

Dirty tricks for statistical mechanics

Martin Grant
Physics Department, McGill University

© MG, August 2004, version 0.91

Preface

These are lecture notes for PHYS 559, Advanced Statistical Mechanics, which I've taught at McGill for many years. I'm intending to tidy this up into a book, or rather the first half of a book. This half is on equilibrium, the second half would be on dynamics.

These were handwritten notes which were heroically typed by Ryan Porter over the summer of 2004, and may someday be properly proof-read by me. Even better, maybe someday I will revise the reasoning in some of the sections. Some of it can be argued better, but it was too much trouble to rewrite the handwritten notes. I am also trying to come up with a good book title, amongst other things. The two titles I am thinking of are "Dirty tricks for statistical mechanics", and "Valhalla, we are coming!". Clearly, more thinking is necessary, and suggestions are welcome.

While these lecture notes have gotten longer and longer until they are almost self-sufficient, it is always nice to have real books to look over. My favorite modern text is "Lectures on Phase Transitions and the Renormalisation Group", by Nigel Goldenfeld (Addison-Wesley, Reading Mass., 1992). This is referred to several times in the notes. Other nice texts are "Statistical Physics", by L. D. Landau and E. M. Lifshitz (Addison-Wesley, Reading Mass., 1970) particularly Chaps. 1, 12, and 14; "Statistical Mechanics", by S.-K. Ma (World Science, Phila., 1986) particularly Chaps. 12, 13, and 28; and "Hydrodynamic Fluctuations, Broken Symmetry, and Correlation Functions", by D. Forster (W. A. Benjamin, Reading Mass., 1975), particularly Chap. 2. These are all great books, with Landau's being the classic. Unfortunately, except for Goldenfeld's book, they're all a little out of date. Some other good books are mentioned in passing in the notes.

Finally, at the present time while I figure out what I am doing, you can read these notes as many times you like, make as many copies as you like, and refer to them as much as you like. But please don't sell them to anyone, or use these notes without attribution.

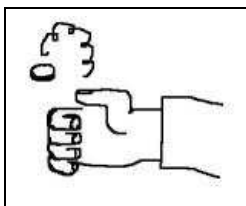
Contents

1	Motivation	1
2	Review of Stat Mech Ideas	5
3	Independence of Parts	7
3.1	Self Averaging	9
3.2	Central Limit Theorem	11
3.3	Correlations in Space	15
4	Toy Fractals	21
4.1	Von Koch Snowflake	22
4.2	Sierpinski Triangle	24
4.3	Correlations in Self-Similar Objects	27
5	Molecular Dynamics	31
6	Monte Carlo and the Master Equation	39
7	Review of Thermodynamics	49
8	Statistical Mechanics	55
9	Fluctuations	59
9.1	Thermodynamic Fluctuations	62
10	Fluctuations of Surfaces	67
10.1	Lattice Models of Surfaces	70
10.2	Continuum Model of Surface	72
10.3	Impossibility of Phase Coexistence in $d=1$	88
10.4	Numbers for $d=3$	90
11	Broken Symmetry and Correlation Functions	93
11.1	Proof of Goldstone's Theorem	96

12 Scattering	99
12.1 Scattering from a Flat Interface	103
12.2 Roughness and Diffuseness	106
12.3 Scattering From Many Flat Randomly-Oriented Surfaces	108
13 Fluctuations and Crystalline Order	113
14 Ising Model	121
14.1 Lattice-Gas Model	126
14.2 One-dimensional Ising Model	127
14.3 Mean-Field Theory	130
15 Landan Theory and Ornstien-Zernicke Theory	139
15.1 Landan Theory	144
15.2 Ornstein — Zernicke Theory of Correlations	145
16 Scaling Theory	151
16.1 Scaling With ξ	155
17 Renormalization Group	161
17.1 ϵ -Expansion RG	170

Chapter 1

Motivation



Why is a coin toss random?

Statistical Mechanics: Weighting for averages is by “equal a priori probability.” So, all many—body quantum states with the same energy (for a given fixed volume) have equal weight for calculating probabilities.

This is a mouthful, but at least it gives the right answer. There are two microscopic states, *HEADS* and *TAILS*, which are energetically equivalent. So

the result is random.

However nice it is to get the right answer, and however satisfying it is to have a prescription for doing other problems, this is somewhat unsatisfying. It is worth going a little deeper to understand how microscopic states (HEADS or TAILS) become macroscopic observations (randomly HEADS or TAILS).

To focus our thinking it is useful to do an experiment. Take a coin and flip it, always preparing it as heads. If you flip it at least 10–20cm in the air, the result is random. If you flip it less than this height, the result is not random. Say you make N attempts, and the number of heads is n_{heads} and the number of tails is n_{tails} , for different flipping heights h . Then you obtain something like this:

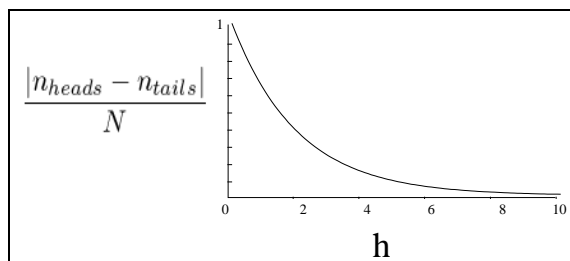


Figure 1.1: Bias of Flips With Height h

There are a lot of other things one could experimentally measure. In fact, my cartoon should look exponential, for reasons we will see later on in the course.

Now we have something to think about: Why is the coin toss biased for flips of small height? In statistical mechanics language we would say, “why is the coin toss *correlated* with its initial state?” Remember the coin was always prepared as heads. (Of course, if it was prepared randomly, why flip it?) There is a length scale over which those correlations persist. For our example, this scale is $\sim 2\text{--}3\text{cm}$, which we call ξ , the correlation length. The tailing off of the graph above follows:

$$\text{Correlations} \sim e^{-\xi/h}$$

which we need more background to show. But if we did a careful experiment, we could show this is true for large h . It is sensible that

$$\xi \geq \text{coin diameter}$$

as a little thinking will clarify.

This is still unsatisfying though. It leads to the argument that coin tosses are random because we flip the coin much higher than its correlation length. Also, we learn that true randomness requires $\xi/h \rightarrow 0$, which is analogous to requiring an infinite system in the thermodynamic limit. [In fact, in passing, note that for $\xi/h = \mathcal{O}(1)$ the outcome of the flip (a preponderance of heads) does not reflect the microscopic states (HEADS and TAILS equivalent). To stretch a bit, this is related to broken symmetry in another context. That is, the macroscopic state has a lower symmetry than the microscopic state. This happens in phase transitions, where $\xi/L = \mathcal{O}(1)$ because $\xi \approx L$ as $L \rightarrow \infty$, where L is the edge length of the system. It also happens in quantum mesoscopic systems, where $\xi/L = \mathcal{O}(1)$ because $L \approx \xi$ and L is somewhat small. This latter case is more like the present example of coin tossing.]

But what about Newton’s laws? Given an initial configuration, and a *precise* force, we can calculate the trajectory of the coin if we neglect inhomogeneities in the air. The reason is that your thumb cannot apply a precise force for flipping to within one correlation length of height. The analogous physical problem is that macroscopic operating conditions (a fixed pressure or energy, for example) do not determine microscopic states. Many such microscopic states are consistent with one macroscopic state.

These ideas, correlations on a microscopic scale, broken symmetries, and correlation lengths, will come up throughout the rest of the course. The rest of the course follows this scenario:

- Fluctuations and correlations: What can we do *without* statistical mechanics?
- What can we do *with* thermodynamics and statistical mechanics?
- Detailed study of interfaces.
- Detailed study of phase transitions.

The topics we will cover are as current as any issue of Physical Review E. Just take a look.

We'll begin the proper course with a review of, and a reflection on, some basic ideas of statistical mechanics.

Chapter 2

Review of Stat Mech Ideas

In statistical physics, we make a few assumptions, and obtain a distribution to do averages. A famous assumption is that of “equal *a priori* probability.” By this it is meant that all microscopic states, consistent with a particular fixed total energy E and number of particles N , are equally likely. This gives the microcanonical ensemble. It can then be checked that thermodynamic quantities have exceedingly small fluctuations, and that one can derive — from the microcanonical ensemble — the canonical ensemble, where temperature T and volume V are fixed. It is this latter ensemble which is most commonly used. Microscopic states, consistent with T and V are *not* equally likely, they are weighted with the probability factor

$$e^{-E_{state}/k_B T} \quad (2.1)$$

where E_{state} is the energy of the state, and k_B is Boltzmann’s Constant. The probability of being in a state is

$$\rho = \frac{1}{Z} e^{-E_{state}/k_B T} \quad (2.2)$$

where the partition function is

$$Z = \sum_{\{states\}} e^{-E_{state}/k_B T} \quad (2.3)$$

The connection to thermodynamics is from the Helmholtz free energy

$$F = E - TS \quad (2.4)$$

where S is the entropy, and

$$e^{-F/k_B T} = Z \quad (2.5)$$

Note that since F is extensive, that is $F \sim \mathcal{O}(N)$, we have

$$Z = (e^{\bar{f}/k_B T})^N$$

where $F = fN$. This exponential factor of N arises from the sum over states which is typically

$$\sum_{\{states\}} = e^{\mathcal{O}(N)} \quad (2.6)$$

For example, for a magnetic dipole, which can have each constituent up or down, there are evidently

$$dipole \#1 \times dipole \#2 \dots dipole \#N = 2 \times 2 \times \dots 2 = 2^N = e^{N \ln 2} \quad (2.7)$$

total states. The total number of states in a volume V is

$$\frac{V^N}{N!} \approx \frac{V^N}{N^N} = e^{N \ln \frac{V}{N}} \quad (2.8)$$

where the factor $N!$ accounts for indistinguishability.

These numbers are huge in a way which cannot be imagined. Our normal scales sit somewhere between microscopic and astronomical. Microscopic sizes are $\mathcal{O}(\text{\AA})$ with times $\mathcal{O}(10^{-12} \text{sec})$. Astronomical sizes, like the size and age of the universe are $\mathcal{O}(10^{30} m)$ and $\mathcal{O}(10^{18} s)$. The ratio of these numbers is

$$\frac{10^{30} m}{10^{-10} m} \sim 10^{40} \quad (2.9)$$

This is in no way comparable to (!)

$$10^{10^{23}} \quad (2.10)$$

In fact, it seems a little odd (and unfair) that a theorist must average over e^N states to explain what an experimentalist does in one experiment.

Chapter 3

Independence of Parts

It is worth thinking about experiments a bit to make this clear. Say an experimentalist measures the speed of sound in a sample, as shown. The experi-

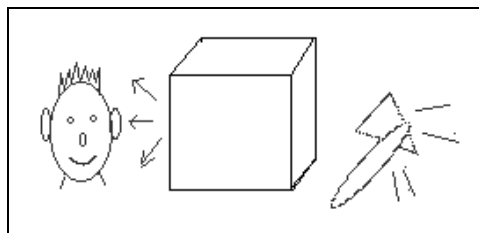


Figure 3.1: Experiment 1

mentalist obtains some definite value of c , the speed of sound. Afterwards, the experimentalist might cut up the sample into many parts so that his or her friends can do the same experiment. That is as shown in Fig. 3.2. Each of

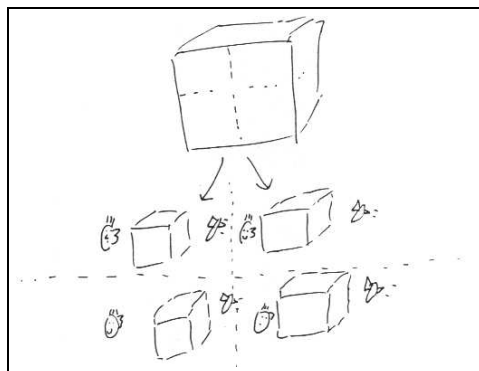


Figure 3.2: Experiment 2

these new groups also measure the sound speed as shown. It would seem that the statistics improve because 4 experiments are now done?! In fact one can continue breaking up the system like this, as shown in Fig. 3.3.

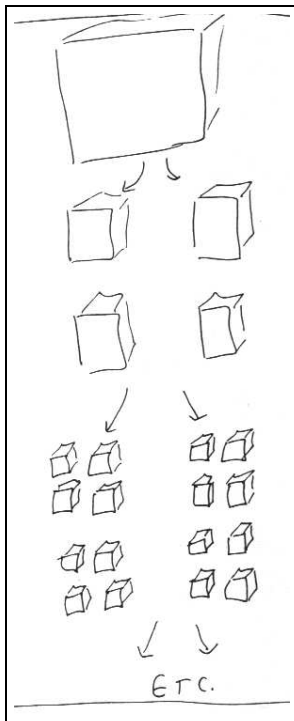


Figure 3.3: Experiment 3

Of course, 1 experiment on the big system must be equivalent to many experiments on individual parts of a big system, as long as those individual parts are independent. By this it is meant that a system “self-averages” all its independent parts.

The number of independent parts is (to use the same notation as the number of particles)

$$N = \mathcal{O}\left(\frac{L^3}{\xi^3}\right) \quad (3.1)$$

for a system of edge length L . The new quantity introduced here is the correlation length ξ , the scale over which things are correlated, so they are not independent. The correlation length must be at least as large as an atom so

$$\xi \geq \mathcal{O}(\text{\AA}) \quad (3.2)$$

and indeed usually $\xi \approx \mathcal{O}(\text{\AA})$. Hence for a macroscopic length, the number of independent points is

$$N \sim \mathcal{O}(10^{23}) \quad (3.3)$$

of course. Now, each independent part has a few degrees of freedom, such as quantum numbers. Hence, the total number of states in a typical system are

$$\begin{array}{ccc} \int dq_1 & \int dq_2 & \dots & \int dq_N \\ \text{quantum} & \text{quantum} & & \text{quantum} \\ \text{states} & \text{states} & & \text{states} \\ \text{of part 1} & \text{of part 2} & & \text{of part } N \end{array}$$

Say $\int dq_i = 4$ for an example, so that the total number of states is

$$e^{\mathcal{O}(N)} \quad (3.4)$$

as it should. Now, however, we have a feeling for the origin of some of the assumptions entering statistical physics.

We can go a little farther with this, and prove the central-limit theorem of probability theory.

3.1 Self Averaging

Consider an extensive quantity X . If a system consists of N independent parts as shown, then

i = 1	i = 2	i = 3
i = 4	i = 5	i = 6
i = 7	i = 8	i = N

Figure 3.4: N independent parts

$$X = \sum_{i=1}^N X_i \quad (3.5)$$

where X_i is the value of X in independent part i . The average value of X is

$$\langle X \rangle = \sum_{i=1}^N \langle X_i \rangle \quad (3.6)$$

But since $\langle X_i \rangle \neq \langle X_i \rangle(N)$, we have

$$\langle X \rangle = \mathcal{O}(N) \quad (3.7)$$

It is handy to introduce the intensive quantity

$$x = \frac{X}{N} \quad (3.8)$$

then

$$\langle x \rangle = \mathcal{O}(1)$$

This is not very amazing yet. Before continuing, it is worth giving an example of what we mean by the average. For example, it might be a microscopically long, but macroscopically short time average

$$\langle X \rangle = \frac{\int_{\tau} dt X(t)}{\tau}$$

as drawn. As also indicated, the quantity $\langle (X - \langle X \rangle)^2 \rangle^{1/2}$ is a natural one to

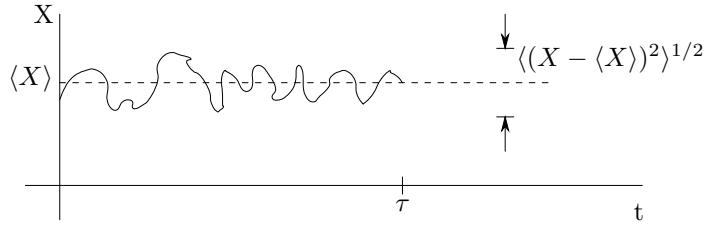


Figure 3.5: Time Average

consider. Let the deviation of X from its average value be

$$\Delta X = X - \langle X \rangle \quad (3.9)$$

Then clearly

$$\langle \Delta X \rangle = \langle X \rangle - \langle X \rangle = 0$$

as is clear from the cartoon. The handier quantity is

$$\begin{aligned} \langle (\Delta X)^2 \rangle &= \langle (X - \langle X \rangle)^2 \rangle \\ &= \langle X^2 - 2X\langle X \rangle + \langle X \rangle^2 \rangle \\ &= \langle X^2 \rangle - \langle X \rangle^2 \end{aligned} \quad (3.10)$$

Giving the condition, in passing,

$$\langle X^2 \rangle \geq \langle X \rangle^2 \quad (3.11)$$

since $\langle (\Delta X)^2 \rangle \geq 0$. Splitting the system into independent parts, again, gives

$$\langle (\Delta X)^2 \rangle = \left\langle \sum_{i=1}^N \Delta X_i \sum_{j=1}^N \Delta X_j \right\rangle$$

Now using that i may or may not equal j

$$\langle (\Delta X)^2 \rangle = \left\langle \sum_{\substack{i=1 \\ (i=j)}}^N (\Delta X_i)^2 \right\rangle + \left\langle \sum_{\substack{i=1 \\ j=1 \\ (i \neq j)}}^N (\Delta X_i)(\Delta X_j) \right\rangle \quad (3.12)$$

But box i is independent of box j , so,

$$\begin{aligned}\langle \Delta X_i \Delta X_j \rangle &= \langle \Delta X_i \rangle \langle \Delta X_j \rangle \\ &= 0 \times 0 \\ &= 0\end{aligned}\tag{3.13}$$

and,

$$\langle (\Delta X)^2 \rangle = \sum_{i=1}^N \langle (\Delta X_i)^2 \rangle$$

or

$$\langle (\Delta X)^2 \rangle = \mathcal{O}(N)\tag{3.14}$$

Also, the intensive $x = X/N$ satisfies

$$\langle (\Delta x)^2 \rangle = \mathcal{O}\left(\frac{1}{N}\right)\tag{3.15}$$

This implies that fluctuations in quantities are very small in systems of many ($N \rightarrow \infty$) independent parts. That is, because

$$\begin{aligned}X &\approx \langle X \rangle \left\{ 1 + \mathcal{O}\left(\frac{\langle (\Delta X)^2 \rangle^{1/2}}{\langle X \rangle}\right) \right\} \\ &\approx \langle X \rangle \left\{ 1 + \mathcal{O}\left(\frac{1}{N^{1/2}}\right) \right\}\end{aligned}\tag{3.16}$$

So, in essence, $X \approx \langle X \rangle$. The result

$$\frac{\langle (\Delta X)^2 \rangle}{\langle X \rangle^2} = \mathcal{O}\left(\frac{1}{N}\right)\tag{3.17}$$

that relative fluctuations are very small, if N is large, is called the self-averaging lemma. Note that, since $N = V/\xi^3$,

$$\frac{\langle (\Delta X)^2 \rangle}{\langle X \rangle^2} = \mathcal{O}\left(\frac{\xi^3}{V}\right)\tag{3.18}$$

so fluctuations measure *microscopic* correlations, and become appreciable for $\frac{\xi^3}{V} \sim \mathcal{O}(1)$. This occurs in quantum mesoscopic systems (where V is small), and phase transitions (where ξ is large), for example.

3.2 Central Limit Theorem

We will now prove (to a physicist's rigor) the central-limit theorem for systems if many independent parts. In the same manner as above we will calculate all the moments, such as the n^{th} moment,

$$\langle (\Delta X)^n \rangle\tag{3.19}$$

or at least their order with N .

For simplicity we will assume that all the odd moments vanish

$$\langle (\Delta X)^{2n+1} \rangle = 0$$

for integer n . This means the distribution ρ is symmetric about the mean $\langle X \rangle$ as shown. It turns out that one can show that the distribution function *is* symmetric

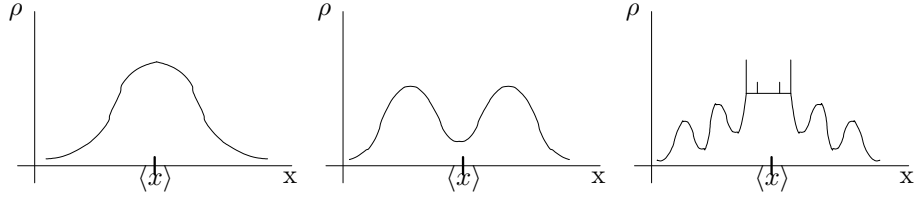


Figure 3.6: Symmetric distribution functions $\rho(x)$

as $(N \rightarrow \infty)$, but let's just assume it so the ideas in the algebra are clear. It can be shown (see below) that we can then reconstruct the distribution $\rho(X)$ if we know all the even moments

$$\langle (\Delta X)^{2n} \rangle$$

for positive integer n . We know $n = 1$, so let's do $n = 2$, as we did $\langle (\Delta X)^2 \rangle$ above

$$\begin{aligned} \langle (\Delta X)^4 \rangle &= \left\langle \sum_{i=1}^N \Delta X_i \sum_{j=1}^N \Delta X_j \sum_{k=1}^N \Delta X_k \sum_{l=1}^N \Delta X_l \right\rangle \\ &= \left\langle \sum_{\substack{i=1 \\ (i=j=k=l)}}^N (\Delta X_i)^4 \right\rangle + 3 \left\langle \sum_{\substack{i=1 \\ j=1 \\ (i \neq j)}}^N (\Delta X_i)^2 (\Delta X_j)^2 \right\rangle + \mathcal{O}(\langle \Delta X_i \rangle^4) \quad (3.20) \\ &= \mathcal{O}(N) + \mathcal{O}(N^2) \end{aligned}$$

So as $(N \rightarrow \infty)$

$$\langle (\Delta X)^4 \rangle = 3 \left\langle \sum_{i=1}^N (\Delta X_i)^2 \right\rangle \left\langle \sum_{j=1}^N (\Delta X_j)^2 \right\rangle$$

or

$$\langle (\Delta X)^4 \rangle = 3 \langle (\Delta X)^2 \rangle^2 = \mathcal{O}(N^2) \quad (3.21)$$

Note that the largest term (as $(N \rightarrow \infty)$) came from pairing up all the ΔX_i 's appropriately. A little thought makes it clear that the largest term in $\langle (\Delta X)^{2n} \rangle$ will result from this sort of pairing giving,

$$\langle (\Delta X)^{2n} \rangle = \text{const.} \cdot \langle (\Delta X)^2 \rangle^n = \mathcal{O}(N^n) \quad (3.22)$$

We will now determine the *const.* in Eq. 3.22. Note that

$$\langle (\Delta X)^{2n} \rangle = \langle \Delta X \Delta X \Delta X \dots \Delta X \Delta X \rangle$$

There are $(2n - 1)$ ways to pair off the first ΔX . After those two ΔX 's are gone, there are then $(2n - 3)$ ways to pair off the next free ΔX . And so on, until finally, we get to the last two ΔX 's which have 1 way to pair off. Hence

$$\langle (\Delta X)^{2n} \rangle = (2n - 1)(2n - 3)(2n - 5) \dots (1) \langle (\Delta X)^2 \rangle^n \equiv (2n - 1)!! \langle (\Delta X)^2 \rangle^n$$

or more conventionally,

$$\langle (\Delta X)^{2n} \rangle = \frac{2n!}{2^n n!} \langle (\Delta X)^2 \rangle^n \quad (3.23)$$

Now we know all the even moments in terms of the first two moments $\langle X \rangle$ and $\langle (\Delta X)^2 \rangle$. We can reconstruct the distribution with a trick. First, let us be explicit about ρ . By an average we shall mean,

$$\langle Q(X) \rangle_X = \int_{-\infty}^{\infty} \rho(X) Q(X) \quad (3.24)$$

where $Q(X)$ is some quantity. Note that

$$\langle \delta(X - X') \rangle_{X'} = \rho(X) \quad (3.25)$$

where $\delta(X)$ is the Dirac delta function! Recall the Fourier representation

$$\delta(X - X') = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{-ik(X-X')} \quad (3.26)$$

Averaging over X' gives

$$\rho(X) = \langle \delta(X - X') \rangle_{X'} = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{-ikX} \langle e^{ikX'} \rangle_{X'} \quad (3.27)$$

In statistics, $\langle e^{ikX} \rangle$ is called the characteristic function. But we can now expand the exponential (for convenience, let $\langle X \rangle = 0$ for now, one only has to add and subtract it in $e^{-ik(X-X')} = e^{-ik(\Delta X - \Delta X')}$ but it makes the algebra messier) as

$$e^{-ikX'} = \sum_{n=0}^{\infty} \frac{(ik)^n}{n!} (X')^n$$

So that we have,

$$\rho(X) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{ikX} \sum_{n=0}^{\infty} \frac{(ik)^n}{n!} \langle X^n \rangle \quad (3.28)$$

where the prime superscript has been dropped. Note that Eq. 3.28 means that $\rho(X)$ can be reconstructed from all the $\langle X^n \rangle$, in general. For our case, all the odd moments vanish, so

$$\rho(X) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{ikX} \sum_{n=0}^{\infty} \frac{(-k^2)^n}{(2n)!} \langle X^{2n} \rangle$$

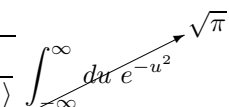
But $\langle X^{2n} \rangle = \frac{(2n)!}{2^n n!} \langle X^2 \rangle^n$ from above, so,

$$\rho(X) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{ikX} \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{-k^2}{2} \langle X^2 \rangle \right)^n$$

which is the exponential again, so,

$$\rho(X) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{ikX} e^{\frac{-k^2}{2} \langle X^2 \rangle} \quad (3.29)$$

(In passing, it is worth remembering that $\langle e^{ikX} \rangle = e^{\frac{-k^2}{2} \langle X^2 \rangle}$ for a Gaussian distribution.) The integral can be done by completing the square, adding and subtracting $\frac{-X^2}{2\langle X^2 \rangle}$, so that

$$\begin{aligned} \rho(X) &= \frac{1}{2\pi} e^{\frac{-X^2}{2\langle X^2 \rangle}} \int_{-\infty}^{\infty} dk e^{-\left(\frac{k\langle X^2 \rangle^{1/2}}{2^{1/2}} + \frac{iX}{2\langle X^2 \rangle^{1/2}}\right)^2} \\ &= \frac{1}{2\pi} e^{\frac{-X^2}{2\langle X^2 \rangle}} \sqrt{\frac{2}{\langle X^2 \rangle}} \int_{-\infty}^{\infty} du e^{-u^2} \end{aligned}$$


using an obvious substitution. Hence, on putting $\langle X \rangle$ as non-zero again,

$$\rho(X) = \frac{1}{\sqrt{2\pi\langle(\Delta X)^2\rangle}} e^{\frac{-(X-\langle X \rangle)^2}{2\langle(\Delta X)^2\rangle}} \quad (3.30)$$

which is the Gaussian distribution. This is the central-limit theorem for a system of many independent parts. Its consequences are most transparent if one deals with the intensive quantity $x = \frac{X}{N}$, then recall

$$\begin{aligned} \langle x \rangle &= \mathcal{O}(1) \\ \langle(\Delta x)^2\rangle &= \mathcal{O}(1/N) \end{aligned} \quad (3.31)$$

Letting

$$\langle(\Delta x)^2\rangle \equiv \frac{a^2}{N} \quad (3.32)$$

where $a^2 = \mathcal{O}(1)$ we have, up to normalization,

$$\rho(x) \propto e^{-N \frac{(x-\langle x \rangle)^2}{2a^2}} \quad (3.33)$$

so the width of the Gaussian is exceedingly small and it most represents a spike.

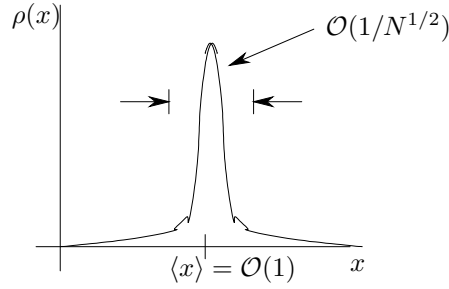


Figure 3.7: Gaussian (Width Exaggerated)

Which is to say again that fluctuations are small, and average values are well

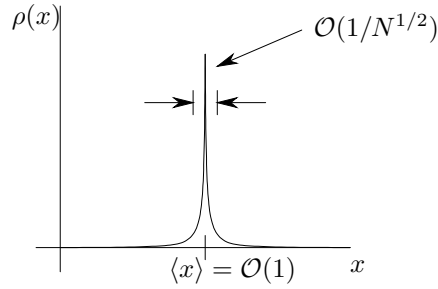


Figure 3.8: Gaussian (Width (sort of) to scale)

defined for a system of many ($N \rightarrow \infty$) independent parts.

3.3 Correlations in Space

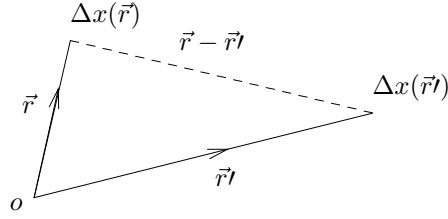
The results so far apply to a large system, and perhaps it is not surprising that the global fluctuations in a large system are small. These results for the global fluctuations have implications for the correlations of quantities, though. Say the intensive variable x has different values at points \vec{r} in space. (Precisely, let $d\vec{r}$, the volume element, be microscopically large, but macroscopically small). Consider the correlation function

$$\langle \Delta x(\vec{r}) \Delta x(\vec{r}') \rangle \equiv C(\vec{r}, \vec{r}') \quad (3.34)$$

We can say a few things about $C(\vec{r}, \vec{r}')$ from general considerations.

1. If space is translationally invariant, then the origin of the coordinate system for \vec{r} and \vec{r}' is arbitrary. It can be placed at point \vec{r} or \vec{r}' . The physical length of interest is $(\vec{r} - \vec{r}')$ or $(\vec{r}' - \vec{r})$. Hence

$$C(\vec{r}, \vec{r}') = C(\vec{r} - \vec{r}') = C(\vec{r}' - \vec{r}) \quad (3.35)$$

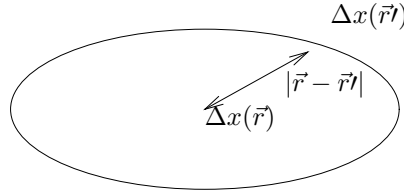
Figure 3.9: Translational invariance (only $\vec{r} - \vec{r}'$ is important)

and, of course,

$$\langle \Delta x(\vec{r}) \Delta x(\vec{r}') \rangle = \langle \Delta x(\vec{r} - \vec{r}') \Delta x(0) \rangle \quad (3.36)$$

etc.

2. If space is isotropic, then the direction of $\Delta x(\vec{r})$ from $\Delta x(\vec{r}')$ is unimportant, and only the magnitude of the distance $|\vec{r} - \vec{r}'|$ is physically

Figure 3.10: Isotropy (only $|\vec{r} - \vec{r}'|$ is important)

relevant. Hence

$$C(\vec{r}, \vec{r}') = C(|\vec{r} - \vec{r}'|) \quad (3.37)$$

So what was a function of 6 numbers (\vec{r}, \vec{r}') is now a function of only 1, $|\vec{r} - \vec{r}'|$.

3. Usually correlations fall off if points are arbitrarily far apart. That is,

$$\begin{aligned} \langle \Delta x(r \rightarrow \infty) \Delta x(0) \rangle &= \langle \Delta x(r \rightarrow \infty) \rangle \langle \Delta x(0) \rangle \\ &= 0 \end{aligned}$$

or

$$C(r \rightarrow \infty) = 0 \quad (3.38)$$

The scale of the correlation is set by the correlation length.

Now we know

$$C(|\vec{r} - \vec{r}'|) = \langle \Delta x(\vec{r}) \Delta x(\vec{r}') \rangle \quad (3.39)$$

Subject to these conditions. A constraint on $C(|\vec{r} - \vec{r}'|)$ is provided by our previous result

$$\langle (\Delta x)^2 \rangle = \mathcal{O}\left(\frac{1}{N}\right)$$

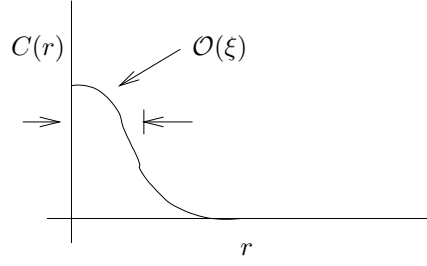


Figure 3.11: Typical correlation function

for a global fluctuation. Clearly,

$$\begin{aligned}\Delta x &= \frac{\int d\vec{r} \Delta x(\vec{r})}{\int d\vec{r}} \\ &= \frac{1}{V} \int d\vec{r} \Delta x(\vec{r})\end{aligned}\tag{3.40}$$

So we have

$$\begin{aligned}\frac{1}{V} \int d\vec{r} \int d\vec{r}' \langle \Delta x(\vec{r}) \Delta x(\vec{r}') \rangle &= \langle (\Delta x)^2 \rangle \\ &= \mathcal{O}\left(\frac{1}{N}\right)\end{aligned}\tag{3.41}$$

or,

$$\int d\vec{r} \int d\vec{r}' C(\vec{r} - \vec{r}') = \mathcal{O}\left(\frac{V^2}{N}\right)$$

Let $\vec{R} = \vec{r} - \vec{r}'$, $\vec{R}' = \frac{1}{2}(\vec{r} + \vec{r}')$. The Jacobian of the transformation is unity, and we obtain

$$\begin{aligned}\int d\vec{R}' \int d\vec{R} C(\vec{R}) &= \mathcal{O}\left(\frac{V^2}{N}\right) \\ &= V \int d\vec{R} C(\vec{R})\end{aligned}\tag{3.42}$$

So letting $\vec{R} \rightarrow \vec{r}$,

$$\int d\vec{r} C(\vec{r}) = \mathcal{O}\left(\frac{V}{N}\right)$$

But $N = \frac{V}{\xi^3}$, or in d dimensions, $N = \frac{L^d}{\xi^d}$, so

$$\int d\vec{r} C(\vec{r}) = \mathcal{O}(\xi^3)$$

or

$$\int d\vec{r} \langle \Delta x(\vec{r}) \Delta x(0) \rangle = \mathcal{O}(\xi^3)\tag{3.43}$$

for a system of $N \rightarrow \infty$ independant parts. This provides a constraint on correlations. It is a neat result in Fourier space. Let

$$\hat{C}(\vec{k}) \equiv \int d\vec{r} e^{i\vec{k} \cdot \vec{r}} C(\vec{r})$$

and finally

$$C(\vec{r}) = \int \frac{d\vec{k}}{(2\pi)^d} e^{-i\vec{k} \cdot \vec{r}} \hat{C}(\vec{k}) \quad (3.44)$$

and,

$$\langle |\Delta \hat{x}(k)|^2 \rangle \equiv \int d\vec{r} e^{i\vec{k} \cdot \vec{r}} \langle \Delta x(\vec{r}) \Delta x(0) \rangle$$

Hence the constraint of Eq. 3.43 is

$$\lim_{k \rightarrow 0} \hat{C}(k) = \mathcal{O}(\xi^3)$$

or

$$\lim_{k \rightarrow 0} \langle |\Delta \hat{x}(k)|^2 \rangle = \mathcal{O}(\xi^3) \quad (3.45)$$

Many functions satisfy the constraint of Eq. 3.43 (see Fig. 3.11). It turns out that usually,

$$\langle \Delta x(\vec{r}) \Delta x(0) \rangle \sim e^{-r/\xi} \quad (3.46)$$

for large r , which satisfies Eq. 3.43. This is obtained from analyticity of $(\hat{C}(k))^{-1}$ near $k = 0$. It can be argued,

$$\frac{1}{\hat{C}(k)} \approx (const.) + (const.)k^2 + \dots$$

so that $\langle |\Delta \hat{x}(k)|^2 \rangle = \frac{const.}{const. + k^2}$ near $k = 0$. Fourier transformation gives the result of Eq. 3.46 above.

(Note: In the previous notation with boxes, as in Fig. 3.4, we had

$$\Delta X = \sum_i \Delta X_i$$

or

$$\Delta x = \frac{1}{N} \sum_i \Delta X_i$$

This means

$$\frac{1}{V} \int d\vec{r} \Delta x(\vec{r}) = \frac{1}{N} \sum_i \Delta X_i \quad (3.47)$$

This is satisfied by

$$\Delta x(r) = \frac{V}{N} \sum_{i=1}^N \delta(\vec{r} - \vec{r}_i) \Delta X_i$$

where \vec{r}_i is a vector pointing to box i , and $\Delta X_i = \Delta X(\vec{r}_i)$.

Let me emphasize again that all these results follow from a system having $N \rightarrow \infty$ independent parts. To recapitulate, for $x = X/N$, where X is extensive, $N = V/\xi^3$ is the number of independent parts. V is the volume, and ξ is the correlation length:

$$\langle x \rangle = \mathcal{O}(1) \tag{3.7}$$

$$\langle (\Delta x)^2 \rangle = \mathcal{O}\left(\frac{1}{N}\right) \tag{3.15}$$

$$\rho(x) \propto e^{\frac{(x - \langle x \rangle)^2}{2\langle (\Delta x)^2 \rangle}} \tag{3.30}$$

$$\int d\vec{r} \langle \Delta x(r) \Delta x(0) \rangle = \mathcal{O}(\xi^3) \tag{3.43}$$

The purpose of this course is to examine the cases where $V/\xi^3 = \mathcal{O}(1)$ and these results break down! It is worthwhile to give some toy examples now where, in particular, Eq. 3.43 breaks down.

Chapter 4

Toy Fractals

Toy Fractals are mathematical constructs that seem to exist in fractional dimensions between 1 and 2, or, 2 and 3, or what have you. They can provide a useful starting point for thinking about real systems where $V/\xi^3 = \mathcal{O}(1)$, or if $V = L^3$, $\xi/L = \mathcal{O}(1)$.

We'll briefly discuss self-affine, or surface fractals, and volume fractals. This

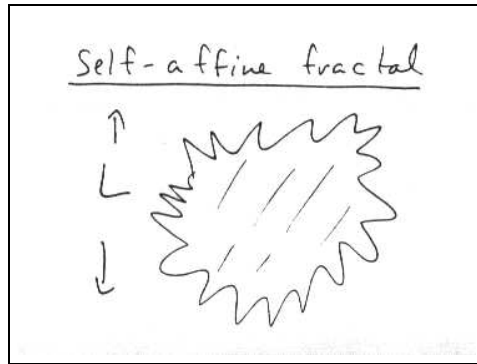


Figure 4.1: Self-Affine Fractal

is a self-affine fractal where the surface is very jagged, or as is said in physics, *rough*. The self-affine fractal has the property that its perimeter length P (or surface area in three dimensions) satisfies,

$$P \sim L^{d_s} \quad , \text{ as } L \rightarrow \infty \quad (4.1)$$

where d_s is the surface self-affine fractal exponent. For a circle of diameter L

$$P_{\text{circle}} = \pi L$$

while for a square of width L

$$P_{\text{square}} = 4L$$

In fact, for run-of-the-mill surfaces in d dimensions

$$P_{nonfractal} \sim L^{d-1} \quad (4.2)$$

Usually

$$d_s > d - 1 \quad (4.3)$$

I'll give an example below. This is a volume fractal, note all the holes with

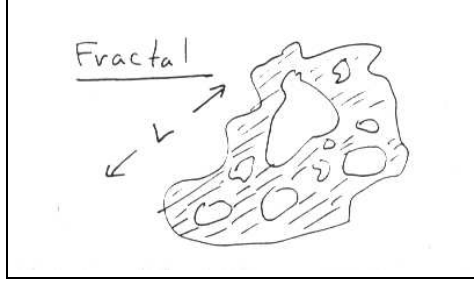


Figure 4.2: Volume Fractal

missing stuff. The fractal has the property that its mass M satisfies

$$M \sim L^{d_f} \text{ as } L \rightarrow \infty \quad (4.4)$$

where d_f is the fractal dimension. For a circle $M = \pi/4 L^2$, in units of the density for a diameter L , while for a square $M = L^3$. For run-of-the-mill objects

$$M_{nonfractal} \sim L^d \quad (4.5)$$

in d dimensions. Usually

$$d_f < d \quad (4.6)$$

Note that since the mass density is M/L^d , this implies that

$$\text{DENSITY OF A FRACTAL} \sim \frac{1}{L^{d-d_f}} \rightarrow 0 \quad (4.7)$$

as $L \rightarrow \infty$, implying fractals float (!?). Think about this.

Two toy examples show that $d_s > d - 1$ and $d_f < d$ are possible.

4.1 Von Koch Snowflake

To draw the von Koch snowflake, follows the steps shown. To find the self-affine

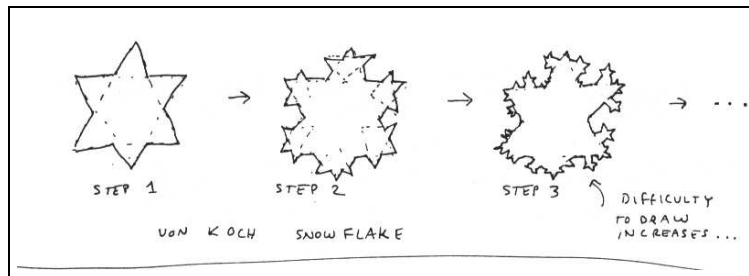


Figure 4.3: Von Koch Snowflake

fractal exponent for the perimeter, imagine I had (somehow) drawn one of these fractals out a very very large number of steps. Say the smallest straight link is length l . We'll calculate the perimeter as we go through increasing numbers of links in a large complex fractal. Hence

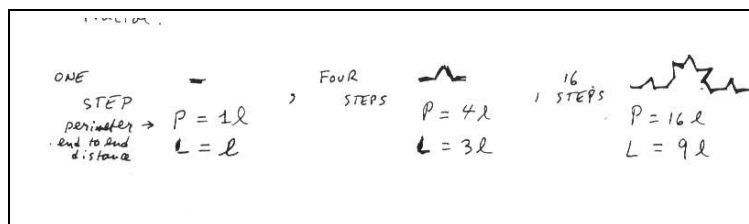


Figure 4.4:

$$P_n = 4^n l$$

when

$$L_n = 3^n l \quad (4.8)$$

Taking the logarithm

$$\ln P/l = n \ln 4$$

and

$$\ln L/l = n \ln 3$$

Dividing these into each other

$$\frac{\ln P/l}{\ln L/l} = \frac{\ln 4}{\ln 3}$$

or

$$P/l = (L/l)^{d_s} \quad (4.9)$$

where

$$\begin{aligned} d_s &= \frac{\ln 4}{\ln 3} \approx 1.26 \\ &> d - l = 1 \end{aligned}$$

It is common in the physics literature to write

$$P(L) = w(L)L^{d-1} \quad (4.10)$$

where $w(L)$ is the width or thickness of a rough interface, and

$$w(L) \sim L^x, \text{ as } L \rightarrow \infty \quad (4.11)$$

where x is the roughening exponent, and clearly, ($x \approx 0.26$ for the snowflake),

$$d_s = x + d - 1 \quad (4.12)$$

4.2 Sierpinski Triangle

To draw the Sierpinski triangle, follow the steps shown. To find the fractal

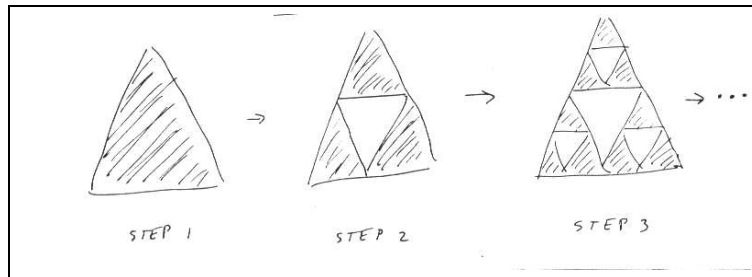


Figure 4.5: Sierpinski Triangle

exponent for the mass, imagine I had (somehow) drawn one of these fractals out to a very large number of steps. Say the smallest filled triangle had mass m and volume equal to v . Then we'll calculate the mass as we go through increasing the numbers of triangles. Hence

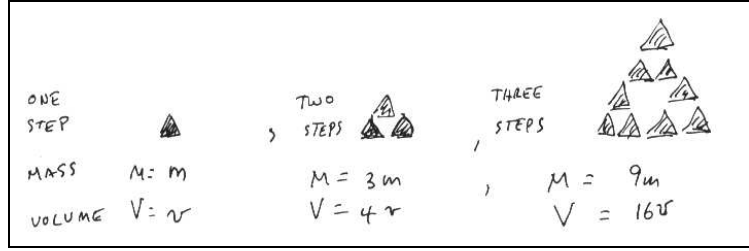


Figure 4.6:

$$M_n = 3^n m$$

when

$$V_n = 4^n v$$

(4.13)

or since $V_n = \frac{1}{2} L_n^2$ and $v = \frac{1}{2} l^2$ we have,

$$L_n^2 = 4^n l^2 \quad \text{or} \quad L_n = 2^n l$$

Taking logarithms,

$$\ln \frac{M}{m} = n \ln 3$$

and

$$\ln \frac{L}{l} = n \ln 2$$

or, taking the ratio

$$\frac{\ln M/m}{\ln L/l} = \frac{\ln 3}{\ln 2}$$

and

$$\frac{M}{m} = \left(\frac{L}{l} \right)^{d_f} \quad (4.14)$$

where

$$d_f = \frac{\ln 3}{\ln 2} \approx 1.585 \quad (4.15)$$

So $d_f < d$ where $d = 2$ here.

This is not the common notation in the physics literature. Below, but near a second-order phase transition, the excess density satisfies

$$n \propto |T - T_c|^\beta$$

where T_c is the critical temperature. However the correlation length satisfies

$$\xi \propto |T - T_c|^{-\nu}$$

Or, since ξ 's largest value is limited by the edge length L of the entire system

$$\xi \sim L \sim |T - T_c|^{-\nu}$$

Using this above gives

$$n \propto [L^{-1/\nu}]^\beta$$

or, Excess density near T_c

$$n \propto \frac{1}{L^{\beta/\nu}}$$

where β and ν are critical exponents. But, for a fractal

$$n \propto \frac{1}{L} d - d_f$$

so

$$d_f = d - \beta/\nu \quad (4.16)$$

where $\beta/\nu \approx 0.415$ for the Sierpinski snowflake.

For a real system (in the Ising universality class in two dimensions),

$$x = 1/2, \quad \text{so} \quad d_s = 1 \frac{1}{2} \quad \text{and} \quad \beta/\nu = 1/8, \quad \text{so} \quad d_f = 1 \frac{7}{8}$$

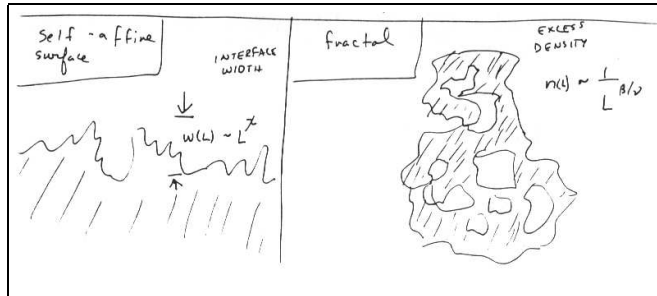


Figure 4.7:

4.3 Correlations in Self-Similar Objects

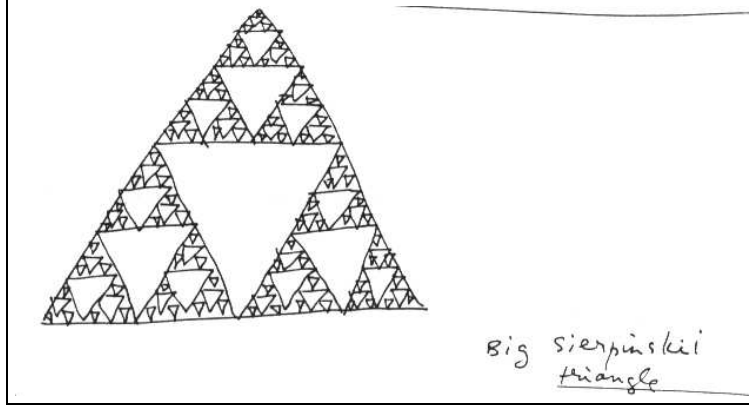


Figure 4.8:

A striking thing about a fractal is that it “looks the same” on different length scales. This is to say that your eye sees similar (or, for these toy fractals, identical features) correlations if you look on length scales r or r/b where $b > 1$. So far as density correlations go, let

$$C(r) = \langle \Delta n(r) \Delta n(0) \rangle$$

(where the average is, say, over all orientations in space for convenience). Then the evident *self-similarity* of fractals implies

$$C(r) \propto C(r/b)$$

or

$$C(r) = f(b) C(r/b) \tag{4.17}$$

It is easy to show that

$$f(b) = b^{-p} \tag{4.18}$$

where p is a constant. Take the derivative with respect to b of Eq. 4.17, letting

$$r^k \equiv r/b \tag{4.19}$$

Then

$$\frac{dC(r)}{db} = 0 = \frac{df}{db} C(r^*) + f \frac{dC}{dr^*} \frac{dr^*}{db}$$

or,

$$\frac{df}{db}C(r^*) = +f \frac{r^*}{b} \frac{dC}{dr^*}$$

and

$$\frac{d \ln f(b)}{d \ln b} = \frac{d \ln C(r^*)}{d \ln r^*} \quad (4.20)$$

Because one side only depends on b , and the other only on r^* . Hence

$$\frac{d \ln f(b)}{d \ln b} = \text{const.} \equiv -p \quad (4.21)$$

So

$$f(b) \propto b^{-p} \quad (4.22)$$

and the invariance relation is

$$C(r) = b^{-p} C(r/b) \quad (4.23)$$

(up to a multiplicative constant).

If one lets

$$b = 1 + \epsilon, \quad \epsilon \ll 1 \quad (4.24)$$

$$\begin{aligned} C(r) &= (1 + \epsilon)^{-p} C\left(\frac{r}{1 + \epsilon}\right) \\ &\approx (1 - p\epsilon) \left\{ C(r) - \epsilon r \frac{\partial C}{\partial r} \right\} + \dots \end{aligned}$$

Terms of order ϵ give

$$-p C(r) = r \frac{\partial C}{\partial r}$$

The solution of which is

$$C \propto \frac{1}{r^p} \quad (4.25)$$

Hence, self similarity implies homogeneous correlation functions

$$C(r) = b^{-p} C(r/b)$$

and power-law correlation

$$C(r) \propto \frac{1}{r^p} \quad (4.26)$$

Usually, these apply for large r . The sign of p chosen to make $C(r \rightarrow \infty) \rightarrow 0$. Note that this is quite different from what occurs for a system of many

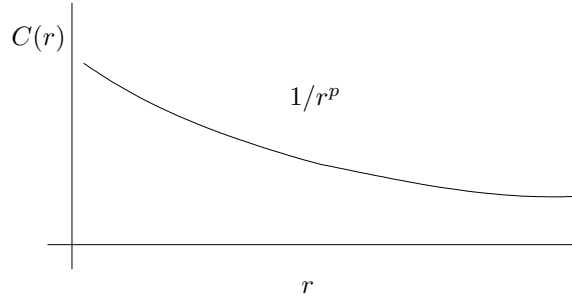


Figure 4.9: Decay of Correlations in a Self-Similar Object

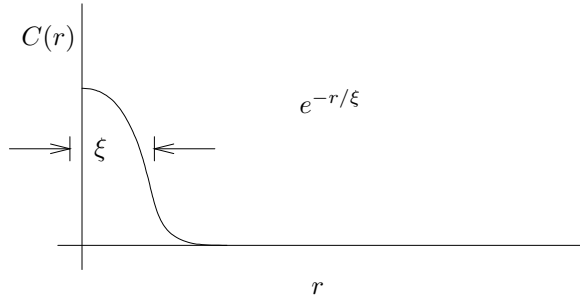


Figure 4.10: Decay of Correlations in a System of Many Independent Parts

independent parts where (we said)

$$C(r) \sim e^{-r/\xi}$$

The restriction

$$\int d\vec{r} C(r) = \mathcal{O}(\xi^d)$$

in d -dimensions is evidently *unsatisfied* for

$$C(r) \sim \frac{1}{r^p}$$

if $d > p$, for then the integral diverges!

Usually,

$$p = 2 - \eta \tag{4.27}$$

and $\eta \geq 0$, so $d > p$ in $d = 2$ and 3 .

So, fractals, self-affine (rough) surfaces, and second-order phase transition fluctuations *do not* satisfy the central limit theorem.

In fact, rough surfaces and phase transitions will be the focus for the rest of the course.

Chapter 5

Molecular Dynamics

At this point, one might say, computers are so big and fierce, why not simply solve the microscopic problem completely! This is the approach of molecular dynamics. It is a very important numerical technique, which is very very conservative. It is used by physicists, chemists, and biologists.

Imagine a classical solid, liquid or gas composed of many atoms or molecules. Each atom has a position \vec{r}_i and a velocity \vec{v}_i (or momentum $m\vec{v}_i$). The Hamil-

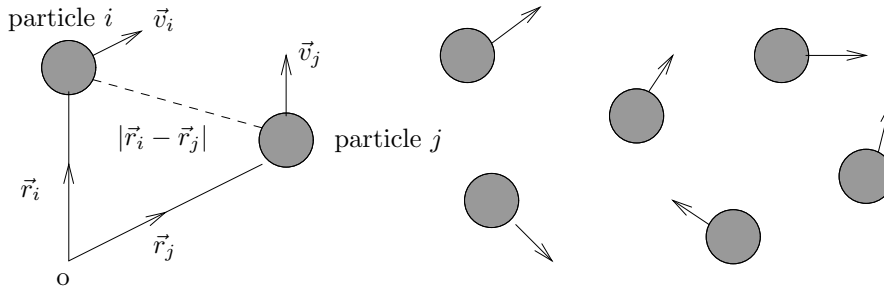


Figure 5.1:

tonian is (where m is the atom's mass)

$$H = \sum_{i=1}^N \frac{m\vec{v}_i^2}{2} + V(\vec{r}_1, \dots, \vec{r}_N)$$

Usually, the total potential can be well approximated as the sum of pairwise forces, i.e.

$$V(r_1, \dots, r_N) = \sum_{\substack{i=1, j=1 \\ (i>j)}}^N V(|\vec{r}_i - \vec{r}_j|)$$

which only depends on the distance between atoms. This potential usually has

a form like that drawn. Where the depth of the well must be

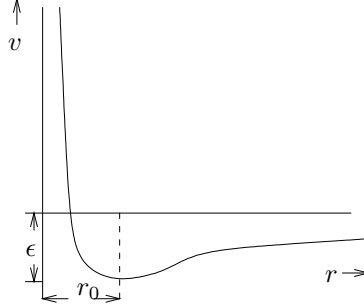


Figure 5.2:

$$\epsilon \sim (100' s K) k_B$$

and

$$r_o \sim (a \text{ few } \text{\AA})$$

This is because, in a vapor phase, the density is small and the well is unimportant; when the well becomes important, the atoms condense into a liquid phase. Hence $\epsilon \sim$ boiling point of a liquid, and r_o is the separation of atoms in a liquid. Of course, say the mass of an atom is

$$m \sim 10 \times 10^{-27} \text{ kg}$$

$$r_o \sim 2 \text{\AA}$$

then the liquid density is

$$\begin{aligned} \rho &\sim \frac{m}{r_o^3} = \frac{10 \times 10^{-27} \text{ kg}}{10 \text{\AA}^3} \\ &= \frac{10^{-27} \times (10^3 \text{ gm})}{(10^{-8} \text{ cm})^3} \\ \rho &\sim 1 \text{ gm/cm}^3 \end{aligned}$$

of course. To convince oneself that quantum effects are unimportant, simply estimate the de Broglie wavelength

$$\begin{aligned} \lambda &\sim \frac{\hbar}{\sqrt{mk_B T}} \\ &\sim 1 \text{\AA} / \sqrt{T} \end{aligned}$$

where T is measured in Kelvin. So this effect is unimportant for most liquids. The counter example is Helium. In any case, quantum effects are usually

marginal, and lead to no quantitative differences of importance. This is simply because the potentials are usually parameterized fits such as the Lennard - Jones potential

$$V(r) = 4\epsilon[(\frac{\sigma}{r})^{12} - (\frac{\sigma}{r})^6]$$

where

$$r_o = 2^{1/6}\sigma \approx 1.1\sigma$$

and

$$V(r_o) = -\epsilon$$

by construction. For Argon, a good “fit” is

$$\begin{aligned} r_o &= 3.4\text{\AA} \\ \epsilon &= 120K \end{aligned}$$

Numerically it is convenient that $V(2r_o) \ll \epsilon$, and is practically zero. Hence one usually considers

$$V(r) = \begin{cases} \text{Lennard - Jones} & , r < 2r_o \\ 0 & , r > 2r_o \end{cases}$$

where $r_o = 2^{1/6}\sigma$. Lennard-Jones has a run-of-the-mill phase diagram which is representative of Argon or Neon, as well as most simple run-of-the-mill pure substances. Simpler potentials like the square well also have the same properties.

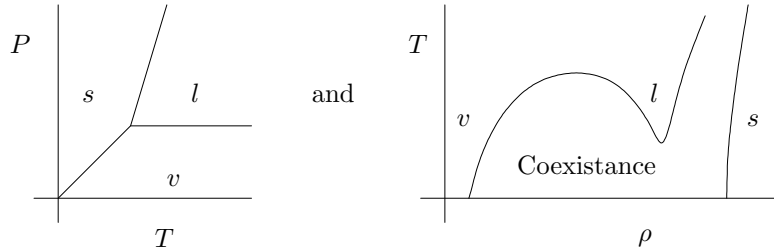


Figure 5.3: Lennard Jousium (The solid phase is fcc as I recall)

But the hard sphere potential has a more simple-minded phase diagram. This

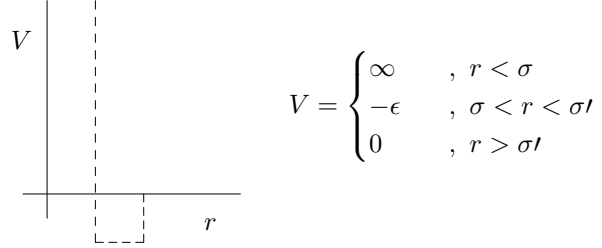
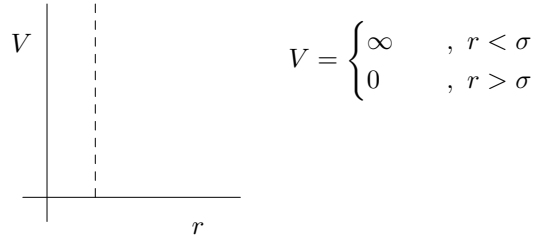


Figure 5.4:



where,

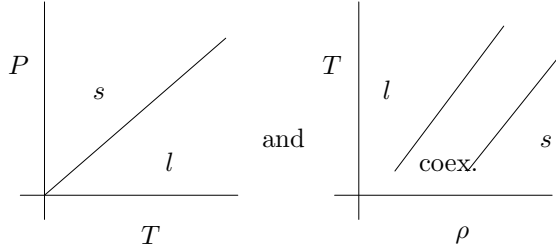


Figure 5.5:

is because of the lack of an attractive interaction.

Finally, let's get to the molecular dynamics method. Newton's equations (or Hamilton's) are,

$$\frac{d\vec{r}_i}{dt} = \vec{v}_i$$

$$\text{and } \vec{a}_i = \frac{\vec{F}_i}{m}$$

$$\frac{d\vec{v}_i}{dt} = \frac{1}{m} \sum_{\substack{j=1 \\ (i \neq j)}}^N \frac{\partial}{\partial \vec{r}_i} V(|\vec{r}_i - \vec{r}_j|)$$

Clearly, we can simply solve these numerically given all the positions and velocities initially, i.e. $\{\vec{r}_i(t=0), \vec{v}_i(t=0)\}$. A very simple (in fact too simple) way

to do this is by

$$\frac{d\vec{r}_i}{dt} = \frac{\vec{r}_i(t + \Delta t) - \vec{r}_i}{\Delta t}$$

for small Δt . This gives

$$\vec{r}_i(t + \Delta t) = \vec{r}_i(t) + \Delta t \vec{v}_i(t)$$

and

$$\vec{r}_i(t + \Delta t) = \vec{v}_i(t) + \Delta t \frac{1}{m} \vec{F}_i$$

where $\vec{F}_i(t)$ is determined by $\{\vec{r}_i(t)\}$ only. A better (but still simple) method is due to Verlet. It should be remembered that the discretization method has a little bit of black magic in it.

Valet notes that,

$$\vec{r}_i(t \pm \Delta t) = \vec{r}_i(t) \pm \Delta t \vec{v}_i + \frac{(\Delta t)^2}{2} \vec{a}_i(t) + \dots$$

where $\vec{a}_i = \vec{F}_i/m = \frac{d^2 \vec{r}_i}{dt^2}$. Add this to itself

$$\vec{r}_i(t + \Delta t) = 2\vec{r}_i(t) - \vec{r}_i(t - \Delta t) + (\Delta t)^2 \frac{1}{m} \vec{F}_i + \mathcal{O}(\Delta t)$$

And us the simple rule

$$\vec{v}_i(t + \Delta t) = \frac{\vec{r}_i(t + \Delta t) - \vec{r}_i(t - \Delta t)}{2\Delta t}$$

Thats all. You have to keep one extra set of \vec{r}_i 's, but it is the common method, and works well. To do statistical mechanics, you do time averages, after the system equilibrates. Or ensemble averages over different initial conditions. The ensemble is microcanonical since energy is fixed. Usually one solves it for a fixed volume, using periodic boundary conditions. A common practice is to do a

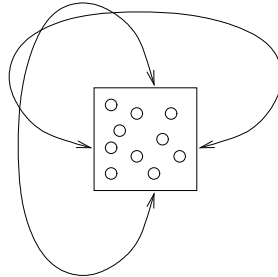


Figure 5.6: Periodic Boundary Conditions

“canonical” ensemble by fixing temperature through the sum rule

$$\frac{1}{2}m\langle v^2 \rangle = \frac{1}{2}k_B T$$

or,

$$T = \frac{\langle v^2 \rangle}{mk_B}$$

This is done by constantly rescaling the magnitude of the velocities to be

$$|v| \equiv \langle v^2 \rangle^{1/2} \equiv \sqrt{mk_B T}$$

What about microscopic reversibility and the arrow of time?

Recall the original equations

$$\frac{\partial \vec{r}_i}{\partial t} = \vec{v}_i$$

$$\frac{\partial \vec{v}_i}{\partial t} = \frac{1}{m} \sum_{\substack{j \\ (j \neq i)}} \frac{\partial}{\partial \vec{r}_i} V(|r_i - r_j|)$$

The equations are invariant under time reversal $t \rightarrow -t$ if

$$\begin{aligned} \vec{r} &\rightarrow \vec{r} \\ \vec{v} &\rightarrow -\vec{v} \end{aligned}$$

How then can entropy increase? This is a deep and uninteresting question. Entropy does increase; it's a law of physics. The real question is how to calculate how *how fast* entropy increases; e.g. calculate a viscosity or a diffusion constant, which set the time scale over which entropy increases.

Nevertheless, here's a neat way to fool around with this a bit. Take a gas of N atoms in a box of size L^d where energy is fixed such that the average equilibrium separation between atoms

$$s_{eq} = \mathcal{O}(3 \text{ or } 4)r_o$$

where r_o is a parameter giving the rough size of the atoms. Now initialize the atoms in a corner of the box with

$$s(t=0) = \mathcal{O}(r_o)$$

or smaller (see fig. 5.7). Now, at time $t = \textit{later}$, let all $\vec{v} \rightarrow -\vec{v}$. By microscopic reversibility you should get fig. 5.8. Try it! You get fig. 5.9. It is round off

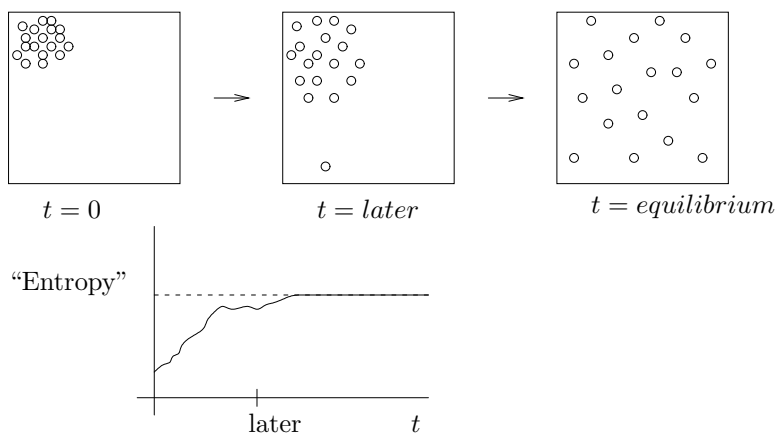


Figure 5.7:

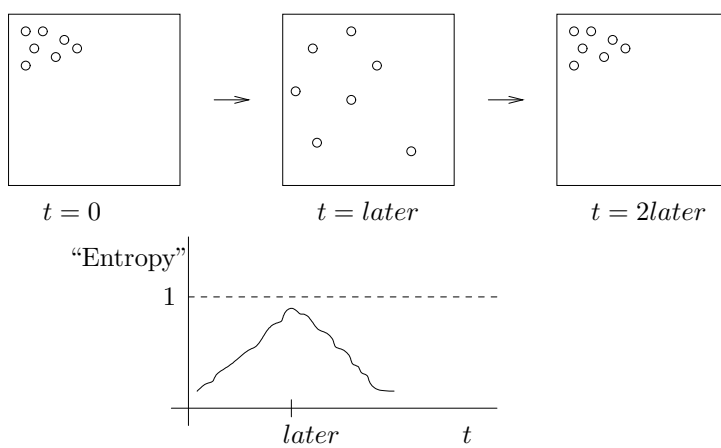


Figure 5.8:

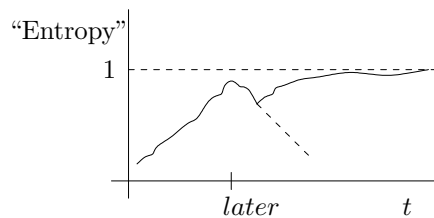


Figure 5.9:

error, but try to improve your precision. *You will always get the same thing.* Think about it. The second law is a law, not someone's opinion.

Chapter 6

Monte Carlo and the Master Equation

An important description for nonequilibrium statistical mechanics is provided by the master equation. This equation also provides the motivation for the Monte Carlo method, the most important numerical method in many-body physics.

The master equation, despite the pretentious name, is phenomenology following from a few reasonable assumptions. Say a system is defined by a state $\{S\}$. This state might, for example, be all the spins ± 1 that an Ising model has at each site. That is $\{S\}$ are all the states appearing in the

$$Z = \sum_{\{S\}} e^{-E\{S\}/k_B T} \quad (6.1)$$

partition function above. We wish to know the probability of being in some state $\{S\}$ at a time t ,

$$p(\{S\}, t) = ? \quad (6.2)$$

from knowledge of probabilities of states at earlier times. In the canonical ensemble (fixed T , fixed number of particles)

$$P_{eq}(\{S\}) \propto e^{-E\{S\}/k_B T} \quad (6.3)$$

of course. For simplicity, let's drop the $\{ \}$ brackets and let

$$\{S\} \equiv S \quad (6.4)$$

As time goes on, the probability of being in state S *increases* because one makes transitions *into* this state from other (presumably nearby in some sense) states S' . Say the probability of going from $S \rightarrow S'$, in a unit time is

$$W(S, S') \equiv_{(\text{Read As})} W(S \leftarrow S') \quad (6.5)$$

This is a transition rate as shown in fig. 6.1. Hence

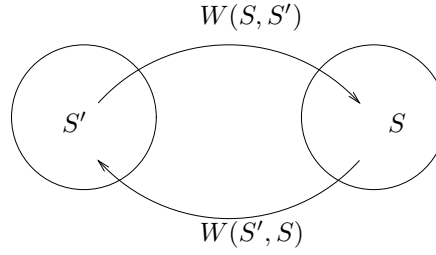


Figure 6.1:

$$W(S, S') P(S') \quad (6.6)$$

are contributions increasing $P(S)$.

However, $P(S)$ *decreases* because one makes transitions *out* of S to the states S' with probability

$$W(S', S) \underset{(\text{Read As})}{\equiv} W(S' \leftarrow S) \quad (6.7)$$

So that

$$W(S', S) P(S) \quad (6.8)$$

are contributions decreasing $P(S)$ as time goes on.

The master equation incorporates both these processes as

$$\frac{\partial P(S, t)}{\partial t} = \sum_{S'} [W(S, S') P(S') - W(S', S) P(S)] \quad (6.9)$$

which can be written as

$$\frac{\partial P(S, t)}{\partial t} = \sum_{S'} L(S, S') P(S') \quad (6.10)$$

where the inappropriately named Liouvillian is

$$L(S, S') = W(S, S') - \left(\sum_{S''} W(S'', S) \right) \delta_{S, S'} \quad (6.11)$$

Note that the master equation

1. has no memory of events further back in time than the differential element “ dt ”. This is called a Markov process
2. is linear.
3. is not time reversal invariant as $t \rightarrow -t$ (since it is first order in time).

In equilibrium,

$$\frac{\partial P_{eq}(S)}{\partial t} = 0 \quad (6.12)$$

since we know this to be the definition of equilibrium. Hence in Eq. 6.9,

$$W(S, S') P_{eq}(S') = W(S', S) P_{eq}(S) \quad (6.13)$$

which is called detailed balance of transitions or fluctuations. For the canonical ensemble therefore,

$$\frac{W(S, S')}{W(S', S)} = e^{-(E(S) - E(S'))/k_B T} \quad (6.14)$$

This gives a surprisingly simple result for the transitions W , it is necessary (but not sufficient) that they satisfy Eq. 6.14. In fact, if we are only concerned with equilibrium properties, *any* W which satisfies detailed balance is allowable. We need only pick a convenient W !

The two most common W 's used in numerical work are the metropolis rule

$$W_{Metropolis}(S, S') = \begin{cases} e^{-\Delta E/k_B T} & , \Delta E > 0 \\ 1 & , \Delta E \leq 0 \end{cases} \quad (6.15)$$

and the Glauber rule

$$W_{Glauber}(S, S') = \frac{1}{2} (1 - \tanh \frac{\Delta E}{2k_B T}) \quad (6.16)$$

where

$$\Delta E = E(S) - E(S') \quad (6.17)$$

Let us quickly check that $W_{Metropolis}$ satisfies detailed balance. If $E(S) > E(S')$

$$\frac{W(S, S')}{W(S', S)} = \frac{e^{-(E(S) - E(S'))/k_B T}}{1} = e^{-\Delta E/k_B T}$$

Now if $E(S) < E(S')$

$$\frac{W(S, S')}{W(S', S)} = \frac{1}{e^{-(E(S) - E(S'))/k_B T}} = e^{-\Delta E/k_B T}$$

where $\Delta E = E(S) - E(S')$ as above. So Detailed balance works. You can check yourself that the Glauber rule works. The form of the rule is shown in figures 6.2 and 6.3. To do numerical work via Monte Carlo, we simply make transitions from state $\{S'\} \rightarrow \{S\}$ using the probability $W(\{S\}, \{S'\})$. A last bit of black magic is that one usually wants the states $\{S'\}$ and $\{S\}$ to be “close”. For something like the Ising model, a state $\{S\}$ is close to $\{S'\}$ if the two states differ by no more than one spin flip, as drawn in fig. 6.4. Let us do this explicitly for the 2 dimensional Ising model in zero field, the same thing that Onsager solved!

$$E_{STATE} = -J \sum_{\langle ij \rangle} S_i S_j \quad (6.18)$$

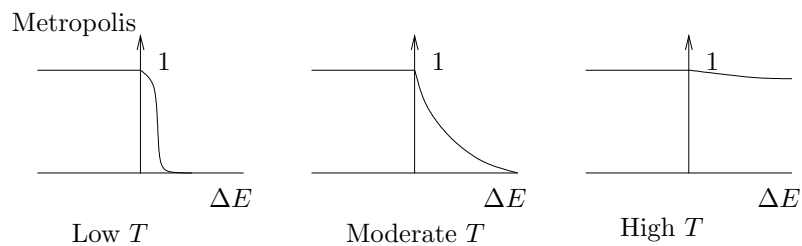


Figure 6.2:

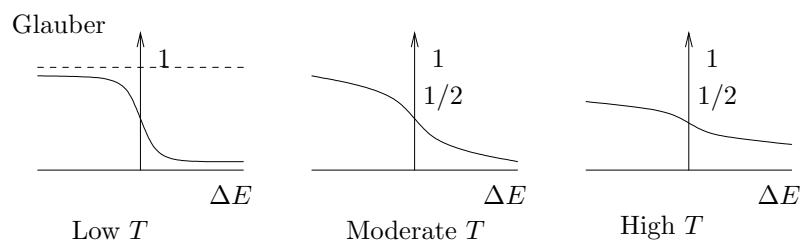


Figure 6.3:

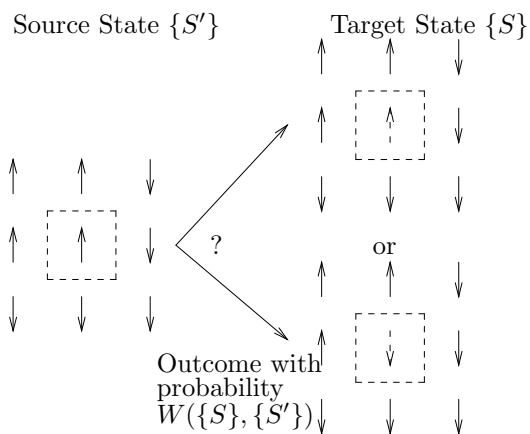


Figure 6.4:

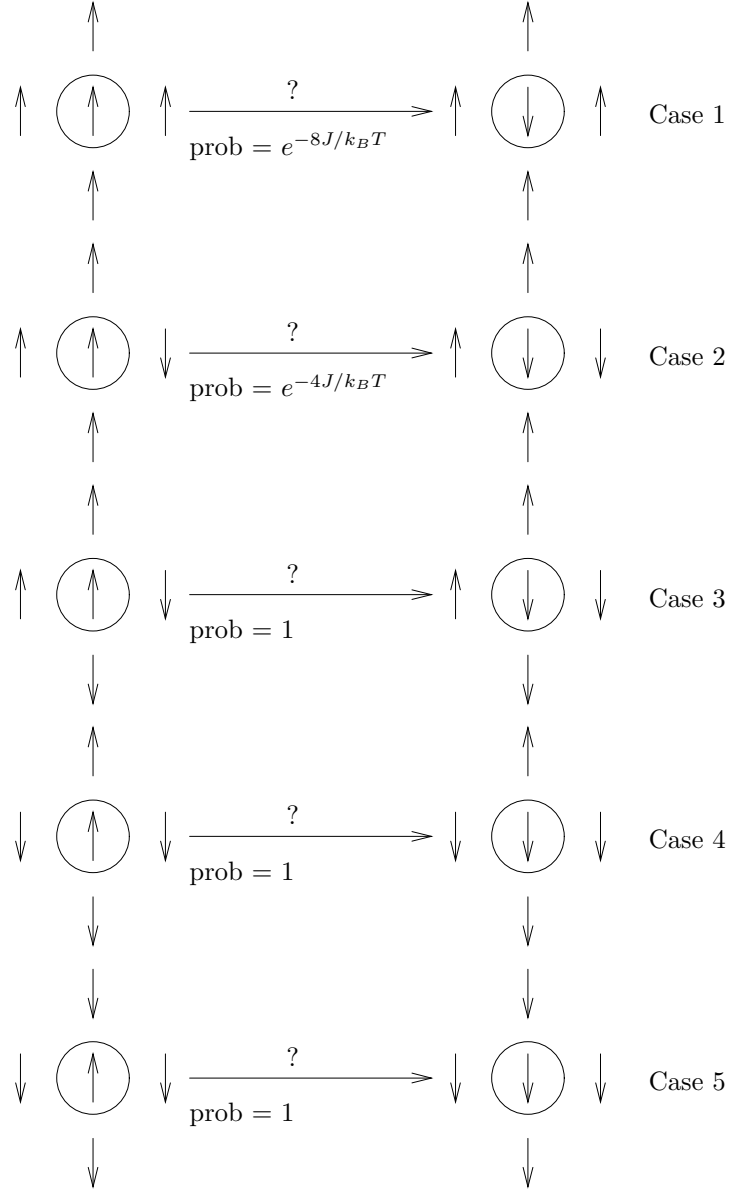


Figure 6.5:

where $S_i = \pm 1$ are the spins on sites $i = 1, 2, \dots, N$ spins (not to be confused with our notation for a state $\{S\}$!), and $\langle ij \rangle$ means only nearest neighbours are summed over. The positive constant J is the coupling interaction. *If only one spin can possibly flip, note there are only five cases to consider as shown in fig. 6.5.*

To get the probabilities (Metropolis)

For Case 1:

$$\begin{aligned} E_{Before} &= -4J \quad , \quad E_{After} = +4J \\ \Delta E &= 8J \\ W(S, S') &= \text{Prob. of flip} = e^{-8J/k_B T} \end{aligned}$$

For Case 2:

$$\begin{aligned} E_{Before} &= -2J \quad , \quad E_{After} = +2J \\ \Delta E &= 4J \\ W(S, S') &= \text{Prob. of flip} = e^{-4J/k_B T} \end{aligned}$$

For Case 3:

$$\begin{aligned} E_{Before} &= 0 \quad , \quad E_{After} = 0 \\ \Delta E &= 0 \\ W(S, S') &= \text{Prob. of flip} = 1 \end{aligned}$$

For Case 4:

$$\begin{aligned} E_{Before} &= +2J \quad , \quad E_{After} = -2J \\ \Delta E &= -4J \\ W(S, S') &= \text{Prob. of flip} = 1 \end{aligned}$$

For Case 5:

$$\begin{aligned} E_{Before} &= +4J \quad , \quad E_{After} = -4J \\ \Delta E &= -8J \\ W(S, S') &= \text{Prob. of flip} = 1 \end{aligned}$$

This is easy if $W = 1$, if $W = e^{-4J/k_B T}$, then you have to know the temperature in units of J/k_B . Let's say, for example,

$$W = 0.1$$

You know the spin flips 10% of the time. You can enforce this by comparing $W = 0.1$ to a random number uniformly between 0 and 1. If the random number is less than or equal to W , then you flip the spin.

Here is an outline of a computer code. If you start at a large temperature and

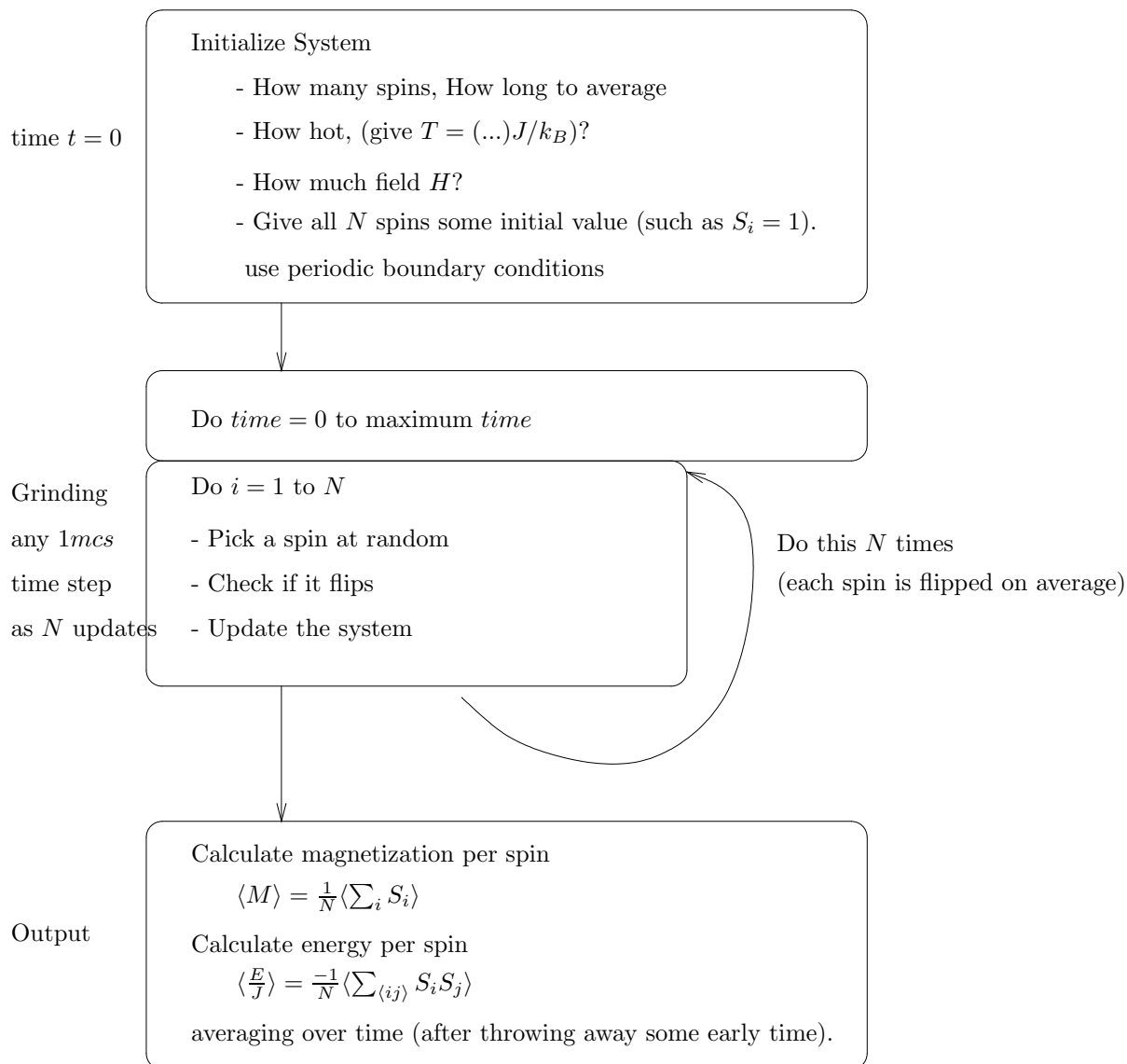
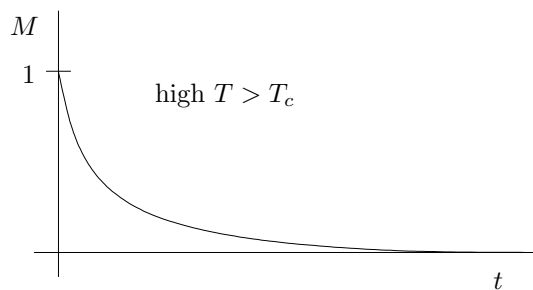


Figure 6.6:



t in Monte Carlo step 1, 1 unit of time
corresponds to attempting N spin flips

Figure 6.7:

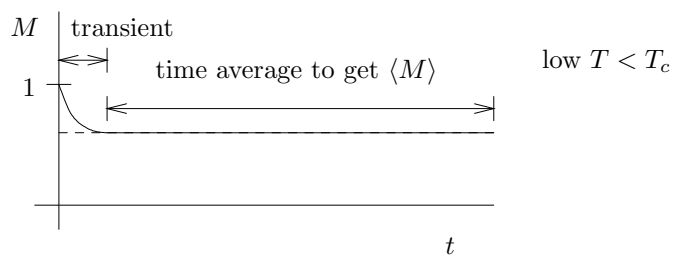


Figure 6.8:

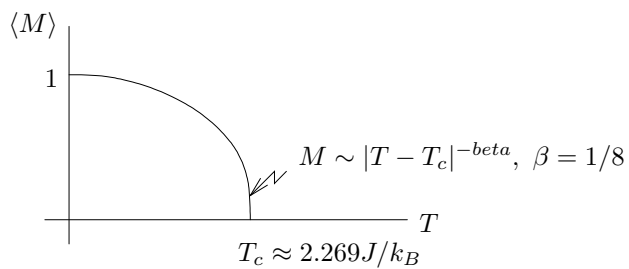


Figure 6.9:

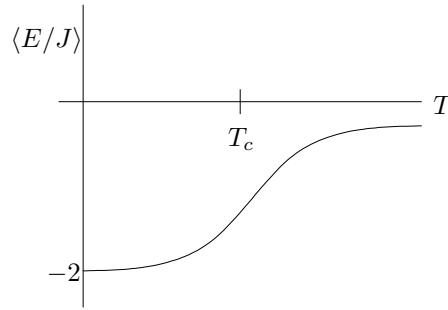


Figure 6.10:

initially $S_i = 1$, the magnetization quickly relaxes to zero (see fig. 6.7). Below T_c , the magnetization quickly relaxes to its equilibrium value (see fig. 6.8). Doing this for many temperatures, for a large enough lattice gives fig. 6.9. Doing this for the energy per spin gives fig. 6.10. which is pretty boring until one plots the heat capacity per spin

$$C \equiv \frac{1}{J} \frac{\partial \langle E \rangle}{\partial T}$$

Try it. You'll get very close to Ousager's results with fairly small systems

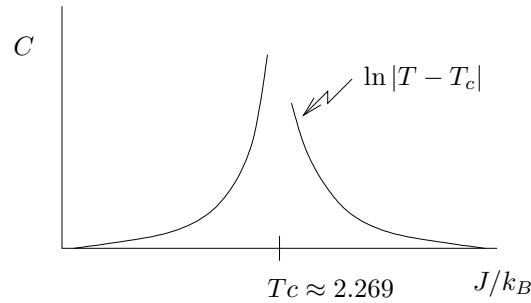


Figure 6.11:

$N = 10 \times 10$, or 100×100 over a thousand or so time steps, and you'll do much better than the mean field theory we did in class.

As an exercise, think how would you do this for the solid-on-solid model

$$Z = \sum_{\{\text{States}\}} e^{-E_{\text{State}}/k_B T}$$

where

$$E_{\text{State}} = J \sum_{i=1}^N |h_i - h_{i+1}|$$

where h_i are heights at sites i .

Chapter 7

Review of Thermodynamics

There are a lot of good thermodynamics books, and a lot of bad ones too. The ones I like are “Heat and Thermodynamics” by Zemansky and Dittman (but beware their mediocre discussion of statistical mechanics, and their confusing discussion of phase transitions), and Callin’s book, “Thermostatistics”, (which is overly formal).

A thermodynamic state of a system is described by a finite number of variables, usually at least two. One needs at least two because of conservation laws for energy E and number of particles N . So one needs at least (E, N) to specify a system, or two independent variables which could give (E, N) . Some systems need more than two, and for some systems, it is a useful approach to only consider one variable, but for now, we will consider that two are enough.

There is a deep question lurking here. Thermodynamics describes systems in the thermodynamic limit: $N \rightarrow \infty$, volume $V \rightarrow \infty$, such that number density $n = N/V = \text{const.}$ Each classical particle has, in three dimensions, 6 independent degrees of freedom describing its position and momentum. So there are not 2, but $\sim 10^{23}$ degrees of freedom! This is a very deep question. Many people have worried about it. But let’s take the physics view. We need *at least* two variables, so let’s use two. Experimentally, this describes thermodynamic systems on large length and time scales, far from, for example, phase transitions.

2 out of $T, P, S, E, V,$ N, \dots are independent

Consider a pure system which can be in a gas, liquid, or simple solid phase. It can be described by temperature T , pressure P , entropy S , in addition to E, V , and N . Temperature is a “good” thermodynamic variable by the *zeroth* law of thermodynamics (two systems in thermal equilibrium with a third system are in thermal equilibrium with each other). Entropy is a “good” thermodynamic variable by the second law of thermodynamics (the entropy of an isolated system never decreases).

Variables can be classified by whether they are

Figure 7.1: Our System

extensive or intensive. Say one considers a system of two independent parts as drawn in fig. 7.2.

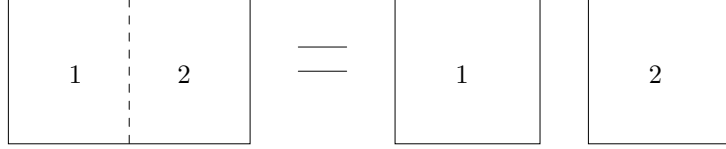


Figure 7.2: A system with two independent parts.

An *extensive* variable satisfies

$$\begin{aligned} V_{total} &= V_1 + V_2 \\ E_{total} &= E_1 + E_2 \\ S_{total} &= S_1 + S_2 \end{aligned}$$

while an *intensive* variable satisfies

$$\begin{aligned} T &= T_1 = T_2 \\ P &= P_1 = P_2 \end{aligned}$$

From mechanics, the work done on a system to displace it by dx is

$$-F dx$$

where F is the force. Acting on an area, we have

$$\underbrace{\bar{d}W}_{\substack{\text{Small} \\ \text{amount} \\ \text{of work}}} = \underbrace{-P}_{\text{Pressure}} \underbrace{dV}_{\substack{\text{Differential} \\ \text{volume} \\ \text{change}}}$$

Two new concepts arise in thermodynamics. Temperature, which is something systems in thermal equilibrium share; and heat, which is a form of energy exchanged between systems at different temperatures.

The law of energy conservation (first law of thermodynamics) is

$$\underbrace{\bar{d}E}_{\substack{\text{change in} \\ \text{internal} \\ \text{energy}}} = \underbrace{\bar{d}W}_{\substack{\text{work} \\ \text{done on} \\ \text{system}}} + \underbrace{\bar{d}Q}_{\substack{\text{heat} \\ \text{transferred}}}$$

The second law of thermodynamics is

$$\bar{d}Q = TdS$$

where S is entropy. Heat always flows from the hot system to the cold system. In fig. 7.3, both sides change due to transfer of heat Q . For the hot side

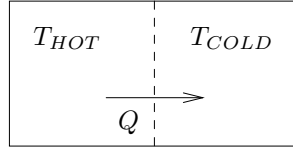


Figure 7.3: System with hot and cold parts exchanging heat Q .

$$(\Delta S)_{HOT} = \frac{-Q}{T_{HOT}}$$

for the cold side

$$(\Delta S)_{COLD} = \frac{-Q}{T_{COLD}}$$

Hence

$$(\Delta S)_{TOTAL} = Q \left[\frac{1}{T_{COLD}} - \frac{1}{T_{HOT}} \right]$$

So

$$(\Delta S)_{TOTAL} \geq 0$$

This is the usual way the second law is presented. Note that this is the only law of physics that distinguishes past and future.

As an aside, it is interesting to note that S must be a “good” thermodynamics variable, or else the second law is violated. If we assume the opposite (that two distinct states have the same S) we can create a perpetual motion machine, as drawn in fig. 7.4. Blobbing up the 1st and 2nd laws gives

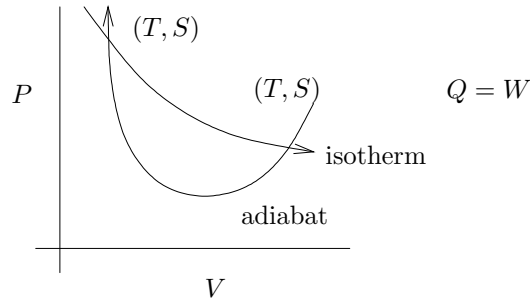


Figure 7.4: Perpetual Motion Machine

$$dE = -P dV + T dS$$

where only 2 out of E , P , V , T and S are independent. This is the most important and useful result of thermodynamics.

To get further, we need equations of state, which relate the variables. Within thermodynamics, these are determined experimentally. For example, the ideal

gas equation of state is

$$P = P(V, T) = N k_B T/V$$

and

$$E = E(T)$$

where usually

$$E = C T$$

and C is a constant determined experimentally. Then

$$C dT = -\frac{N k_B T}{V} dV + T dS$$

So, for an adiabatic process (wherein no heat is exchanged, so $dS = 0$), a change in volume causes a change in temperature.

In general, equations of state are not known. Instead one obtains differential equations of state

$$dP = \left(\frac{\partial P}{\partial V}\right)_T dV + \left(\frac{\partial P}{\partial T}\right)_V dT$$

and

$$dE = \left(\frac{\partial E}{\partial V}\right)_T dV + \left(\frac{\partial E}{\partial T}\right)_V dT$$

Clearly, if we know all the (T, V) dependence of these derivatives, such as $(\frac{\partial E}{\partial T})_V$, we know the equations of state. In fact, most work focuses on the derivatives. Some of these are so important they are given their own names.

Isothermal Compressibility	$k_T = -\frac{1}{V} \left(\frac{\partial V}{\partial P}\right)_T$
Adiabatic Compressibility	$k_s = -\frac{1}{V} \left(\frac{\partial V}{\partial P}\right)_s$
Volume Expansivity	$\alpha = \frac{1}{V} \left(\frac{\partial V}{\partial T}\right)_P$
Heat Capacity at Constant Volume	$C_V = T \left(\frac{\partial S}{\partial T}\right)_V$
Heat Capacity at Constant Pressure	$C_P = T \left(\frac{\partial S}{\partial T}\right)_P$

Of course, $k_T = k_T(T, V)$, etc.

These results rest on such firm foundations, and follow from such simple principles, that any microscopic theory must be consistent with them. Furthermore, a microscopic theory should be able to address the unanswered questions of thermodynamics, e.g., calculate k_T or C_V from first principles.

These quantities can have very striking behavior. Consider the phase diagram of a pure substance (fig. 7.5). The lines denote 1st order phase transi-

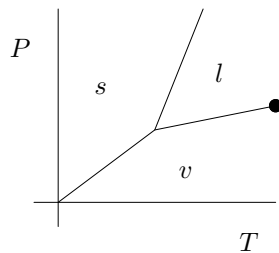


Figure 7.5: Solid, liquid, vapor phases of pure substance

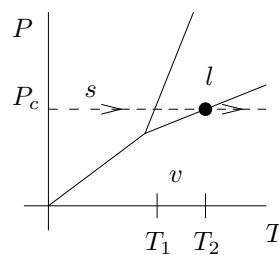


Figure 7.6:

tions. The dot denotes a *continuous* phase transition, often called second order. Imagine preparing a system at the pressure corresponding to the continuous transition, and increasing temperature from zero through that transition (see fig. 7.6).

Figure 7.7 shows the behavior of k_T and C_V : At the first-order transition, k_T has a delta-function divergence while C_V has a cusp. At the continuous transition, both C_V and k_T have power-law divergences. We will spend a lot of time on continuous, so-called second order transitions in the course.

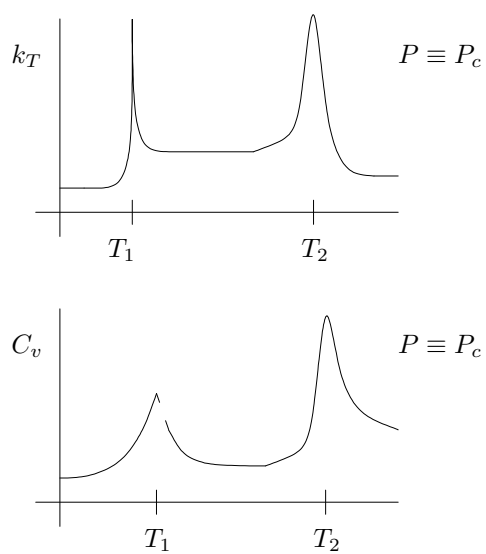


Figure 7.7:

Chapter 8

Statistical Mechanics

We will develop statistical mechanics with “one eye” on thermodynamics. We know that matter is made up of a large number of atoms, and so we expect thermodynamic quantities to be averages of these atomic properties. To do the averages, we need a probability weight.

We will assume that, in a isolated system, the probability of being in a state is a function of the entropy alone. This is because, in thermodynamics, the equilibrium state is the one with the maximum entropy. Hence, the probability weight

$$\rho = \rho(S)$$

To get the form of S we need to use

1. S is extensive.
2. S is maximized in equilibrium.
3. S is defined up to a constant.

Since S is extensive, two independent systems of respective entropies S_1 and S_2 have total entropy $(S_1 + S_2)$ (see fig. 8.1). But

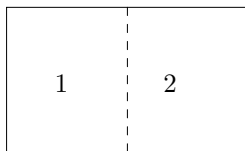


Figure 8.1:

$$\rho(S_{total}) = \rho(S_1)\rho(S_2)$$

Since they are independent (and note I will not worry about normalization for now). Since S is extensive we have

$$\rho(S_1 + S_2) = \rho(S_1)\rho(S_2)$$

or

$$\ln \rho(S_1 + S_2) = \ln \rho(S_1) + \ln \rho(S_2)$$

This is of the form

$$f(S_1 + S_2) = f(S_1) + f(S_2)$$

where f is some function. The solution is clearly

$$f(S) = AS$$

where A is a constant, so

$$\rho = e^{AS}$$

But S is defined up to a constant so

$$\rho = Be^{S/k_B}, \quad A = 1/k_B$$

where the constant B is for normalization, and our earlier carelessness with normalization has no consequence. This is the probability weight for the *micro-canonical* ensemble since $S = S(E, V)$ in thermodynamics from

$$T dS = dE + P dV$$

That is, it is the ensemble for given (E, V) . The more handy ensemble, for given (T, V) is the canonical one. Split the system into a reservoir and a system. They are still independent, but the reservoir is much bigger than the system (fig. 8.2).

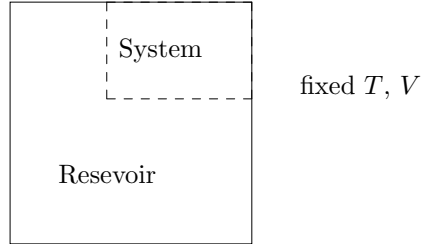


Figure 8.2:

Then

$$\begin{aligned} S(E_{\text{reservoir}}) &= S(E_{\text{total}} - E) \\ &\approx S(E_{\text{total}}) - \left(\frac{\partial S}{\partial E}\right)_{\text{total}} E \end{aligned}$$

and

$$S(E_{\text{reservoir}}) = \text{const} - \frac{E}{T}$$

So, for the canonical ensemble

$$\rho_{state} = \text{const } e^{-\frac{E_{state}}{k_B T}}$$

The normalization factor is

$$Z = \sum_{states} e^{-E_{state}/k_B T}$$

The connection to thermodynamics is from the Helmholtz free energy $F = E - TS$, where

$$Z = e^{-F/k_B T}$$

We haven't shown this, but it is done in many books.

Chapter 9

Fluctuations

Landan and Lifshitz are responsible for “organizing” the theory of fluctuations. They attribute the theory to Einstein in their book on statistical physics. I am following their treatment. First, lets do toy systems, where entropy only depends on one variable, x .

The probability weight

$$\rho \sim e^{S(x)/k_B}$$

so the probability of being in $x \rightarrow x + dx$ is

$$const e^{S(x)/k_B} dx$$

where the constant gives normalization. At equilibrium, S is maximized, and x is at its mean value of $\langle x \rangle$ as drawn in fig. 9.1. Near the maximum we have

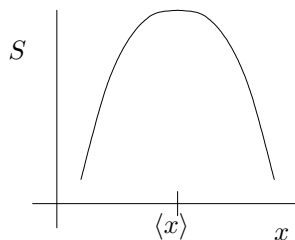


Figure 9.1:

$$S(x) = S(\langle x \rangle) + \left. \frac{\partial S}{\partial x} \right|_{x=\langle x \rangle} (x - \langle x \rangle) + \frac{1}{2} \left. \frac{\partial^2 S}{\partial x^2} \right|_{x=\langle x \rangle} (x - \langle x \rangle)^2 + \dots$$

But $S(\langle x \rangle)$ is just a constant, $\partial S / \partial \langle x \rangle = 0$, and $\partial^2 S / \partial \langle x \rangle^2$ is negative definite, since S is a maximum. Hence,

$$\frac{S(x)}{k_B} = \frac{1}{2\sigma^2} (x - \langle x \rangle)^2$$

where

$$\sigma^2 \equiv -k_B/(\partial^2 S/\partial \langle x \rangle^2)$$

So the probability of being in $x \rightarrow x + dx$ is

$$\text{const. } e^{-(x-\langle x \rangle)^2/2\sigma^2}$$

which is simply a Gaussian distribution. Since it is sharply peaked at $x = \langle x \rangle$, we will let integrals vary from $x = -\infty$ to $x = \infty$. Then

$$\langle (x - \langle x \rangle)^2 \rangle = \sigma^2$$

i.e.,

$$\langle (x - \langle x \rangle)^2 \rangle = \frac{-k_B}{\partial^2 S/\partial \langle x \rangle^2}$$

This is called a fluctuation-dissipation relation of the first kind. Fluctuations in x are related to changes in entropy. They are also called correlation–response relations. As an aside, the “second kind” of fluctuation–dissipation relations turn up in nonequilibrium theory.

Since S is extensive, we have

$$\langle (\Delta x)^2 \rangle = \mathcal{O}(1/N)$$

again, if x is intensive. Note that, from before we said that

$$\langle (\Delta x)^2 \rangle = \mathcal{O}(\xi^3/L^3)$$

So that fluctuations involved microscopic correlation lengths. Now we have

$$\frac{-k_B}{\partial^2 S/\partial \langle x \rangle^2} = \mathcal{O}(\xi^3/L^3)$$

So thermodynamic derivatives are determined by microscopic correlations. This is why we need statistical mechanics to calculate thermodynamic derivatives. We need more degrees of freedom than x to do anything all that interesting, but we have enough to do something.

Consider isothermal toys. These are things like balls, pencils, ropes, and pendulums, in thermal equilibrium. They have a small number of degrees of freedom. For constant temperature, using thermodynamics

$$\begin{array}{rcl} T \Delta S & = & \underbrace{\Delta E}_{\equiv 0} - \underbrace{W}_{\text{work done to system}} \\ & & \text{for simple systems} \end{array}$$

Hence $\Delta S = -W/T$. Since entropy is only defined up to a constant, this is all we need.

This is how it works. Consider a simple pendulum (fig. 9.2). The amount of

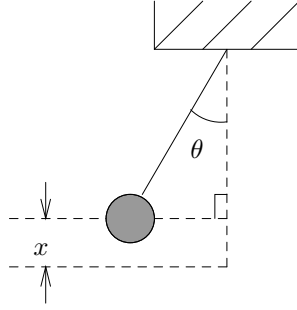


Figure 9.2: Simple pendulum of mass m , length l , displaced an angle θ .

work done by the thermal background to raise the pendulum a height x is

$$mgx$$

where g is the gravitational acceleration. From the picture

$$\begin{aligned} x &= l - l \cos \theta \\ &\approx l \frac{\theta^2}{2} \end{aligned}$$

so

$$W = \frac{mgl\theta^2}{2}$$

and $S = -\frac{mgl\theta^2}{2} + \text{const.}$ so, since $\langle \theta \rangle = 0$, we have

$$\langle \theta^2 \rangle = \frac{k_B T}{mgl}$$

This is a very small fluctuation for a pendulum like in the physics labs. When the mass is quite small, the fluctuations are appreciable. This is the origin, for example, of fluctuations of small particles in water, called Brownian motion, which was figured out by Einstein. There are many other isothermal toys.

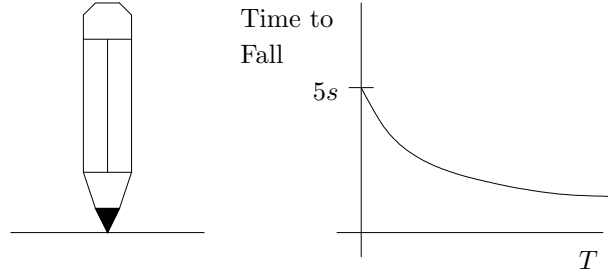


Figure 9.3:

Landan and Lifshitz give several examples with solutions. My favorite, not given in Landan's book though, is the pencil balanced on its tip. It is interesting to estimate how long it takes to fall as a function of temperature (at zero temperature, one needs to use the uncertainty principle).

9.1 Thermodynamic Fluctuations

We can use this approach to calculate the corrections to thermodynamics. Again imagine we have two weakly interacting systems making up a larger system. Due to a fluctuation, there is a change in entropy of the total system due to each subsystem: “1”, the system we are interested in, and “2”, the reservoir

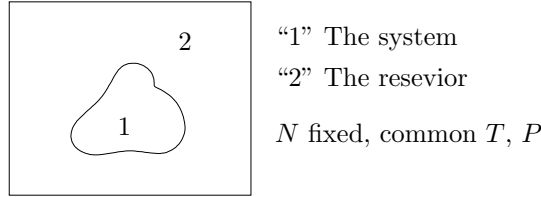


Figure 9.4:

$$\begin{aligned}\Delta S &= \Delta S_1 + \Delta S_2 \\ &= \Delta S_1 + \underbrace{\frac{dE_2 + PdV_2}{T}}_{\text{thermodynamics for the reservoir}}\end{aligned}$$

But when the reservoir increases its energy, the system decreases its energy. The same thing for the volume. However, they have the same P (for mechanical equilibrium) and T (for thermal equilibrium). So we have

$$\Delta S = \Delta S_1 - \frac{\Delta E_1}{T} - \frac{P}{T} \Delta V_1$$

The probability weight is $e^{\Delta S/k_B}$, so we have

$$Probability \propto e^{-\frac{1}{k_B T}(\Delta E - T\Delta S + P\Delta V)}$$

where I have dropped the “1” subscript. There are still only 2 relevant variables, so, for example:

$$\begin{aligned}\Delta E &= \Delta E(S, V) \\ &= \frac{\partial E}{\partial S} \Delta S + \frac{\partial E}{\partial V} \Delta V + \\ &\quad + \frac{1}{2} \left[\frac{\partial^2 E}{\partial S^2} (\Delta S)^2 + 2 \frac{\partial^2 E}{\partial S \partial V} (\Delta S \Delta V) + \frac{\partial^2 E}{\partial V^2} (\Delta V)^2 \right]\end{aligned}$$

to second order. By $\partial E / \partial V$, I mean $(\partial E / \partial V)_S$, but it is already pretty messy notation. Rearranging:

$$\Delta E = \frac{\partial E}{\partial S} \Delta S + \frac{\partial E}{\partial V} \Delta V + \frac{1}{2} \left[\Delta \frac{\partial E}{\partial S} \Delta S + \Delta \frac{\partial E}{\partial V} \Delta V \right]$$

But

$$\left(\frac{\partial E}{\partial S} \right)_V = T$$

and

$$\left(\frac{\partial E}{\partial V} \right)_S = -P$$

so we have

$$\Delta E = T\Delta S - P\Delta V + \frac{1}{2}(\Delta T\Delta S - \Delta P\Delta V)$$

The term in brackets gives the first-order correlation to thermodynamics. In the probability weight

$$Probability \propto e^{-(\Delta T\Delta S - \Delta P\Delta V)/2k_B T}$$

There are still only two independent variables, so, for example

$$\begin{aligned}\Delta S &= \left(\frac{\partial S}{\partial T} \right)_V \Delta T + \left(\frac{\partial S}{\partial V} \right)_T \Delta V \\ &= \frac{C_V}{T} \Delta T + \left(\frac{\partial P}{\partial T} \right)_T \Delta V\end{aligned}$$

Likewise

$$\begin{aligned}\Delta P &= \left(\frac{\partial P}{\partial T} \right)_V \Delta T + \left(\frac{\partial P}{\partial V} \right)_T \Delta V \\ &= \left(\frac{\partial P}{\partial T} \right)_V \Delta T - \frac{1}{\kappa_T V} \Delta V\end{aligned}$$

Putting these expressions into the probability weight, and simplifying, gives

$$Probability \propto e^{[\frac{-C_V}{2k_B T^2}(\Delta T)^2 - \frac{1}{2k_B T \kappa_T V}(\Delta V)^2]}$$

Evidently the probabilities for fluctuations in T and V are independent since

$$Probability = Prob(\Delta T) \cdot Prob(\Delta V)$$

Also, all the averages are gaussians, so by inspection:

$$\begin{aligned}\langle(\Delta T)^2\rangle &= \frac{k_B T^2}{C_V} \\ \langle(\Delta V)^2\rangle &= k_B T \kappa_T V\end{aligned}$$

and

$$\langle\Delta T \Delta V\rangle = 0$$

If we had chosen ΔS and ΔP as our independent variables, we would have obtained

$$Probability \propto e^{[\frac{-1}{2k_B C_P}(\Delta S)^2 - \frac{\kappa_S V}{2k_B T}(\Delta P)^2]}$$

so that

$$\begin{aligned}\langle(\Delta S)^2\rangle &= k_B C_p \\ \langle(\Delta P)^2\rangle &= \frac{k_B T}{\kappa_S V} \\ \langle\Delta S \Delta P\rangle &= 0\end{aligned}$$

Higher-order moments are determined by the Gaussian distributions. So, for example,

$$\begin{aligned}\langle(\Delta S)^4\rangle &= 3\langle(\Delta S)^2\rangle^2 \\ &= 3(k_B C_P)^2\end{aligned}$$

Note that any thermodynamic quantity is a function of only two variables. Hence, its fluctuations can be completely determined from, for example $\langle(\Delta T)^2\rangle$, $\langle(\Delta V)^2\rangle$, and $\langle\Delta T \Delta V\rangle$. For example

$$\begin{aligned}\Delta Q &= \Delta Q(T, V) \\ &= \left(\frac{\partial Q}{\partial T}\right)_V \Delta T + \left(\frac{\partial Q}{\partial V}\right)_T \Delta V\end{aligned}$$

so

$$\langle(\Delta Q)^2\rangle = \left(\frac{\partial Q}{\partial T}\right)_V^2 \langle(\Delta T)^2\rangle + \dots \langle\Delta T \Delta V\rangle^0 + \left(\frac{\partial Q}{\partial V}\right)_T^2 \langle(\Delta V)^2\rangle$$

and

$$\langle(\Delta Q)^2\rangle = \left(\frac{\partial Q}{\partial T}\right)_V^2 \frac{k_B T^2}{C_V} + \left(\frac{\partial Q}{\partial V}\right)_T^2 k_B T \kappa_T V$$

Check this, for example, with $Q = S$ or $Q = P$. Note the form of these relations is

$$\langle(\Delta x)^2\rangle = \mathcal{O}(1/N)$$

for intensive variables, and

$$\langle(\Delta X)^2\rangle = \mathcal{O}(N)$$

for extensive variables. Also note the explicit relationship between corrections to thermodynamics ($\langle(\Delta T)^2\rangle$), thermodynamic derivatives ($C_V = T(\partial S/\partial T)_V$), and microscopic correlations ($1/N = \xi^3/L^3$).

The most important relationship is

$$\langle(\Delta V)^2\rangle = k_B T \kappa_T V$$

although not in this form. Consider the number density $n = N/V$. Since N is fixed

$$\Delta n = \Delta \frac{N}{V} = -\frac{N}{V^2} \Delta V$$

Hence

$$\langle(\Delta V)^2\rangle = \frac{V^4}{N^2} \langle(\Delta n)^2\rangle$$

so,

$$\langle(\Delta n)^2\rangle = \frac{N^2}{V^4} k_B T \kappa_T V$$

or

$$\langle(\Delta n)^2\rangle = \frac{n^2 k_B T \kappa_T}{V}$$

for the fluctuations in the number density. This corresponds to the total fluctuation of number density in our system. That is, if $\Delta n(\vec{r})$ is the fluctuation of number density at a point in space \vec{r} ,

$$\Delta n = \frac{1}{V} \int d\vec{r} \Delta n(\vec{r})$$

(Sorry for the confusing looking notation.) Hence

$$\langle(\Delta n)^2\rangle = \frac{1}{V^2} \int d\vec{r} \int d\vec{r}' \langle \Delta n(\vec{r}) \Delta n(\vec{r}') \rangle$$

Again, for translational invariance of space, and isotropy of space

$$\langle \Delta n(\vec{r}) \Delta n(\vec{r}') \rangle = C(|\vec{r} - \vec{r}'|)$$

In the same way as done earlier we have

$$\langle (\Delta n)^2 \rangle = \frac{1}{V} \int d\vec{r} C(r)$$

where the dummy integral “ \vec{r} ” is “ $\vec{r} - \vec{r}'$ ” of course. But $\langle (\Delta n)^2 \rangle = n^2 k_B T \kappa_T / V$, so we have the thermodynamic “sum rule”:

$$\int d\vec{r} C(r) \equiv \int d\vec{r} \langle \Delta n(r) \Delta n(0) \rangle = n^2 k_B T \kappa_T$$

If

$$\hat{C}(k) \equiv \int d\vec{r} e^{i\vec{k} \cdot \vec{r}} C(\vec{r})$$

we have

$$\lim_{k \rightarrow 0} \hat{C}(k) = n^2 k_B T \kappa_T$$

The quantity $\hat{C}(k)$ is called the structure factor. It is directly observable by x-ray, neutron, light, etc. scattering. For a liquid, one typically obtains something like this The point at $k = 0$ is fixed by the result above.

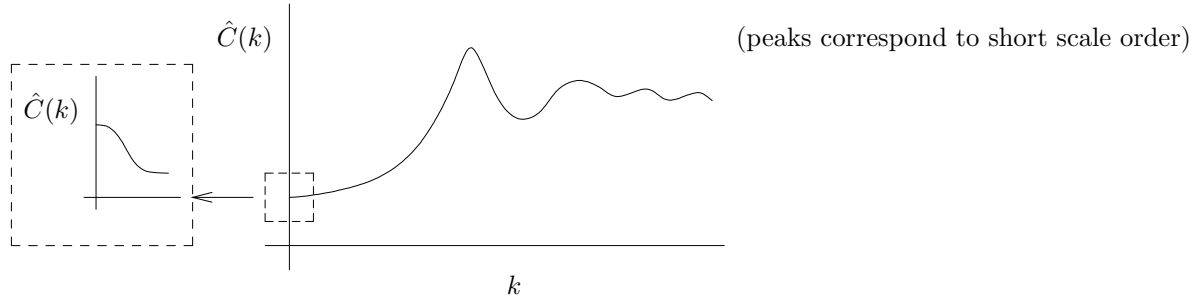


Figure 9.5: Typical structure factor for a liquid.

Chapter 10

Fluctuations of Surfaces

A striking fact is that, although most natural systems are inhomogeneous, with complicated form and structure, almost all systems in thermodynamic equilibrium must have no structure, by Gibbs' phase rule.

As an example, consider a pure system which can be in the homogeneous phases of solid, liquid, or vapour. Each phase can be described by a Gibbs energy $g_s(T, P)$, $g_l(T, P)$, or $g_v(T, P)$. In a phase diagram, the system chooses its lowest Gibbs free energy at each point (T, P) . Hence, the only possibility for inhomogeneous phases are along lines and points:

$$g_s(T, P) = g_l(T, P)$$

gives the solidification line

$$g_l(T, P) = g_v(T, P)$$

gives the vaporization line, while

$$g_s(T, P) = g_l(T, P) = g_v(T, P)$$

gives the triple point of solid–liquid–vapour coexistence.

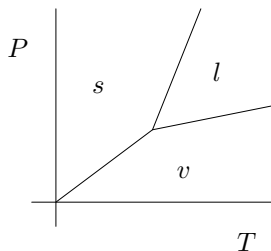


Figure 10.1:

Inhomogeneous states exist on lines in the phase diagram, a set of measure zero compared to the area taken up by the P - T plane.

Much modern work in condensed-matter physics concerns the study of inhomogeneous states. To get a basis for this, we shall consider the most simple such states, equilibrium coexisting phases, separated by a surface.

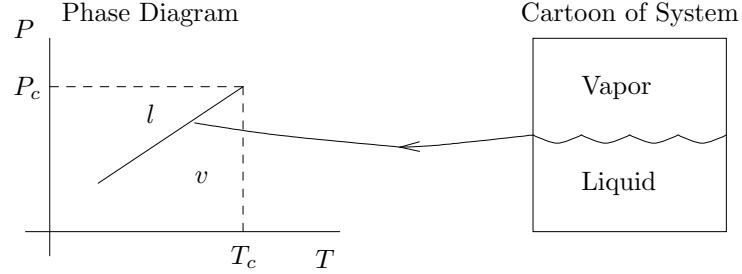


Figure 10.2:

As every one knows, in a finite volume droplet, droplets are spheres while much larger droplets are locally flat. This is the same for liquid-vapour systems, oil-water systems, magnetic domain walls, as well as many other systems. The reason for this is simple thermodynamics. A system with a surface has extra free energy proportional to the area of that surface. Using the Helmholtz free energy now, for convenience,

$$F = (\text{Bulk Energy}) + \underbrace{\sigma \int d\vec{x}}_{\text{Surface Area}} \quad (10.1)$$

where the positive constant σ is the surface tension. Note that the extra free

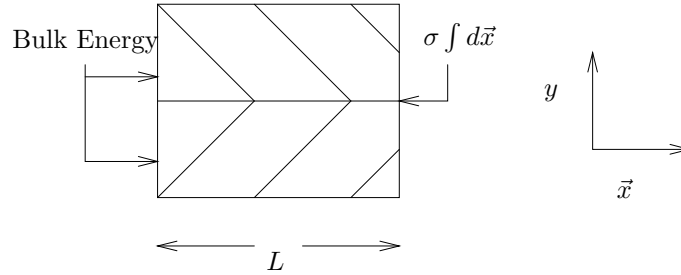


Figure 10.3:

energy due to the surface is minimized by minimizing the surface area.

The coordinate system is introduced in the cartoon above. Note that y is perpendicular to the surface, while \vec{x} is a $(d-1)$ -dimensional vector parallel to the surface. Furthermore, the edge length of the system is L , so its total volume

$$V = L^d$$

while the total surface area

$$A = L^{d-1}$$

Equation 10.1, $\Delta F = \sigma L^{d-1}$, is enough to get us started at describing fluctuations of a surface at a non-zero temperature. Of course, Eq. 10.1 is thermodynamics, while in statistical mechanics, we need a probability function, or equivalently, the partition function.

$$Z = \sum_{\{states\}} e^{-E_{state}/k_B T} \quad (10.2)$$

to calculate such fluctuations strangely enough, using Eq. 10.1 we can construct models for E_{state} in Eq. 10.2! This sounds suspicious since E_{state} should be microscopic, but it is a common practice.

We shall make the decision to limit our investigations to long-length-scale phenomena. So it is natural to expect that phenomena on small length scales, like the mass of a quark or the number of gluons, do not affect our calculations. In particle physics they call this renormalizability: if we consider length scales

$$r > \text{a few Angstroms}$$

all the physics on scales below this only gives rise to coupling constants entering our theory on those larger length scales. The main such coupling constant we shall consider is simply the radius of an atom

$$r_o \approx 5\text{\AA} \quad (10.3)$$

so that we consider length scales

$$r \geq r_o$$

Usually this appears as a restriction, and introduction of coupling constant, in wavenumber k , so that we consider

$$k \leq \Lambda$$

where $\Lambda \equiv 2\pi/r_o$. In particle physics this is called an ultraviolet cutoff. I emphasize that for us Λ or r_o is a physically real cutoff giving the smallest length scale in which our theory is well defined.

This is not too dramatic so far. An important realization in statistical mechanics is that *many* microscopic systems can give rise to the same behaviour on long length scales. In the most simple sense, this means that large-length-scale physics is independent of the mass of the gluon. We can choose any convenient mass we want, and get the same large-length-scale behaviour. In fact, it is commonplace to go much much further than this, and simply invent microscopic models — which are completely unrelated to the true microscopics in question! — which are consistent with thermodynamics. The rule of thumb is to construct models which are easy to solve, but still retain the correct long-length-scale physics.

This insight, which has dramatically changed condensed-matter physics over the last twenty years, is called “universality”. Many microscopic models have the same long-length-scale physics, and so are said to be in the same “universality class”. The notion of a universality class can be made precise in the context of the theory of phase transitions. Even without precision, however, it is commonly applied throughout modern condensed-matter physics!

10.1 Lattice Models of Surfaces

To get a sense of this, let us reconsider our system of interest, an interface. We will construct a “microscopic” model, as shown. We have split the \vec{x} and y axes

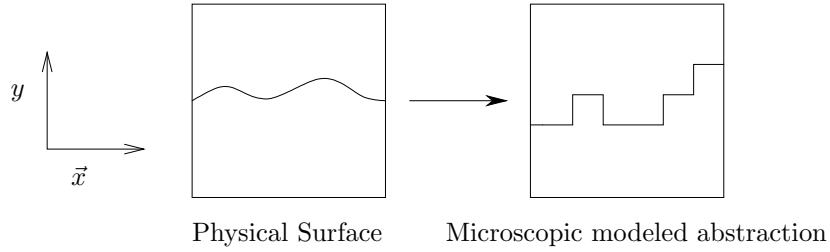


Figure 10.4:

into discrete elements. Instead of a continuous variable where

$$-L/2 \leq x \leq L/2 \quad \text{and} \quad -L/2 \leq y \leq L/2 \quad (10.4)$$

Say (in $d = 2$), we have a discrete sum along the x direction

$$i = -\frac{L}{2}, \dots, -3, -2, -1, 0, 1, 2, 3, \dots, \frac{L}{2}$$

(where L is measured in units of r_o , the small length scale cutoff), and a similar discrete sum along the y direction

$$j = -\frac{L}{2}, \dots, -3, -2, -1, 0, 1, 2, 3, \dots, \frac{L}{2} \quad (10.5)$$

Actually, for greater convenience, let the index along the x axis vary as

$$i = 1, 2, 3, \dots, L \quad (10.6)$$

All the possible states of the system then correspond to the height of the interface at each point i , called h_i , where h_i is an integer $-L/2 \leq h_i \leq L/2$ as shown in fig. 10.5. Hence the sum in the partition function is

$$\sum_{\{states\}} = \sum_{\{h_i\}} = \prod_{i=1}^L \sum_{h_i=-L/2}^{L/2} \quad (10.7)$$

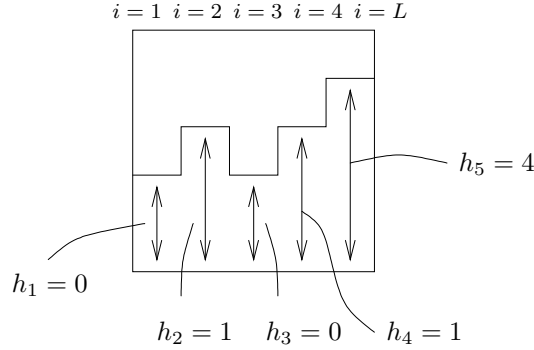


Figure 10.5:

This is a striking simplification of the microscopics of a true liquid–vapour system, but as $L \rightarrow \infty$, the artificial constant imposed by the i, j discrete grid will become unimportant.

Now let us invent a “microscopic” energy $E_{state} = E\{h_i\}$ which seems physically reasonable, and consistent with the free energy $\Delta f = \sigma L^{d-1}$. Physically, we want to discourage surfaces with too much area, since the free energy is minimized by a flat surface. It is natural for a microscopic energy to act with a

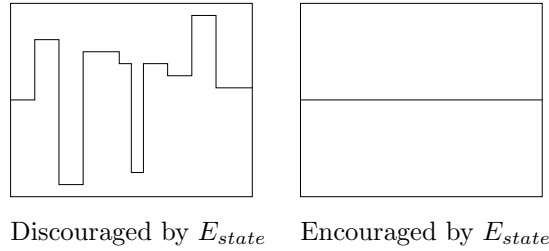


Figure 10.6:

microscopic range. So we enforce flatness, by enforcing local flatness

$$local E_{state} = +(h_i - h_{i+1})^2 \quad (10.8)$$

So the energy increases if a system does not have a locally flat interface. In total

$$E_{state} = J \sum_{i=1}^L (h_i - h_{i+1})^2 \quad (10.9)$$

where J is a positive constant. This is called the discrete Gaussian solid–on–solid model. Equally good models(!) are the generalized solid–on–solid models

$$E_{state}^n = J \sum_{i=1}^L |h_i - h_{i+1}|^n \quad (10.10)$$

where $n = 1$ is the solid-on-solid model, and $n = 2$ is the discrete Gaussian model. With Eq. 10.9, the partition function is

$$Z = \left(\prod_i \sum_{h_i} \right) e^{-\frac{J}{k_B T} \sum_{i'=1}^L |h_{i'} - h_{i'+1}|^2} \quad (10.11)$$

This is for $d = 2$, but the generalization to $d=3$ is easy. Note that

$$Z = Z\left(\frac{k_B T}{J}, L\right) \quad (10.12)$$

One can straight forwardly solve Z numerically. Analytically it is a little tough. It turns out to be easier to solve for the continuum version of Eq. 10.11 which we shall consider next.

10.2 Continuum Model of Surface

Instead of using a discrete mesh, let's do everything in the continuum limit. Let $h(\vec{x})$ be the height of the interface at any point \vec{x} as shown below. This

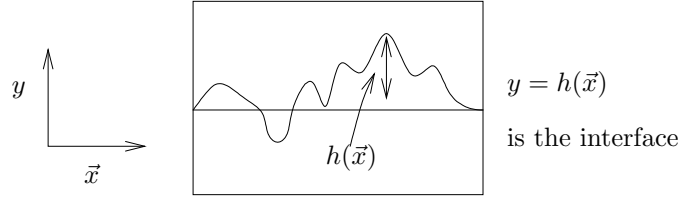


Figure 10.7:

is a useful description if there are no overhangs or bubbles, which cannot be described by $y = h(\vec{x})$. Of course, although it was not mentioned above, the

These give a multivalued $h(\vec{x})$, as shown

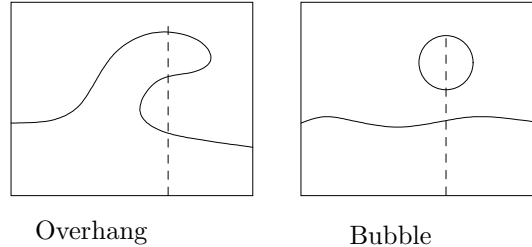


Figure 10.8:

“lattice model” (fig. 10.5) also could not have overhangs or bubbles. The states

the system can have any value of $h(\vec{x})$ at any point \vec{x} . The continuum analogue of Eq. 10.7 is

$$\begin{aligned} \sum_{\{states\}} &= \sum_{\{h_{\vec{x}}\}} = \prod_{\vec{x}} \int dh(\vec{x}) \\ &\equiv \int \int \mathcal{D}h(\cdot) \end{aligned} \quad (10.13)$$

where the last step just shows some common notation for $\prod_{\vec{x}} \int dh(\vec{x})$. Things are a little subtle because \vec{x} is a continuous variable, but we can work it through. One other weaselly point is that the sum over states is dimensionless, while $\prod_{\vec{x}} \int dh(\vec{x})$ has dimensions of $(\prod_x L)$. This constant factor of dimensionality does not affect our results, so we shall ignore it.

For the energy of a state, we shall simply make it proportional to the amount of area in a state $h(\vec{x})$, as shown in fig. 10.7.

$$E_{state}\{h(\vec{x})\} \propto (\text{Area of surface } y = h(\vec{x}))$$

or

$$E_{state} = \sigma(L')^{d-1} \quad (10.14)$$

where $(L')^{d-1}$ is the area of the surface $y = h(\vec{x})$, and σ is a positive constant.

It is worth stopping for a second to realize that this is consistent with our definition of excess free energy $\Delta F = \sigma L^{d-1}$. In fact, it almost seems a little too consistent, being essentially equal to the excess free energy!

This blurring of the distinction between microscopic energies and macroscopic free energies is common in modern work. It means we are considering a “coarse-grained” or *mesoscopic* description. Consider a simple example

$$e^{-F/k_B T} \equiv Z = \sum_{\{h\}} e^{-E_h/k_B T}$$

where h labels a state. Now say one has a large number of states which have the same energy E_h . Call these states, with the same E_h , H . Hence we can rewrite

$$\sum_{\{h\}} e^{-E_h/k_B T} = \sum_{\{H\}} g_H e^{-E_H/k_B T}$$

where

$$g_H = \# \text{ of states } h \text{ for state } H$$

Now let the local entropy

$$S = k_B \ln g_H$$

So we have

$$e^{-F/k_B T} = \sum_{\{H\}} e^{-F_H/k_B T}$$

where $F_H = E_H - TS_H$. This example is partially simple since we only partially summed states with the same energy. Nevertheless, the same logic applies if one

does a partial sum over states which might have different energies. this long digression is essentially to rationalize the common notation (instead of Eq. 10.14) of

$$E_{state} = F_{state} = \sigma(L')^{d-1} \quad (10.15)$$

Let us now calculate $(L')^{d-1}$ for a given surface $y = h(\vec{x})$

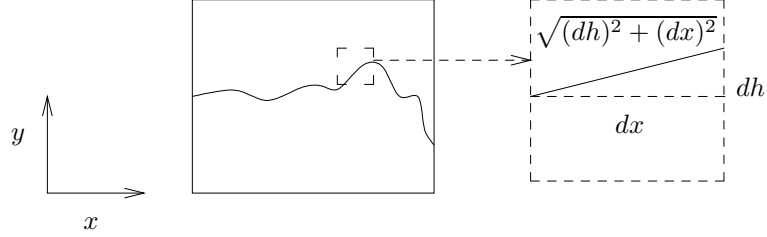


Figure 10.9:

As shown above, each element (in $d = 2$) of the surface is

$$\begin{aligned} dL' &= \sqrt{(dh)^2 + (dx)^2} \\ &= dx \sqrt{1 + \left(\frac{dh}{dx}\right)^2} \end{aligned} \quad (10.16)$$

In d -dimensions one has

$$d^{d-1}L' = d^{d-1}x \sqrt{1 + \left(\frac{\partial h}{\partial \vec{x}}\right)^2} \quad (10.17)$$

Integrating over the entire surface gives

$$(L')^{d-1} = \int d^{d-1}x \sqrt{1 + \left(\frac{\partial h}{\partial \vec{x}}\right)^2}$$

so that

$$E_{state} = \sigma \int d^{d-1}\vec{x} \sqrt{1 + \left(\frac{\partial h}{\partial \vec{x}}\right)^2} \quad (10.18)$$

Of course, we expect

$$\left(\frac{\partial h}{\partial \vec{x}}\right)^2 \ll 1 \quad (10.19)$$

(that is, that the surface is fairly flat), so expanding to lowest order gives

$$E_{state} = \sigma L^{d-1} + \underbrace{\frac{\sigma}{2} \int d^{d-1}\vec{x} \left(\frac{\partial h}{\partial \vec{x}}\right)^2}_{\text{Extra surface area due to fluctuation}} \quad (10.20)$$

The partition function is therefore,

$$Z = \sum_{\{\text{states}\}} e^{-E_{state}/k_B T}$$

or,

$$Z = e^{-\frac{\sigma L^{d-1}}{k_B T}} \left(\prod_{\vec{x}} \int dh(x) \right) \exp \frac{-\sigma}{2k_B T} \int d^{d-1}\vec{x} \left(\frac{\partial h}{\partial \vec{x}'} \right)^2 \quad (10.21)$$

This looks quite formidable now, but it is straightforward to evaluate. In particle physics it is called a free-field theory. Note the similarity to the discrete partition function of Eq. 10.11. The reason Eq. 10.21 is easy to evaluate is that it simply involves a large number of Gaussian integrals

$$\int dh e^{-ah^2}$$

As we know, Gaussian integrals only have two important moments,

$$\langle h(x) \rangle$$

and

$$\langle h(\vec{x}) h(\vec{x}') \rangle$$

All properties can be determined from knowledge of these two moments! However, space is translationally invariant and isotropic in the \vec{x} plane (the plane of the interface), so properties involving two points do not depend upon their

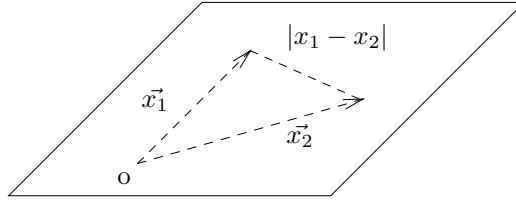


Figure 10.10: \vec{x} plane

exact positions, only the distance between them. That is

$$\langle h(x) \rangle = \text{const.}$$

and

$$\langle h(\vec{x}_1) h(\vec{x}_2) \rangle = G(|\vec{x}_1 - \vec{x}_2|) \quad (10.22)$$

For convenience, we can choose

$$\langle h \rangle = 0 \quad (10.23)$$

then our only task is to evaluate the correlation function

$$\begin{aligned} G(x) &= \langle h(\vec{x}) h(0) \rangle \\ &= \langle h(\vec{x} + \vec{x}') h(\vec{x}') \rangle \end{aligned} \quad (10.24)$$

To get $G(x)$, we use the partition function from Eq. 10.21, from which the probability of a state is

$$\rho_{state} \propto e^{\frac{-\sigma}{2k_B T} \int d^{d-1} \vec{x} (\frac{\partial h}{\partial \vec{x}})^2} \quad (10.25)$$

The subtle thing here is the gradients which mix the probability of $h(\vec{x})$ with that of a very near by $h(\vec{x}')$. Eventually this means we will use fourier transforms. To see this mixing, consider

$$\begin{aligned} \int d\vec{x} (\frac{\partial h}{\partial \vec{x}})^2 &= \int d\vec{x} [\frac{\partial}{\partial \vec{x}} \int d\vec{x}' h(\vec{x}') \delta(\vec{x} - \vec{x}')]^2 \\ &= \int d\vec{x} \int d\vec{x}' \int d\vec{x}'' h(\vec{x}) h(\vec{x}'') (\frac{\partial}{\partial \vec{x}} \delta(\vec{x} - \vec{x}')) \cdot (\frac{\partial}{\partial \vec{x}'} \delta(\vec{x} - \vec{x}')) \end{aligned}$$

Using simple properties of the Dirac delta function. Now let

$$\begin{aligned} x &\rightarrow x'' \\ x' &\rightarrow x \\ x'' &\rightarrow x' \end{aligned}$$

Since they are just dummy indices, giving

$$\begin{aligned} &= \int d\vec{x} \int d\vec{x}' \int d\vec{x}'' h(\vec{x}) h(\vec{x}') (\frac{\partial}{\partial \vec{x}''} \delta(\vec{x}'' - \vec{x})) \cdot (\frac{\partial}{\partial \vec{x}'} \delta(\vec{x}' - \vec{x})) \\ &= \int d\vec{x} \int d\vec{x}' \int d\vec{x}'' h(\vec{x}) h(\vec{x}') \frac{\partial}{\partial \vec{x}} \cdot \frac{\partial}{\partial \vec{x}'} (\delta(\vec{x}'' - \vec{x}) \delta(\vec{x}' - \vec{x})) \\ &= \int d\vec{x} \int d\vec{x}' h(\vec{x}) h(\vec{x}') \frac{\partial}{\partial \vec{x}} \cdot \frac{\partial}{\partial \vec{x}'} \delta(\vec{x} - \vec{x}') \\ &= \int d\vec{x} \int d\vec{x}' h(\vec{x}) h(\vec{x}') [-\frac{\partial^2}{\partial \vec{x}^2} \delta(\vec{x} - \vec{x}')] \end{aligned}$$

(you can do this by parts integration, and there will be terms involving surfaces at ∞ , which all vanish.) Hence,

$$\int d^{d-1} \vec{x} (\frac{\partial h}{\partial \vec{x}})^2 = \int d^{d-1} \vec{x} \int d^{d-1} \vec{x}' h(\vec{x}) M(\vec{x} - \vec{x}') h(\vec{x}') \quad (10.26)$$

where the interaction matrix,

$$M(\vec{x}) = -\frac{\partial^2}{\partial \vec{x}^2} \delta(\vec{x}) \quad (10.27)$$

which has the form shown in fig. 10.11. This “Mexican hat” interaction matrix appears in many different contexts. It causes $h(\vec{x})$ to interact with $h(\vec{x}')$.

In fourier space, gradients $\partial/\partial \vec{x}$ become numbers, or rather wave numbers \vec{k} . We will use this to simplify our project. It is convenient to introduce continuous

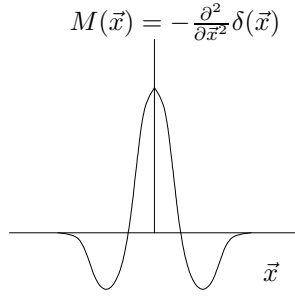


Figure 10.11:

and discrete fourier transforms. One or another is more convenient at a given time. We have,

$$h(\vec{x}) = \begin{cases} \int \frac{d^{d-1}\vec{k}}{(2\pi)^{d-1}} e^{i\vec{k}\cdot\vec{x}} \hat{h}(\vec{k}) & (Continuous) \\ \frac{1}{L^{d-1}} \sum_{\vec{k}} e^{i\vec{k}\cdot\vec{x}} \hat{h}_{\vec{k}} & (Discrete) \end{cases} \quad (10.28)$$

where

$$\begin{aligned} \hat{h}(\vec{k}) &= \int d^{d-1}\vec{x} e^{-i\vec{k}\cdot\vec{x}} h(\vec{x}) \\ &= \hat{h}_{\vec{k}} \end{aligned} \quad (10.29)$$

Closure requires

$$\delta(\vec{x}) = \begin{cases} \int \frac{d^{d-1}\vec{k}}{(2\pi)^{d-1}} e^{i\vec{k}\cdot\vec{x}} \\ \frac{1}{L^{d-1}} \sum_{\vec{k}} e^{i\vec{k}\cdot\vec{x}} \end{cases} \quad (10.30)$$

and

$$\delta(\vec{k}) = \int \frac{d^{d-1}\vec{x}}{(2\pi)^{d-1}} e^{i\vec{k}\cdot\vec{x}}$$

but

$$\delta_{\vec{k},0} = \frac{1}{L^{d-1}} \int d^{d-1}\vec{x} e^{i\vec{k}\cdot\vec{x}} \quad (10.31)$$

is the Kronecker delta

$$\delta_{\vec{k},0} = \begin{cases} 1 & , \text{ if } \vec{k} = 0 \\ 0 & , \text{ otherwise} \end{cases} \quad (10.32)$$

not the Dirac delta function.

We have used the density of states in \vec{k} space as

$$\left(\frac{L}{2\pi}\right)^{d-1}$$

so that the discrete \rightarrow continuum limit is

$$\sum_{\vec{k}} \rightarrow \left(\frac{L}{2\pi}\right)^{d-1} \int d^{d-1}\vec{k}$$

and

$$\delta_{\vec{k},0} \rightarrow \left(\frac{L}{2\pi}\right)^{d-1} \delta(\vec{k}) \quad (10.33)$$

Using the Fourier transforms we can simplify the form of E_{state} in Eq. 10.20. Consider

$$\begin{aligned} \int d^{d-1}\vec{x} \left(\frac{\partial h}{\partial \vec{x}}\right)^2 &= \\ &= \int d\vec{x} \left(\frac{\partial}{\partial \vec{x}} \int \frac{d\vec{k}}{(2\pi)^{d-1}} e^{i\vec{k}\cdot\vec{x}} \hat{h}(\vec{k})\right)^2 \\ &= \int d\vec{x} \int \frac{d\vec{k}}{(2\pi)^{d-1}} \int \frac{d\vec{k}'}{(2\pi)^{d-1}} (-\vec{k} \cdot \vec{k}') \hat{h}(\vec{k}) \hat{h}(\vec{k}') e^{i(\vec{k}+\vec{k}')\cdot\vec{x}} \\ &= \int \frac{d\vec{k}}{(2\pi)^{d-1}} \int \frac{d\vec{k}'}{(2\pi)^{d-1}} (-\vec{k} \cdot \vec{k}') \hat{h}(\vec{k}) \hat{h}(\vec{k}') \left[\int d\vec{x} e^{i(\vec{k}+\vec{k}')\cdot\vec{x}} \right] \\ &= \int \frac{d\vec{k}}{(2\pi)^{d-1}} k^2 \hat{h}(\vec{k}) \hat{h}(-\vec{k}) \end{aligned}$$

after doing \vec{k}' integral. But $h(\vec{x})$ is real, so $\hat{h}^*(\vec{k}) = \hat{h}(-\vec{k})$ (from Eq. 10.29), and we have

$$\int d^{d-1}\vec{x} \left(\frac{\partial h}{\partial \vec{x}}\right)^2 = \int \frac{d^{d-1}\vec{k}}{(2\pi)^{d-1}} k^2 |\hat{h}(\vec{k})|^2 \quad (\text{Continuum})$$

or

$$\int d^{d-1}\vec{x} \left(\frac{\partial h}{\partial \vec{x}}\right)^2 = \frac{1}{L^{d-1}} \sum_{\vec{k}} k^2 |\hat{h}_{\vec{k}}|^2 \quad (\text{Discrete}) \quad (10.34)$$

Hence modes of $\hat{h}_{\vec{k}}$ are *uncoupled* in fourier space; the matrix $M(\vec{x})$ has been replaced by the number k^2 , (see Eqs. 10.26 and 10.27). Therefore we have

$$E_{state} = \sigma L^{d-1} + \frac{\sigma}{2L^{d-1}} \sum_{\vec{k}} k^2 |\hat{h}_{\vec{k}}|^2 \quad (10.35)$$

For the partition function we need to sum out all the states $\sum_{\{states\}}$ using the fourier modes $\hat{h}_{\vec{k}}$ rather than $h(\vec{x})$. This is a little subtle because $h(\vec{x})$ is real while $\hat{h}_{\vec{k}}$ is complex. Of course not all the components are independent since $\hat{h}_{\vec{k}} = \hat{h}_{-\vec{k}}^*$, so we double count each mode. To do the sum properly we have

$$\sum_{\{states\}} = \prod_{\vec{x}} \int dh(\vec{x}) = \prod_{\vec{k}}' \int d^2 \hat{h}_{\vec{k}} \quad (10.36)$$

where $d^2 \hat{h}_k = d \Re \hat{h}_k d \Im \hat{h}_k$

and $\prod'_{\vec{k}} \leftarrow$ Restriction to half of \vec{k} space to avoid double counting

For the most part, this prime plays no role in the algebra below. Recall that

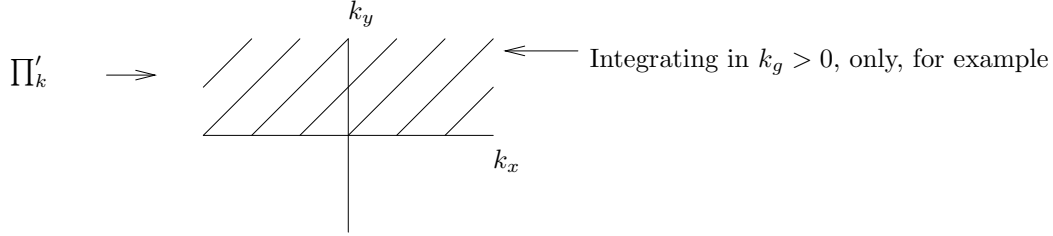


Figure 10.12:

the only nontrivial moment is

$$\begin{aligned} G(x) &= \langle h(\vec{x}) h(0) \rangle \\ &= \langle h(\vec{x} + \vec{x}') h(\vec{x}') \rangle \end{aligned}$$

from Eq. 10.23 above. This has a simple consequence for the Fourier mode correlation function

$$\langle \hat{h}(\vec{k}) \hat{h}(\vec{k}') \rangle$$

Note

$$\begin{aligned} \langle \hat{h}(\vec{k}) \hat{h}(\vec{k}') \rangle &= \int d^{d-1} \vec{x} \int d^{d-1} \vec{x}' e^{-i\vec{k} \cdot \vec{x} - i\vec{k}' \cdot \vec{x}'} \langle h(\vec{x}) h(\vec{x}') \rangle \\ &= \int d\vec{x} \int d\vec{x}' e^{-i\vec{k} \cdot \vec{x} - i\vec{k}' \cdot \vec{x}'} G(\vec{x} - \vec{x}') \end{aligned}$$

Let,

$$\begin{aligned} \vec{y} &= \vec{x} - \vec{x}' & \vec{x} &= \vec{y}' + \frac{1}{2}\vec{y} \\ \vec{y} &= \frac{1}{2}(\vec{x} + \vec{x}') & \vec{x}' &= \vec{y}' + \frac{1}{2}\vec{y} \end{aligned}$$

so that $d\vec{x}d\vec{x}' = d\vec{y}d\vec{y}'$, and

$$\langle \hat{h}(\vec{k}) \hat{h}(\vec{k}') \rangle = \int d\vec{y} e^{-i(\vec{k} - \vec{k}') \cdot \frac{\vec{y}}{2}} G(y) \times \int d\vec{y}' e^{-i(\vec{k} + \vec{k}') \cdot \vec{y}'}$$

so,

$$\begin{aligned} \langle \hat{h}(\vec{k}) \hat{h}(\vec{k}') \rangle &= \int d\vec{y} e^{-i\vec{k} \cdot \vec{y}} G(y) (2\pi)^{d-1} \delta(\vec{k} + \vec{k}') \\ &= \left[\int d\vec{x} e^{-i\vec{k} \cdot \vec{x}} G(\vec{x}) \right] (2\pi)^{d-1} \delta(\vec{k} + \vec{k}') \quad , \text{ letting } \vec{y} \rightarrow \vec{x} \end{aligned}$$

as

$$\langle \hat{h}(\vec{k}) \hat{h}(\vec{k}') \rangle = \begin{cases} \hat{G}(\vec{k}) (2\pi)^{d-1} \delta(\vec{k} + \vec{k}') & (Continuum) \\ \hat{G}_{\vec{k}} L^{d-1} \delta_{\vec{k}+\vec{k}', 0} & (Discrete) \end{cases} \quad (10.37)$$

Hence, in Fourier space, the only non-zero Fourier cos is

$$\langle \hat{h}_{\vec{k}} \hat{h}_{-\vec{k}} \rangle = \langle |\hat{h}_{\vec{k}}|^2 \rangle$$

which is directly related to $\hat{G}_{\vec{k}}$. Using the probability

$$\rho_{state} \propto e^{-\frac{\sigma}{2k_B T L^{d-1}} \sum_{\vec{k}} k^2 |\hat{h}_{\vec{k}}|^2}$$

we have,

$$\langle |\hat{h}_{\vec{k}}|^2 \rangle = \frac{\prod_{\vec{k}'} \int d^2 \hat{h}_{\vec{k}'} |\hat{h}_{\vec{k}}|^2 e^{-\frac{\sigma}{2k_B T L^{d-1}} \sum_{\vec{k}'} k'^2 |\hat{h}_{\vec{k}'}|^2}}{\prod_{\vec{k}'} \int d^2 \hat{h}_{\vec{k}'} e^{-\frac{\sigma}{2k_B T L^{d-1}} \sum_{\vec{k}'} k'^2 |\hat{h}_{\vec{k}'}|^2}} \quad (10.38)$$

Except for the integrals involving $\hat{h}_{\vec{k}}$, everything on the numerator and denominator cancels out. We have to be careful with real and imaginary parts though. Consider

$$\sum_k k^2 |\hat{h}_k|^2 = \sum_k k^2 \hat{h}_k \hat{h}_{-k}$$

For the particular term involving \hat{h}_{15} , for example

$$(15)^2 \hat{h}_{15} \hat{h}_{-15} + (-15)^2 \hat{h}_{-15} \hat{h}_{-15} = 2(15)^2 \hat{h}_{15} \hat{h}_{-15}$$

giving a factor of 2. In any case,

$$\langle |\hat{h}_{\vec{k}}|^2 \rangle = \frac{\int d^2 \hat{h}_{\vec{k}} |\hat{h}_{\vec{k}}|^2 e^{-\frac{2\sigma}{2k_B T L^{d-1}} k^2 |\hat{h}_{\vec{k}}|^2}}{\int d^2 \hat{h}_{\vec{k}} e^{-\frac{2\sigma}{2k_B T L^{d-1}} k^2 |\hat{h}_{\vec{k}}|^2}}$$

(These steps are given in Goldenfeld's book in Section 6.3, p.174.) Now let

$$\hat{h}_{\vec{k}} = R e^{i\theta}$$

so that

$$\langle |\hat{h}_k|^2 \rangle = \frac{\int_0^\infty R dR R^2 e^{-R^2/a}}{\int_0^\infty R dR e^{-R^2/a}}$$

where

$$a = \frac{k_B T}{\sigma k^2} \frac{1}{L^{d-1}} \quad (10.39)$$

A little fiddling gives

$$\langle |\hat{h}_k|^2 \rangle = a = \frac{k_B T}{\sigma k^2} \frac{1}{L^{d-1}}$$

Hence we obtain, from Eq. 10.37

$$\langle \hat{h}(\vec{k}) \hat{h}(\vec{k}') \rangle = \hat{G}(\vec{k}) (2\pi)^{d-1} \delta(\vec{k} + \vec{k}')$$

where

$$\hat{G}(\vec{k}) = \frac{k_B T}{\sigma} \frac{1}{k^2}$$

and

$$G(x) = \langle h(\vec{x}) h(0) \rangle \quad (10.40)$$

From this correlation function we will obtain the width of the surface. We will

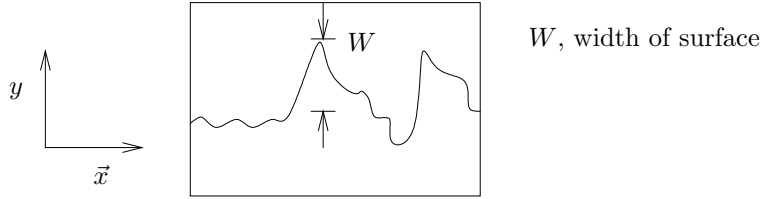


Figure 10.13:

define the width to be the root-mean-square of the height fluctuations averaged over the \vec{x} plane. That is

$$w^2 = \frac{1}{L^{d-1}} \int d^{d-1} \vec{x} \langle (h(\vec{x}) - \langle h(\vec{x}) \rangle)^2 \rangle \quad (10.41)$$

which is overcomplicated for our case, since $\langle h(\vec{x}) \rangle = 0$, and $\langle h(\vec{x}) h(\vec{x}) \rangle = G(0)$, from above. So we have the simpler form

$$w^2 = G(0) \quad (10.42)$$

From the definition of fourier transform,

$$G(0) = \int \frac{d^{d-1} \vec{k}}{(2\pi)^{d-1}} \frac{k_B T}{\sigma k^2} \quad (10.43)$$

which has a potential divergence as $\vec{k} \rightarrow 0$, or $\vec{k} \rightarrow \infty$. We shall always restrict our integrations so that

$$\underbrace{\frac{2\pi}{L}}_{\text{System size}} \leq |\vec{k}| \leq \underbrace{\Lambda}_{\text{Lattice constant}} \quad (10.44)$$

Then it is easy to do the integral of Eq. 10.43. First in $d = 2$,

$$\begin{aligned} G(0) &= \frac{k_B T}{\sigma} \frac{1}{2\pi} \int_{-\Lambda}^{\Lambda} dk \frac{1}{k^2} \\ &= \frac{k_B T}{\pi \sigma} \int_{\frac{2\pi}{L}}^{\Lambda} dk \frac{1}{k^2} \end{aligned}$$

and

$$G(0) = \frac{k_B T}{2\pi^2 \sigma} L, \quad d = 2$$

so

$$w = \sqrt{\frac{k_B T}{2\pi^2 \sigma}} L^{1/2}, \quad d = 2 \quad (10.45)$$

Similarly, in $d = 3$,

$$G(0) = \frac{k_B T}{\sigma} \frac{1}{4\pi^2} 2\pi \int_{\frac{2\pi}{L}}^{\Lambda} \frac{k dk}{k^2}$$

and

$$G(0) = \frac{k_B T}{2\pi \sigma} \ln(L\Lambda/2\pi), \quad d = 3$$

so,

$$w = \sqrt{\frac{k_B T}{2\pi \sigma}} \left(\ln \frac{L\Lambda}{2\pi} \right)^{1/2}, \quad d = 3 \quad (10.46)$$

It is easy to obtain L dependences as dimension of space varies. One obtains,

$$w = \begin{cases} \text{undefined}, & d \leq 1 \\ L^X, & 1 < d < 3 \\ (\ln L)^{1/2}, & d = 3 \\ \text{const.}, & d > 3 \end{cases} \quad (10.47)$$

where

$$X = \frac{3-d}{2} \quad (10.48)$$

is called the roughening exponent. The fact that w is undefined for $d \leq 1$ is explained below.

Eq. 10.47 implies the interface is rough and increases with L for $d \leq 3$. However, note that the width (for $d > 1$) is not large compared to the system

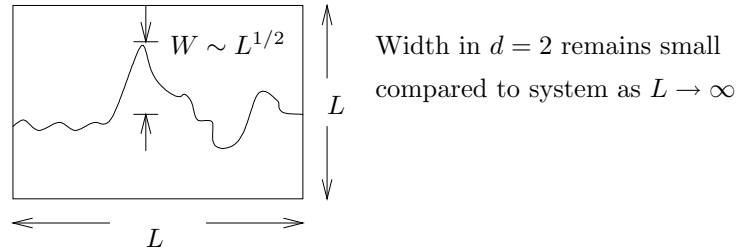


Figure 10.14:

itself, as $L \rightarrow \infty$.

$$\frac{w}{L} = \frac{1}{L^{1-X}} = \frac{1}{L^{(d-1)/2}} \quad (10.49)$$

Hence the width is small for $d > 1$. In $d = 1$, it appears the width is comparable to the system size! This means the interface — and indeed phase coexistence

itself — does not exist in $d = 1$. This is one signature of a *lower critical dimension*, at and below which, phase coexistence cannot occur at any nonzero temperature.

In general, one defines the roughening exponent X by

$$w \sim L^X \quad (10.50)$$

as $L \rightarrow \infty$. This is equivalent to the definition of the surface correlation exponent η_s

$$G(\vec{k}) \sim \frac{1}{k^{2-\eta_s}} \quad (10.51)$$

as $k \rightarrow 0$. (Here $\eta_s = 0$, of course.) The two exponents are related by the formula

$$X = \frac{3-d}{2} - \frac{\eta_s}{2} \quad (10.52)$$

The surface turns out to be self-affine, in the way we discussed earlier for the Von Koch snowflake, since correlation functions like $\hat{G}(\vec{k})$ are power-law like. The self-affine dimension of the rough surface is

$$d_s = X + d - 1 \quad (10.53)$$

To obtain the correlation function in real space is a little awkward because of the divergence of $G(X \rightarrow 0)$ as $L \rightarrow 0$. For convenience, consider

$$g(x) \equiv \langle (h(\vec{x}) - h(0))^2 \rangle \quad (10.54)$$

Multiplying this out gives

$$g(x) = 2(\langle h^2 \rangle - \langle h(\vec{x})h(0) \rangle)$$

This is simply

$$g(x) = \begin{cases} 2(G(0) - G(x)) \\ \text{or} \\ 2(w^2 - G(x)) \end{cases} \quad (10.55)$$

The interpretation of $g(x)$ from Eq. 10.54 is simple, it is the (squared) difference in the heights between two points a distance x apart, as shown in fig. 10.15. Not surprisingly, we expect (for large L)

$$g(x \rightarrow L) = 2w^2 \quad (10.56)$$

since, $G(L) = \langle h(L) h(0) \rangle = \langle h(L) \rangle \langle h(0) \rangle$ because the points 0 and L are so far apart so as to be uncorrelated, so $G(L) = 0$.

In any case, we have

$$\begin{aligned} g(x) &= 2 \frac{k_B T}{\sigma} \frac{1}{(2\pi)^{d-1}} \int d^{d-1} \vec{k} \left(\frac{1 - e^{i\vec{k} \cdot \vec{x}}}{k^2} \right) \\ &= 2 \frac{k_B T}{\sigma} \frac{1}{(2\pi)^{d-1}} \int d^{d-1} \vec{k} \left(\frac{1 - \cos \vec{k} \cdot \vec{x}}{k^2} \right) \end{aligned} \quad (10.57)$$

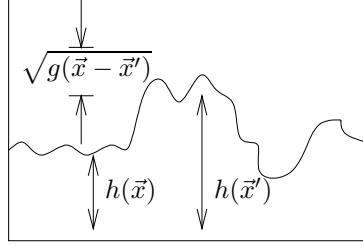


Figure 10.15:

Consider $d = 2$. Let

$$\begin{aligned} g(x) &= 2 \frac{k_B T}{\sigma} \frac{1}{2\pi} \int_{-\Lambda}^{\Lambda} dk \left(\frac{1 - \cos kx}{k^2} \right) \\ &= 2 \frac{k_B T}{\pi \sigma} \int_{2\pi/L}^{\Lambda} dk \frac{1 - \cos kx}{k^2} \end{aligned}$$

let $u = kx$

$$g(x) = 2 \frac{k_B T}{\pi \sigma} x \left\{ \int_{2\pi x/L}^{\Lambda x} du \frac{1 - \cos u}{u^2} \right\}$$

let $\Lambda \rightarrow \infty$ and $L \rightarrow \infty$ in the integral (for convenience)

$$g(x) = 2 \frac{k_B T}{\pi \sigma} x \underbrace{\left(\int_0^{\infty} du \frac{1 - \cos u}{u^2} \right)}_{\pi/2}$$

so,

$$g(x) = \frac{k_B T}{\sigma} x, \quad d = 2 \quad (10.58)$$

In $d = 3$,

$$g(x) = 2 \frac{k_B T}{\sigma} \frac{1}{4\pi^2} \int_0^{2\pi} d\theta \int_{2\pi/L}^{\Lambda} k dk \left(\frac{1 - \cos kx \cos \theta}{k^2} \right)$$

This is a little more involved, and I will just quote the answer

$$g(x) = \frac{k_B T}{4\pi \sigma} \ln x \Lambda, \quad d = 3 \quad (10.59)$$

In general, one obtains

$$g(x) = \begin{cases} \text{undefined}, & d = 1 \\ x^{2X}, & 1 < d < 3 \\ \ln x, & d = 3 \\ \text{const.}, & d > 3 \end{cases} \quad (10.60)$$

where, again $X = (3 - d)/2$. This confirms our earlier claim that correlations are power — law like.

(There is a subtlety here involving constraints which deserves a minute of time. The limit $g(x \rightarrow L) \sim L^{2X}$, but the constraint is not the same as $2w^2$ as derived above. Reconsider the integral of Eq. 10.58:

$$g(x) = 2 \frac{k_B T}{\pi \sigma} x \left\{ \int_{\frac{2\pi x}{L}}^{\Lambda x} du \frac{1 - \cos u}{u^2} \right\}$$

Now *just* let $\Lambda = \infty$. Rewriting gives

$$\begin{aligned} g(x) &= 2 \frac{k_B T}{\pi \sigma} x \left\{ \int_0^\infty du \frac{1 - \cos u}{u^2} - \int_0^{\frac{2\pi x}{L}} du \frac{1 - \cos u}{u^2} \right\} \\ &= 2 \frac{k_B T}{\pi \sigma} x \left\{ \frac{\pi}{2} - \int_0^{\frac{2\pi x}{L}} du \frac{1 - \cos u}{u^2} \right\} \end{aligned}$$

or, for $d = 2$

$$g(x, L) = \frac{k_B T}{\sigma} x f(x/L) \quad (10.61)$$

where

$$f(x) = 1 - \frac{2}{\pi} \int_0^{2\pi x} du \frac{1 - \cos u}{u^2} \quad (10.62)$$

Note that

$$f(x \rightarrow \infty) = 0$$

but

$$f(x \rightarrow 0) = 1$$

This is called a scaling function. End of digression.)

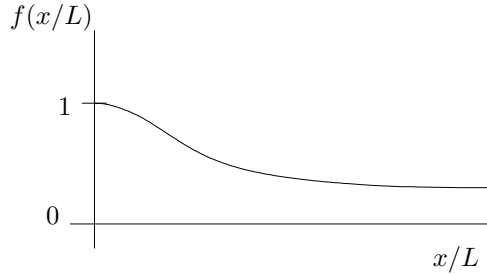


Figure 10.16:

We are pretty well done now, except we shall calculate the “renormalized” surface tension from the partition function. Recall from Eq. 10.21

$$Z = e^{-\frac{\sigma L^{d-1}}{k_B T}} \left(\prod_{\vec{x}} \int dh(x) \exp \frac{-\sigma}{2k_B T} \int d^{d-1} \vec{x} \left(\frac{\partial h}{\partial \vec{x}'} \right)^2 \right)$$

Switching to Fourier space using Eqs. 10.34 and 10.36 gives

$$\begin{aligned} Z &= e^{-\frac{\sigma L^{d-1}}{k_B T}} \left(\prod_{\vec{k}} \int d^2 \hat{h}_{\vec{k}} \right) \exp \frac{-\sigma}{2k_B T} \frac{1}{L^{d-1}} \sum_k k^2 |\hat{h}_k|^2 \\ &\equiv e^{-F/k_B T} \end{aligned} \quad (10.63)$$

So calculating Z gives the thermodynamic free energy F . We will call the “renormalized”, or more properly the thermodynamic surface tension σ^*

$$\sigma^* \equiv \left(\frac{\partial F}{\partial A} \right)_{T, V, N}$$

or,

$$\sigma^* = \left(\frac{\partial F}{\partial L^{d-1}} \right)_{T, V, N} \quad (10.64)$$

going back to Eq. 10.63

$$Z = e^{-\frac{\sigma L^{d-1}}{k_B T}} \prod_{\vec{k}} \int d^2 \hat{h}_{\vec{k}} e^{-\frac{\sigma}{2k_B T} \frac{k^2}{L^{d-1}} |\hat{h}_k|^2}$$

Since all modes are independent. But

$$\int d^2 \hat{h}_{\vec{k}} = \int_{-\infty}^{\infty} d\Re \hat{h}_{\vec{k}} \int_{-\infty}^{\infty} d\Im \hat{h}_{\vec{k}}$$

or if $\hat{h}_{\vec{k}} = R e^{i\theta}$

$$\int d^2 \hat{h}_k = \int_0^{2\pi} d\theta \int_0^{\infty} R dR \quad (10.65)$$

Hence

$$Z = e^{-\frac{\sigma L^{d-1}}{k_B T}} \prod_{\vec{k}} \int_0^{2\pi} d\theta \int_0^{\infty} R dR e^{-(\frac{\sigma}{2k_B T} \frac{k^2}{L^{d-1}}) R^2}$$

$$\text{let } u = \left(\frac{\sigma}{2k_B T} \frac{k^2}{L^{d-1}} \right) R^2$$

$$\begin{aligned} Z &= e^{-\frac{\sigma L^{d-1}}{k_B T}} \prod_{\vec{k}} \int_0^{2\pi} d\theta \frac{1}{2 \left(\frac{\sigma}{2k_B T} \frac{k^2}{L^{d-1}} \right)} \int_0^{\infty} du e^{-u} \\ &= e^{-\frac{\sigma L^{d-1}}{k_B T}} \prod_{\vec{k}} \int_0^{2\pi} d\theta \frac{2\pi k_B T L^{d-1}}{\sigma k^2} \\ &= e^{-\frac{\sigma L^{d-1}}{k_B T}} e^{\ln \prod_{\vec{k}} \frac{2\pi k_B T L^{d-1}}{\sigma k^2}} \\ &= \exp \left(-\frac{\sigma L^{d-1}}{k_B T} + \sum_{\vec{k}} \ln \frac{2\pi k_B T L^{d-1}}{\sigma k^2} \right) \end{aligned}$$

$$Z = \exp\left(-\frac{\sigma L^{d-1}}{k_B T} + \frac{1}{2} \sum_{\vec{k}} \ln \frac{2\pi k_B T L^{d-1}}{\sigma k^2}\right) \quad (10.66)$$

But, $Z = \exp(-F/k_B T)$, up to a constant, so

$$F = \sigma L^{d-1} - \frac{k_B T}{2} \sum_{\vec{k}} \ln \frac{2\pi k_B T L^{d-1}}{\sigma k^2} \quad (10.67)$$

is the free energy due to the surface. The second term gives the contribution due to fluctuations. To take the derivative with respect to $\partial/\partial L^{d-1}$, it is important to note that the \vec{k} 's get shifted with L 's, and that it is the (kL) combination which is *not* a function of L . So to take the derivative, we rewrite

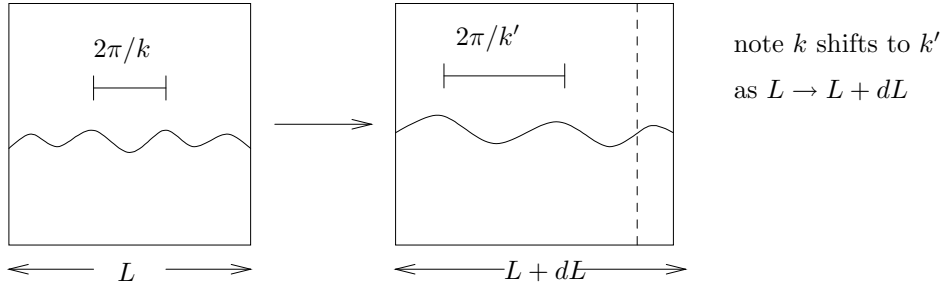


Figure 10.17:

$$F = \sigma L^{d-1} - \frac{k_B T}{2} \sum_{\vec{k}} \ln \frac{2\pi k_B T L^{d-1}}{\sigma (kL)^2}$$

or,

$$F = \sigma L^{d-1} - \frac{k_B T}{2} \sum_{\vec{k}} \ln(L^{d-1})^{(d+1)/(d-1)} - \frac{k_B T}{2} \sum_{\vec{k}} \ln \frac{2\pi k_B T}{\sigma (kL)^2}$$

or,

$$F = \sigma L^{d-1} - \frac{k_B T}{2} \left(\frac{d+1}{d-1}\right) (\ln L^{d-1}) \sum_{\vec{k}} - \frac{k_B T}{2} \sum_{\vec{k}} \ln \frac{2\pi k_B T}{\sigma (kL)^2} \quad (10.68)$$

The last term is independent of L^{d-1} , so taking the derivative $(\frac{\partial F}{\partial L^{d-1}})_{T,V,N} = \sigma^*$ gives,

$$\sigma^* = \sigma - \frac{k_B T}{2} \frac{d+1}{d-1} \left(\frac{1}{L^{d-1}} \sum_{\vec{k}} \right)_{(Discrete)}$$

or,

$$\sigma^* = \sigma - \frac{k_B T}{2} \frac{d+1}{d-1} \left(\int \frac{d^{d-1} \vec{k}}{(2\pi)^{d-1}} \right)_{(Continuum)} \quad (10.69)$$

is the renormalized surface tension. Note that fluctuations decrease the value of σ to the lower σ^* .

$$\text{In } d = 2, \int \frac{d^{d-1}\vec{k}}{(2\pi)^{d-1}} = \frac{1}{2\pi} \int_{-\Lambda}^{\Lambda} dk = \frac{\Lambda}{\pi} \quad (10.70)$$

$$\text{In } d = 3, \int \frac{d^{d-1}\vec{k}}{(2\pi)^{d-1}} = \frac{1}{4\pi^2} \int_0^{2\pi} d\theta \int_0^{\Lambda} k dk = \frac{1}{4\pi^2} \pi \Lambda^2 = \frac{\Lambda^2}{4\pi} \quad (10.71)$$

$$\text{In } d = 1, \int \frac{d^{d-1}\vec{k}}{(2\pi)^{d-1}} = 1 \quad (10.72)$$

Hence we have,

$$\sigma^* = \sigma - \frac{3k_B T}{2\pi} \Lambda \quad , d = 2$$

$$\sigma^* = \sigma - \frac{k_B T}{4\pi} \Lambda^2 \quad , d = 3$$

and since $d = 1$ is odd, let $d = 1 + \epsilon$, for very small ϵ , and

$$\sigma^* = \sigma - \frac{k_B T}{\epsilon} \quad , d = 1 + \epsilon, \epsilon \ll 1 \quad (10.73)$$

(It is worth noting that this is the first term in the $d = 1 + \epsilon$ expansion, where we have only considered the leading term in $E_{State} = \sigma \int d^{d-1}\vec{x} \sqrt{1 + (\frac{\partial h}{\partial \vec{x}})^2} \approx \sigma L^{d-1} + \frac{\sigma}{2} \int d^{d-1}\vec{x} (\frac{\partial h}{\partial \vec{x}})^2 + \dots$)

10.3 Impossibility of Phase Coexistence in $d=1$

Note that, again, the renormalized surface tension vanishes as $d \rightarrow 1$ (as $\epsilon \rightarrow 0$). In our simple calculation σ^* actually becomes negative infinity at $d = 1$.

In fact, the lack of phase coexistence in $d = 1$ is a theorem due to Landau and Pekar, which is worth quickly proving. (See Goldenfeld's book, p.46.)

A homogeneous phase has an energy

$$E_{Bulk} \propto +L^{d-1} \quad (10.74)$$

Putting in a surface gives an extra energy

$$E_{Surface} \propto +L^{d-1} \quad (10.75)$$

So the extra energy one gets due to adding a surface is

$$\begin{aligned} \Delta E_{1Surface} &\propto +L^{d-1} \\ \text{In } d = 1, \Delta E_{1Surface} &\propto +\mathcal{O}(1) \end{aligned} \quad (10.76)$$

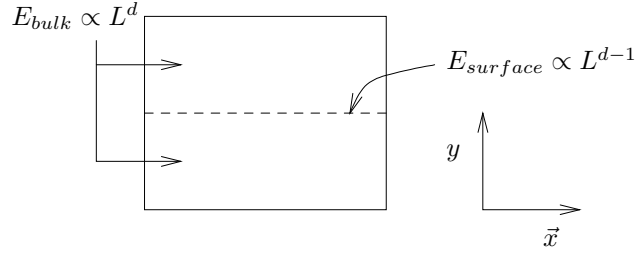


Figure 10.18:

Basically, this looks bad for surfaces.

Since the energy increases, it looks like one wants to have the fewest possible surfaces, so the picture on the previous page is that one: two coexisting phases with one well-defined surface.

However, the system's state is chosen by *free energy* minimization not energy minimization, where

$$F = E - TS$$

Surfaces are *good* for entropy because they can occur anywhere along the y axis, as shown. The total number of such ways is L along the y axis. Of course there

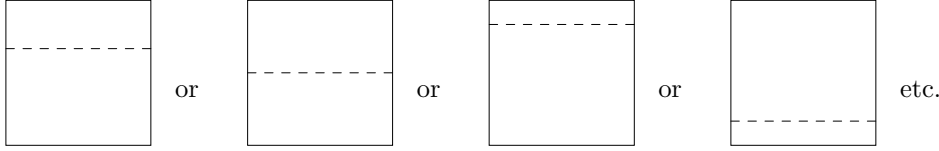


Figure 10.19:

are all rotations and upside down interfaces, so that in the end there are L^d ways. The entropy is

$$\Delta S = k_B \ln(\text{number of states})$$

so we obtain, for a surface,

$$\Delta S = +k_B \ln L^d$$

or,

$$\Delta S = +k_B d(\ln L) \quad (10.77)$$

is the increase in entropy due to having a surface. Hence, the total change in the free energy due to a surface is

$$\Delta F = +\mathcal{O}(L^{d-1}) - k_B T \mathcal{O}(\ln L) \quad (10.78)$$

So, for $d > 1$, $L \rightarrow \infty$, the free energy increases if a surface is added, but for $d = 1$ (at $T > 0$) the free energy *decreases* if a surface is added!

This means coexisting phases are stable for $d > 1$, but in $d = 1$ they are unstable. Since one interface decreases F , we can add another and another and another, and keep on decreasing F until there are no longer coexisting phases (of bulk energy $E \sim L^d$ as $L \rightarrow \infty$), just a mishmash of tiny regions of phase fluctuations. Hence, for systems with short-ranged forces, at $T > 0$, there can be no phase coexistence in $d = 1$. (Short-ranged forces were assumed when we said the surface energy $E_{\text{Surface}} \sim L^{d-1}$).

10.4 Numbers for d=3

Some of these things may seem a little academic, since we have constantly done everything in arbitrary d . Our results for $d = 2$, apply to films absorbed on liquid surfaces and to ledges in crystal surfaces, for example. Experiments have

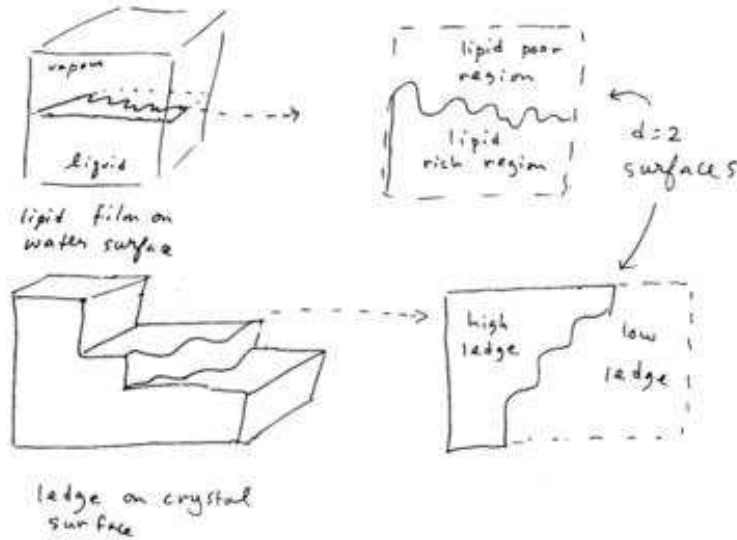


Figure 10.20:

been done on these systems, confirming our $d = 2$ results, such as $w \sim L^{1/2}$.

In $d = 3$, it is worth looking again at two of our main results, Eq. 10.47

$$w = \sqrt{\frac{k_B T}{2\pi\sigma}} \left(\ln \frac{L\Lambda}{2\pi} \right)^{1/2} \quad (10.79)$$

$$\sigma^* = \sigma - \frac{k_B T}{4\pi} \Lambda^2$$

or

$$\sigma^* = \sigma \left(1 - \frac{4\pi k_B T}{\sigma} \left(\frac{\Lambda}{2\pi} \right)^2 \right)$$

or

$$\sigma^* = \sigma(1 - T/T^*) \quad (10.80)$$

where

$$T^* \equiv \frac{\sigma}{4\pi k_B (\Lambda/2\pi)^2} \quad (10.81)$$

Now, reconsider the phase diagram of a pure substance. Our analysis describes,

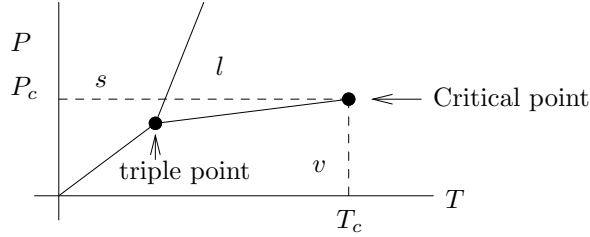


Figure 10.21: Phase Diagram

for example, the behavior along the liquid-vapour coexistence line. Near the triple point of a typical fluid, the surface tension is such that

$$\sqrt{\frac{k_B T_t}{2\pi\sigma}} \approx 1\text{\AA} \quad (10.82)$$

where T_t is the triple point temperature. Similarly, one roughly has

$$\frac{\Lambda}{2\pi} \approx 10\text{\AA} \quad (10.83)$$

Putting these in Eq. 10.79 above implies

$$w = 1\text{\AA} (\ln \frac{L}{10\text{\AA}})^{1/2} \quad (10.84)$$

Putting in numbers is amusing. The width divergence is *very* weak.

$$w = \begin{cases} 2.6\text{\AA}, & L = 10000\text{\AA} \text{ (light wavelength)} \\ 4.3\text{\AA}, & L = 10\text{cm} \text{ (coffee cup)} \\ 5.9\text{\AA}, & L = 1000\text{km} \text{ (ocean)} \\ 9.5\text{\AA}, & L = 10^{30}\text{m} \text{ (universe)} \end{cases}$$

(Of course, in a liquid-vapour system, gravity plays a role. (Certainly it plays a role if one had a liquid-vapour system the size of the universe.) The extra energy due to gravity is

$$E_{Gravity} = \frac{1}{2} \int d^{d-1} \vec{x} (\rho_l - \rho_v) g h^2(x)$$

where ρ is the mass density of the liquid or vapour, and g is gravity. This changes our results in a way you should work out for yourself.)

Putting in numbers for the renormalized surface tension gives the result

$$\sigma^*(T = T_t) \approx \sigma(0.3)$$

which is a much bigger result!

Note that σ^* vanishes (assume $\sigma \neq \sigma(T)$) at $T = T_t$, as T is increased. It is tempting, although a little too rough to interpret

$$T^* = T_c$$

putting in numbers gives $T^* \approx 1.4T_t$, which is T_c to the right order. More precisely, it is known that

$$\sigma^* = \text{Const.}(1 - T/T_c)^\mu$$

for $T \rightarrow T_c$, where $\mu \approx 1.24$ in $d = 3$. So our theory, and (over)–interpretation gives $\mu = 1$, instead. This is not too bad, considering we only went to lowest order in the $E_{State} = \int d\vec{x} \sqrt{1 + (\partial h / \partial \vec{x})^2}$ expansion, and did not consider bulk properties at all.

Chapter 11

Broken Symmetry and Correlation Functions

A major concept in modern work is symmetry. These results for surfaces are an example of more general properties of correlations in systems with broken symmetries.

By a broken symmetry, we mean the following. Note that in the partition function

$$Z = \sum_{\{States\}} e^{-E_{States}/k_B T}$$

the energy of a state has a “symmetry”

$$E_{State}(h + Const.) = E_{State}(h) \quad (11.1)$$

Since

$$E_{State} = \sigma L^{d-1} + \frac{\sigma}{2} \int d^{d-1} \vec{x} \left(\frac{dh}{\partial \vec{x}} \right)^2$$

Naively, since the sum $\sum_{\{States\}}$ does not break this symmetry, we would expect it to be respected by any quantity obtained from Z . The two main quantities we have obtained are the moments $\langle h \rangle$ and $\langle hh \rangle$. As we found before,

$$\langle h \rangle = 0 \quad (11.2)$$

In fact we could have the average value be any result, without changing our algebra. Note that h 's expectation value is a definite number. In particular

$$\langle h + Const. \rangle \neq \langle h \rangle$$

This is a common signature of a broken symmetry variable: Its correlations and averages do not respect that full symmetry of Z or E_{State} .

It is also instructive to think of the symmetry of Eq. 11.1 in a more abstract way. Note that homogeneous systems, whether liquid or vapour, are translationally invariant and rotationally invariant. However, a system with a surface,

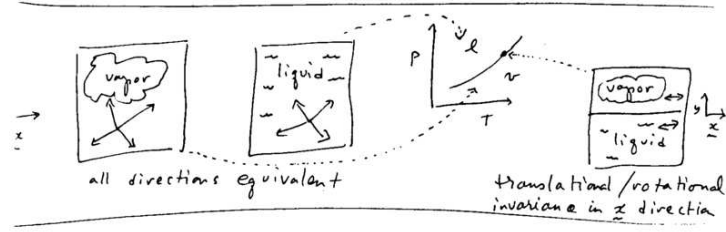


Figure 11.1:

as drawn in fig. 11.1, only has translational invariance and isotropy *in the \vec{x} plane of the interface*; There is no invariance in the y direction. Hence the system with a surface has a lower symmetry (a broken symmetry) compared to the liquid or vapour system.

This perfect symmetry of a homogeneous system is lost precisely because h , the broken symmetry variable, has a definite value, like $\langle h \rangle = 0$.

It is instructive to see that the fact Z and E_{State} have this symmetry implies the form of the correlation functions of the broken symmetry variable h . Consider $d = 2$ only. Since a surface can be anywhere for a given system (it only has a definite value for a given system), it can be different places for different systems. In a very large system, part of $\langle h \rangle$ might have one value, while far far

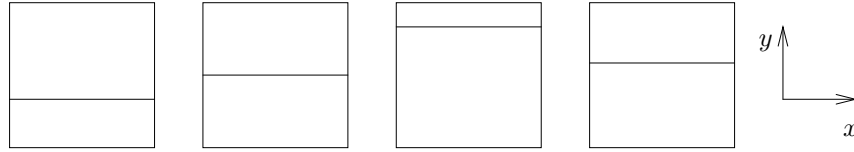


Figure 11.2:

away on the x axis it might want to have another value. The free energy cost of such a fluctuation, if it involves very large length scales (equivalent to $k \rightarrow 0$) is infinitesimal.

$$\Delta E_{State}(k) = \frac{\sigma}{2} k^2 |\hat{h}_k|^2$$

as $k \rightarrow 0$. Indeed, if one has a very very large system encompassing the four shown in fig. 11.2, the four subsystems behave like independent systems which

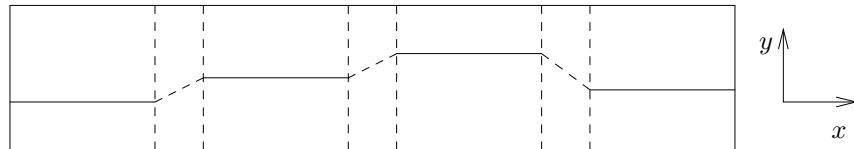


Figure 11.3: Surface of a very big system

each have a definite value for $\langle h \rangle$. Since the energy cost of such fluctuations is infinitesimal as $k \rightarrow 0$, we can set up long-wavelength sloshing modes with infinitesimal $k_B T$. Roughly,

$$\frac{k_B T}{2} = \frac{\sigma}{2} k^2 |\hat{h}_k|^2$$

and

$$\langle |\hat{h}_k|^2 \rangle \sim \frac{k_B T}{\sigma} \frac{1}{k^2}$$

as we recall. These sloshing modes are *capillary waves* driven by thermal noise.

But we need not know the form of E_{State} to obtain $\langle |\hat{h}_k|^2 \rangle \sim 1/k^2$. We only need to know that it breaks a symmetry of Z and E_{State} . As shown in fig. 11.3, fluctuations occur, say $\pm \Delta y$, over a long distance Δx . The values of Δy (which could be $\sim \text{\AA}$) and Δx (which could be \sim meters) are not important for the argument. Let us invent a model where, after one goes a distance Δx to the right, one goes up or down Δy , at random, as shown in fig. 11.4. Of

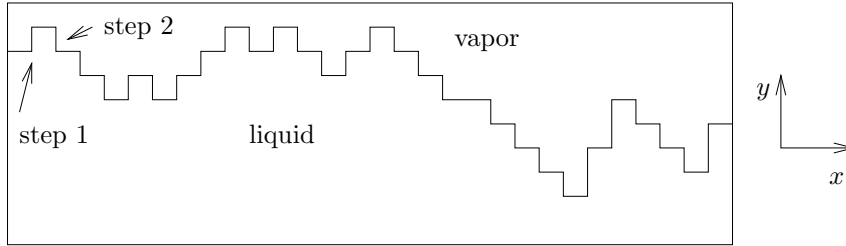


Figure 11.4: Interface (in units of Δx and Δy)

course, this is a one-dimensional random walk along the y axis in “time” x . The root-mean-square deviation of y from its initial value is

$$y_{RMS} \sim x^{1/2}$$

for large “time” x . Of course y_{RMS} is related to the correlation function $(g(x))^{1/2}$ from Eq. 10.54, so we have

$$g(x) \sim x$$

Fourier transforming, and fiddling via Eqs. 10.55 and 10.56, gives $\hat{G}(k) \sim 1/k^2$ for small k so that,

$$\langle |\hat{h}_k|^2 \rangle \sim 1/k^2$$

as expected. Our argument was only in $d = 2$, but it can be generalized to other dimensions.

Note that it was the fact that h was a broken symmetry variable which gave rise to the $1/k^2$; no special details of the interaction were necessary.

Before obtaining the general result, note that the system reacts to a symmetry breaking variable by “shaking it” in such a way as to try and restore the

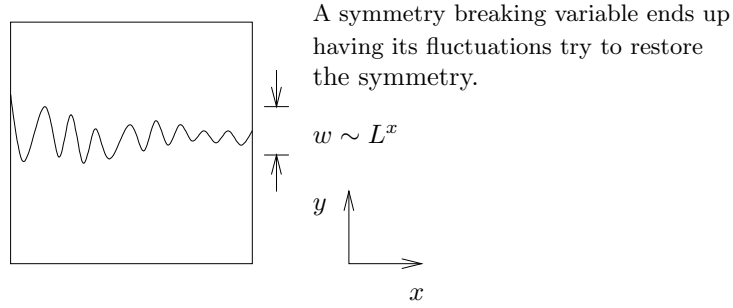


Figure 11.5:

broken symmetry. The rough width of a surface is this attempt to restore the full translational invariance in the y direction, as shown in fig. 11.5. In fact, in $d = 1$, the fluctuations are so strong that the full symmetry *is* restored, and one has translational invariance in the y direction (since coexisting phases cannot exist). In, $1 < d \leq 3$, the fluctuations do not restore the symmetry, but they make the surface rough, or diffuse, as $L \rightarrow \infty$. In $d > 3$, fluctuations do not affect the flat interface.

Finally, it should be noted that the cartoon on p. 64 is self-similar and self affine, in the same sense as the von-koch snowflake we discussed below. Note the important new ingredient: *Broken symmetry* \Leftrightarrow power-law correlations \Leftrightarrow self-similar structure.

11.1 Proof of Goldstone's Theorem

We shall now prove Goldstone's theorem. The theorem is, if B is a broken symmetry variable, where $B(\vec{r})$ is B 's local variable, (\vec{r} is a field point in d dimensions) then

$$\langle \hat{B}(\vec{k}) \hat{B}(\vec{k}') \rangle = (2\pi)^d \delta(\vec{k} + \vec{k}') \hat{G}(k)$$

where

$$\lim_{k \rightarrow 0} \hat{G}(k) \rightarrow \infty$$

and usually

$$\hat{G}(k) \propto 1/k^2$$

We will give not so much a proof of this as an indication of how to work it out for whatever system one is interested in. To make things easier, say

$$\langle B \rangle = 0$$

If this is not true, redefine $B \Rightarrow B - \langle B \rangle$. Assuming one has identified B as an important broken symmetry variable, then the free energy's dependence on B can be found by a Taylor expansion:

$$F(B) = F(0) + \frac{\partial F}{\partial B} B + \frac{1}{2} \frac{\partial^2 F}{\partial B^2} B^2 + \dots$$

But $F(0)$ is just a constant, and $\frac{\partial F}{\partial B} = 0$ at equilibrium, so

$$F(B) = \frac{1}{2} \frac{\partial^2 F}{\partial B^2} B^2 + \dots \quad (11.3)$$

Of course, by F we mean E_{State} or F_{State} in the sense of Eq. 10.10. Now, it turns out that it is important to deal with the local value of $B(\vec{r})$, or the Fourier component $\hat{B}(\vec{k})$. If we use the *local* value for the expansion in Eq. 11.3 we have,

$$F(B) = \frac{1}{2} \int d^d \vec{r} \int d^d \vec{r}' \frac{\partial^2 F}{\partial B(\vec{r}) \partial B(\vec{r}')} B(\vec{r}) B(\vec{r}') \quad (11.4)$$

which looks worse than it is. Since $\frac{\partial^2 F}{\partial B(\vec{r}) \partial B(\vec{r}')}$ is a thermodynamic derivative of some kind, in a system which is (we assume) rotationally invariant and translationally invariant,

$$\frac{\partial^2 F}{\partial B(\vec{r}) \partial B(\vec{r}')} \equiv C(\vec{r}, \vec{r}') = C(\vec{r} - \vec{r}') \quad (11.5)$$

In fourier space

$$F(B) = \frac{1}{2} \int \frac{d^d \vec{k}}{(2\pi)^d} \int \frac{d^d \vec{k}'}{(2\pi)^d} \frac{\partial^2 F}{\partial \hat{B}(\vec{k}) \partial \hat{B}(\vec{k}')} \hat{B}(\vec{k}) \hat{B}(\vec{k}') \quad (11.6)$$

But translational invariance gives

$$\frac{\partial^2 F}{\partial \hat{B}(\vec{k}) \partial \hat{B}(\vec{k}')} = (2\pi)^d \delta(\vec{k} + \vec{k}') \hat{C}(\vec{k}) \quad (11.7)$$

and, if $B(\vec{x})$ is real, so $\hat{B}^*(-k) = \hat{B}(k)$, we obtain (finally)

$$F(B) = \begin{cases} \frac{1}{2} \int \frac{d^d k}{(2\pi)^d} \hat{C}(k) |\hat{B}(k)|^2 & (Continuum) \\ \frac{1}{2L^d} \sum_{\vec{k}} \hat{C}_{\vec{k}} |\hat{B}_{\vec{k}}|^2 & (Discrete) \end{cases} \quad (11.8)$$

We can obtain the form of $\hat{C}(\vec{k})$ from expanding close to $\vec{k} = 0$.

$$\hat{C}(\vec{k}) = (Const.) + (Const.)k^2 + (Const.)k^4 + \dots \quad (11.9)$$

where we have assumed everything is analytic near $k = 0$. Hence terms like $\vec{k} \cdot \vec{k} = k^2$ and $(\vec{k} \cdot \vec{k})^2 = k^4$ appear, but *not* nonanalytic terms like $(\vec{k} \cdot \vec{k})|\vec{k}| = k^3$, (which has the absolute value sign).

Finally, we use that B *breaks* a symmetry of F . Hence F can have *no* dependence on B in the thermodynamic limit $\vec{k} = 0$. Therefore

$$\lim_{k \rightarrow 0} \hat{C}(k) = 0 \quad (11.10)$$

and, if we have analyticity

$$\hat{C}(k) = (Const.)k^2 \quad (11.11)$$

for a broken-symmetry variable near $\vec{k} = 0$. This gives

$$F(B) = (Const.) \sum_{\vec{k}} k^2 |\vec{B}_k|^2 \quad (11.12)$$

This is the same as Eq. 10.35 so we have

$$\langle \hat{B}(\vec{k}) \hat{B}(\vec{k}') \rangle = (2\pi)^d \delta(\vec{k} + \vec{k}') \hat{G}(k)$$

where

$$\hat{G}(k) \propto \frac{1}{k^2}, \quad \vec{k} = 0$$

as claimed. If $\hat{C}(k)$ is not analytic as assumed in Eq. 11.9 we get the weaker condition

$$\lim_{k \rightarrow 0} \hat{G}(k) \rightarrow \infty$$

Finally, it is amusing to note that if B is *not* a broken symmetry variable, from Eq. 11.9 we have

$$F(B) = (Const.) \sum_{\vec{k}} (k^2 + \xi^{-2}) |\hat{B}_k|^2$$

where ξ^{-2} is a constant. This gives $\langle \hat{B}(\vec{k}) \hat{B}(\vec{k}') \rangle = (2\pi)^d \delta(\vec{k} + \vec{k}') \hat{G}(k)$ where $\hat{G}(k) \propto \frac{1}{k^2 + \xi^{-2}}$ as $k \rightarrow 0$. Fourier transforming gives

$$\langle B(\vec{r}) B(0) \rangle = G(r) \propto e^{-r/\xi}$$

as $r \rightarrow \infty$, as we anticipated earlier, when correlations were discussed.

Chapter 12

Scattering

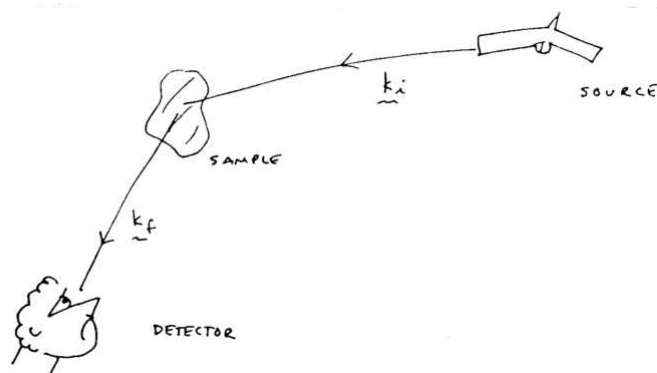


Figure 12.1:

Scattering experiments evidently involve the change in momentum of some incident particle or radiation by a sample. The momentum transferred to the sample (in units of \hbar) is

$$\vec{q} = \vec{k}_i - \vec{k}_f \quad (12.1)$$

by conservation of momentum. This, along with the intensity of the scattered radiation

$$I(\vec{q}) \quad (12.2)$$

is measured at the detector. We shall only consider elastic scattering where energy, w in units of \hbar , is unchanged. Furthermore, we are implicitly considering the first Born approximation and thereby neglecting multiple scattering.

The intensity of the scattered radiation might be due to an electromagnetic field. That is

$$I(\vec{q}) \propto |\hat{E}(q)|^2 \quad (12.3)$$

where \hat{E} is the electric field. Clearly the electric field variations with \vec{q} are due to variations in the dielectric properties of the sample, i.e.,

$$\hat{E}(q) \propto \delta\hat{\epsilon}(q) \quad (12.4)$$

where $\delta\hat{\epsilon}$ is the dielectric variability. Now, if $\delta\hat{\epsilon}$ is a function of only a small number of local thermodynamic variables, like density and temperature

$$\delta\hat{\epsilon} = \left(\frac{\partial\epsilon}{\partial\rho}\right)\delta\hat{\rho} + \left(\frac{\partial\epsilon}{\partial T}\right)\delta\hat{T} \quad (12.5)$$

Usually $\partial\epsilon/\partial T$ is exceedingly small, so we have

$$I(q) \propto |\hat{\rho}(q)|^2 \quad (12.6)$$

A similar argument can be constructed if the scattering is not electromagnetic, giving the same result. We'll not focus on $|\hat{\rho}(q)|^2$ letting

$$S(q) \equiv |\hat{\rho}(q)|^2 \quad (12.7)$$

be the structure factor.

I Should note that, although I've skimmed this, the details giving Eq. 12.6 are often experimentally important. But that's their job, ours is to figure out $S(q)$. We will deal with two cases for $\hat{\rho}(q)$

1. ρ corresponds to the order parameter, a local density or local concentration.
2. The order parameter is *not* the local density, but instead corresponds to a sublattice concentration or magnetization, or to a Bragg peak's intensity.

The first of these is easier. Consider a system of uniform density (we'll do interfaces later), then

$$\langle \rho(x) \rho(0) \rangle = \rho^2 \quad (12.8)$$

of course, so the fourier transform gives

$$S(q) \propto \delta(q) \quad (12.9)$$

A cartoon (fig. 12.2) may make this clear. The second case is more subtle. An example is the sublattice concentration of a binary alloy. Say atoms are A and B in the alloy, and that locally, A wants to be neighbours with B but not with another A. This makes the ordered state look like a checker-board as shown in fig. 12.3.

Assuming the scattering is different from A or B atoms this sort of order can be seen by looking at $S(q)$. However, it must be realized that the ordering of, say, the A atoms is on a sublattice, which has twice the lattice spacing of the original system as shown in fig. 12.4. Hence the density of A atoms is uniform on a sublattice with twice the lattice constant of the original system, and the

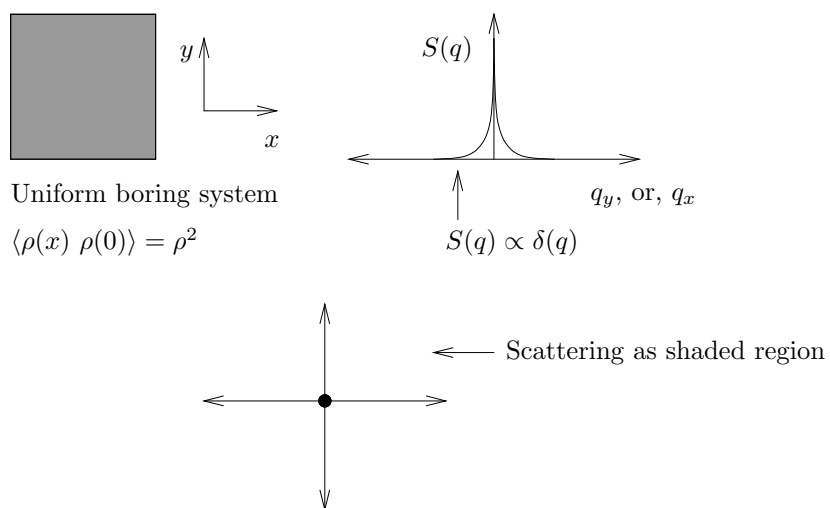


Figure 12.2:

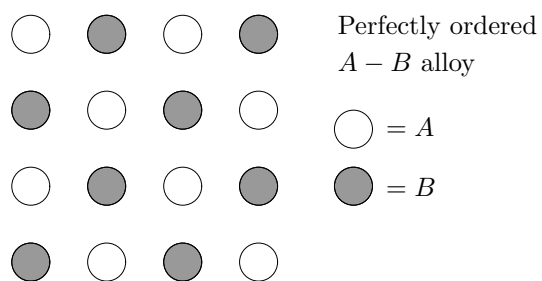


Figure 12.3:

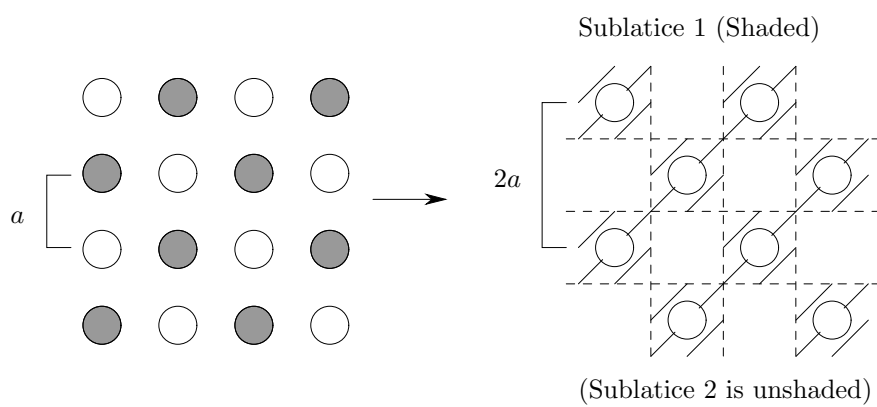


Figure 12.4:

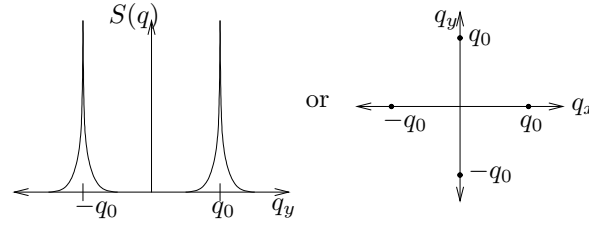


Figure 12.5:

scattering will show peaks *not* at $\vec{q} = 0$ but at $\vec{q} = \vec{q}_0$ corresponding to that structure

$$S(\vec{q}) \propto \sum_{\vec{q}_0} \delta(\vec{q} - \vec{q}_0) \quad (12.10)$$

where $q_0 = 2\pi/(2a)$, and a is the original lattice spacing.

The same thing happens in a crystal, where Bragg peaks form at specific positions in \vec{q} space. By monitoring the height and width of such peaks, the degree of order can be determined. A further complication is that usually the sample will not be so well aligned with the incident radiation that one gets spots on the q_x and q_y axes as shown. Instead they appear at some random orientation (fig. 12.6). This might seem trivial (one could just realign the sam-

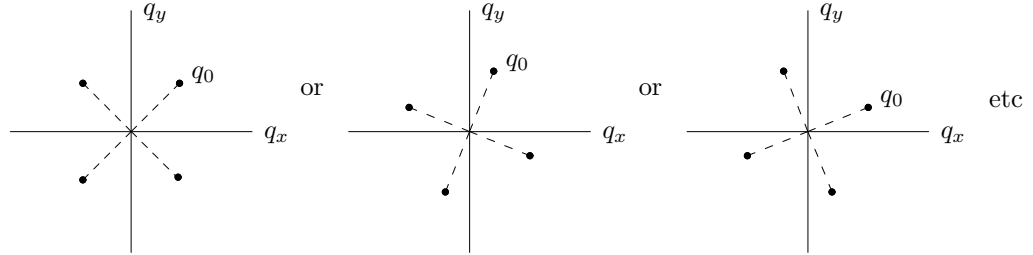


Figure 12.6:

ple), except that often one has many different crystallite orientations scattering simultaneously, so that all the orientations indicated above are smeared into a ring of radius q_0 .

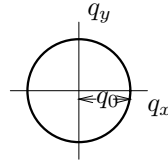


Figure 12.7: Average of Many Crystallites

12.1 Scattering from a Flat Interface

Consider a flat interface without roughening, as drawn in fig. 12.8. For simplicity we will usually consider an interface embedded in dimension $d = 2$. If the density

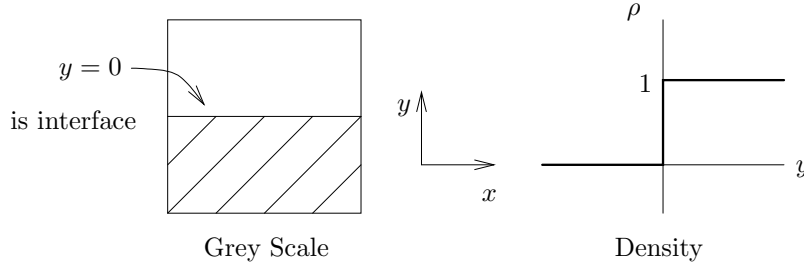


Figure 12.8:

is uniform, it can be written as

$$\rho(x, y) = A \theta(y) + B \quad (12.11)$$

where $\theta(y)$ is the Heaviside step function, and A and B are constants. For simplicity, we will let $A = 1$ and $B = 0$ from now on:

$$\rho(x, y) = \theta(y) \quad (12.12)$$

To get $S(q)$ we need the fourier transform of θ . Let

$$\theta = H + C$$

where C is a to-be-determined constant. Taking a derivative gives

$$\frac{d\theta}{dy} = \frac{dH}{dy} + 0$$

or,

$$\frac{dH}{dy} = \delta(y)$$

Fourier transform gives

$$\hat{H}(k_y) = \frac{1}{ik_y}$$

so that

$$\hat{\theta}(k_y) = \frac{1}{ik_y} + C \delta(k_y)$$

To determine C , note that

$$\theta(x) + \theta(-x) = 1$$

so

$$\hat{\theta}(k_y) + \hat{\theta}(-k_y) = \delta(k_y)$$

Hence $C = 1/2$ and,

$$\hat{\theta}(k_y) = \frac{1}{ik_y} + \frac{1}{2}\delta(k_y) \quad (12.13)$$

The two-dimensional Fourier transform of $\rho(x, y) = \theta(y)$ is therefore

$$\hat{\rho}(\vec{k}) = \frac{1}{ik_y}\delta(k_x) + \frac{1}{2}\delta(\vec{k}) \quad (12.14)$$

From this we obtain the scattering from a flat interface $y = 0$ as the modulus,

$$S(k) = |\hat{\rho}(k)|^2 \propto \frac{1}{k_y^2}\delta(k_x) \quad (12.15)$$

where we have neglected the trivial bulk factor delta function. If the unit vector normal to the interface is \hat{n} and that along the surface is \hat{n}_\perp , then

$$S(k) \propto \frac{1}{(\vec{k} \cdot \hat{n})^2} \delta(\vec{k} \cdot \hat{n}_\perp) \quad (12.16)$$

As a cartoon (fig. 12.10), the scattering appears as a streak pointed in the

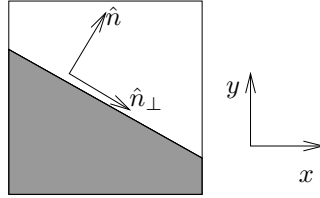


Figure 12.9:

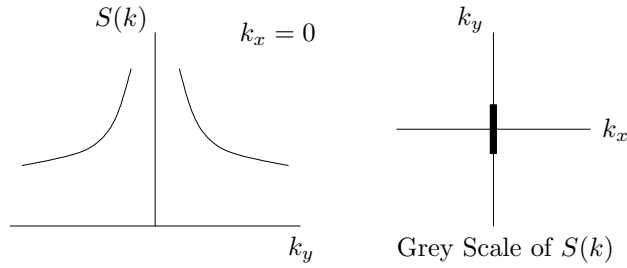


Figure 12.10: Scattering from interface $y = 0$

direction of \hat{n} (\hat{y} in fig. 12.10).

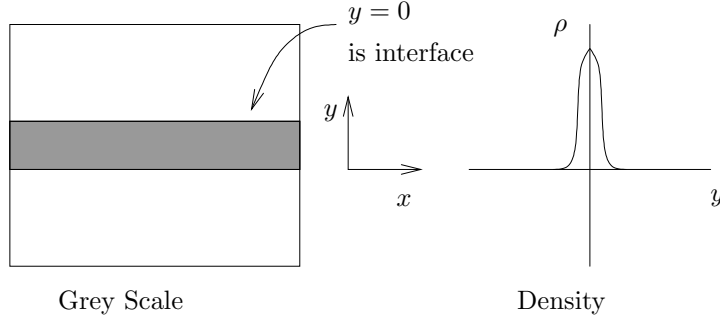


Figure 12.11:

Lets also consider the scattering from a line defect, which is similar, but has important differences. In this case,

$$\rho(x, y) = \delta(y) \quad (12.17)$$

Then

$$\hat{\rho} = \delta(k_x) \quad (12.18)$$

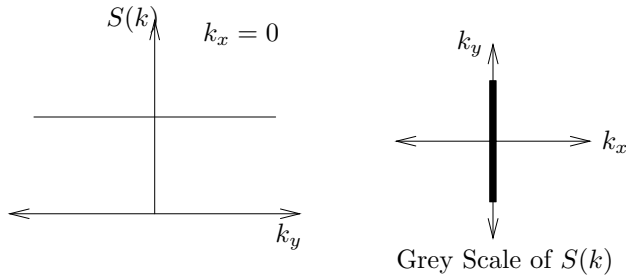
and

$$S(k) \propto \delta(k_x) \quad (12.19)$$

or if the unit vector normal to the surface is \hat{n} , and that along the surface is \hat{n}_\perp ,

$$S(k) \propto \delta(\vec{k} \cdot \hat{n}_\perp) \quad (12.20)$$

for a line defect. Now, if ρ is *not* the order parameter (like the second case

Figure 12.12: Scattering from a line defect at $y = 0$

discussed above) so that the density doesn't vary across the interface, but an order parameter corresponding to sublattice concentration or crystalline orientation does, we have to incorporate the wavenumber \vec{k}_0 corresponding to that structure. For simplicity, we will assume that all one has to do is modulate the structure up to that wavenumber. That is, for an interface

$$S(k) \propto \frac{1}{|(\vec{k} - \vec{k}_0) \cdot \hat{n}|^2} \delta((\vec{k} - \vec{k}_0) \cdot \hat{n}_\perp) \quad (12.21)$$

while for a line defect

$$S(k) \propto \delta((\vec{k} - \vec{k}_0) \cdot \hat{n}_\perp) \quad (12.22)$$

12.2 Roughness and Diffuseness

So far this is *only* geometry. And even as geometry, we have not considered a diffuse interface. Before doing roughness, let's consider a diffuse interface as shown in fig. 12.13. One example of this is

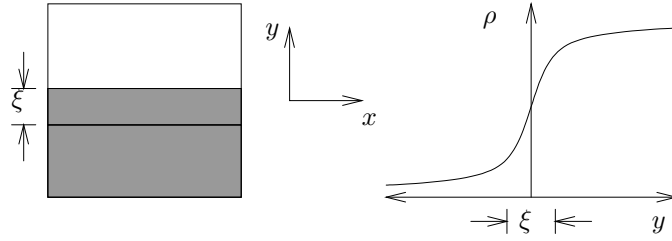


Figure 12.13: A diffuse interface, with no roughness

$$\rho_{diffuse}(x, y) \sim \tanh(y/\xi) \quad (12.23)$$

So that

$$S(k) = |\hat{\rho}(k)_{diffuse}|^2 \sim \frac{\delta(k_x)}{k_y^2} (1 + \mathcal{O}(\xi k_y)^2) \quad (12.24)$$

for an interface at $y = 0$. So the diffuseness only effects scattering at $k_x = 0$

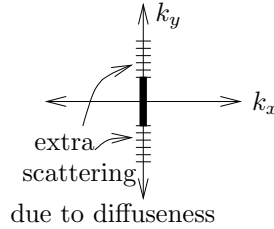


Figure 12.14:

and large k_y . On the other hand, let us now consider a rough interface, with no diffuseness, where the interface is given by

$$y = h(x) \quad (12.25)$$

and

$$\rho_{Rough}(x, y) = \theta(y - h(x)) \quad (12.26)$$

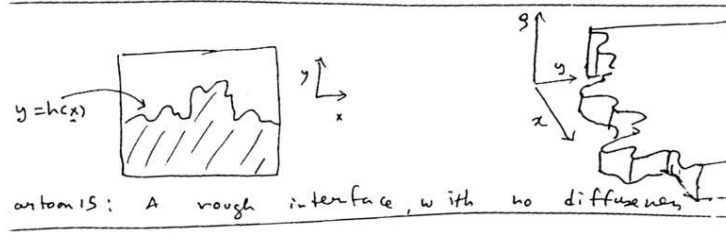


Figure 12.15: A very rough interface, with no diffuseness

If $h(x)$ involves a small variation, we can accordingly expand to obtain

$$\rho_{\text{Rough}} \approx \theta(y) + h(x)\delta(y) + \dots \quad (12.27)$$

since $d\theta/dy = \delta(y)$.

Taking the fourier transform and squaring this gives

$$S(k) = |\hat{\rho}_{\text{Rough}}(k)|^2 \sim \frac{\delta(k_x)}{k_y^2} + |\hat{h}(k_x)|^2 \quad (12.28)$$

Averaging over thermal fluctuations using

$$\langle |\hat{h}(k_x)|^2 \rangle \sim \frac{1}{k_x^{2-\eta}} \quad (12.29)$$

which defines η , gives

$$S_{\text{Rough}}(\vec{k}) \sim \frac{\delta(k_x)}{k_y^2} + \frac{T}{k_x^{2-\eta}} \quad (12.30)$$

where T is a constant. It would appear, then, that one could deconvolute the

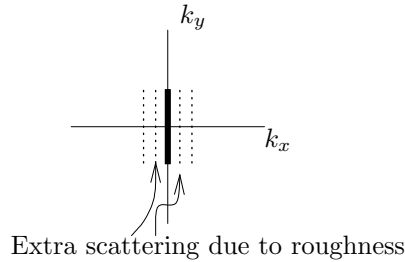


Figure 12.16:

geometry $\frac{\delta(k_x)}{k_y^2}$ from the diffuseness $\mathcal{O}(\xi k_y)^2$, from the roughness $1/k_x^{2-\eta}$, since they have different signatures in k space. For excellent samples this is in fact true, but if there are many randomly oriented interfaces present, these signatures are smeared out, and washed away.

12.3 Scattering From Many Flat Randomly–Oriented Surfaces

Evidently the scattering from many randomly–oriented surfaces gives a pinwheel–like cluster of streaks as shown in fig. 12.17. Angularly averaging over all orien-

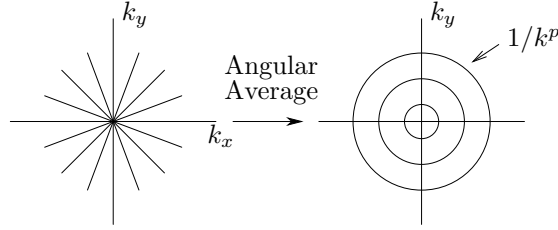


Figure 12.17:

tations of the interface gives

$$S = \frac{\int d\hat{n} |\hat{\rho}_k|^2}{\int d\hat{n}} \quad (12.31)$$

where \hat{n} is the unit vector normal to the surface. We will write

$$\hat{n} = -\sin \theta \hat{x} + \cos \theta \hat{y} \quad (12.32)$$

and

$$\hat{n}_\perp = \cos \theta \hat{x} + \sin \theta \hat{y} \quad (12.33)$$

with

$$\vec{k} = k \cos \beta \hat{x} + k \sin \beta \hat{y} \quad (12.34)$$

Then, for scattering from a surface where $|\hat{\rho}_k|^2 = \frac{\delta(\vec{k} \cdot \hat{n}_\perp)}{|\vec{k} \cdot \hat{n}|^2}$ we have (in $d = 2$)

$$\begin{aligned} S &= \frac{1}{2\pi} \int_0^{2\pi} d\theta \frac{\delta[k \cos \beta \cos \theta + k \sin \beta \sin \theta]}{|-k \cos \beta \sin \theta + k \sin \beta \cos \theta|^2} \\ &= \frac{1}{2\pi} \frac{1}{k^3} \int_0^{2\pi} d\theta \frac{\delta[\cos(\theta - \beta)]}{|\sin(\theta - \beta)|^2} \\ &= \frac{1}{2\pi} \frac{1}{k^3} \int_0^{2\pi} d\theta \frac{\delta[\cos \theta]}{|\sin \theta|^2} \end{aligned}$$

But $\cos \theta = 0$ at $\theta = \pi/2, 3\pi/2$ where $|\sin \theta|^2 = 1$, so

$$S_{Surfaces} = \frac{1}{\pi} \frac{1}{k^3} \quad (12.35)$$

Similarly, in arbitrary dimension one has

$$S_{Surfaces} \propto \frac{1}{k^{d+1}} \quad (12.36)$$

for scattering from a surface. For scattering from a line defect which has $|\hat{\rho}(k)|^2 \propto \delta(k_x)$, one similarly obtains

$$S_{\text{linedefects}} \propto \frac{1}{k^{d-1}} \quad (12.37)$$

We shall call these results, Porod's Law. They involve *only* geometry.

The case where ρ is not the order, and a modulated structure exists is more subtle. If the wavenumber of that structure is \vec{k}_0 then this introduces a new angle via

$$\vec{k}_0 = k_0 \cos \alpha \hat{x} + k_0 \sin \alpha \hat{y} \quad (12.38)$$

Now there are two possible averages

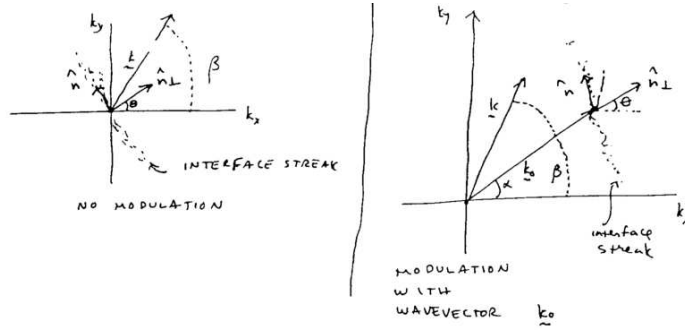


Figure 12.18:

1. α or β , average over crystallite orientation, or, angles in a non-angularly resolved detector. These are equivalent (look at the cartoon for a few seconds), and so we only need to do one of these two averages.
2. θ , averages over surface orientation.

If we average over α or β , for a fixed \hat{n} , it is easy to anticipate the answer as shown in fig. 12.19, where in the cartoon the regions around $(k_x = k_0, k_y = 0)$

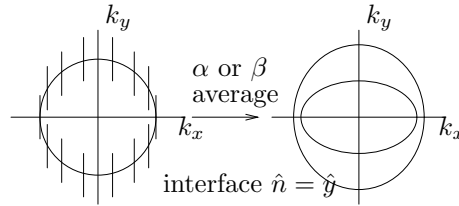


Figure 12.19:

retain the singularity of the original $\delta(k_x - k_0)/(k_y - k_0)^2$.

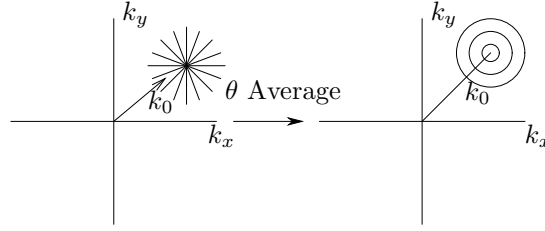


Figure 12.20:

This is basically a detector problem, so instead we will consider averaging θ first. In fact it is obvious what the result of such an average is, it must be the previous result, but shifted from $\vec{k} = 0$ to $\vec{k} = \vec{k}_0$, that is

$$\begin{aligned} S_{surface}(k) &\propto \frac{1}{|\vec{k} - \vec{k}_0|^{d+1}} \\ S_{line}(k) &\propto \frac{1}{|\vec{k} - \vec{k}_0|^{d-1}} \end{aligned} \quad (12.39)$$

If a further average is done over crystallite orientation, we can take these expressions and angularly average them over the angle between \vec{k} and \vec{k}_0 . Clearly after averaging we will have

$$\bar{S} = S\left(\frac{|\vec{k}| - |\vec{k}_0|}{|\vec{k}_0|}\right) \equiv S(\Delta k) \quad (12.40)$$

First, let

$$\phi = \beta - \alpha \quad (12.41)$$

so that

$$(\vec{k} - \vec{k}_0)^2 = k^2 - 2kk_0 \cos \phi + k_0^2$$

or, on using

$$k \equiv k_0(\Delta k + 1) \quad (12.42)$$

we have

$$(\vec{k} - \vec{k}_0)^2 = k_0^2 \{(\Delta k)^2 + 2(1 - \cos \phi)(\Delta k + 1)\} \quad (12.43)$$

Hence in two dimensions, for a surface, we have

$$\bar{S}(\Delta k) = \frac{1}{2\pi k_0^3} \int_0^{2\pi} d\phi \frac{1}{|(\Delta k)^2 + 2(1 - \cos \phi)(\Delta k + 1)|^{3/2}} \quad (12.44)$$

which is some weird integral. We will work out its asymptotic limits $\Delta k \gg 1$ and $\Delta k \ll 1$.

Firstly, if $\Delta k \gg 1$, then

$$\bar{S}(\Delta k) = \frac{1}{2\pi k_0^3} \frac{1}{(\Delta k)^3} \int_0^{2\pi} d\phi + \dots$$

or (in arbitrary d)

$$\bar{S}(\Delta k) \sim \frac{1}{(\Delta k)^{d+1}}, \Delta k \gg 1 \quad (12.45)$$

If $\Delta k \ll 1$, consider

$$\bar{S}(\Delta k \rightarrow 0) \stackrel{?}{=} \frac{1}{2\pi k_0^3} \int_0^{2\pi} d\phi \frac{1}{|2(1 - \cos \phi)|^{3/2}}$$

which *diverges* due to the $\phi = 0$ (and $\phi = \pi$) contribution. Oops. So we had better keep a little Δk . For convenience, we will also expand ϕ around 0.

$$\bar{S}(\Delta k) \approx \frac{1}{2\pi k_0^3} \int_0^{2\pi} d\phi \frac{1}{|(\Delta k)^2 + \phi^2|^{3/2}} + \dots$$

Now let $u = \phi/\Delta k$

$$\bar{S}(\Delta k) = \frac{1}{2\pi k_0^3} \frac{1}{(\Delta k)^2} \int_0^{2\pi/\Delta k} du \frac{1}{|1 + u^2|^{3/2}}$$

so,

$$\bar{S}(\Delta k) \sim \frac{1}{(\Delta k)^2} \text{ as } \Delta k \rightarrow 0 \quad (12.46)$$

for *all* dimensions d .

If one does this in d dimensions with

$$S = \frac{1}{|\vec{k} - \vec{k}_0|^\gamma} \quad (12.47)$$

then

$$\bar{S} = \begin{cases} \frac{1}{(k-k_0)^{\gamma-(d-1)}}, & \frac{k-k_0}{k_0} \ll 1 \\ \frac{1}{(k-k_0)^\gamma}, & \frac{k-k_0}{k_0} \gg 1 \end{cases} \quad (12.48)$$

If $\gamma - (d-1) > 0$. For the case $\gamma = d-1$, one obtains

$$\bar{S} \sim \begin{cases} -\ln(k-k_0), & \frac{k-k_0}{k_0} \ll 1 \\ \frac{1}{(k-k_0)^{d-1}}, & \frac{k-k_0}{k_0} \gg 1 \end{cases} \quad (12.49)$$

It should be noted that the actual scattering is given by a weird boring integral

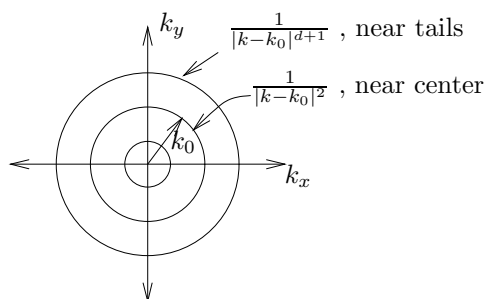


Figure 12.21:

like Eq. 12.44, which is pretty close to an elliptical integral.

Finally, note the big mess of power laws, and we have not considered diffuseness or roughness! Everything is just geometry.

Chapter 13

Fluctuations and Crystalline Order

Usually it is safer to apply Goldstone's theorem by “re-deriving” it for whatever system one is interested in. However, to give an interesting and concrete example, consider crystalline order in a solid. The density ρ clearly depends on

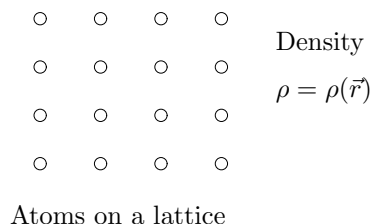


Figure 13.1:

positions \vec{r} , as shown, in a regular crystal. A crystal, although it looks symmetric in a common sense way, actually has very little symmetry compared to a liquid or a vapour.

A liquid or a vapour are isotropic in *all* directions of space and traditionally invariant in *all* directions. A crystalline solid has a very restricted symmetry. In figure 13.1, the system is invariant under (e.g.)

$$\vec{r} \rightarrow \vec{r} + r_0(n\hat{x} + m\hat{y}) \quad (13.1)$$

where n and m are integers (positive or negative), and r_0 is the distance between atoms. Clearly then a crystal has very low symmetry, and—indeed—breaks the translational and rotational invariance of space!

This means there should be a broken symmetry variable, associated with this, which satisfies Goldstone's theorem. Clearly this variable is associated with the \vec{r} dependence of

$$\rho(\vec{r}) \quad (13.2)$$

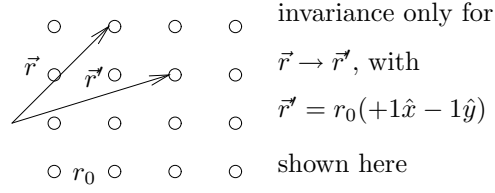


Figure 13.2:

Let us then introduce a variable $\vec{u}(\vec{r})$ which describes the local variations in the density. That is, due to some fluctuation,

$$\vec{r} \rightarrow \vec{r} + \vec{u}(\vec{r})$$

and

$$\rho(\vec{r}) \rightarrow \rho(\vec{r} + \vec{u}(\vec{r})) \quad (13.3)$$

where $\vec{u}(\vec{r})$ is the local displacement vector as shown in fig. 13.3.

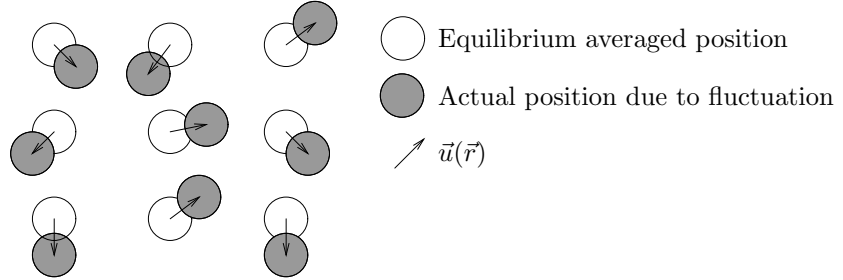


Figure 13.3: Jiggled atoms due to thermal fluctuations

Aside: (What does it mean to be translationally invariant

$$\rho(\vec{r}) \rightarrow \rho(\vec{r} + \vec{u})$$

Local variable corresponding to broken symmetry

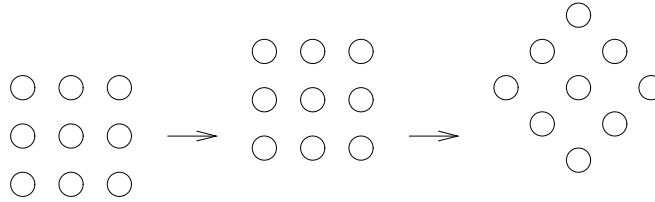


Figure 13.4:

$$\vec{u}(\vec{r})$$

A definite value for $\langle u(\vec{r}) \rangle = 0$, say, breaks space symmetry (translational and rotational invariance).)

By construction

$$\langle \vec{u}(\vec{r}) \rangle = 0 \quad (13.4)$$

In Fourier space, if all directions of fluctuation are equally likely,

$$\langle \hat{u}(\vec{k}) \hat{u}(\vec{k}') \rangle \propto (2\pi)^d \delta(\vec{k} + \vec{k}') \left(\frac{1}{k^2} \right) \quad (13.5)$$

It is a little hard to see what this does to a density, since \vec{u} appears inside $\rho(\vec{r} + \vec{u}(\vec{r}))$. The thing to do is to expand in a Fourier series, but to get an idea, consider $d = 1$. Atoms don't look like cosines, but this gives the sense of it.

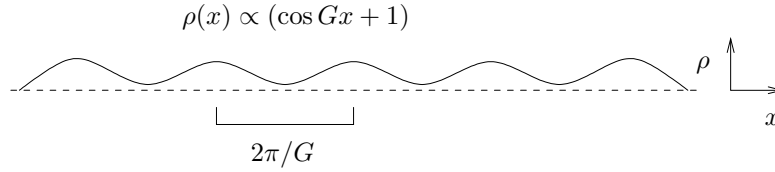


Figure 13.5: Rough depiction of a one-dimensional variation of density via cosine

More precisely,

$$\rho(x) = \sum_G e^{iGx} \hat{a}_G \quad (13.6)$$

where there might be several wavenumbers G giving the density, and \hat{a}_G is a Fourier coefficient giving the atomic form. In arbitrary dimension we write

$$\rho(x) = \sum_G e^{iGx} \hat{a}_G$$

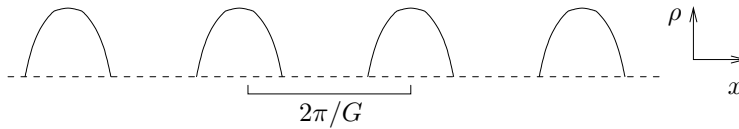


Figure 13.6: Better depiction of a one-dimensional variation. But the cosine is essentially the right idea.

$$\rho(\vec{r}) = \sum_{\vec{G}} e^{i\vec{G} \cdot \vec{r}} \hat{a}_{\vec{G}} \quad (13.7)$$

where \vec{G} are called the reciprocal lattice vectors. Then we have a fluctuation represented by

$$\rho(\vec{r} + \vec{u}(\vec{r})) = \sum_{\vec{G}} e^{i\vec{G} \cdot (\vec{r} + \vec{u}(\vec{r}))} \hat{a}_{\vec{G}} \quad (13.8)$$

$$\langle \rho(\vec{r} + \vec{u}(\vec{r})) \rangle = \sum_{\vec{G}} e^{i\vec{G} \cdot \vec{r}} \hat{a}_{\vec{G}} \langle e^{i\vec{G} \cdot \vec{u}(\vec{r})} \rangle \quad (13.9)$$

To calculate the average we use the fact that $\vec{u}(\vec{r})$ is a Gaussian variable. Therefore

$$\begin{aligned} \langle e^{i\vec{G} \cdot \vec{u}} \rangle &= e^{-\frac{1}{2} \langle \vec{G} \cdot \vec{u} \rangle^2} \\ &= e^{-\frac{1}{2} \vec{G} \vec{G} : \langle \vec{u} \vec{u} \rangle} \end{aligned} \quad (13.10)$$

But \vec{u} is only correlated to \vec{u} if they are in the same direction, we expect. That is

$$\langle u_\alpha u_\beta \rangle \propto \delta_{\alpha\beta}$$

so

$$\langle e^{-i\vec{G} \cdot \vec{u}} \rangle \approx e^{\frac{1}{2} G^2 \langle u^2 \rangle}$$

But, from Eq. 13.5 we have,

$$\langle u^2 \rangle = \int \frac{d^d k}{(2\pi)^d} \frac{1}{k^2}$$

or

$$\langle u^2 \rangle = \begin{cases} \text{Const.} & d = 3 \\ \ln L & d = 2 \\ L & d = 1 \end{cases} \quad (13.11)$$

So in $d = 3$

$$\langle \rho \rangle = \text{Const.} \sum_{\vec{G}} e^{i\vec{G} \cdot \vec{r}} \hat{a}_{\vec{G}} \quad (13.12)$$

where the $\text{Const.} = \exp -\frac{1}{2} G^2 (\text{Const.})$ is called the deBye–Waller factor. In $d = 2$ we have

$$\langle \rho \rangle = \sum_{\vec{G}} e^{i\vec{G} \cdot \vec{r}} \hat{a}_{\vec{G}} e^{\underbrace{-\frac{1}{2} G^2 \ln L}_{\frac{1}{L} \text{const.} G^2 / 2}} \quad (13.13)$$

which is a marginal case. In $d = 1$ we have

$$\langle \rho \rangle = \sum_{\vec{G}} e^{i\vec{G} \cdot \vec{r}} \hat{a}_{\vec{G}} e^{-\frac{1}{2} G^2 L} \quad (13.14)$$

which is a large effect, much greater than the periodicity as $L \rightarrow \infty$. This is the content of the original Landau–Peierls theorem, which states that a one dimensional variation in density, $\rho(x)$ (whether crystal structure or phase coexistence) is impossible: The only G which can survive in Eq. 13.14 is $G = 0$ as $L \rightarrow \infty$ which gives $\rho(x) \rightarrow \text{Constant}$, due to $\vec{u}(\vec{r})$ fluctuations.

It is more revealing to do these things for the *scattering intensity* $I(k)$, where

$$I(k) \propto |\hat{\rho}_k|^2 \quad (13.15)$$

(Scattering is from coherent density fluctuations). From Eq. 13.7,

$$\begin{aligned}\hat{\rho}_k &= \int d^d r e^{-i\vec{k}\cdot\vec{r}} \left\{ \sum_{\vec{G}} e^{i\vec{G}\cdot\vec{r}} \hat{a}_{\vec{G}} \right\} \\ &= \sum_{\vec{G}} \hat{a}_{\vec{G}} \int d^d r e^{-i(\vec{k}-\vec{G})\cdot\vec{r}}\end{aligned}$$

or

$$\hat{\rho}_k = \sum_{\vec{G}} \hat{a}_{\vec{G}} \delta(\vec{k} - \vec{G}) \quad (13.16)$$

so,

$$\hat{\rho}_k^* = \sum_{\vec{G}} \hat{a}_{\vec{G}}^* \delta(\vec{k} - \vec{G})$$

and

$$|\hat{\rho}_k|^2 = \sum_{\vec{G}, \vec{G}'} \hat{a}_{\vec{G}} \hat{a}_{\vec{G}'}^* \delta(\vec{k} - \vec{G}) \delta(\vec{k} - \vec{G}')$$

or,

$$|\hat{\rho}_k|^2 \propto \sum_{\vec{G}} |\hat{a}_{\vec{G}}|^2 \delta(\vec{k} - \vec{G}) \quad (13.17)$$

where $|\hat{a}_{\vec{G}}|^2$ is an atomic form factor. This gives well defined peaks where $\vec{k} = \vec{G}$ as shown in fig. 13.7. When thermal fluctuations are present

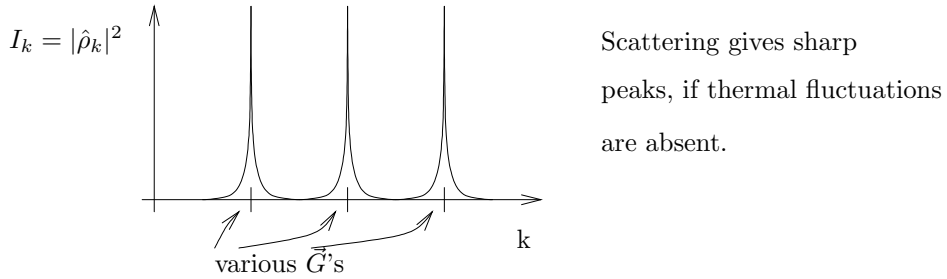


Figure 13.7:

$$\hat{\rho}_k = \int d^d \vec{r} e^{-i\vec{k}\cdot\vec{r}} \left(\sum_{\vec{G}} e^{i\vec{G}\cdot\vec{r}} \hat{a}_{\vec{G}} e^{i\vec{G}\cdot\vec{u}(\vec{r})} \right)$$

or,

$$\hat{\rho}_k = \sum_{\vec{G}} \hat{a}_{\vec{G}} \int d^d \vec{r} e^{-i(\vec{k}-\vec{G})\cdot\vec{r}} e^{i\vec{G}\cdot\vec{u}(\vec{r})} \quad (13.18)$$

so

$$\hat{\rho}_k^* = \sum_{\vec{G}'} \hat{a}_{\vec{G}'}^* \int d^d \vec{r}' e^{i(\vec{k}-\vec{G}')\cdot\vec{r}'} e^{-i\vec{G}'\cdot\vec{u}(\vec{r}')}$$

Hence

$$\langle |\hat{\rho}_k|^2 \rangle = \sum_{\vec{G}, \vec{G}'} \hat{a}_G \hat{a}_G^* \int d^d \vec{r} \int d^d \vec{r}' e^{-i(\vec{k}-\vec{G}) \cdot \vec{r}} e^{i(\vec{k}-\vec{G}') \cdot \vec{r}'} \langle e^{i\vec{G} \cdot u(\vec{r}) - i\vec{G}' \cdot u(\vec{r}')} \rangle \quad (13.19)$$

This is the same integral as on page so we can simply use that

$$\langle e^{i\vec{G} \cdot u(\vec{r}) - i\vec{G}' \cdot u(\vec{r}')} \rangle = f(\vec{r} - \vec{r}') \quad (13.20)$$

and then,

$$\langle |\hat{\rho}_k|^2 \rangle = \sum_{\vec{G}, \vec{G}'} \hat{a}_G \hat{a}_G^* \int d^d \vec{r} e^{-i(\vec{k}-\vec{G}) \cdot \vec{r}} (2\pi)^d \delta(\vec{G} - \vec{G}') f(\vec{r}) \quad (13.21)$$

where

$$\begin{aligned} f(\vec{r}) &= \langle e^{i\vec{G} \cdot u(\vec{r}) - i\vec{G} \cdot u(0)} \rangle \\ &= e^{-\frac{G^2}{2} \langle (u(\vec{r}) - u(0))^2 \rangle} \\ &= \text{Const.} e^{-G^2 \langle u(\vec{r}) u(0) \rangle} \end{aligned} \quad (13.22)$$

so,

$$\langle |\hat{\rho}_k|^2 \rangle = \text{const.} \sum_{\vec{G}} |\hat{a}_{\vec{G}}|^2 \int d^d r e^{-i(\vec{k}-\vec{G}) \cdot \vec{r}} e^{-G^2 \langle u(\vec{r}) u(0) \rangle} \quad (13.23)$$

The correlation function $\langle u(\vec{r}) u(0) \rangle$ can be worked out as before (knowing $\langle |\hat{u}(k)|^2 \rangle \sim \frac{1}{k^2}$),

$$\langle u(\vec{r}) u(0) \rangle = \begin{cases} \text{Const.}, & d = 3 \\ \ln r, & d = 2 \\ r, & d = 1 \end{cases} \quad (13.24)$$

In $d = 3$,

$$\langle |\hat{\rho}_k|^2 \rangle = \text{Const.} \sum_{\vec{G}} |\hat{a}_G|^2 \delta(\vec{k} - \vec{G})$$

with the same delta function peaks.

In $d = 2$, the marginal case,

$$\begin{aligned} \langle |\hat{\rho}_k|^2 \rangle &= \text{Const.} \sum_{\vec{G}} |\hat{a}_G|^2 \int d^d r e^{-i(\vec{k}-\vec{G}) \cdot \vec{r}} e^{-\text{Const.} \ln r} \\ &= \text{Const.} \sum_{\vec{G}} |\hat{a}_G|^2 \underbrace{\int d^d r \frac{e^{-i(\vec{k}-\vec{G}) \cdot \vec{r}}}{r^{\text{Const.}}}}_{\text{this is the fourier transform of a power}} \\ &= \text{Const.} \sum_{\vec{G}} |\hat{a}_G|^2 \frac{1}{|\vec{k} - \vec{G}|^{d-\text{Const.}}} \end{aligned}$$

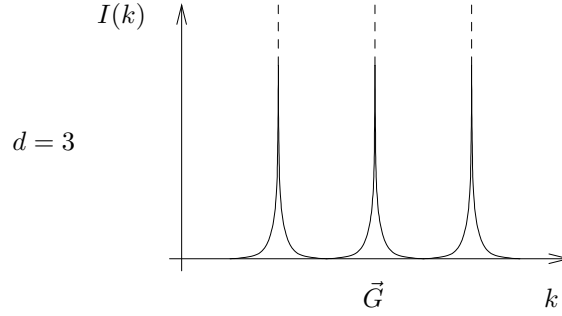


Figure 13.8: Delta Function Singularities

or, letting $(d - Const.) \equiv 2 - \eta$, we have

$$\langle |\hat{\rho}_k|^2 \rangle = Const. \sum_{\vec{G}} |\hat{a}_G|^2 \frac{1}{|\vec{k} - \vec{G}|^{2-\eta}} \quad (13.25)$$

in $d = 2$. The thermal fluctuations change the delta function singularities to power-law singularities! Sometimes they are called η -line shapes (eta-line shapes).

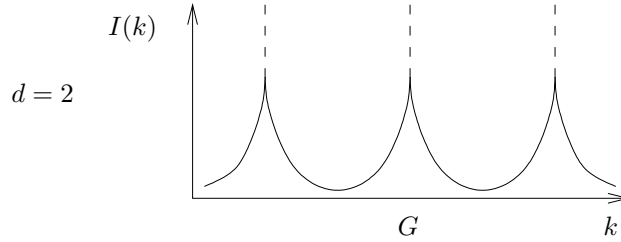


Figure 13.9: Power Law Singularities

In $d = 1$, we have

$$\langle |\hat{\rho}_k|^2 \rangle = Const. \sum_{\vec{G}} |\hat{a}_G|^2 \underbrace{\int d^d r e^{-i(\vec{k}-\vec{G}) \cdot \vec{r}} e^{-(Const.)r}}_{\substack{\text{this is the Fourier} \\ \text{transform of an} \\ \text{exponential. But} \\ \int d^d r e^{i\vec{k}' \cdot \vec{r}} e^{-r/\xi} \propto \frac{1}{(k')^2 + \xi^{-2}} \\ \text{for small } k'}}$$

So, for $d = 1$,

$$\langle |\hat{\rho}_k|^2 \rangle = Const. \sum_{\vec{G}} |\hat{a}_G|^2 \frac{1}{(\vec{k} - \vec{G})^2 + \xi^{-2}} \quad (13.26)$$

where ξ^{-2} is a constant. The thermal fluctuations remove all singularities as $\vec{k} = \vec{G}$ now! This is another sign that $\rho(x)$ is impossible, and there can be no

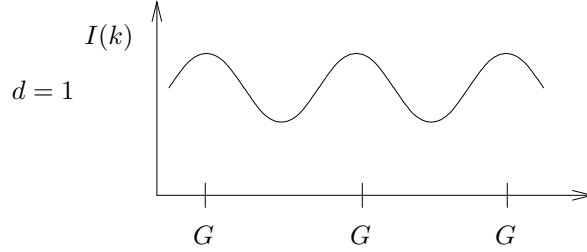


Figure 13.10: No Singularities

$1 - d$ crystalline order (or $1 - d$ phase coexistence).

There are two other things associated with broken symmetry which I will not discuss, but you can look up for yourself.

1. Broken symmetries are usually associated with a new (gapless, soft, or acoustic) mode. For surfaces, these propagating modes are capillary waves. For crystalline solids, they are shear-wave phonons.
2. Broken symmetries are associated with a “rigidity”. For crystalline solids, this is simply their rigidity: it takes a lot of atoms to break a symmetry of nature, and they all have to work together.

Chapter 14

Ising Model

There are many phenomena where a large number of constituents interact to give some global behavior. A magnet, for example, is composed of many interacting dipoles. A liquid is made up of many atoms. A computer program has many interacting commands. The economy, the psychology of road rage, plate tectonics, and many other things all involve phenomena where the global behavior is due to interactions between a large number of entities. One way or another, almost all such systems end up being posed as an "Ising model" of some kind. It is so common, that it is almost a joke.

The Ising model has a large number of spins which are individually in microscopic states $+1$ or -1 . The value of a spin at a site is determined by the spin's (usually short-range) interaction with the other spins in the model. Since I have explained it so vaguely, the first, and major import, of the Ising model is evident: It is the most simple nontrivial model of an interacting system. (The only simpler model would have only one possible microscopic state for each spin. I leave it as an exercise for the reader to find the exact solution of this trivial model.) Because of this, whenever one strips away the complexities of some interesting problem down to the bone, what is left is usually an Ising model of some form.

In condensed-matter physics, a second reason has been found to use the simplified-to-the-bone Ising model. For many problems where one is interested in large length scale phenomena, such as continuous phase transitions, the additional complexities can be shown to be unimportant. This is called universality. This idea has been so successful in the study of critical phenomena, that it has been tried (or simply assumed to be true) in many other contexts.

For now, we will think of the Ising model as applied to anisotropic magnets. The magnet is made up of individual dipoles which can orient up, corresponding to spin $+1$ or down corresponding to spin -1 . We neglect tilting of the spins,

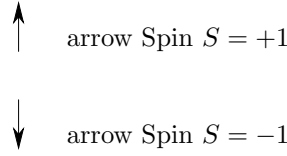
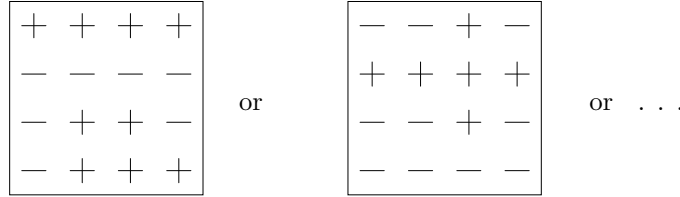


Figure 14.1:

such as \nwarrow , or \nearrow , or \searrow , and so on. For concreteness, consider dimension $d = 2$, putting all the spins on a square lattice, where there are $i = 1, 2, 3, \dots, N$ spins. In the partition function, the sum over states is

Figure 14.2: Possible configurations for a $N = 4^2 = 16$ system.

$$\begin{aligned} \sum_{\{\text{States}\}} &= \sum_{s_1=-1}^{+1} \sum_{s_2=-1}^{+1} \dots \sum_{s_N=-1}^{+1} \\ &= \prod_{i=1}^N \sum_{s_i=-1}^{+1} \end{aligned}$$

where each sum has only two terms, of course. The energy in the $\exp - E_{\text{State}}/k_B T$ now must be prescribed. Since we want to model a magnet, where all the spins are $+1$ or -1 , by and large, we make the spin of site i want to oriented the same as its neighbors. That is, for site i

$$\begin{aligned} \text{Energy of Site } i &= \sum_{\text{neighbors of } i} (S_i - S_{\text{neighbors}})^2 \\ &= S_i^2 - 2 \sum_{\text{neighbors of } i} S_i S_{\text{neigh.}} + S_{\text{neigh.}}^2 \end{aligned}$$

But $S_i^2 = 1$ always, so up to a constant (additive and multiplicative)

$$\text{Energy of Site } i = - \sum_{\text{neighbors of } i} S_i S_{\text{neighbors of } i}$$

On a square lattice, the neighbors of “ i ” are the spin surrounding it (fig. 14.3).

As said before, the idea is that the microscopic constituents -the spins- don’t know how big the system is, so they interact locally. The choice of the 4 spins

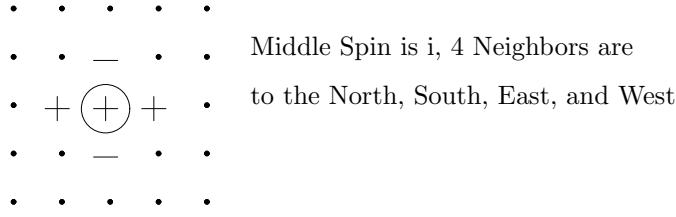


Figure 14.3: Four nearest neighbors on a square lattice.

above, below and to the *right* and *left* as “nearest neighbors” is a convention. The spins on the diagonals (NW, NE, SE, SW) are called “next-nearest neighbors”. It doesn’t actually matter how many are included, so long as the interaction is local. The total energy for all spins is

$$E_{State} = -J \sum_{\langle ij \rangle} S_i S_j$$

where J is a positive interaction constant, and $\langle ij \rangle$ is a sum over all i and $j = 1, 2, \dots, N$, provided i and j are nearest neighbors, with no double counting. Note

$$\sum_{\langle ij \rangle} = \underbrace{\sum_{i=1}^N \sum_{j=1}^N}_{\text{neighbors no double counting}} = N \times \frac{4}{2}$$

If there are “ q ” nearest neighbors, corresponding to a different interaction, or a different lattice (say hexagonal), or another dimension of space

$$\sum_{\langle ij \rangle} = \frac{q}{2} N$$

So the largest (negative) energy is $E = -\frac{qJ}{2} N$. If there is constant external magnetic field H , which tends to align spins, the total energy is

$$E_{State} = -J \sum_{\langle ij \rangle} S_i S_j - H \sum_i S_i$$

and the partition function is

$$\begin{aligned} Z &= \prod_{i=1}^N \sum_{S_i = -1}^{+1} e^{(\frac{J}{k_B T}) \sum_{\langle jk \rangle} S_j S_k + (\frac{H}{k_B T}) \sum_i S_i} \\ &= Z(N, \frac{J}{k_B T}, \frac{H}{k_B T}) \end{aligned}$$

As described above, the sum over q nearest neighbors is $q = 4$ for a two-dimensional square lattice. For three dimensions, $q = 6$ on a simple cubic lattice. The one-dimensional Ising model has $q = 2$.

The one-dimensional model was solved by Ising for his doctoral thesis (we will solve it below). The two-dimensional model was solved by Onsager, for $H = 0$, in one of the premiere works of mathematical physics. We will not give that solution. you can read it in Landau and Lifshitz's "Statistical Physics". They give a readable, physical, but still difficult treatment. The three-dimensional Ising model has not been solved.

The most interesting result of the Ising model is that it exhibits broken symmetry for $H = 0$. In the absence of a field, the energy of a state, and the partition function, have an exact symmetry

$$S_i \leftrightarrow -S_i$$

(The group is Z_2 .) Hence (?)

$$\langle S_i \rangle \leftrightarrow -\langle S_i \rangle$$

and

$$\langle S_i \rangle \stackrel{?}{=} 0$$

In fact, for dimension $d = 2$ and 3 this is not true. The exact result in $d = 2$ is surprisingly simple: The magnetization/spin is

$$m = \langle S_i \rangle = \begin{cases} (1 - \operatorname{cosech}^2 \frac{2J}{k_B T})^{1/8}, & T < T_c \\ 0, & T > T_c \end{cases}$$

The critical temperature $T_c \approx 2.269J/k_B$ is where $(1 - \operatorname{cosech}^2 \frac{2J}{k_B T})^{1/8}$ vanishes. A rough and ready understanding of this transition is easy to get. Consider the

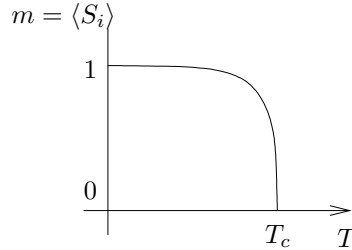


Figure 14.4: Two-dimensional square lattice Ising model magnetization per spin

Helmholtz free energy

$$F = E - TS$$

where S is *now* the entropy. At low temperatures, F is minimized by minimizing E . The largest negative value E has is $-\frac{q}{2}JN$, when all spins are aligned $+1$ or -1 . So at low temperatures (and rigorously at $T = 0$), $\langle S_i \rangle \rightarrow +1$ or $\langle S_i \rangle \rightarrow -1$. At high temperatures, entropy maximization minimizes F . Clearly entropy is maximized by having the spins helter-skelter random: $\langle S_i \rangle = 0$.

The transition temperature between $\langle S_i \rangle = 1$ and $\langle S_i \rangle = 0$ should be when the temperature is comparable to the energy per spin $k_B T = \mathcal{O}(\frac{q}{2}J)$ which is

in the right ball park. There is no hand-waving explanation for the abrupt increase of the magnetization per spin $m = \langle S_i \rangle$ at T_c . Note it is *not* like a paramagnet which shows a gradual decrease in magnetization with increasing temperature (in the presence of an external field H). One can think of it (or

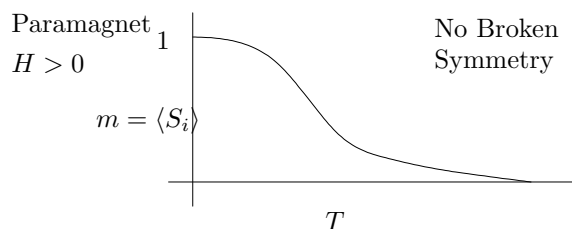


Figure 14.5:

at least remember it) this way: At T_c , one breaks a fundamental symmetry of the system. In some sense, one breaks a symmetry of nature. One can't break such a symmetry in a tentative wishy-washy way. It requires a firm, determined hand. That is, an abrupt change.

The phase diagram for the Ising model (or indeed for any magnet) look like fig. 14.6. It is very convenient that H is so intimately related with the broken

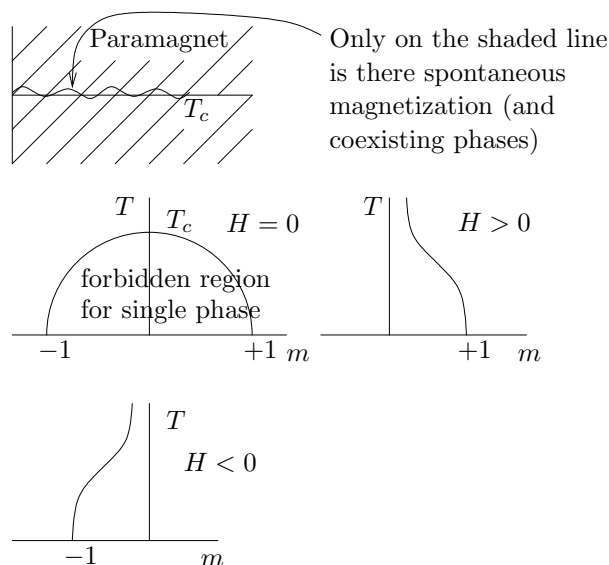


Figure 14.6:

symmetry of the Ising model. Hence it is very easy to understand that broken symmetry is for $H = 0$, not $H \neq 0$. Things are more subtle in the lattice-gas version of the Ising model.

14.1 Lattice–Gas Model

In the lattice–gas model, one still has a lattice labeled $i = 1, 2, \dots, N$, but now there is a concentration variable $c_i = 0$ or 1 at each site. The most simple model of interaction is the lattice–gas energy

$$E^{LG} = -\vartheta \sum_{\langle ij \rangle} c_i c_j - \mu \sum_i c_i$$

where ϑ is an interaction constant and μ is the constant chemical potential.

One can think of this as a model of a binary alloy (where C_i is the local concentration of one of the two phases), or a liquid–vapour system (where C_i is the local density).

This is very similar to the Ising model

$$E^{Ising} = -J \sum_{\langle ij \rangle} S_i S_j - H \sum_i S_i$$

To show they are exactly the same, let

$$S_i = 2C_i - 1$$

then

$$\begin{aligned} E^{Ising} &= -J \sum_{\langle ij \rangle} (2C_i - 1)(2C_j - 1) - H \sum_i (2C_i - 1) \\ &= -4J \sum_{\langle ij \rangle} C_i C_j + 4J \sum_{\langle ij \rangle} C_i + Const. - 2H \sum_i C_i + Const. \\ &= -4J \sum_{\langle ij \rangle} C_i C_j + 4J \frac{q}{2} \sum_i C_i - 2H \sum_i C_i + Const. \end{aligned}$$

And

$$E^{Ising} = -4J \sum_{\langle ij \rangle} C_i C_j - 2(H - J_q) \sum_i C_i + Const.$$

Since the additive constant is unimportant, this is the same as E^{LG} , provided

$$\vartheta = 4J$$

and

$$\mu = 2(H - J_q)$$

Note that the phase transition here is not at some convenient point $\mu = 0$, it is at $H = 0$, which is the awkward point $\mu = 2J_q$. This is one reason why the liquid–gas transition is sometimes said — incorrectly — to involve no change of symmetry (Z_2 group) as the Ising model. But the external field must have a carefully tuned value to get that possibility.

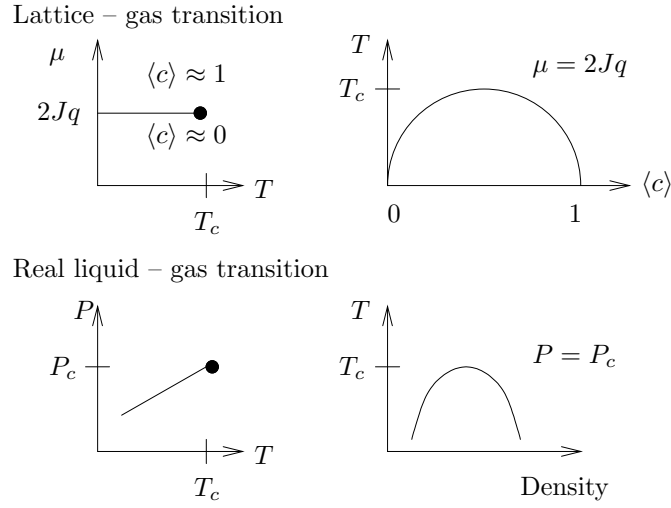


Figure 14.7:

14.2 One-dimensional Ising Model

It is easy to solve the one-dimensional Ising model. Of course, we already know that it cannot exhibit phase coexistence, from the theorem we proved earlier. Hence we expect

$$\langle S_i \rangle = m = 0$$

for all $T > 0$, where m is the magnetization per spin. The result and the method are quite useful, though.

$$Z = \prod_{i=1}^N \sum_{S_i=-1}^{+1} e^{\frac{J}{k_B T} \sum_{j=1}^N S_j S_{j+1} + \frac{H}{k_B T} \sum_{j=1}^N S_j}$$

Note we have simplified the $\sum_{\langle ij \rangle}$ for one dimension. We need, however, a prescription for the spin S_{N+1} . We will use periodic boundary conditions so that $S_{N+1} \equiv S_1$. This is called the Ising ring. You can also solve it, with a little

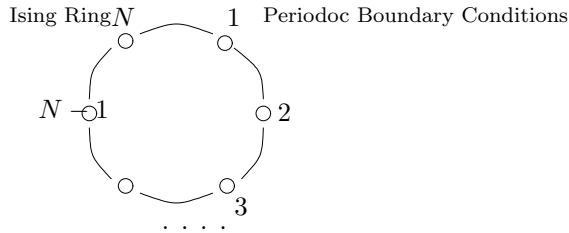


Figure 14.8:

more work, with other boundary conditions. Let the *coupling constant*

$$k \equiv J/k_B T$$

and

$$h = H/k_B T$$

be the external field. Then

$$Z(k, h, N) = \prod_{i=1}^N \sum_{S_i=-1}^{+1} e^{\sum_{j=1}^N (k S_j S_{j+1} + h S_j)}$$

or, more symmetrically,

$$\begin{aligned} Z &= \prod_{i=1}^N \sum_{S_i=-1}^{+1} e^{\sum_j [k S_j S_{j+1} + \frac{1}{2} h (S_j + S_{j+1})]} \\ &= \prod_{i=1}^N \sum_{S_i=-1}^{+1} \prod_{j=1}^N e^{k S_j S_{j+1} + \frac{1}{2} h (S_j + S_{j+1})} \end{aligned}$$

Let matrix elements of P be defined by

$$\langle S_j | P | S_{j+1} \rangle \equiv e^{k S_j S_{j+1} + \frac{1}{2} h (S_j + S_{j+1})}$$

There are only four elements

$$\begin{aligned} \langle +1 | P | +1 \rangle &= e^{k+h} \\ \langle -1 | P | -1 \rangle &= e^{k-h} \\ \langle +1 | P | -1 \rangle &= e^{-k} \\ \langle -1 | P | +1 \rangle &= e^{-k} \end{aligned}$$

Or, in matrix form,

$$P = \begin{bmatrix} e^{k+h} & e^{-k} \\ e^{-k} & e^{k-h} \end{bmatrix}$$

The partition function is

$$\begin{aligned} Z &= \prod_{i=1}^N \sum_{S_i=-1}^{+1} \langle S_1 | P | S_2 \rangle \langle S_2 | P | S_3 \rangle \dots \langle S_N | P | S_1 \rangle \\ &= \sum_{S_1=-1}^{+1} \langle S_1 | P (\sum_{S_2=-1}^{+1} | S_2 \rangle \langle S_2 |) P (\sum_{S_3=-1}^{+1} | S_3 \rangle \langle S_3 |) \dots P | S_1 \rangle \end{aligned}$$

But, clearly, the identity operator is

$$1 \equiv \sum_S | S \rangle \langle S |$$

so,

$$\begin{aligned} Z &= \sum_{S_1=-1}^{+1} \langle S_1 | P^N | S_1 \rangle \\ &= \text{Trace}(P^N) \end{aligned}$$

Say the eigenvalues of P are λ_+ and λ_- . Then,

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

Finally, say $\lambda_+ > \lambda_-$. Hence as $N \rightarrow \infty$

$$Z = e^{-F/k_B T} = \lambda_+^N$$

and

$$F = -Nk_B T \ln \lambda_+$$

It remains only to find the eigenvalues of P .

We get the eigenvalues from

$$\det(\lambda - P) = 0$$

This gives

$$(\lambda - e^{k+h})(\lambda - e^{k-h}) - e^{-2k} = 0$$

or,

$$\lambda^2 - 2\lambda \underbrace{\left(\frac{e^{k+h} + e^{k-h}}{2} \right)}_{e^k \cosh h} + e^{2k} - e^{-2k} = 0$$

The solutions are

$$\begin{aligned} \lambda_{\pm} &= \frac{2e^k \cosh h \pm \sqrt{4e^{2k} \cosh^2 h - 4e^{2k} + 4e^{-2k}}}{2} \\ &= e^k [\cosh h \pm \sqrt{\cosh^2 h - 1 + e^{-4k}}] \end{aligned}$$

So,

$$\lambda_+ = e^k (\cosh h + \sqrt{\sinh^2 h + e^{-4k}})$$

is the larger eigenvalue. The free energy is

$$F = -Nk_B T [k + \ln (\cosh h + \sqrt{\sinh^2 h + e^{-4k}})]$$

where $k = J/k_B T$, and $h = H/k_B T$.

The magnetization per spin is

$$m = -\frac{1}{N} \frac{\partial F}{\partial H} = -\frac{1}{Nk_B T} \frac{\partial F}{\partial h}$$

Taking the derivative and simplifying, one obtains

$$m(h, k) = \frac{\sinh h}{\sqrt{\sinh^2 h + e^{-4k}}}$$

As anticipated, $m(h = 0, k) = 0$. There is magnetization for $h \neq 0$, although there is no broken symmetry of course. There is a phase transition at $T = 0$.

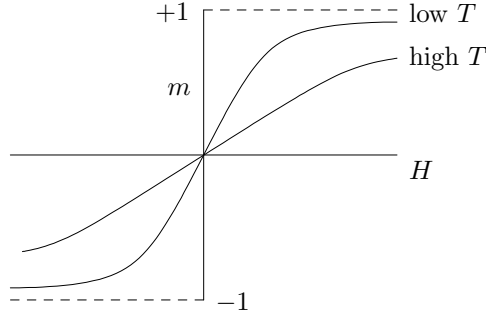


Figure 14.9: Magnetization of 1-d Ising model

14.3 Mean-Field Theory

There is a very useful approximate treatment of the Ising model. We will now work out the most simple such “mean-field” theory, due in this case to Bragg and Williams.

The idea is that each spin is influenced by the local magnetization, or equivalently local field, of its surrounding neighbors. In a magnetized, or demagnetized state, that local field does not fluctuate very much. Hence, the idea is to replace the local field, determined by a local magnetization, with an average field, determined by the average magnetization. “Mean” field means average field, not nasty field.

It is easy to solve the partition function if there is only a field, and no interactions. This is a (very simple) paramagnet. Its energy is

$$E_{State} = -H \sum_i S_i$$

and,

$$\begin{aligned} Z &= \prod_i \sum_{S_i=\pm 1} e^{\frac{H}{k_B T} \sum_j S_j} \\ &= \left(\sum_{S_1=\pm 1} e^{\frac{H}{k_B T} S_1} \right) \dots \left(\sum_{S_N=\pm 1} e^{\frac{H}{k_B T} S_N} \right) \\ &= \left(\sum_{S=\pm 1} e^{\frac{H}{k_B T} S} \right)^N \end{aligned}$$

so,

$$Z = \left(2 \cosh \frac{H}{k_B T} \right)^N$$

and

$$F = -N k_B T \ln 2 \cosh \frac{H}{k_B T}$$

so since,

$$m = -\frac{1}{N} \partial F / \partial H$$

then

$$m = \tanh H / k_B T$$

For the interacting Ising model we need to find the effective local field H_i at

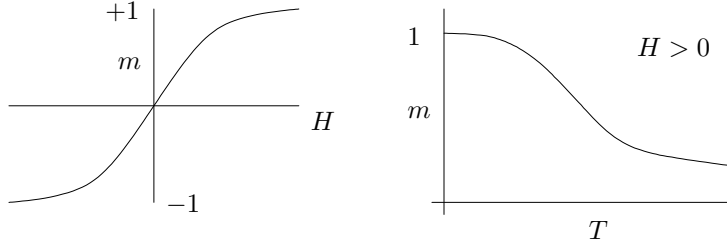


Figure 14.10: Noninteracting paramagnet solution

site i . The Ising model is

$$E_{State} = -J \sum_{\langle ij \rangle} S_i S_j - H \sum_i S_i$$

we want to replace this — in the partition function weighting — with

$$E_{State} = - \sum_i H_i S_i$$

Clearly

$$-\frac{\partial E}{\partial S_i} = H_i$$

Thus, from the Ising model

$$\begin{aligned} -\frac{\partial E}{\partial S_i} &= J \sum_{\langle jk \rangle} (S_j \delta_{k,i} + \delta_{j,i} S_k) + H \\ &= J \cdot 2 \cdot \sum_{\substack{j \text{ neighbors of } i \\ \text{(no double counting)}}} S_j + H \end{aligned}$$

Now let's get rid of the 2 and the double-counting restriction. Then

$$H_i = J \sum_{j \text{ neighbors of } i} S_j + H$$

is the effective local field. To do mean-field theory we let

$$H_i \approx \langle H_i \rangle$$

where

$$\langle H_i \rangle = J \sum_{j \text{ neighbors}} \langle S_j \rangle + H$$

but all spins are the same on average, so

$$\langle H_i \rangle = qJ\langle S \rangle + H$$

and the weighting of states is given by

$$E_{State} = - \underbrace{(qJ\langle S \rangle + H)}_{\text{mean field, } \langle H \rangle} \sum_i S_i$$

For example, the magnetization is

$$m = \langle S \rangle = \frac{\sum_{S=\pm 1} S e^{\frac{\langle H \rangle}{k_B T} S}}{\sum_{S=\pm 1} e^{\frac{\langle H \rangle}{k_B T} S}}$$

where $\langle S \rangle = \langle S_1 \rangle$, and contributions due to $S_2, S_3 \dots$ cancel from top and bottem. So

$$m = \tanh \frac{\langle H \rangle}{k_B T}$$

or

$$m = \tanh \frac{qJm + H}{k_B T}$$

This gives a critical temperature

$$k_B T_c = qJ$$

as we will now see.

Consider $H = 0$.

$$m = \tanh \frac{m}{T/T_c}$$

Close to T_c , m is close to zero, so we can expand the tanh function:

$$m = \frac{m}{T/T_c} - \frac{1}{3} \frac{m^3}{(T/T_c)^3} + \dots$$

One solution is $m = 0$. This is stable above T_c , but unstable for $T < T_c$. Consider $T < T_c$, where m is small, but nonzero. Dividing by m gives

$$1 = \frac{1}{T/T_c} - \frac{1}{3} \frac{m^2}{(T/T_c)^3} + \dots$$

Rearranging

$$m^2 = 3\left(\frac{T}{T_c}\right)^2\left(1 - \frac{T}{T_c}\right) + \dots$$

Or,

$$m = \left(3\frac{T}{T_c}\right)^{1/2}\left(1 - \frac{T}{T_c}\right)^{1/2}$$

for small m , $T < T_c$. Note that this solution does not exist for $T > T_c$. For $T < T_c$, but close to T_c , we have

$$m \propto \left(1 - \frac{T}{T_c}\right)^\beta$$

where the critical exponent, $\beta = 1/2$. All this is really saying is that m is a

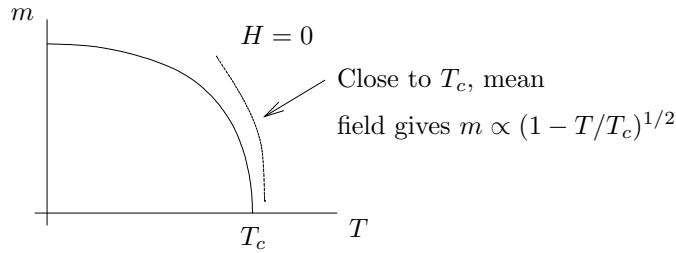


Figure 14.11:

parabola near T_c , as drawn in fig. 14.12 (within mean field theory). This is

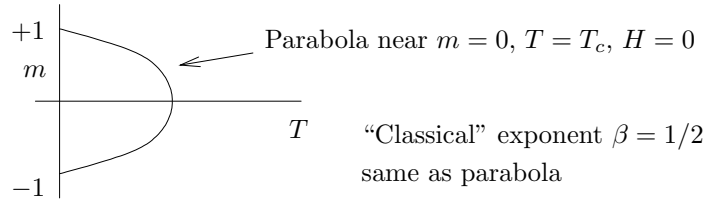


Figure 14.12:

interesting, and useful, and easy to generalize, but only qualitatively correct. Note

$$k_B T_c = qJ = \begin{cases} 2J & , d = 1 \\ 4J & , d = 2 \\ 6J & , d = 3 \end{cases}$$

The result for $d = 1$ is wildly wrong, since $T_c = 0$. In $d = 2$, $k_B T_c \approx 2.27J$, so the estimate is poor. Also, if we take the exact result from Onsager's solution,

$m \propto (1 - T/T_c)^{1/8}$, so $\beta = 1/8$ in $d = 2$. In three dimensions $k_B T_c \sim 4.5J$ and estimates are $\beta = 0.313\dots$ Strangely enough, for $d \geq 4$, $\beta = 1/2$ (although there are logarithmic corrections in $d = 4$)!

For $T \equiv T_c$, but $H \neq 0$, we have

$$m = \tanh\left(m + \frac{H}{k_B T_c}\right)$$

or

$$H = k_B T_c (\tanh^{-1} m - m)$$

Expanding giving

$$H = k_B T_c \left(m + \frac{1}{3}m^3 + \dots - m\right)$$

so

$$H = \frac{k_B T_c}{3} m^3$$

for m close to zero, at $T = T_c$. This defines another critical exponent

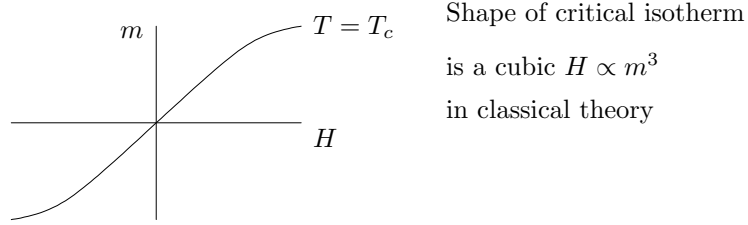


Figure 14.13:

$$H \propto m^\delta$$

for $T = T_c$, where $\delta = 3$ in mean-field theory.

The susceptibility is

$$X_T \equiv \left(\frac{\partial m}{\partial H}\right)_T$$

This is like the compressability in liquid-gas systems

$$\kappa_T = -\frac{1}{V} \left(\frac{\partial V}{\partial P}\right)_T$$

Note the correspondence between density and magnetization per spin, and between pressure and field. The susceptibility “asks the question”: how easy is it to magnetize something with some small field. A large X_T means the material is very “susceptible” to being magnetized.

Also, as an aside, the “sum rule” for liquids

$$\int d\vec{r} \langle \Delta n(r) \Delta n(0) \rangle = n^2 k_B T \kappa_T$$

has an analog for, magnets

$$\int d\vec{r} \langle \Delta m(r) \Delta m(0) \rangle \propto X_T$$

We do not need the proportionality constant.

From above, we have

$$m = \tanh \frac{k_B T_c m + H}{k_B T}$$

or,

$$H = k_B T \cdot \tanh^{-1} m - k_B T_c m$$

and

$$\left(\frac{\partial H}{\partial m} \right)_T = k_B T \frac{1}{1 - m^2} - k_B T_c$$

so

$$X_T = \frac{1}{T - T_c}$$

close to T_c . Or

$$X_T \propto |T - T_c|^{-\gamma}$$

where $\gamma = 1$. This is true for $T \rightarrow T_c^+$, or, $T \rightarrow T_c^-$. To get the energy, and so

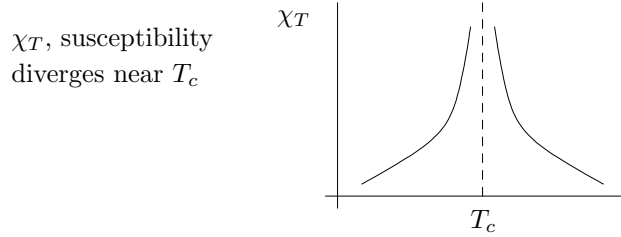


Figure 14.14:

the heat capacity, consider

$$\langle E \rangle = \left\langle -J \sum_{\langle ij \rangle} S_i S_j - H \sum_i S_i \right\rangle$$

All the spins are independent, in mean-field theory, so

$$\langle S_i S_j \rangle = \langle S_i \rangle \langle S_j \rangle$$

Hence

$$\langle E \rangle = -\frac{Jq}{2} N m^2 - H N m$$

The specific heat is

$$c = \frac{1}{N} \left(\frac{\partial E}{\partial T} \right)_N$$

Near T_c , and at $H = 0$,

$$m^2 \propto \begin{cases} (T_c - T) & T < T_c \\ 0 & T > T_c \end{cases}$$

So,

$$\langle E \rangle = \begin{cases} -(Const.)(T_c - T), & T < T_c \\ 0, & T > T_c \end{cases}$$

and

$$c = \begin{cases} Const. & T < T_c \\ 0 & T > T_c \end{cases}$$

If we include temperature dependence away from T_c , the specific heat looks more like a saw tooth, as drawn in fig. 14.15. This discontinuity in c , a second

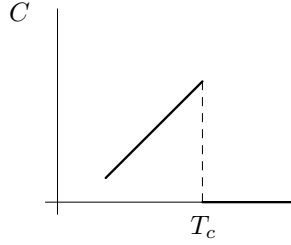


Figure 14.15: Discontinuity in C in mean field theory.

derivative of the free energy is one reason continuous phase transitions were called “second order”. Now it is known that this is an artifact of mean field theory.

It is written as

$$c \propto |T - T_c|^{-\alpha}$$

where the critical exponent

$$\alpha = o(disc.)$$

To summarize, near the critical point,

$$\begin{aligned} m &\sim (T_c - T)^\beta, & T < T_c, H = 0 \\ H &\sim m^\delta, & T = T_c \\ X_T &\sim |T - T_c|^{-\gamma}, & H = 0 \end{aligned}$$

and

$$c \sim |T - T_c|^{-\alpha}, \quad H = 0$$

The exponents are:

	mean field	$d = 2$ (exact)	$d = 3$ (num)	$d \geq 4$
α	0 (disc.)	0 (log)	0.125...	0 (disc.)
β	1/2	1/8	0.313...	1/2
γ	1	7/4	1.25...	1
δ	3	15	5.0...	3

Something is happening at $d = 4$. We will see later that mean-field theory works for $d \geq 4$. For now we will leave the Ising model behind.

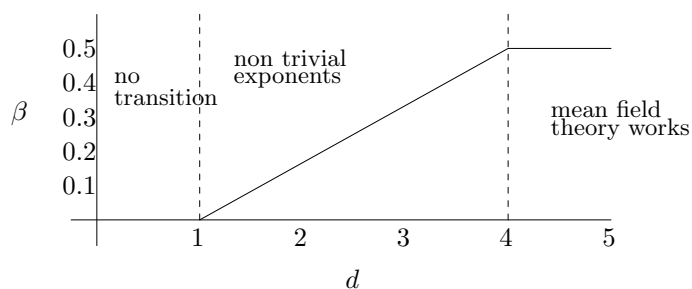


Figure 14.16: Variation of critical exponent β with spatial dimension

Chapter 15

Landan Theory and Ornstien–Zernicke Theory

Nowadays, it is well established that, near the critical point, one can either solve the Ising model

$$E = -J \sum_{\langle ij \rangle} S_i S_j - H \sum_i S_i$$

with

$$Z = \prod_i \sum_{S_i = \pm 1} e^{-E/k_B T}$$

or, the ψ^4 field theory

$$F = \int d^d x \left\{ \frac{k}{2} (\nabla \psi)^2 + \frac{r}{2} \psi^2 + \frac{u}{4} \psi^4 \right\} - H \int d^d x \psi(\vec{x})$$

where k , u , and H are constants, and

$$r \propto T - T_c$$

with the partition function

$$Z = \int_{-}^{+} \int_{-}^{+} \mathcal{D} \psi(\cdot) e^{-F}$$

where the $k_B T$ factor has been set to $k_B T_c$, and incorporated into k , r , u and H (this is ok near T_c).

Look back in the notes for the discussion of the discrete Gaussian model versus the continuous drumhead model of interfaces. Although it may not look like it, the field theory model is a very useful simplification of the Ising model, wherein no important features have been lost.

To motivate their similarities, consider the following. First, the mapping between peices involving external field only require getting used to notation:

$$\begin{array}{ccc} \text{Ising Model} & & \psi^4 \text{ Field Theory} \\ -H \sum_{i=1}^N S_i & \Longleftrightarrow & -H \int d^d x \psi(\vec{x}) \end{array}$$

One is simply the continuum limit of the other. The square gradient term is likewise straightforward, as we now see.

$$\begin{array}{ccc} \text{Ising Model} & & \psi^4 \text{ Field Theory} \\ \frac{J}{2} \sum_{i=1}^N \sum_{j \text{ neighbors of } i} (S_i - S_j)^2 & \Longleftrightarrow & K \int d^d x (\nabla \psi)^2 \\ \downarrow \text{using } S_i^2 = 1 & & \\ \text{const.} - J \sum_{\langle ij \rangle} S_i S_j & & \end{array}$$

The terms harder to see are the $\int d^d \vec{x} (\frac{r}{2} \psi^2 + \frac{u}{4} \psi^4)$ terms. Let us define

$$f(\psi) = \frac{r}{2} \psi^2 + \frac{u}{4} \psi^4$$

This is the free energy for a *homogeneous* system in the ψ^4 field theory (that is, where $\vec{\nabla} \psi \equiv 0$), in the absence of an external field, per unit volume. Note

$$\begin{array}{l} r > 0, \quad T > T_c \\ r < 0, \quad T < T_c \end{array}$$

So we have the double well form. This may look familiar, it was originally due

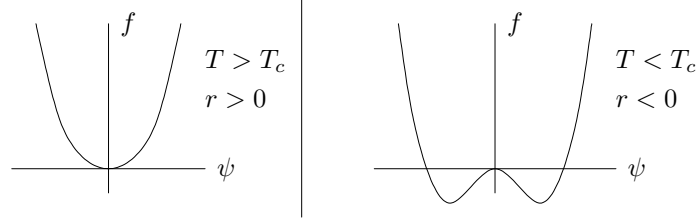


Figure 15.1: “free energy”/volume. No $\vec{\Delta} \psi$, no H

to Van der Waals. (In fact, the whole ψ^4 theory is originally due to Van der Waals.)

Note that, since ψ is determined by the minimum of F , this single or double well restricts ψ to be within a small range. Otherwise, ψ could have any value, since

$$\int_{-}^{+} \int \mathcal{D}\psi(\cdot) = \prod_{\vec{x}} \int_{-\infty}^{\infty} d\psi(\vec{x})$$

These unphysical values of ψ are suppressed in the free energy. In the Ising model, the restriction is $S_i = \pm 1$. We can lift this restriction with a kronecker delta

$$\begin{aligned} \prod_i \sum_{S_i=-1}^{+1} &= \prod_i \sum_{S_i=-\infty}^{\infty} \delta_{S_i^2-1} \\ &= \prod_i \sum_{S_i=-\infty}^{\infty} e^{-\ln(1/\delta_{S_i^2-1})} \end{aligned}$$

So, although this is awkward and artificial, it is sort of the same thing. Expanding it out, as $S_i^2 \rightarrow 0$,

$$-\ln(1/\delta_{S_i^2-1}) \approx S_i^2 - \mathcal{O}(S_i^4)$$

the double-well form. This can be done in a very precise way near T_c . Here I am just trying to give an idea of what is going on.

So the ψ^4 theory is similar to the Ising model, where $\psi(\vec{x})$ is the local magnetization. We will be concerned with the critical exponents $\alpha, \beta, \gamma, \delta, \eta$, and ν defined as follows

$$F(T) \sim |T - T_c|^{2-\alpha}, \quad H = 0$$

So that $C = \frac{\partial^2 F}{\partial T^2} \sim |T - T_c|^{-\alpha}$

$$\begin{aligned} \psi &\sim (T_c - T)^\beta, & H = 0, T < T_c \\ \left(\frac{\partial \psi}{\partial H}\right)_T &\sim X_T \sim |T - T_c|^{-\gamma}, & H = 0 \\ \psi &\sim H^{1/\delta}, & T = T_c \\ \langle \psi(\vec{x})\psi(0) \rangle &\sim \frac{1}{x^{d-2+\eta}}, & T = T_c, H = 0 \\ \langle \psi(\vec{x})\psi(0) \rangle &\sim \begin{cases} e^{-x/\xi} & , T > T_c \\ \langle \psi \rangle^2 + \mathcal{O}(e^{-x/\xi}) & , T < T_c \end{cases} \end{aligned}$$

and

$$\xi \sim |T - T_c|^{-\nu}, \quad H = 0$$

This is the first time we have mentioned η and ν in this context.

Landau motivates the ψ^4 theory in a more physical way. He asserts that a continuous phase transition involves going from a disordered phase, with lots of symmetry, to an ordered phase, with a lower symmetry. To describe the transition we need a variable, called the *order parameter* which is nonzero in the ordered phase only. For example, in a magnet, the order parameter is the magnetization. In general, identifying the order parameter, and the symmetry

broken at T_c can be quite subtle. This will take us too far afield, so we will only consider a scalar order parameter, where the symmetry broken is Z_2 , that is, where $\psi \leftrightarrow -\psi$ symmetry is spontaneously broken. (Note that a magnetic spin confined to dimension $d = 2$, wherein all directions are equivalent, has an order parameter $\vec{\psi} = \psi_x \hat{x} + \psi_y \hat{y}$. The group is $u(1)$. A spin in $d = 3$ has $\vec{\psi} = \psi_x \hat{x} + \psi_y \hat{y} + \psi_z \hat{z}$, if all directions are equivalent; the group is $O(3)$ I think.)

So Landan argues that the free energy to describe the transition must be an analytic function of ψ . Hence

$$F = \int d^d x f(\psi)$$

where

$$f(\psi) = \frac{r}{2} \psi^2 + \frac{u}{4} \psi^4 + \dots$$

No odd terms appear, so that the

$$\psi \leftrightarrow -\psi$$

Symmetry is preserved. In the event an external field breaks the symmetry, another term is added to the free energy, namely

$$-H \int d^d x \psi$$

At high temperatures, in the disordered phase, we want $\psi = 0$. Hence

$$r > 0$$

for $T > T_c$. However, at low temperatures, in the ordered phase, we must have $\psi > 0$. Hence

$$r < 0$$

(To make sure everything is still well defined, we need $u > 0$. There is no need for u to have any special dependence in T near T_c , so we take it to be a constant. Or, if you prefer

$$u(T) \approx u(T_c) + \overset{\text{unimportant}}{(\text{Const.})}(T - T_c) + \dots$$

near T_c . So we simply set u to be constant.)

To find r 's evident T dependence, we Taylor expand

$$r = r_0(T - T_c) + \dots$$

keeping only the lowest-order term, where $r_0 > 0$.

Finally, we know that spatial fluctuations should be suppressed. Expanding in powers of $\vec{\nabla}$ and ψ , the lowest order nontrivial term is

$$\frac{1}{2} K \int d^d r |\nabla \psi|^2$$

where $K > 0$. Similar terms like $\int \psi \nabla^2 \psi$ turn into this after parts integration. Note that this term means inhomogeneities ($\vec{\nabla} \psi \neq 0$) add free energy.

Now we're done. The Landau free energy is

$$F = \int d^d x \left\{ f(\psi) + \frac{k}{2} |\nabla \psi|^2 \right\} - H \int d^d x \psi$$

with

$$f(\psi) = \frac{r}{2} \psi^2 + \frac{u}{4} \psi^4$$

In high-energy physics, this is called a ψ^4 field theory. The ∇^2 term gives kinetic energy; $r \propto$ mass of particle; and ψ^4 gives interactions. Van der Waals wrote down the same free energy (the case with $k \equiv 0$ is what is given in the text books). He was thinking of liquid-gas transitions, so $\psi = n - n_c$, where n is density and n_c is the density at the critical point. Finally, from the theory of thermodynamic fluctuations, we should expect that (in the limit in which the square gradient and quartic term can be neglected)

$$r \propto 1/\kappa_T$$

or equivalently

$$r \propto 1/X_T$$

It is worth emphasizing that the whole machinery of quadratic, quartic, and square gradient terms is necessary. For example, note that the critical point is *defined* thermodynamically as

$$\left(\frac{\partial P}{\partial V} \right)_{T_c} = 0$$

and

$$\left(\frac{\partial^2 P}{\partial V^2} \right)_{T_c} = 0$$

in a liquid-gas system. In a magnet:

$$\left(\frac{\partial H}{\partial m} \right)_{T_c} = 0, \text{ and } \left(\frac{\partial^2 H}{\partial m^2} \right)_{T_c} = 0$$

Trying to find the equation of state for $T \equiv T_c$ in a liquid we expand:

$$P(n, T_c) = P_c + \left(\frac{\partial P}{\partial n} \right)_{T_c} (n - n_c) + \frac{1}{2} \left(\frac{\partial^2 P}{\partial n^2} \right)_{T_c} (n - n_c)^2 + \frac{1}{6} \left(\frac{\partial^3 P}{\partial n^3} \right)_{T_c} (n - n_c)^3 + \dots$$

For a magnet

$$H(m, T_c) = H_c + \left(\frac{\partial H}{\partial m} \right)_{T_c} (m - m_c) + \frac{1}{2} \left(\frac{\partial^2 H}{\partial m^2} \right)_{T_c} m^2 + \frac{1}{2} \left(\frac{\partial^3 H}{\partial m^3} \right)_{T_c} m^3 + \dots$$

So, at the critical point for both systems — *from thermodynamics and analyticity only* — we have

$$(P - P_c) \propto (n - n_c)^\delta$$

and

$$H \propto m^\delta$$

where

$$\delta = 3$$

Experimentally, $\delta \approx 5.0\dots$ in three dimensions, while $\delta \equiv 15$ in $d = 2$. So we have a big subtle problem to solve!

15.1 Landan Theory

We now solve for the homogeneous states of the system. In that case $\vec{\nabla}\psi = 0$, so we need only consider

$$\begin{aligned} F/V &= f(\psi) - H\psi \\ &= \frac{r_0}{2}(T - T_c)\psi^2 + \frac{u}{4}\psi^4 - H\psi \end{aligned}$$

For $H = 0$, we find the solution by minimizing F/V (or f):

$$\frac{\partial f}{\partial \psi} = 0 = r_0(T - T_c)\psi + u\psi^3$$

There are 3 solutions

$$\begin{aligned} \psi &= 0, \text{ (stable for } T > T_c) \\ \psi &= \pm \sqrt{\frac{r_0(T_c - T)}{u}}, \text{ (stable for } T < T_c) \end{aligned}$$

(To check stability, consider $\partial^2 f / \partial \psi^2$.) for $H \neq 0$, only one solution is stable for $T < T_c$. You can check this yourself.

$$\frac{\partial}{\partial \psi}(f(\psi) - H\psi) = 0 = \frac{\partial f}{\partial \psi} - H$$

or

$$H = r_0(T - T_c)\psi + u\psi^3$$

Note that this gives

$$1/X_T = \left(\frac{\partial H}{\partial \psi}\right)_T = r_0(T - T_c)$$

close to the critical point, or

$$X_T \propto |T - T_c|^{-1}$$

Also, at $T = T_c$, this gives

$$H \propto \psi^3$$

Finally, note that (at $H = 0$)

$$f = \frac{1}{2}r\psi^2 + \frac{1}{4}u\psi^4$$

but

$$\psi^2 = \begin{cases} 0 & , T > T_c \\ -\frac{r}{u} & , T < T_c \end{cases}$$

so,

$$f = \begin{cases} 0 & , T > T_c \\ -\frac{1}{4}\frac{r^2}{u} & , T < T_c \end{cases}$$

or

$$f \propto \begin{cases} 0 & , T > T_c \\ (T_c - T)^2 & , T < T_c \end{cases}$$

but $c \propto \partial^2 f / \partial T^2$, so the heat capacity

$$c \propto \begin{cases} 0 & , T > T_c \\ \text{Const.} & , T < T_c \end{cases}$$

In summary, these are the *same* results as we obtained from the mean field theory of the Ising model. The relationship between analyticity of equations of state, and mean-field theory, is worth thinking about.

$$\begin{array}{ll} \alpha & 0 \text{ (disc.)} \\ \beta & 1/2 \\ \gamma & 1 \\ \delta & 3 \end{array}$$

15.2 Ornstein — Zernicke Theory of Correlations

Ornstein and Zernicke give a physically plausible approach to calculating correlation functions. It works well (except near the critical point).

In a bulk phase, say $T \gg T_c$, and we shall only consider $H = 0$, ψ has an average value of zero:

$$\langle \psi \rangle = 0$$

So we expect

$$\psi^4 \ll \psi^2$$

and, above T_c ,

$$F \approx \int d^d x \left[\frac{k}{2} (\nabla \psi)^2 + \frac{r}{2} \psi^2 \right]$$

Far below T_c , we expect ψ to be close to its average value

$$\langle \psi \rangle = \pm \sqrt{-\frac{r}{u}}$$

where $r = r_0(T - T_c)$ = negative below T_c . Let

$$\Delta\psi \equiv \psi - \langle \psi \rangle$$

Then

$$\vec{\nabla}\psi = \vec{\nabla}(\Delta\psi - \langle \psi \rangle) = \vec{\nabla}(\Delta\psi)$$

Also,

$$\psi^2 = (\Delta\psi - \langle \psi \rangle)^2 = (\Delta\psi)^2 - 2\langle \psi \rangle \Delta\psi + \langle \psi \rangle^2$$

and

$$\begin{aligned} \psi^4 &= (\Delta\psi - \langle \psi \rangle)^4 \\ &= (\Delta\psi)^4 - 4\langle \psi \rangle (\Delta\psi)^3 + 6\langle \psi \rangle^2 (\Delta\psi)^2 - 4\langle \psi \rangle^3 (\Delta\psi) + \langle \psi \rangle^4 \end{aligned}$$

Hence, grouping terms by powers of $\Delta\psi$,

$$\frac{r}{2}\psi^2 + \frac{u}{4}\psi^4 = (-r\langle \psi \rangle - u\langle \psi \rangle^3)\Delta\psi + \left(\frac{r}{2} + \frac{3u}{2}\langle \psi \rangle^2\right)(\Delta\psi)^2 + \mathcal{O}(\Delta\psi)^3$$

But $\langle \psi \rangle = \pm\sqrt{-r/u}$, so

$$\frac{r}{2}\psi^2 + \frac{u}{4}\psi^4 = -r(\Delta\psi)^2 + \mathcal{O}((\Delta\psi)^3)$$

But $(\Delta\psi)^3 \ll (\Delta\psi)^2$, so, below T_c

$$F \approx \int d^d x \left[\frac{k}{2} (\nabla \Delta\psi)^2 + |r| (\Delta\psi)^2 \right]$$

where we have used the fact that r is negative definite below T_c to write

$$-r = |r|$$

Both approximate free energies are pretty much the same as the free energy for the interface model:

$$\frac{\sigma}{2} \int d^{d-1} x \left(\frac{\partial h}{\partial x} \right)^2$$

They can be solved exactly the same way. The definitions of Fourier transforms are the same (except for $d^{d-1}x \rightarrow d^d x$, and $L^{d-1} \rightarrow L^d$, for example), and the arguments concerning translational invariance of space, and isotropy of space still hold. Let

$$C(x) \equiv \langle \psi(x) \psi(0) \rangle$$

Then we have (putting the $k_B T$ back in $e^{\frac{-F}{k_B T}}$)

$$\hat{C}(k) = \frac{k_B T}{Kk^2 + r}, \quad T > T_c$$

and since

$$\langle \Delta\psi(x) \Delta\psi(0) \rangle = \langle \psi(x) \psi(0) \rangle - \frac{|r|}{u}$$

using $\langle \psi \rangle^2 = |r|/u$, we have

$$\hat{C}(k) = \frac{k_B T}{K k^2 + 2|r|} - \frac{|r|}{u} (2\pi)^d \delta(k) \quad , \quad T < T_c$$

These can be Fourier transformed back to real space.

As an aside, remember the Green function for the Helmholtz equation satisfies

$$(\nabla^2 - \xi^{-2})G(x) = -\delta(x)$$

Hence

$$\hat{G}(k) = 1/(k^2 + \xi^{-2})$$

The Green function is given in many books

$$G(x) = \begin{cases} \xi/2 e^{-x/\xi}, & , d = 1 \\ \frac{1}{2\pi} K_0(x/\xi) & , d = 2 \\ \frac{1}{4\pi x} e^{-x/\xi} & , d = 3 \end{cases}$$

for example. Recall the modified Bessel function of *zero*th order satisfies $K_0(y \rightarrow 0) = -\ln y$, and $K_0(y \rightarrow \infty) = (\pi/2y)^{1/2} e^{-y}$. This is more than we need. In discussions of critical phenomena, most books state that if

$$\hat{G}(k) \propto \frac{1}{k^2 + \xi^{-2}}$$

then

$$G(x) \propto \frac{1}{x^{d-2}} e^{-x/\xi}$$

The reason for the seemingly sloppiness is that proportionality constants are unimportant, and it is only long-length-scale phenomena we are interested in. Indeed, we should only trust out results for large x , or equivalently small k (since we did a gradient expansion in $\vec{\nabla}\psi$).

In fact, so far as the Fourier transforms go, here is what we can rely on:

If

$$\hat{G}(k) \propto \frac{1}{k^2 + \xi^{-2}}$$

then

$$G(x) \propto e^{-x/\xi}$$

unless $\xi = \infty$, in which case

$$G(x) \propto \frac{1}{x^{d-2}}$$

for large x .

Using these results, we finally obtain the correlation function up to a constant of proportionality,

$$C(x) = \begin{cases} e^{-x/\xi} & , T > T_c \\ \frac{1}{x^{d-2}} & , T = T_c \\ \langle \psi \rangle^2 + \mathcal{O}(e^{-x/\xi}) & , T < T_c \end{cases}$$

where the *correlation length*:

$$\xi = \begin{cases} \sqrt{\frac{k}{r}} & , T > T_c \\ \sqrt{\frac{k}{2|r|}} & , T < T_c \end{cases}$$

Note that the result for $T < T_c$

$$C(x) = \langle \psi(x) \psi(0) \rangle = \langle \psi \rangle^2 + \mathcal{O}(e^{-x/\xi})$$

is a rewriting of

$$\langle \Delta\psi(x) \Delta\psi(0) \rangle = e^{-x/\xi}$$

In cartoon form: figure 15.2. Note the correlation length satisfies $\xi \sim 1/\sqrt{|r|}$

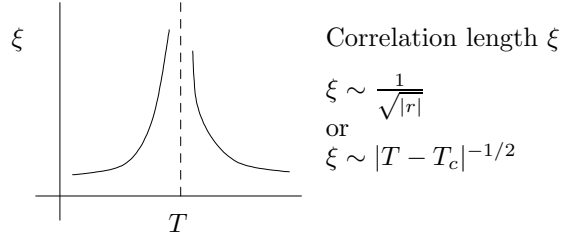
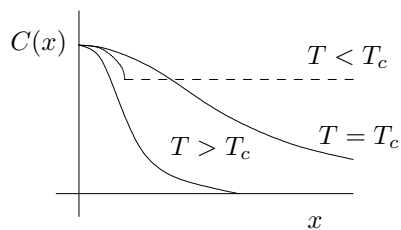


Figure 15.2:

above and below T_c , so

$$\xi \sim |T - T_c|^{-1/2}$$

in Ornstein–Zernicke theory. The correlation function becomes power-law-like at T_c .

Figure 15.3: Correlation function above at and below T_c

$$C(x) = \frac{1}{x^{d-2}} \quad , \quad T = T_c, \quad (H = 0)$$

Since

$$\xi \sim |T - T_c|^\nu$$

and

$$C(x) \sim \frac{1}{x^{d-2+\eta}}$$

we have $\nu = 1/2$ and $\eta = 0$.

This completes the list of critical exponents.

In summary:

	Mean Field Ising	Landau 0 - Z	$d = 2$ Ising	$d = 3$ Ising	$d \geq 4$ Ising
α	0 (disc.)	0 (disc.)	0 (log)	0.125...	0
β	1/2	1/2	1/8	0.313...	1/2
γ	1	1	7/4	1.25...	1
δ	3	3	15	5.0...	3
η	0	0	1/4	0.041...	0
ν	1/2	1/2	1	0.638...	1/2

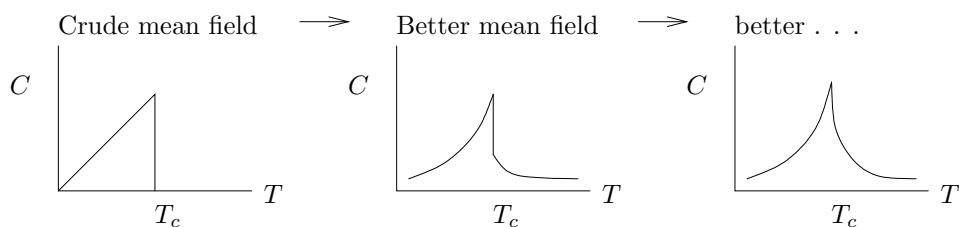
So something new is needed.

Chapter 16

Scaling Theory

Historically, the field of phase transitions was in a state of turmoil in the 1960's. Sensible down-to-earth theories like the mean-field approach, Landau's approach, and Ornstein-Zernicke theory were not working. A lot of work was put into refining these approaches, with limited success.

For example, the Bragg-Williams mean field theory is quite crude. Refinements can be made by incorporating more local structure. Pathria's book on Statistical Mechanics reviews some of this work. Consider the specific heat of the two-dimensional Ising model. Since Onsager's solution, this was known exactly. It obeys $C \sim \ln |T - T_c|$, near T_c . The refinements make for much better predictions of C *away from* T_c . At T_c , *all* mean field theories give C to be discontinuous. Comparing to the exact result shows this as well. (This



All mean theories give $C \sim$ discontinuous

Figure 16.1:

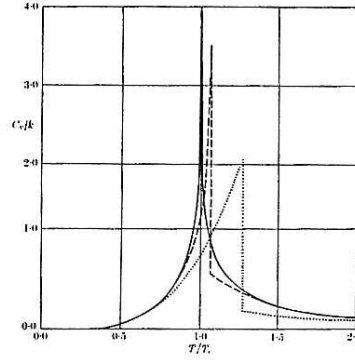


FIG. 1.13. The solid curve shows the specific heat of the two-dimensional Ising model as obtained from the exact solution of Onsager (solid curve), from the Bethe approximation (dotted curve), and the Kramers-Wannier and Kikuchi approximation (broken curve). After Domb (1960).

Figure 16.2:

is taken from Stanley's book, "Introduction to Phase Transitions and Critical Phenomena".)

Experiments consistently did not (and do not of course) agree with Landau's prediction for the liquid-gas or binary liquid case. Landau gives

$$n - n_c \sim (T_c - T)^\beta$$

with $\beta = 1/2$. Three-dimensional experiments gave $\beta \approx 0.3$ (usually quoted as $1/3$ historically), for many liquids. The most famous paper is by Guggenheim in J. Chem. Phys. 13, 253 (1945).

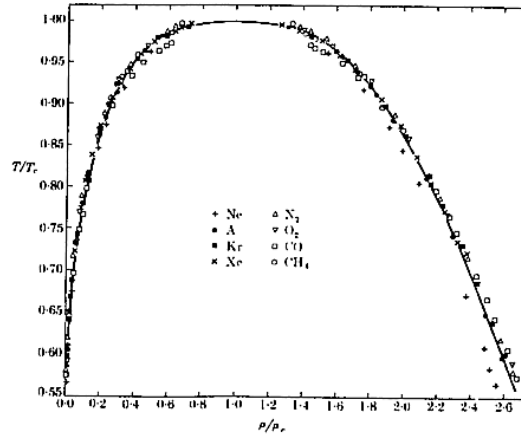


FIG. 1.8. Measurements on eight fluids of the coexistence curve (a reflection of the $P\rho T$ surface in the ρT plane analogous to Fig. 1.3). The solid curve corresponds to a fit to a cubic equation, i.e. to the choice $\beta = 1/3$, where $\rho - \rho_c \sim (-t)^{1/3}$. From Guggenheim (1945).

Figure 16.3:

This is taken from Stanley's book again. The phenomena at T_c , where

$\xi \rightarrow \infty$ is striking.

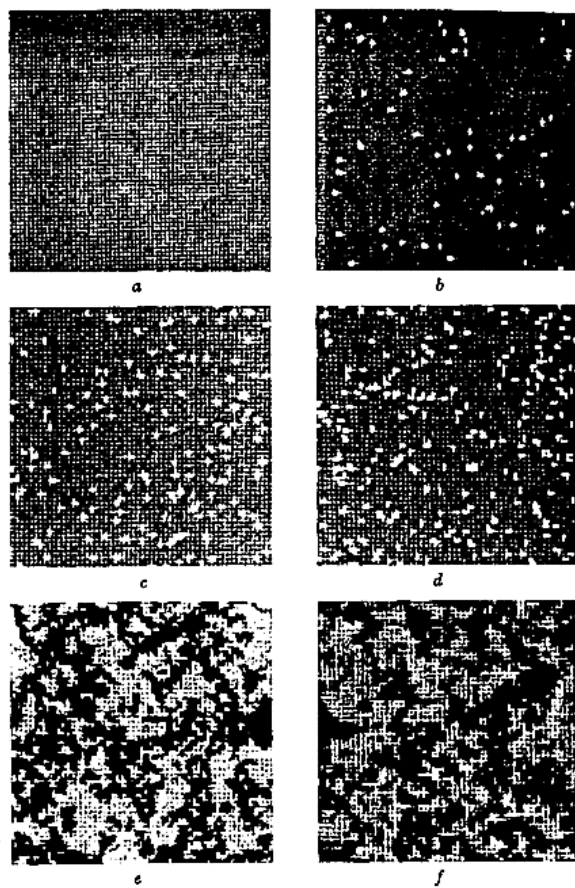


FIG. 1.6. Schematic indication of the lattice-gas model of a fluid system. Fig. 1.4 (which contains 8×20 cells) should be interpreted as representing a relatively small portion of this lattice (which contains 64×64 cells). (a) is the completely ordered state (which exists only at $T = 0$); (b), $T \approx \frac{1}{2}T_c$; (c), $T \approx \frac{1}{3}T_c$; (d), $T \approx \frac{1}{4}T_c$; (e), $T \approx T_c$; (f), $T \geq T_c$. This illustration and the associated temperatures are to be regarded as purely schematic. In fact, the figure was constructed from a computer simulation of the time-dependent aspects of the two-dimensional Ising model and actually represents rather different phenomena (cf. Appendix E). After Ogita *et al.* (1969).

Figure 16.4:

Here I show some (quite old) simulations of the Ising model from Stanley's book. Note the "fractally" structure near T_c . You can easily program this yourself if you are interested. Nowadays, a PC is much much faster than one needs to study the two-dimensional Ising model. Strange things are also seen experimentally.

For a fluid, imagine we heat it through T_c , fixing the pressure so that it

corresponds to phase coexistence (i.e., along the liquid–vapor line in the P–T diagram as drawn). The system looks quite different depending on temperature.

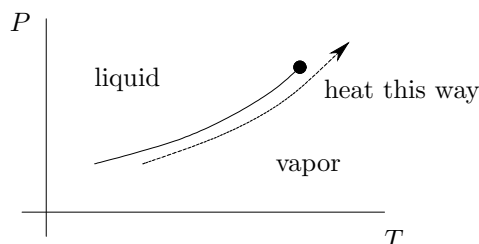


Figure 16.5:

It is transparent, except for the meniscus separating liquid and vapour for $T <$

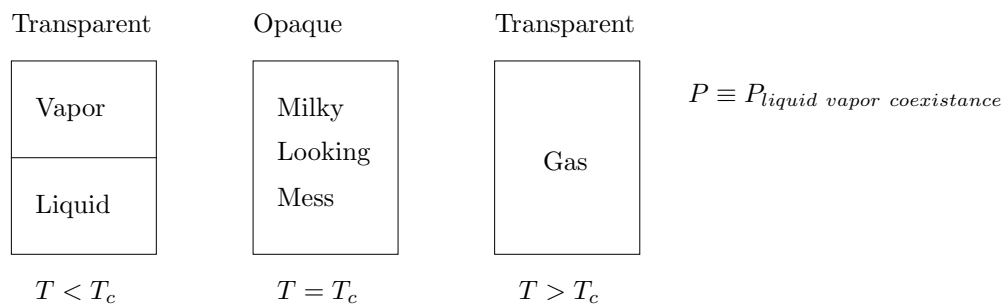


Figure 16.6:

T_c , for $T < T_c$ and $T > T_c$. But at T_c , it is milky looking. This is called “critical opalescence”. The correlation length ξ , corresponding to fluctuations, is so big it scatters light.

Rigorous inequalities were established with the six critical exponents. (I should mention that I am assuming any divergence is the same for $T \rightarrow T_c^+$ or $T \rightarrow T_c^-$, which is nontrivial, but turns out to be true.) For example, Rushbrooke proved

$$\alpha + 2\beta + \gamma \geq 2$$

Griffiths proved

$$\alpha + \beta(1 + \delta) \geq 2$$

These are rigorous. Other inequalities were proved with some plausible assumptions. These all work for the exponents quoted previously — as *equalities*! In fact, all the thermodynamic inequalities turned out to be obeyed as equalities. These inequalities were very useful for ruling out bad theories and bad experiments.

Here is a list of all the independent relations between exponents

$$\begin{aligned}\alpha + 2\beta + \gamma &= 2 \\ \gamma &= \nu(2 - \eta) \\ \gamma &= \beta(\delta - 1)\end{aligned}$$

and

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

So it turns out there are only two independent exponents. The last equation is called the “hyperscaling” relation. It is the only one involving spatial dimension d . The others are called “scaling” relations. This is because the derivation of them requires scaling arguments.

16.1 Scaling With ξ

Note that near T_c , ξ diverges. Also, correlation functions become power-law like at T_c according to the approximate Ornstein–Zernicke theory as well as the exact result for the Ising model in two dimensions.

So let us make the bold assertion that, since $\xi \rightarrow \infty$ at T_c , the only observable information on a large length scale involves quantities proportional to ξ . That is, for example, all lengths scale with ξ .

Let us now derive the hyperscaling relation. The free energy is extensive, and has units of energy, so

$$F = k_B T \frac{V}{\xi^d}$$

One can think of V/ξ^d as the number of independent parts of the system. But

$$\xi \sim |T - T_c|^{-\nu}$$

so

$$F = |T - T_c|^{\nu d}$$

But the singular behavior in F is from the heat capacity exponent α via

$$F = |T - T_c|^{2-\alpha}$$

Hence

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

This back-of-the-envelope argument is the fully rigorous (or as rigorous as it comes) argument! The hyperscaling relation works for all systems, except the mean field solutions must have $d = 4$.

To derive the three scaling relations, we assert that correlations are scale invariant:

$$C(x, T, H) = \frac{1}{x^{d-2+\eta}} \tilde{C}(x/\xi, H\xi^\eta)$$

Note that, since blocks of spin are correlated on scale ξ , this means only a little H is needed to magnetize the system. So the effect of H is *magnified* by ξ . The factor is ξ^y , where y must be determined. In a finite-size system we would have

$$C(x, T, H L) = \frac{1}{x^{d-2+\eta}} \tilde{C}(x/\xi, H\xi^y, L/\xi)$$

In any case, in the relation above

$$\xi \sim |T - T_c|^{-\nu}$$

For ease of writing, let

$$t \equiv |T - T_c|$$

So that $\xi \sim t^{-\nu}$. Then we have

$$C(x, T, H) = \frac{1}{x^{d-2+\eta}} \tilde{C}(xt^\nu, H/t^\Delta)$$

where

$$\Delta = y\nu$$

is called the “gap exponent”. But the magnetic susceptibility satisfies the sum rule

$$\begin{aligned} X_T &\propto \int d^d x C(x, T, H) \\ &= \int d^d x \frac{1}{x^{d-2+\eta}} \tilde{C}(xt^\nu, H/t^\Delta) \\ &= (t^\nu)^{-2+\eta} \int d^d (xt^\nu) \frac{\tilde{C}(xt^\nu, H/t^\Delta)}{(xt^\nu)^{d-2+\eta}} \end{aligned}$$

so,

$$X_T(T, H) = t^{(-2+\eta)\nu} \tilde{X}(H/t^\Delta)$$

where

$$\int_{-\infty}^{\infty} d^d (xt^\nu) \frac{1}{(xt^\nu)^{d-2+\eta}} \tilde{C}(xt^\nu, H/t^\Delta) \equiv \tilde{X}\left(\frac{H}{t^\Delta}\right)$$

But

$$X(T, H = 0) = t^{-\gamma}$$

Hence

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

called Fisher’s law.

Now note

$$\frac{\partial \psi}{\partial H}(T, H) = X_T(T, H)$$

so integrating, we have the indefinite integral

$$\begin{aligned}\psi(T, H) &= \int dH X(T, H) \\ &= \int dH t^{-\gamma} \tilde{X}\left(\frac{H}{t^\Delta}\right) \\ &= t^{\Delta-\gamma} \int d\left(\frac{H}{t^\Delta}\right) \tilde{X}\left(\frac{H}{t^\Delta}\right) \\ &= t^{\Delta-\gamma} \tilde{\psi}(H/t^\Delta)\end{aligned}$$

But

$$\psi(T, H=0) = t^\beta$$

So

$$\beta = \Delta - \gamma$$

and the gap exponent

$$\Delta = \beta + \gamma$$

or equivalently, $y = (\beta + \gamma)/\nu$. So we have

$$\psi(T, H) = t^\beta \tilde{\psi}\left(\frac{H}{t^{\beta+\gamma}}\right)$$

Here is a neat trick for changing the dependence “out front” from t to H . We don’t know what the function $\tilde{\psi}$ is. All we know is that it is a function of $\frac{H}{t^{\beta+\gamma}}$.

Any factor of $H/t^{\beta+\gamma}$ that we multiply or divide into $\tilde{\psi}$ leaves another unknown function, which is still a function of only $H/t^{\beta+\gamma}$. In particular, a crafty choice can cancel the t^β factor:

$$\psi(T, H) = t^\beta \left(\frac{H}{t^{\beta+\gamma}}\right)^{\frac{\beta}{\beta+\gamma}} \left(\frac{H}{t^{\beta+\gamma}}\right)^{\frac{-\beta}{\beta+\gamma}} \tilde{\psi}\left(\frac{H}{t^{\beta+\gamma}}\right)$$

We call the last term on the right-hand side $\tilde{\tilde{\psi}}$, and obtain

$$\psi(T, H) = H^{\beta/\beta+\gamma} \tilde{\tilde{\psi}}\left(\frac{H}{t^{\beta+\gamma}}\right)$$

Note that $H^{-\beta/\beta+\gamma} \tilde{\tilde{\psi}}(H) = \tilde{\tilde{\psi}}(H)$. Now

$$\psi(T_c, H) = H^{1/\delta}$$

is the shape of the critical isotherm, so $\delta = 1 + \gamma/\beta$, or

$$\gamma = \beta(\delta - 1)$$

called the Widom scaling law.

Note that $\psi = \partial F / \partial H$, so

$$\begin{aligned} F &= \int dH \psi \\ &= \int dH t^\beta \tilde{\psi}\left(\frac{H}{t^{\beta+\gamma}}\right) \\ &= t^{2\beta+\gamma} \int d\left(\frac{H}{t^{\beta+\gamma}}\right) \tilde{\psi}\left(\frac{H}{t^{\beta+\gamma}}\right) \end{aligned}$$

or,

$$F(T, H) = t^{2\beta+\gamma} \tilde{F}\left(\frac{H}{t^{\beta+\gamma}}\right)$$

But

$$F(T, 0) = t^{2-\alpha}$$

so $2 - \alpha = 2\beta + \gamma$, or

$$\alpha + 2\beta + \gamma = 2$$

which is Rushbrooke's law.

The form for the free energy is

$$F(T, H) = t^{2-\alpha} \tilde{F}\left(\frac{H}{t^{\beta+\gamma}}\right)$$

Note that

$$\lim_{H \rightarrow 0} \frac{\partial^n F}{\partial H^n} = t^{2-\alpha} t^{-n(\beta+\gamma)}$$

or

$$F^{(n)} = t^{2-\alpha-n(\beta+\gamma)}$$

So all derivatives of F are separated by a constant “gap” exponent $(\beta + \gamma)$.

$$y_n = 2 - \alpha - n(\beta + \gamma)$$

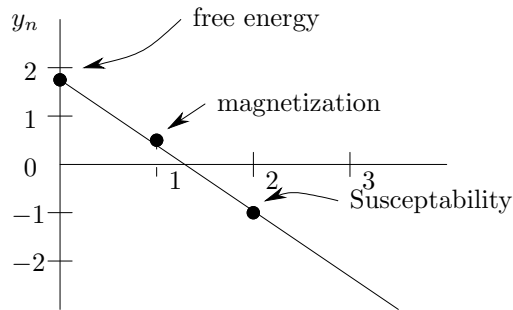


Figure 16.7: Gap Exponent

Finally, we will give another argument for the hyperscaling relation. It is not, however, as compelling as the one above, but it is interesting.

Near T_c , we assert that fluctuations in the order parameter are of the same order as ψ itself. Hence

$$\frac{\langle(\Delta\psi)^2\rangle}{\langle\psi\rangle^2} = \mathcal{O}(1)$$

This is the weak part of the argument, but it is entirely reasonable. Clearly

$$\langle\psi\rangle^2 \sim t^{2\beta}$$

To get $\langle(\Delta\psi)^2\rangle$, recall the sum rule

$$\int d^d x \langle \Delta\psi(x) \Delta\psi(0) \rangle = X_T$$

The integral is only nonzero for $x \leq \xi$. In that range,

$$\langle \Delta\psi(x) \Delta\psi(0) \rangle \sim \mathcal{O}(\langle(\Delta\psi)^2\rangle)$$

Hence

$$\int d^d x \langle \Delta\psi(x) \Delta\psi(0) \rangle = \xi^d \langle(\Delta\psi)^2\rangle = X_T$$

or,

$$\langle(\Delta\psi)^2\rangle = X_T \xi^{-d}$$

But

$$X_T \sim t^{-\gamma}$$

and

$$\xi \sim t^{-\nu}$$

so,

$$\langle(\Delta\psi)^2\rangle = t^{\nu d - \gamma}$$

The relative fluctuation therefore obeys

$$\frac{\langle(\Delta\psi)^2\rangle}{\langle\psi\rangle^2} = t^{\nu d - \gamma - 2\beta}$$

But if, as we argued

$$\frac{\langle(\Delta\psi)^2\rangle}{\langle\psi\rangle^2} = \mathcal{O}(1)$$

we have

$$\nu d = \gamma + 2\beta$$

which is the hyperscaling relation. (To put it into the same form as before, use Rushbrooke scaling law of $\alpha + 2\beta + \gamma = 2$. This gives $\nu d = 2 - \alpha$.)

As a general feature

$$\frac{\langle(\Delta\psi)^2\rangle}{\langle\psi\rangle^2}$$

measures how important fluctuations are. We discussed this at the beginning of the course. Mean field theory works best when fluctuations are small. So we say that mean-field theory is correct (or at least self-consistent) if its predictions give

$$\frac{\langle(\Delta\psi)^2\rangle}{\langle\psi\rangle^2} \ll 1$$

This is called the Ginzburg criterion for the correctness of mean-field theory. But from above

$$\frac{\langle(\Delta\psi)^2\rangle}{\langle\psi\rangle^2} = t^{\nu d - \gamma - 2\beta}$$

Putting in the results from mean-field theory ($\nu = 1/2$, $\gamma = 1$, $\beta = 1/2$) gives

$$\frac{\langle(\Delta\psi)^2\rangle}{\langle\psi\rangle^2} = t^{\frac{1}{2}(d-4)}$$

Hence, mean-field theory is self consistent for $d > 4$ as $T \rightarrow T_c$ ($t \rightarrow 0$). Mean-field theory doesn't work for $d < 4$.

This is called the *upper critical dimension*; the dimension at which fluctuations can be neglected.

Chapter 17

Renormalization Group

The renormalization group is a theory that builds on scaling ideas. It gives a way to calculate critical exponents. The idea is that phase transition properties depend only on a few long wavelength properties. These are the symmetry of the order parameter and the dimension of space, for systems without long-range forces or quenched impurities. We can ignore the short length scale properties, but we have to do that “ignoring” in a controlled careful way. That is what the renormalization group does.

The basic ideas were given by Kadanoff in 1966, in his “block spin” treatment of the Ising model. A general method, free of inconsistencies, was figured out by Wilson (who won the Nobel prize for his work). Wilson and Fisher figured out a very useful implementation of the renormalization group in the $\epsilon = 4 - d$ expansion, where the small parameter ϵ was eventually extrapolated to unity, and so $d \rightarrow 3$.

At a critical point, Kadanoff reasoned that spins in the Ising model act together. So, he argued one could solve the partition function recursively, by progressively coarse-graining on scales smaller than ξ . From the recursion relation describing scale changes, one could solve the Ising model. This would only be a solution on large length scales, but that is what is of interest for critical phenomena.

The partition function is

$$Z(N, K, h) = \sum_{\{\text{States}\}} e^{K \sum_{\langle ij \rangle} S_i S_j + h \sum_i S_i}$$

where $K = J/k_B T$ and $h = H/k_B T$.

Now, let us (imagine we) reorganize all of these sums as follows.

1. Divide the lattice into blocks of size b^b (we’ll let $b = 2$ in the cartoon below).
2. Each block will have an internal sum done on it leading to a block spin

$S = \pm 1$. For example, the majority-rule transformation is

$$S_i = \begin{cases} +1, & \sum_{i \in b^d} S_i > 0 \\ -1, & \sum_{i \in b^d} S_i < 0 \\ \pm 1_{\text{randomly}}, & \sum_{i \in b^d} S_i = 0 \end{cases}$$

Clearly, if N is the number of spins originally, N/b^d is the number of spins after blocking.

It is easy to see that this is a good idea. Any feature correlated on a finite length scale l before blocking becomes l/b after blocking once. Blocking m times means this feature is now on the small scale l/b^m . In particular, consider the correlation length ξ . The renormalization group (RG) equation for ξ is

$$\xi^{(m+1)} = \xi^{(m)}/b$$

If the RG transformation has a fixed point (denoted by $*$), which we always assume is true here,

$$\xi^* = \xi^*/b$$

The only solutions are

$$\xi^* = \begin{cases} 0, & \text{Trivial : } T = 0, \text{ Fixed Point : } T = \infty \\ \infty, & \text{Nontrivial fixed point : } T = T_c \end{cases}$$

The critical fixed point is unstable, while the $T = 0$ and $T = \infty$ fixed points

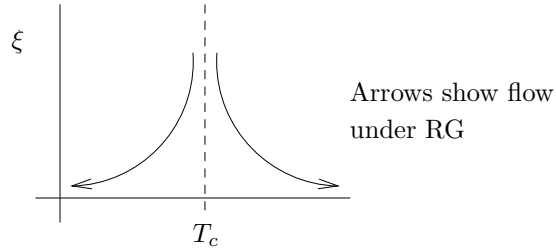


Figure 17.1:

are stable.

Note that the coarse-graining eliminates information on small scales in a controlled way. Consider figure 17.2. Let us consider a blocking transformation

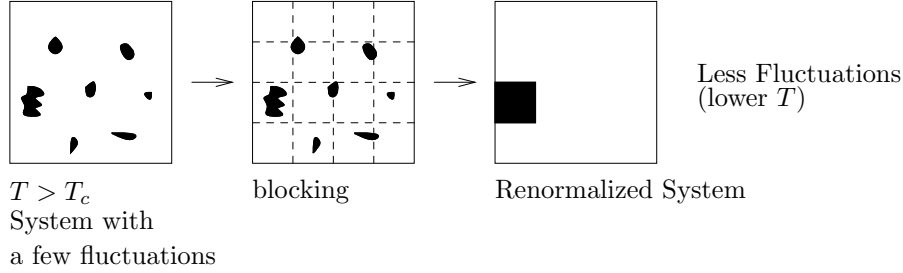


Figure 17.2:

(fig. 17.3). We know the states in the new renormalized system, $S_i^{(m+1)} = \pm 1$,

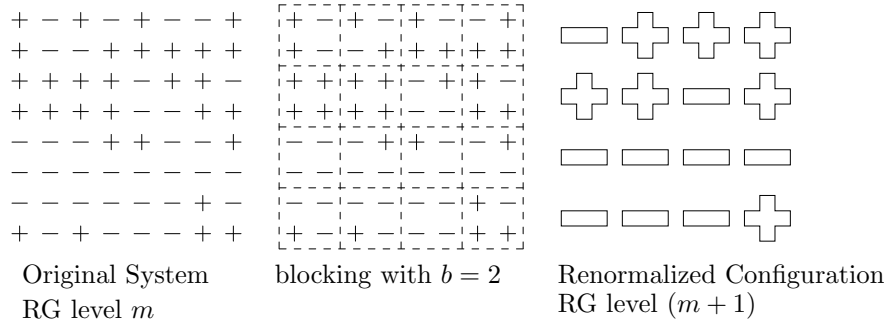


Figure 17.3:

for the renormalized spins on sites

$$i = 1, 2, \dots, N^{(m+1)}$$

Note,

$$N^{(m+1)} N^{(m)} / b^d$$

In the cartoon, $N^{(m+1)} = 16$, $b^d = 2^2$, and $N^{(m)} = 64$.

This is all pretty clear: we can easily label all the states in the sum for Z . But how do the spins interact after renormalization? It is clear that, if the original system has short-ranged interactions, the renormalized system will as well.

At this point, historically, Kadanoff simply said that if $E_{State}^{(m)}$ was the Ising model, so was $E_{State}^{(m+1)}$. The coupling constants $K^{(m)}$ and $h^{(m)}$ would depend on the level of transformation, but that's all. This is not correct as we will see shortly, but the idea and approach is dead on target.

In any case, assuming this is true

$$Z^{(m)}(N^{(m)}, K^{(m)}, h^{(m)}) = \sum_{\{States\}} e^{K^{(m)} \sum_{\langle ij \rangle} S_i^{(m)} S_j^{(m)} + h^{(m)} \sum_i S_i^{(m)}}$$

But

$$Z^{(m)} = e^{-N^{(m)} f^{(m)}(h^{(m)}, t^{(m)})}$$

where $t^{(m)}$ is the reduced temperature obtained from $K^{(m)}$, and $f^{(m)}$ is the free energy per unit volume. But we must have $Z^{(m)} = Z^{(m+1)}$, since we obtain $Z^{(m+1)}$ by doing some of the sums in $Z^{(m)}$. Hence

$$N^{(m)} f^{(m)}(h^{(m)}, t^{(m)}) = N^{(m+1)} f^{(m+1)}(h^{(m+1)}, t^{(m+1)})$$

or,

$$f^{(m)}(h^{(m)}, t^{(m)}) = b^{-d} f^{(m+1)}(h^{(m+1)}, t^{(m+1)})$$

The only reason h and t vary under RG is due to B . We assume

$$t^{(m+1)} = b^A t^{(m)}$$

and

$$h^{(m+1)} = b^B t^{(m)}$$

So (and dropping the annoying superscripts from t and h) we have

$$f^{(m)}(h, t) = b^{-d} f^{(m+1)}(hb^B, tb^A)$$

Now we use Kadanoff's bold, and unfortunately incorrect, assertion that $f^{(m)}$ and $f^{(m+1)}$ are both simply the Ising model. Hence

$$f(h, t) = b^{-d} f(hb^B, tb^A)$$

This is one form of the scaling assumption, which is due to Widom. Widom argued that the free energy density was a *homogeneous* function. This allows for the possibility of nontrivial exponents (albeit built into the description), and one can derive the scaling relations.

For example, the factor b is not fixed by our arguments. Let

$$tb^A \equiv 1$$

Then

$$f(h, t) = t^{d/A} f(ht^{-B/A}, 1)$$

or,

$$f(h, t) = t^{d/A} \tilde{f}(ht^{-B/A})$$

which is of the same form as the scaling relation previously obtained for F .

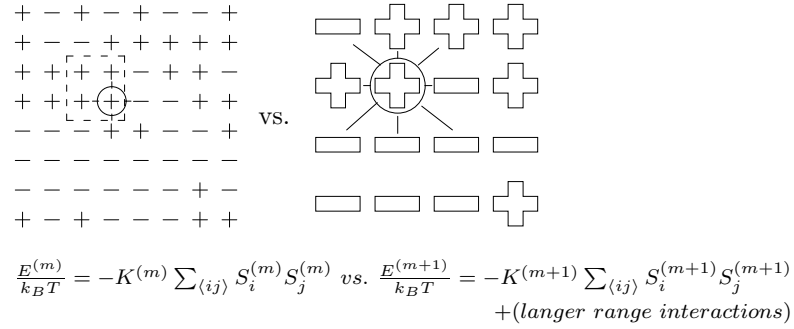


Figure 17.4:

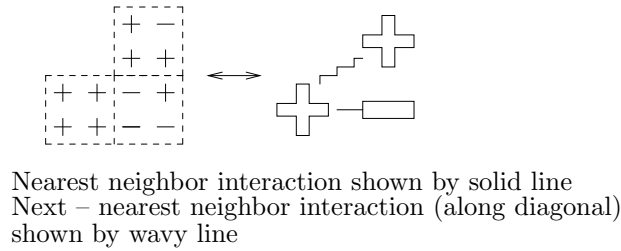


Figure 17.5:

Kadanoff also presented results for the correlation function, but we'll skip that.

Let us consider the interactions between spins again—before and after application of RG. For simplicity, let $H = 0$. The original system has interactions to its nearest neighbors as drawn in fig. 17.4. The blocked spin treats 4 spins as interacting together, with identical strength with blocked spins. The nearest neighbor interaction is as a cartoon, in fig. 17.5. Note that the direct part of the nearest-neighbor interaction to the right in the renormalized system has 4 spins interacting equivalently with 4 other spins. There are 16 cases, which are replaced by one global effective interaction. Those 16 cases have: All $K^{(m+1)}$ contributions treated equally in block spin near neighbor interaction.

$$\left\{ \begin{array}{l} 2 \text{ interactions separated by 1 lattice constant} \\ 6 \text{ interactions separated by 2 lattice constants} \\ 8 \text{ interactions separated by 3 lattice constants} \\ 2 \text{ interactions separated by 4 lattice constants} \end{array} \right.$$

Let's compare this to the *next*-nearest neighbor interaction, along the diagonals. There are again 16 cases which have: All next nearest neighbor contributions

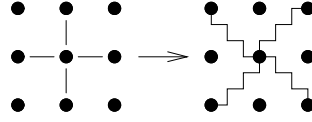


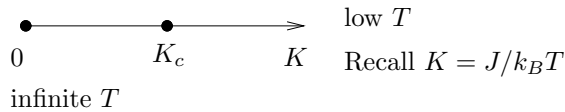
Figure 17.6: Nearest neighbor interaction vs. Next-nearest neighbor

treated equally in block spin representation.

$$\left\{ \begin{array}{l} 1 \text{ interactions separated by 2 lattice constants} \\ 4 \text{ interactions separated by 3 lattice constants} \\ 6 \text{ interactions separated by 4 lattice constants} \\ 4 \text{ interactions separated by 5 lattice constants} \\ 1 \text{ interactions separated by 6 lattice constants} \end{array} \right.$$

Let us call the block-spin next-nearest neighbor coupling $K_{nn}^{(m+1)}$, which lumps all of this into one effective interaction. It is clear from looking at the interactions underlying the block spins that $K^{(m+1)}$ (nearest neighbor) and $K_{nn}^{(m+1)}$ (next nearest neighbor) are quite similar, although it looks like $K_{nn}^{(m+1)} < K^{(m+1)}$. However, Kadanoff's step of $K_{nn}^{(m+1)} \equiv 0$ looks *very* drastic, in fact with the benefit of hindsight, we know it is too drastic.

One can actually numerically track what happens to the Ising model under the block-spin renormalization group. It turns out that longer-range interactions between two spins are generated (as we have described in detail above) as well as three spin interactions (like $\sum S_i S_j S_k$), and other such interactions. Although the interactions remain short ranged, they become quite complicated. It is useful to consider the flow of the Energy of a state under application of RG. For $H = 0$, for the Ising model, the phase diagram is as in fig. 17.7. For

Figure 17.7: Usual Ising model $H = 0$

the Ising model *with only* next nearest neighbor interactions ($J \equiv 0$) one gets the same phase diagram at $H = 0$ (fig. 17.8). The phase diagram for an Ising

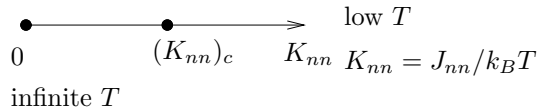


Figure 17.8: Ising model, no near neighbor interaction

model with *both* K and K_{nn} namely,

$$\frac{E}{k_B T} = -K \sum_{\langle ij \rangle} S_i S_j - K_{nn} \sum_{\langle\langle ij \rangle\rangle} S_i S_j$$

for $H = 0$, is drawn in fig. 17.9. The flow under RG is hence as drawn in

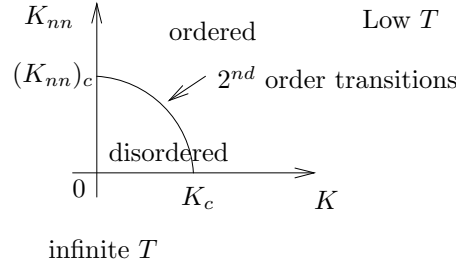


Figure 17.9:

fig. 17.10. Of course there are many other axes on this graph corresponding to

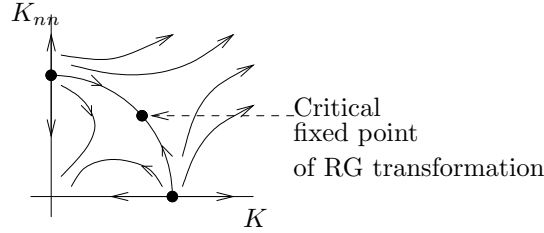


Figure 17.10:

different interactions. One should also note that the interactions developed by the RG, depend on the RG. For example, $b = 3$ would give a different sequence of interactions, and a different fixed result, than $b = 2$. But, it turns out, the exponents would all be the same. This is an aspect of universality: All models with the same symmetry and spatial dimension share the same critical exponents.

The formal way to understand these ideas was figured out by Wilson. The RG transform takes us from one set of coupling constants $K^{(m)}$ to another $K^{(m+1)}$ via a length scale change of b entering a recursion relation: $K^{(m+1)} = R(k^{(m)})$.

At the fixed point

$$K^* = R(k^*)$$

Let us now imagine we are close enough to the fixed point that we can linearize around it. In principle this is clear. In practice, because of all the coupling constants generated on our way to the fixed point, it is a nightmare. Usually it is done by numerical methods, or by the $\epsilon = 4 - d$ expansion described later.

Using subscripts, which correspond to different coupling constants

$$K_\alpha^{(m+1)} - K_\alpha^* = R(K^{(m)}) - R(K^*)$$

or,

$$\delta K_\alpha^{(m+1)} = \sum_\beta M_{\alpha\beta} \delta K_\beta^{(m)}$$

where

$$M_{\alpha\beta} = (\partial R_\alpha / \partial K_\beta)^*$$

is the linearized RG transform through which “ b ” appears, and

$$\delta K = K - K^*$$

All that is left to do is linear algebra to diagonalize $M_{\alpha\beta}$.

Before doing this, let us simply pretend that $M_{\alpha\beta}$ is *already* diagonal. Then we can see everything clearly. If

$$\begin{bmatrix} \lambda^{(1)} & & & 0 \\ & \lambda^{(2)} & & \\ & & \cdots & \\ 0 & & & \cdots \end{bmatrix}$$

the transform becomes

$$\delta K_\alpha^{(m+1)} = \lambda^{(\alpha)} \delta K_\alpha^{(m)}$$

The form of $\lambda^{(\alpha)}$ is easy to get. It is clear that acting with the RG on scale b , followed by acting again by b' , is the same as acting once by $b \cdot b'$. Hence:

$$M(b)M(b') = M(bb')$$

somewhat schematically, or

$$\lambda(b)\lambda(b') = \lambda(bb')$$

The solution of this is

$$\lambda = b^y$$

The renormalization group transformation becomes

$$\delta K_\alpha^{(m+1)} = b^{y_\alpha} \delta K_\alpha^{(m)}$$

If $y_\alpha > 0$, δK_α is *relevant*, getting bigger with RG.

If $y_\alpha = 0$, δK_α is *marginal*, staying unchanged with RG.

If $y_\alpha < 0$, δK_α is *irrelevant*, getting smaller with RG.

The singular part of the free energy near the fixed point satisfies

$$f^*(\delta K_1, \delta K_2, \dots) = b^{-d} f^*(b^{y_1} \delta K_1, b^{y_2} \delta K_2, \dots)$$

This is the homogeneous scaling form, where we now know how to calculate the y_α 's (in principle), and we know why (in principle) only two exponents are necessary to describe the Ising model (presumably the remainder are irrelevant).

In practice, the problem is to find the fixed point. A neat trick is the $\epsilon = 4 - d$ expansion. From our discussion above about the Ginzburg criteria, mean-field theory works for $d > 4$ (it is marginal with logarithmic correlations in $d = 4$ itself). In fact we then expect there to be a very easy-to-find, easy-to-fiddle-with, fixed point in $d \geq 4$. If we are lucky, this fixed point can be analytically continued to $d < 4$ in a controllable $\epsilon = 4 - d$ expansion. We will do this below. I say “lucky” because the fixed point associated with mean-field theory is not always related to the fixed point we are after (usually it is then a fixed point related to a new phase transition at even higher dimensions, i.e., it corresponds to a lower critical dimension).

Finally let us now return to the linear algebra:

$$\delta K_\alpha^{(m+1)} = \sum_\beta M_{\alpha\beta} \delta K_\beta^{(m)}$$

since $M_{\alpha\beta}$ is rarely diagonal. We will assume it is symmetric to make things easier. (If not, we have to define right and left eigenvalues. This is done in Goldenfeld's book.) We define an orthonormal complete set of eigenvectors by

$$\sum_\beta M_{\alpha\beta} e_\beta^{(\gamma)} = \lambda^{(\gamma)} e_\alpha^{(\gamma)}$$

where $\lambda^{(\gamma)}$ are eigenvalues and e are the eigenvectors. We expand δK via

$$\delta K_\alpha = \sum_\gamma \delta \hat{K}_\gamma e_\alpha^{(\gamma)}$$

or,

$$\delta \hat{K}_\gamma = \sum_\alpha \delta K_\alpha e_\alpha^{(\gamma)}$$

where $\delta \hat{K}$ are the “scaling fields” — a linear combination of the original interactions. Now we act on the linearized RG transform with $\sum_\alpha e_\alpha^{(\gamma)}(\dots)$. This gives

$$\sum_\alpha e_\alpha^{(\gamma)}(\delta K_\alpha^{(m+1)}) = \sum_\alpha e_\alpha^{(\gamma)} \left(\sum_\beta M_{\alpha\beta} \delta K_\beta^{(m)} \right)$$

It is a little confusing, but the superscripts on the e 's have nothing to do with the superscripts on the δK 's. We get

$$\delta \hat{K}_\gamma^{(m+1)} = \lambda^{(\gamma)} \sum_\beta e_\beta^{(\gamma)} \delta K_\beta^{(m)}$$

so,

$$\delta \hat{K}_\gamma^{(m+1)} = \lambda^{(\gamma)} \delta \hat{K}_\gamma^{(m)}$$

This is what we had (for δK 's rather than $\delta \hat{K}$'s) a few pages back. Everything follows the same from there.

17.1 ϵ -Expansion RG

Reconsider

$$F[\psi] = \int d^d x \left\{ \frac{K}{2} (\nabla \psi)^2 + \frac{r}{2} \psi^2 + \frac{u}{4} \psi^4 \right\}$$

near T_c , where K and u are constants and

$$r \propto T - T_c$$

Near T_c we can conveniently consider

$$\frac{F}{k_B T} = \int d^d x \left\{ \frac{K}{2} (\nabla \psi)^2 + \frac{r}{2} \psi^2 + \frac{u}{4} \psi^4 \right\}$$

by describing the $k_B T$, which is nonsingular, within K , r , and u . We can remove the constant K by absorbing it into the scale length or letting $r \rightarrow \frac{r}{K}$, $F \rightarrow \frac{F}{K}$ etc., so that

$$\frac{F}{k_B T} = \int d^d x \left\{ \frac{1}{2} (\nabla \psi)^2 + \frac{r}{2} \psi^2 + \frac{u}{4} \psi^4 \right\}$$

Now we can easily construct dimensional arguments for all coefficients. Clearly

$$\left[\frac{F}{k_B T} \right] = 1$$

where $[...]$ means the dimension of the quantity. Hence,

$$\left[\int d^d x (\nabla \psi)^2 \right] = 1$$

or

$$[L^{d-2} \psi^2] = 1, \text{ since } [x] = L$$

or,

$$[\psi] = \frac{1}{L^{\frac{d}{2}-1}}$$

Similarly,

$$\left[\int d^d x \frac{r}{2} \psi^2 \right] = 1$$

or

$$\left[L^d r \frac{1}{L^{d-2}} \right] = 1$$

or

$$[r] = \frac{1}{L^2}$$

Finally,

$$\begin{aligned} \left[\int d^d x \frac{u}{4} \psi^4 \right] &= 1 \\ \text{so } \left[L^d u \frac{1}{L^{2d-4}} \right] &= 1 \\ \text{or } [u] &= \frac{1}{L^{4-d}} \end{aligned}$$

Eventually we will be changing scale by a length rescaling factor, so we need to find out how these quantities are affected by a simple change in scale. These are called *engineering* dimensions.

Note the following,

$$[\psi^2] = \frac{1}{L^{d-2}}$$

But at T_c

$$\langle \psi^2 \rangle = \frac{1}{x^{d-2+\eta}}$$

where η is nonzero! Also, since $r \propto T - T_c$, near T_c

$$[T - T_c] = \frac{1}{L^2}$$

but $[\xi] = L$, so

$$\xi = (T - T_c)^{-1/2}$$

near T_c . But

$$\xi = (T - T_c)^{-\nu}$$

So our results of scaling theory give

$$[\psi]_{after \text{ avg.}} = \frac{1}{L^{(d-2+\eta)/2}} = \frac{1}{L^{\frac{d}{2}-1}} \frac{1}{L^{\eta/2}}$$

and

$$[r]_{after \text{ avg.}} = \frac{1}{L^{1/\nu}} = \frac{1}{L^2} \frac{1}{L^{\frac{1}{\nu}-2}}$$

The extra length which appears on averaging is the lattice constant $a \approx 5\text{\AA}$. Hence

$$\langle \psi^2 \rangle = \frac{1}{x^{d-2+\eta}} a^\eta$$

and $\eta/2$ is called ψ 's anomalous dimension, while $1/\eta - 2$ is the anomalous dimension of r . By doing the averaging, carefully taking account of changes in scale by L , we can calculate ν and η as well as other quantities.

In this spirit, let us measure quantities in units of their engineering dimensions. Let

$$\psi' = \frac{\psi}{[\psi]} = \frac{\psi}{1/L^{\frac{d}{2}-1}}$$

and

$$x' = \frac{x}{[x]} = \frac{x}{L}$$

so that (dropping the $k_B T$ under F),

$$F = \int d^d x' \left\{ \frac{1}{2} (\nabla' \psi')^2 + \left(\frac{r}{2} L^2\right) \psi'^2 + \left(\frac{u}{4} L^{4-d}\right) \psi'^4 \right\}$$

we can of course let

$$r' = \frac{r}{[r]} = \frac{r}{1/L^2}$$

and

$$u' = \frac{u}{[u]} = \frac{u}{1/L^{4-d}} = u L^{4-d}$$

so we have

$$F = \int d^d x' \left\{ \frac{1}{2} (\nabla' \psi')^2 + \frac{r'}{2} \psi'^2 + \frac{u'}{4} \psi'^4 \right\}$$

Equivalently, since we are interested in the behaviour at T_c , let $rL^2 \equiv 1$ (choosing the scale by the distance from T_c), then

$$F = \int d^d x' \left\{ \frac{1}{2} (\nabla' \psi')^2 + \frac{1}{2} \psi'^2 + \frac{u''}{4} \psi'^4 \right\}$$

where

$$u'' = u \left(\frac{1}{r}\right)^{\frac{4-d}{2}}$$

Note that u' is negligible as $L \rightarrow \infty$ for $d > 4$, and *more revealingly*, u'' is negligible as $T \rightarrow T_c$ for $d > 4$.

This is the Ginzburg criterion again, and note the equivalence between $L \rightarrow \infty$ and $T \rightarrow T_c$ (i.e., $r \rightarrow 0$).

It is worthwhile to consider how the coefficient of the quartic term (with its engineering dimension removed) varies as scale L is varied. Consider the "beta function":

$$\begin{aligned} -L \frac{du'}{dL} &= -(4-d) L^{4-d} u \\ &= -(4-d) u' \end{aligned}$$

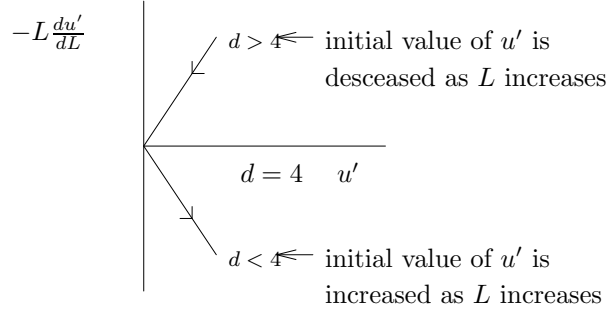


Figure 17.11:

The direction of the arrows is from

$$\frac{du'}{d \ln L} = -(d-4)u' \quad , \quad d > 4$$

$$\frac{du'}{d \ln L} = (4-d)u' \quad , \quad d < 4$$

Note that $u' = 0$ is a *stable* fixed point for $d > 4$ and an *unstable* fixed point for $d < 4$. This sort of picture implicitly assumes u' is small. The renormalization group will give us a more elaborate flow similar to

$$-L \frac{du'}{dL} = -(4-d)u' + \mathcal{O}(u'^2)$$

so that the flow near the fixed pt. determines the dimension of a quantity.

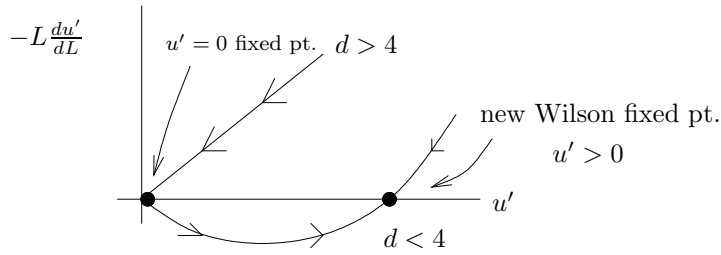


Figure 17.12:

Clearly

$$-L \frac{du'}{dL} = (\text{Dim. of } u) u'$$

and similarly

$$-L \frac{dr'}{dL} = (\text{Dim. of } r) r'$$

So the anomalous dimension can be determined by the “flow” near the fixed point.

The analysis above is OK for $d > 4$, and only good for u' close to zero for $d < 4$. To get some idea of behavior below 4 dimensions, we stay close to $d = 4$ and expand in $\epsilon = 4 - d$. (See note below)

So consider

$$F = \int d^d x \left\{ \frac{K}{2} (\nabla \psi)^2 + \frac{r}{2} \psi^2 + \frac{u}{4} \psi^4 \right\}$$

we will find

	$\epsilon = 1 \ (d = 3)$	Exact $(d = 3)$
$\alpha = \epsilon/6$	0.17	0.125
$\beta = \frac{1}{2} - \frac{\epsilon}{6}$	0.33	0.313
$\gamma = 1 + \epsilon/6$	4	5.0
$\delta = 3 + \epsilon$	1.17	1.25
$\nu = \frac{1}{2} + \epsilon/12$	0.58	0.638
$\eta = 0 + \mathcal{O}(\epsilon^2)$	0	0.041

(Note: Irrelevant Couplings)

Consider

$$\int d^d x \left[\frac{1}{2} (\nabla \psi)^2 + \frac{r}{2} \psi^2 + \frac{u}{4} \psi^4 + \frac{v}{6} \psi^6 + \frac{w}{8} \psi^8 \right]$$

as before,

$$\begin{aligned} [\psi] &= 1/L^{\frac{d}{2}-1} \\ [\nu] &= 1/L^2 \\ [u] &= 1/L^{4-d} \end{aligned}$$

But similarly,

$$\begin{aligned} [v] &= 1/L^{2(3-d)} \\ [w] &= 1/L^{3(2\frac{2}{3}-d)} \end{aligned}$$

so,

$$F = \int d^d x' \left[\frac{1}{2} (\nabla' \psi')^2 + \left(\frac{r}{2} L^2\right) \psi'^2 + \left(\frac{u}{4} L^{4-d}\right) \psi'^4 + \frac{v L^{2(3-d)}}{6} \psi'^6 + w L^{3(2\frac{2}{3}-d)} \psi'^8 \right]$$

Hence, v , w look quite unimportant (at least for the Gaussian fixed point) near $d = 4$. This is called power counting. For the remainder of these notes we will simply assume v , w are irrelevant always.)

We will do a scale transformation in Fourier space,

$$\psi(x) = \int_{|k| < \Lambda} \frac{d^d k}{(2\pi)^d} e^{i\vec{k} \cdot \vec{x}} \hat{\psi}_k$$

where (before we start) $\Lambda = \pi/a$, since we shall need the lattice constant to get anomalous dimensions. We will continually apply the RG to integrate out

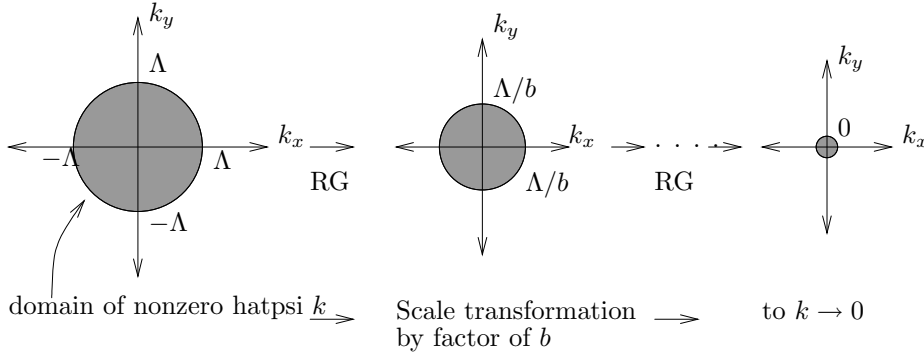


Figure 17.13:

small length scale behavior, which, in Fourier space, means integrating out large wavenumber behaviour. Schematically shown in fig. 17.13. Let

$$\begin{array}{ccccc}
 \psi(x) & = & \bar{\psi}(x) & + & \delta\psi(x) \\
 \text{all degrees of freedom} & & \text{degrees of freedom} & & \text{degrees of freedom} \\
 \text{after several} & & \text{for "slow" modes on} & & \text{for "fast" modes on} \\
 \text{transformations} & & |k| < \Lambda/b & & \Lambda/b < |k| < \Lambda \\
 & & |k| < \Lambda & & \text{to be integrated out}
 \end{array}$$

Also, for convenience,

$$\frac{\Lambda}{b} = \Lambda - \delta\Lambda$$

So that the scale factor

$$b = \frac{\Lambda}{\Lambda - \delta\Lambda} \geq 1$$

Then we explicitly have

$$\bar{\psi}(x) = \int_{|k| < \Lambda - \delta\Lambda} \frac{d^d k}{(2\pi)^d} e^{i\vec{k} \cdot \vec{x}} \hat{\psi}_k$$

and

$$\delta\psi(x) = \int_{\Lambda - \delta\Lambda < |k| < \Lambda} \frac{d^d k}{(2\pi)^d} e^{i\vec{k} \cdot \vec{x}} \hat{\psi}_k$$

Schematically, refer to fig. 17.14. This is called the “momentum-shell” RG. We do this in momentum space rather than real space because it is convenient for factors of

$$\int \underbrace{d^d x (\nabla \psi)^2}_{\text{coupling different } \psi(x)\text{'s}} = \int \underbrace{\frac{d^d k}{(2\pi)^d} k^2 |\hat{\psi}_k|^2}_{\text{fourier modes uncoupled}}$$

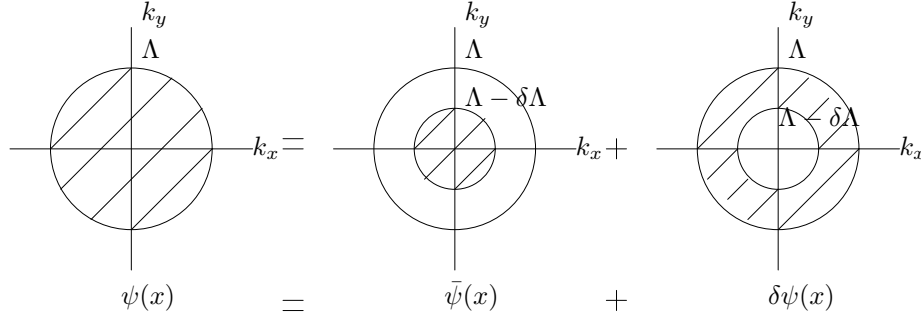


Figure 17.14:

which are trivial in momentum space.

The size of the momentum shell is

$$\int_{\Lambda - \delta\Lambda < |k| < \Lambda} \frac{d^d k}{(2\pi)^d} = a\Lambda^{d-1}\delta\Lambda$$

where a is a number that we will later set to unity for convenience, although $a(d=3) = 1/2\pi^2$, $a(d=2) = 1/2\pi$. Let us now integrate out the modes $\delta\psi$

$$\underbrace{e^{-\bar{F}[\bar{\psi}]}}_{\text{scales } |k| < \Lambda - \delta\Lambda} \equiv \sum_{\substack{\text{States } \delta\psi \\ \text{shell scales}}} \underbrace{e^{-F[\bar{\psi} + \delta\psi]}}_{\text{scales } |k| < \Lambda}$$

Essentially we are — very slowly — solving the partition function by bit-by-bit integrating out degrees of freedom for large k .

We will only do this once, and thus obtain recursion relations for how to change scales.

$$F[\bar{\psi} + \delta\psi] = \int d^d x \left\{ \frac{1}{2} (\nabla(\bar{\psi} + \delta\psi))^2 + \frac{r}{2} [\bar{\psi} + \delta\psi]^2 + \frac{u}{4} [\bar{\psi} + \delta\psi]^4 \right\}$$

We will expand this out, *ignoring odd terms* in $\bar{\psi}$ (which must give zero contribution since they change the symmetry of F), *constant terms* (which only shift the zero of the free energy), and terms of order *higher than quartic* like ψ^6 and ψ^8 (which are irrelevant near $d=4$ by power counting as we did above).

In that case

$$\begin{aligned}
\frac{1}{2}[\nabla(\bar{\psi} + \delta\psi)]^2 &= \frac{1}{2}(\nabla\bar{\psi})^2 + \frac{1}{2}(\nabla\delta\psi)^2 + \text{odd term} \\
\frac{r}{2}[\bar{\psi} + \delta\psi]^2 &= \frac{r}{2}\bar{\psi}^2 + \frac{r}{2}(\delta\psi)^2 + \text{odd term} \\
\frac{u}{4}[\bar{\psi} + \delta\psi]^4 &= \frac{u}{4}\bar{\psi}^4[1 + \frac{\delta\psi}{\bar{\psi}}]^4 \\
&= \frac{u}{4}\bar{\psi}^4[1 + 4\frac{\delta\psi}{\bar{\psi}} + 6\frac{(\delta\psi)^2}{\bar{\psi}^2} + 4\frac{(\delta\psi)^3}{\bar{\psi}^3} + \frac{(\delta\psi)^4}{\bar{\psi}^4}] \\
&= \frac{u}{4}\bar{\psi}^4 + \frac{3u}{2}\bar{\psi}^2(\delta\psi)^2 + \mathcal{O}(\delta\psi)^4 + \text{odd term}
\end{aligned}$$

We will neglect $(\delta\psi)^4$, since these modes have a small contribution to the $k = 0$ behavior. This can be self-consistently checked. Hence

$$F[\bar{\psi} + \delta\psi] = \underbrace{F\{\bar{\psi}\}}_{|k| < \Lambda - \delta\Lambda} + \int d^d x [\frac{1}{2}(\nabla\delta\psi)^2 + \frac{1}{2}(r + 3u\bar{\psi}^2)(\delta\psi)^2]$$

but $\delta\psi = \int_{\Lambda - \delta\Lambda < |\vec{k}| < \Lambda} \frac{d^d k}{(2\pi)^d} e^{i\vec{k} \cdot \vec{x}} \hat{\psi}_k$, so

$$(\nabla\delta\psi)^2 = \Lambda^2(\delta\psi)^2$$

for a thin shell, and

$$F[\bar{\psi} + \delta\psi] = F[\bar{\psi}] + \int d^d x \frac{1}{2}(\Lambda^2 + r + 3u\bar{\psi}^2)(\delta\psi)^2$$

and we have,

$$e^{-\bar{F}[\bar{\psi}]} = e^{-F[\bar{\psi}]} \sum_{\text{States } \delta\psi} e^{-\int d^d x \frac{\Lambda^2 + r + 3u\bar{\psi}^2}{2}(\delta\psi)^2}$$

where

$$\sum_{\text{States } \delta\psi} = \prod_{\text{all } \delta\psi \text{ states}} \int d\delta\psi(x)$$

The integrals are trivially Gaussian in $\delta\psi$, so

$$e^{-\bar{F}[\bar{\psi}]} = e^{-F[\bar{\psi}]} \prod_x \prod_{\text{total no. of } \delta\psi \text{ states}} \sqrt{\frac{2\pi}{\Lambda^2 + r + 3u\bar{\psi}^2}}$$

we know the total number of $\delta\psi$ states because we know the total number of modes in Fourier space in the shell are

$$\int_{\Lambda - \delta\Lambda < |\vec{k}| < \Lambda} \frac{d^d k}{(2\pi)^d} = a\Lambda^{d-1}\delta\Lambda$$

so

$$\prod_x \prod_{\text{total no.}} (...) \rightarrow e^{(\int dx)(a\Lambda^{d-1}\delta\Lambda) \ln (...)}$$

and

$$e^{-\bar{F}[\psi]} = e^{-F[\bar{\psi}]} e^{-\int d^d x \ a\Lambda^{d-1}\delta\Lambda \frac{1}{2} \ln (\Lambda^2 + r + 3u\bar{\psi}^2) + \text{const.}}$$

Now consider

$$\frac{1}{2} \ln [\Lambda^2 + r + 3u\bar{\psi}^2] = \frac{1}{2} \ln \left(1 + \frac{r}{\Lambda^2} + \frac{3u}{\Lambda^2} \bar{\psi}^2\right) + \text{const.}$$

we expand using $\ln(1 + \epsilon) = \epsilon - \frac{\epsilon^2}{2} + \frac{\epsilon^3}{3}$

$$\begin{aligned} &= \frac{1}{2} \left\{ \frac{r}{\Lambda^2} + \frac{3u}{\Lambda^2} \bar{\psi}^2 - \frac{1}{2} \left(\frac{r}{\Lambda^2} + \frac{3u}{\Lambda^2} \bar{\psi}^2 \right)^2 + \dots \right\} \\ &= \frac{1}{2} \left\{ \cancel{\frac{r}{\Lambda^2}} + \frac{3u}{\Lambda^2} \bar{\psi}^2 - \frac{1}{2} \left(\cancel{\frac{r}{\Lambda^2}} \right)^2 - \frac{\cancel{3ur}}{\Lambda^2} \bar{\psi}^2 - \frac{9u^2}{2\Lambda^4} \bar{\psi}^4 + \dots \right\} \\ &= \frac{1}{2} \left(\frac{3u}{\Lambda^2} - \frac{3ur}{\Lambda^4} \right) \bar{\psi}^2 + \frac{1}{4} \left(\frac{-9u^2}{\Lambda^4} \right) \bar{\psi}^4 \end{aligned}$$

So that

$$\bar{F}[\bar{\psi}] = F[\bar{\psi}] + \int d^d x \left\{ \frac{1}{2} \left(\frac{3u}{\Lambda^2} - \frac{3ur}{\Lambda^4} \right) (a\Lambda^{d-1}\delta\Lambda) \bar{\psi}^2 + \frac{1}{4} \left(\frac{-9u^2}{\Lambda^4} \right) (a\Lambda^{d-1}\delta\Lambda) \bar{\psi}^4 \right\}$$

so,

$$\bar{F}[\bar{\psi}] = \int d^d x \left\{ \frac{1}{2} (\nabla \bar{\psi})^2 + \frac{\bar{r}}{2} \bar{\psi}^2 + \frac{\bar{u}}{4} \bar{\psi}^4 \right\}$$

where,

$$\begin{aligned} \bar{r} &= r + \left(\frac{3u}{\Lambda^2} - \frac{3ur}{\Lambda^4} \right) (a\Lambda^{d-1}\delta\Lambda) \\ \bar{u} &= u + \left(\frac{-9u^2}{\Lambda^4} \right) (a\Lambda^{d-1}\delta\Lambda) \end{aligned}$$

But

$$\bar{r} = r(\Lambda - \delta\Lambda) \quad \text{etc,}$$

so

$$\frac{r(\Lambda - \delta\Lambda) - r(\Lambda)}{\delta\Lambda} = \left(\frac{3u}{\Lambda^2} - \frac{3ur}{\Lambda^4} \right) a\Lambda^{d-1}$$

and

$$\frac{u(\Lambda - \delta\Lambda) - u(\Lambda)}{\delta\Lambda} = \left(\frac{-9u^2}{\Lambda^4} \right) a\Lambda^{d-1}$$

or,

$$\begin{aligned} -\frac{1}{a} \frac{\partial r}{\partial \Lambda} &= \left(\frac{3u}{\Lambda^2} - \frac{3ur}{\Lambda^4} \right) \Lambda^{d-1} \\ -\frac{1}{a} \frac{\partial u}{\partial \Lambda} &= \left(\frac{-9u^2}{\Lambda^4} \right) \Lambda^{d-1} \end{aligned}$$

Set $a = 1$ (or let $u \rightarrow u/a$), and set $\Lambda = 1/L$, to obtain

$$\begin{aligned} -L \frac{\partial r}{\partial L} &= -3uL^{2-d} + 3urL^{4-d} \\ -L \frac{\partial u}{\partial L} &= 9u^2L^{4-d} \end{aligned}$$

We need to include the trivial change in length due to the overall rescaling of space, with engineering dimensions $[r] = 1/L^2$, $[u] = 1/L^{4-d}$, by letting $r' = rL^2$, $u' = uL^{4-d}$. Note

$$\frac{dr}{dL} = \frac{d(r'/L^2)}{dL} = \left(\frac{dr'}{dL} - 2\frac{r'}{L}\right) \frac{1}{L^2}$$

so

$$-L \frac{dr}{dL} = (-L \frac{dr'}{dL} + 2r') \frac{1}{L^2}$$

and

$$-L \frac{du}{dL} = (-L \frac{du'}{dL} + (4-d)u') \frac{1}{L^{4-d}}$$

But

$$-3uL^{2-d} = -3u' \frac{L^{2-d}}{L^{4-d}} = -3u' \frac{1}{L^2}$$

and

$$-3urL^{4-d} = -3u'r' \frac{L^{4-d}}{L^{4-d}L^2} = -3u'r' \frac{1}{L^2}$$

and

$$9u^2L^{4-d} = 9u'^2 \frac{L^{4-d}}{(L^{4-d})^2} = 9u'^2 \frac{1}{L^{4-d}}$$

Hence we obtain

$$-L \frac{dr'}{dL} = -2r' - 3u' + 3u'r'$$

and

$$-L \frac{du'}{dL} = -(4-d)u' + 9u'^2$$

Which are the recursion relations, to first order. Evidently there is still a trivial Gaussian fixed point $r^* = u^+ = 0$, but there is a nontrivial “Wilson–Fisher” fixed point which is stable for $d < 4$! From the u equation (using $\epsilon = 4 - d$)

$$u^* = \frac{\epsilon}{9}$$

and in the r equation (using $u^* \sim r^* \sim \epsilon$ with $\epsilon \ll 1$), $2r^* = -3u^*$, so

$$r^* = -\frac{\epsilon}{6}$$

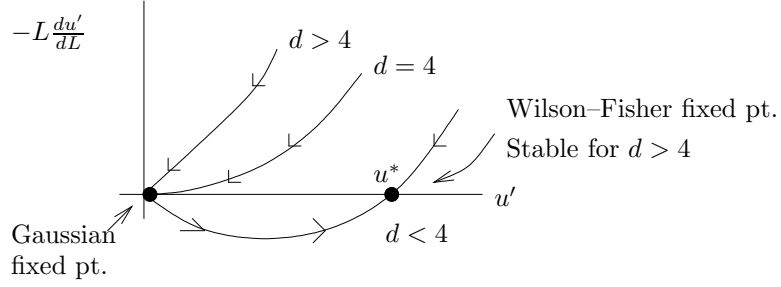


Figure 17.15:

Linearizing around this fixed point gives the dimensions of r and u as that fixed point is approached.

Let $\delta r = r' - r^*$, $\delta u = u' - u^*$. In the u equation,

$$\begin{aligned} -L \frac{d\delta u}{dL} &= -\epsilon u^* - \epsilon \delta u + 9(u^*)^2 \left(1 + \frac{2\delta u}{u^*}\right) \\ &= [18u^* - \epsilon] \delta u \\ -L \frac{d\delta u}{dL} &= \epsilon \delta u \quad (\text{Wilson-Fisher fixed pt.}) \end{aligned}$$

So the dimension of u is ϵ .

In the r equation,

$$-L \frac{d\delta r}{dL} = -2r^* - 3u^* - 2\delta r + 3u^* \delta r$$

but $\delta r \gg \delta u$ as $L \rightarrow \infty$, so,

$$-L \frac{d\delta r}{dL} = (-2 + 3u^*) \delta r$$

or,

$$-L \frac{d\delta r}{dL} = -2\left(1 - \frac{\epsilon}{6}\right) \delta r \quad \text{Wilson-Fisher fixed pt. (stable } d < 4)$$

So the dimension of r is $-2(1 - \epsilon/6)$.

Similarly, near the Gaussian fixed pt. where $u^* = r^* = 0$, we have (as we did before)

$$\begin{aligned} -L \frac{d\delta u}{dL} &= -\epsilon \delta u \\ -L \frac{d\delta r}{dL} &= -2\delta r \end{aligned}$$

Giving the dimensions of r and u there.

To get the physical exponents, recall

$$[r] = 1/L^{1/\nu}$$

Hence

$$\begin{aligned} \nu &= \frac{1}{2}, & \text{Gaussian fixed pt.} \\ \nu &= \frac{1}{2}(1 + \frac{\epsilon}{6}), & \text{Wilson-Fisher fixed pt.} \end{aligned}$$

To get another exponent we use

$$[\psi] = \frac{1}{L^{\frac{d}{2}-1+\eta/2}}$$

or more precisely

$$[\psi] = \frac{1}{L^{\frac{d}{2}-1+\eta/2}} f(rL^{1/\nu}, uL^{-y_u})$$

where

$$y_u = \begin{cases} -\epsilon, & \text{Gaussian} \\ \epsilon, & \text{Wilson-Fisher} \end{cases}$$

is the dimension of u . Choosing the scale

$$rL^{1/\nu} \equiv 1$$

gives

$$[\psi] = r^{\frac{\nu}{2}(d-2+\eta)} f(1, ur^{\nu y_u})$$

for the Wilson-Fisher fixed point this is straight forward. Since $\nu y_u > 0$, the term $ur^{\nu y_u}$ is irrelevant as $r \rightarrow 0$. Hence

$$\begin{aligned} [\psi] &= r^{\frac{\nu}{2}(d-2+\eta)} \cancel{f(1,0)} \rightarrow^{const.} \\ &\equiv r^\beta \end{aligned}$$

Now it turns out $\eta = 0 + \mathcal{O}(\epsilon^2)$ (which can be calculated later via scaling relations), so,

$$\begin{aligned} \beta &= \frac{1}{2} \frac{1}{2} (1 + \frac{\epsilon}{6}) (2 - \epsilon) \\ &= \frac{1}{2} (1 + \frac{\epsilon}{6}) (1 - \frac{\epsilon}{2}) \end{aligned}$$

or

$$\beta = \frac{1}{2} (1 - \frac{\epsilon}{3}) \quad \text{Wilson-Fisher fixed pt.}$$

The Gaussian fixed point is more tricky (although academic since it is for $d > 4$). We again obtain

$$[\psi] = r^{\frac{\nu}{2}(d-2+\eta)} f(1, ur^{\nu y_u})$$

but now $\nu = \frac{1}{2}$, $y_u = -\epsilon = d - 4$, $\eta = 0$, so

$$[\psi] = r^{\frac{d-2}{4}} f(1, ur^{\frac{d-4}{2}})$$

This looks like $\beta = \frac{d-2}{4}$! Not the well-known $\beta = 1/2$. The problem is that u is irrelevant, but dangerous, for $T < T_c$ (where ψ is nonzero). Without this quartic term, the free energy is ill defined. For the equation above, $f(1, u \rightarrow 0)$ is singular. Its form is easy to see. By dimensions $\psi \sim \sqrt{r/u}$ below T_c , so

$$f(1, u \rightarrow 0) = \frac{1}{u^{1/2}}$$

and

$$\begin{aligned} [\psi]_{u \rightarrow 0} &= r^{\frac{d-2}{4}} \frac{1}{(ur^{\frac{d-4}{2}})^{1/2}} \\ &= \frac{r^{\frac{d-2}{4} - \frac{d-4}{4}}}{u^{1/2}} \\ &= \left(\frac{r}{u}\right)^{1/2} \end{aligned}$$

So, $\beta = 1/2$ for the Gaussian fixed point. Hence $d \geq 4$, the Gaussian fixed point is stable and $\alpha = 0$, $\beta = 1/2$, $\delta = 3$, $\gamma = 1$, $\eta = 0$, $\nu = \frac{1}{2}$ (The remainder are obtained *carefully* with the scaling relations).

For $d < 4$, the Wilson–Fisher fixed point is stable and since $\alpha + 2\beta + \gamma = 2$, $\gamma = \nu(2 - \eta)$, $\gamma = \beta(\delta - 1)$, $\nu d = 2 - \alpha$ we have:

	$\epsilon = 1$	Ising $d = 3$
$\alpha = \frac{\epsilon}{6}$	0.17	0.125...
$\beta = \frac{1}{2}(1 - \frac{\epsilon}{3})$	0.33	0.313...
$\gamma = 1 + \frac{\epsilon}{6}$	1.17	1.25...
$\delta = 3(1 + \frac{\epsilon}{3})$	4	5.0...
$\nu = \frac{1}{2}(1 + \frac{\epsilon}{12})$	0.58	0.638...
$\eta = 0 + \mathcal{O}(\epsilon^2)$	0	0.041...