{P(\mathbf{v})Z} e^{\mathbf{b} \cdot \mathbf{v}}$.

Now, we have some data probability distribution $P_{\text{data}}(\mathbf{x})$. We have a model probability distribution $P_{\text{model}}(\mathbf{x}; \theta)$ with model parameters θ . The goal is to find the optimal model parameters, obtained from:

$$\theta_{\text{model}} = \operatorname{argmax}_{\theta} \prod_{i=1}^m P_{\text{model}}(x_i, \theta) = \sum_{i=1}^m \log P_{\text{model}}(x_i, \theta) \tag{15.14}$$

16 Non-equilibrium Dynamics

Today we will discuss the Fokker-Planck equation and how to derive Field theories. We won't have time to discuss anomalous diffusion; how to start with a differential equation that looks like it should scale in an obvious way and then get interesting critical exponents. There's also interesting discussion of analyzing a laser through a statistical field theory lens, which we will not quite get to.

16.1 Diffusion Equation ($V = 0$)

Our basic starting point is an equation of motion of the form:

$$m\ddot{x} = -\frac{\dot{x}}{\mu} - \frac{\partial V}{\partial x} + f_{\text{noise}}(t) \quad (16.1)$$

i.e. a particle moving deterministically in a potential under the influence of drag and stochastic noise. The potential term leads to ballistic motion, while the noise term leads to diffusive motion. We consider the limit of small \dot{x} where we can neglect the acceleration term, thus:

$$\dot{x} = v(x) + \eta(t) \quad (16.2)$$

with:

$$v(x) = -\mu \frac{\partial V}{\partial x} \quad (16.3)$$

and:

$$\langle \eta \rangle = 0 \quad (16.4)$$

$$\langle \eta(t)\eta(t') \rangle = 2D\delta(t-t') \quad (16.5)$$

i.e. Markovian white noise.

In the $V = 0$ case, we have the solution:

$$x(t) = x(0) + \int_0^t dt' \eta(t') \quad (16.6)$$

where:

$$\langle (x(t) - x(0))^2 \rangle = \int_0^t d\tau_1 \int_0^t d\tau_2 \langle \eta(\tau_1)\eta(\tau_2) \rangle = 2Dt. \quad (16.7)$$

We then obtain the propagator of the diffusion equation:

$$P(x, t) = G(x, t, 0, 0) = \frac{1}{(4\pi Dt)^{3/2}} e^{-\frac{x^2}{4Dt}} \quad (16.8)$$

With P the solution of:

$$\frac{\partial P}{\partial t} = -D\nabla^2 P \quad (16.9)$$

this is the familiar diffusion equation.

16.2 Fokker-Planck Equation ($V \neq 0$)

Let's now look at advancing time by some small time ϵ :

$$P(x, t + \epsilon) = \int d^3x' P(x', t) \langle x|T_\epsilon|x' \rangle \quad (16.10)$$

where $\langle x|T_\epsilon|x' \rangle$ represents a transition amplitude.

Splitting up the two parts of the motion:

$$x = x' + v(x)\epsilon + \eta_\epsilon \quad (16.11)$$

where:

$$\eta_\epsilon = \int_t^{t+\epsilon} dt \eta(t) \quad (16.12)$$

which has the properties:

$$\langle \eta_\epsilon \rangle = 0 \quad (16.13)$$

$$\langle \eta_\epsilon^2 \rangle = 6D\epsilon. \quad (16.14)$$

With this, we can define the transition matrix amplitude:

$$\langle x|T_\epsilon|x' \rangle = \frac{1}{(4\pi D\epsilon)^{3/2}} \exp\left(\frac{-(x - x' - \epsilon v(x'))^2}{4D\epsilon}\right) \quad (16.15)$$

Let us explain this. If we had a potential, the particle would have rolled to a specific place deterministically from x' . If the particle instead reaches x , we need the noise to be precisely the correct magnitude for the particle to get to the position we want it to be, which gives us the above expression. Then:

$$P(x, t) = \int d^3 x' \frac{1}{(4\pi D\epsilon)^{3/2}} P(x', t) \exp\left(\frac{-(x - x' - \epsilon v(x'))^2}{4D\epsilon}\right) \quad (16.16)$$

for $|x - x'|$ small we can then replace $P(x', t)$ in the integral with $P(x, t)$, and linearize:

$$\frac{\partial P}{\partial t} + \nabla \cdot \mathbf{J} = 0 \implies \mathbf{J} = \mathbf{v}P - D\nabla P \quad (16.17)$$

The interpretation of this is intuitive. We have a drift current (from the potential) and a diffusion current (from the noise).

The stationary/equilibrium station is:

$$P_{\text{eq}}(x) = e^{-V(x)/k_B T} \quad (16.18)$$

with:

$$\nabla P_{\text{eq}} = \frac{\mathbf{v}}{\mu k_B T} P_{\text{eq}} \implies D = \mu k_B T \quad (16.19)$$

So we identify what the parameter D must be. This is usually called the ‘‘fluctuation-dissipation’’ theorem.

16.3 Generalizing to a field - Model A

We write down the Hamiltonian as a function of a field m :

$$H[m] = \int d^d x \left(\frac{1}{2} r m^2 + \frac{1}{4} u m^4 + \frac{1}{2} \kappa (\nabla m)^2 + \dots \right) \quad (16.20)$$

The analogous EoM is:

$$\frac{\partial m}{\partial t} = -\mu \frac{\delta H}{\delta m} + \eta(t) = -\mu r m - \mu u m^3 + \mu \kappa (\nabla^2 m) + \eta \quad (16.21)$$

In the Gaussian theory (taking $u = 0$) we have:

$$\frac{\partial m(q, t)}{\partial t} = -\mu(r + \kappa q^2)m + \eta \quad (16.22)$$

where:

$$\langle \eta \rangle = 0 \quad (16.23)$$

$$\langle \eta_q \eta_{q'} \rangle = 2D\delta(t-t')\delta(q-q')(2\pi)^d \quad (16.24)$$

There is then a q -mode decay time:

$$\frac{1}{\tau(q)} = \mu(r + \kappa q^2) \quad (16.25)$$

from this we can determine what happens to the fluctuations:

$$\langle m(q, t)m(q', t) \rangle = \frac{D\delta(q+q')}{\mu(v + \kappa q^2)} \quad (16.26)$$

where there is no time-dependence for we have performed an ensemble average. At very long times the distribution approaches the Lorentzian. Note that $\frac{D}{\mu} = k_B T$. This is what is known as “critical slowing down”.

In some sense we pulled the EoM out of a hat, and its not the most general EoM we could have written down. In particular we notice the order parameter is not conserved. This specific model is known as model A. There are models A-F, check out the review article by Honenberg in reviews of modern physics if interested. But for now, we only study A/B.

16.4 Model B

Now consider a model where we do want the order parameter to be conserved:

$$\frac{d}{dt} \int d^d x m(x, t) = 0 \quad (16.27)$$

so we now have the equation of motion:

$$\frac{\partial m}{\partial t} = -\nabla j + \nabla \sigma \quad (16.28)$$

with $j = \mu v$ and $\sigma \sim \eta$. This new model has the constraint that the magnetization is fixed, and this gets rid of low-order derivatives. We thus end up with:

$$\frac{\partial m}{\partial t} = \mu v \nabla^2 m - \mu \kappa \nabla^4 m + \eta \quad (16.29)$$

Thus we have different dynamics associated with the conservation of the order parameter.

16.5 Noise and Optimal Trajectories

The above was all pretty heuristic. Let’s explore something that’s less heuristic, with a similar starting point:

$$\dot{X} = A(X) + B\zeta \quad (16.30)$$

with $\langle \zeta \zeta \rangle = 2D\delta(t-t')$.

We solve this via the identity:

$$1 = Z[\zeta] = \int \mathcal{D}X(t) J(X) \delta[\partial_t X - A(X) - B\zeta] \quad (16.31)$$

where we have stuck in the EoM as a delta function.

We then generate an auxiliary field X^q using an identity:

$$Z[\zeta] = \int \mathcal{D}X(t) \int \mathcal{D}X^q(t) \exp(-2i \int dt X^q(t) (\partial_t X - A - B\zeta)) \quad (16.32)$$

Why we set this up is so that we can integrate over the noise (also let's set $B = 1$):

$$Z = \int \mathcal{D}\zeta e^{-\frac{1}{4} \int dt \zeta^2} Z[\zeta] = \int \mathcal{D}X \mathcal{D}X^q \exp\left(\int dt \left[-2iX^q(\partial_t X - A(X)) - 4(X^q)^2\right]\right) \quad (16.33)$$

We now have an action that is a function of two fields. Faced with something like this, we look at the saddle point equations (here of two variables). This yields two equations:

$$\dot{X} = A(X) + 4iX^q D(X) \quad (16.34)$$

$$i\dot{X}^q = -iX^q A'(X) + 2(X^q)^2 D' \quad (16.35)$$

with $' = \frac{\partial}{\partial X}$. X^q must be imaginary since X is real. Let us do a change of variables $P = 2iX^q$. In this case, it turns out that we can write the equations of motion we saw above as:

$$\dot{X} = \partial_P H[P, X] \quad (16.36)$$

$$\dot{P} = -\partial_X H[P, X] \quad (16.37)$$

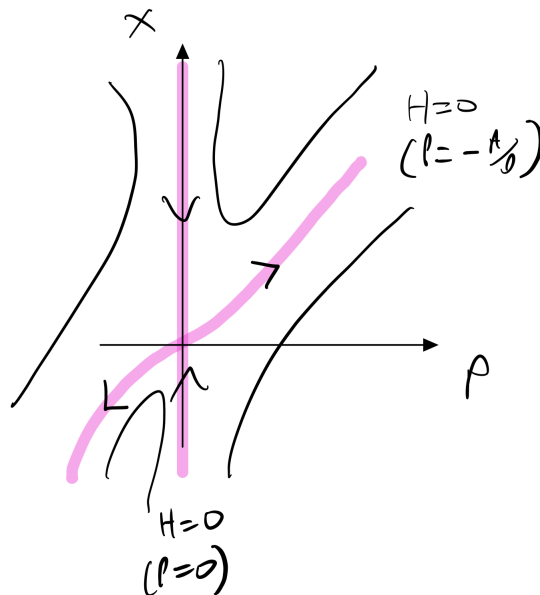
with:

$$H = PA(X) + P^2 D(X). \quad (16.38)$$

These look exactly like Hamilton's equations! X^q "looks like" a conjugate momentum. We can then identify an action:

$$iS[X, P] = - \int dt [P\dot{X} - H] \quad (16.39)$$

and think about optimal trajectories. The simplest one to think about would be to take $A(X) = -\frac{\partial V}{\partial X}$. Since $D(X) \sim T$, we have two stable trajectories where H is zero; when $P = 0$, or when $P = -\frac{A}{D}$. Sketching the flows in phase space:



In equilibrium of this problem, what is $P(X_0)$? Let us write the action:

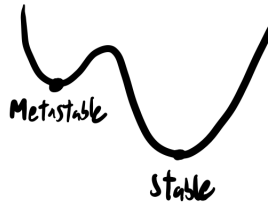
$$iS[X_0] = - \int dt P \dot{X} = - \int_0^{X_0} P(X) \quad (16.40)$$

Then taking $P(X) = -\frac{A}{D} = -\frac{V'}{D}$:

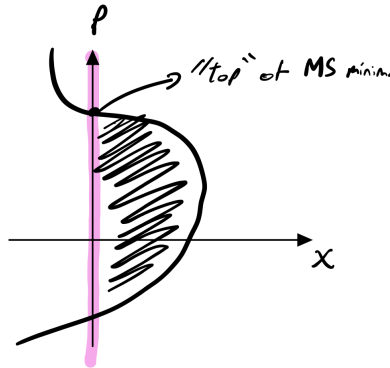
$$iS[X_0] = -\frac{1}{T} \int_0^{X_0} dx V'(X) = \frac{1}{T} (V - V_0) \quad (16.41)$$

which is just the Boltzmann distribution!

A last point. Suppose the potential has a metastable minimum:



Then, what I am interested in is the probability of escape, given by the shaded region:



which gives:

$$S[X_S] = -\frac{V(X_S) - V(0)}{T} \quad (16.42)$$

which is just WKB.

Here, we find an optimal trajectory, which gives us where we want to be. This is of course an approximation. We can go back and add more terms to our theory, making it a dynamical field theory.

Note: The best review of this material comes from Kamenev.

16.6 Where have we come?

There are generic and universal properties of theories. They are made most evident at the point where these theories are hardest to solve. The underlying principle behind this universality is scaling and embodied by the renormalization group. The definition of the interesting point/phase transition is associated with the scale invariant quantities. This doesn't always happen, and almost any real system deviates from

this, but such deviation can itself lead to interesting phenomenology, e.g. in the case of disorder and spin glasses.

If you have the time, next week we'll discuss quantum critical points (adding back the non-linear terms and seeing their effects). We will also study the laser and fluids/porous media, which gives us anomalous scaling/stretched power laws.