

Advanced Equilibrium Statistical Mechanics

Martin Grant and Jorge Viñals
Physics Department, McGill University

© MG (2004); MG, JV (2009), version 1.0

Preface

These are lecture notes for PHYS 559, Advanced Statistical Mechanics, which we have taught at McGill for many years. The plan is to grow them into a book, or rather the first half of a book. This half is on equilibrium, the second half would be on dynamics.

The initial manuscript was based on handwritten notes which were heroically typed by Ryan Porter over the summer of 2004, and were later proof read and extended in 2009. We are then at version 1.0 as of now.

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. Our 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 Unfortunately, except for Goldenfeld’s book, they’re all a little out of date.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, just a note of caution: 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.

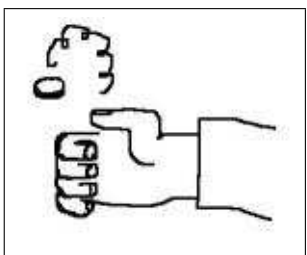
Contents

1	Introduction	1
1.1	Random variables	2
1.1.1	Gaussian distribution	4
1.2	Multivariate distributions	4
1.2.1	Gaussian distribution	5
2	Large systems, and independence of parts	7
2.1	Independence of parts	8
2.2	Self averaging	10
2.3	Central Limit Theorem	11
2.3.1	Probability of the sum of two independent variables	12
2.3.2	Central Limit Theorem	12
2.4	Correlations in Space	13
2.5	(*) Central Limit Theorem through explicit calculation of moments	15
3	Self similarity and Fractals	19
3.1	Von Koch Snowflake	20
3.2	Sierpinski Triangle	21
3.3	Correlations in Self-Similar Objects	22
4	Review of Statistical Mechanics and Fluctuations	25
4.1	Thermodynamics	25
4.2	Statistical Mechanics	27
4.2.1	Closed system	29
4.2.2	System at constant temperature	29
4.2.3	Other ensembles	30
4.3	Fluctuations	30
4.3.1	Simple mechanical example	32
4.4	Thermodynamic Fluctuations	33
4.4.1	Correlation functions and response functions	35
5	Fluctuations of Surfaces	37
5.1	Lattice models of fluctuating surfaces	39
5.2	A continuum model of fluctuating surface	40
5.2.1	Review of Fourier series	42
5.2.2	Height-height correlation function	44
5.2.3	Partition function, free energy, and surface tension	50
5.3	Impossibility of Phase Coexistence in $d=1$	52
5.4	Order of magnitude calculations in $d=3$	53
5.5	(*) A different view of gradient terms	55
6	Broken Symmetry and Correlation Functions	57
6.1	Goldstone's Theorem	59

7	Equilibrium correlation functions and Scattering	63
7.1	Measurement of $g(\vec{r})$ by diffraction	64
7.2	Scattering off a flat interface	67
7.3	Scattering from a line defect	69
7.4	Scattering from rough interfaces	71
7.5	(*) Scattering from many planar and randomly oriented surfaces	72
8	Fluctuations and Crystalline Order	77
8.1	Calculation of $\langle \rho \rangle$ and $\langle \rho(0)\rho(\vec{r}) \rangle$	79
8.1.1	Scattering intensity	81
9	Ising Model	85
9.1	(*) Lattice Gas Model	89
9.2	One dimensional Ising Model. Transfer matrix solution	89
9.2.1	No magnetic field	90
9.2.2	One dimensional Ising model in a magnetic field	92
9.2.3	Pair correlation function. Exact solution	92
9.3	Mean Field Theory	94
10	Continuum theories of the Ising model	101
10.1	Landau theory	103
10.2	Ornstein-Zernicke theory of correlations	106
10.3	Ginzburg criterion	108
11	Scaling Theory	111
11.1	Scaling With ξ	115
11.1.1	Derivation of exponent equalities	115
12	Renormalization Group	119
12.1	Decimation in one dimension	119
12.2	Block scaling renormalization group	121
12.2.1	Monte Carlo Renormalization Group	126
12.2.2	Scaling form, and critical exponents	127
12.3	ϕ^4 model: dimensional analysis	128
12.4	Momentum shell RG. Gaussian fixed point	129
12.5	(*) ϵ expansion	130
13	Master Equation and the Monte Carlo method	139
13.1	Other ensembles	146

Chapter 1

Introduction



Why is a coin toss random?

The response that you would receive in a Statistical Mechanics course is more or less as follows: individual states follow the “equal a priori probability” postulate. All many-body quantum states with the same energy (for a given fixed volume) have equal weight or probability when computing macroscopic averages.

This is a mouthful, but at least it gives the right answer. There are two microscopic states, namely *HEADS* and *TAILS*, and since they are energetically equivalent, they appear with equal probability after a coin toss. So the result of the toss is “random”.

However nice it is to get the right answer, and however satisfying it is to use this prescription in the many other problems which we will cover in this class, it 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:

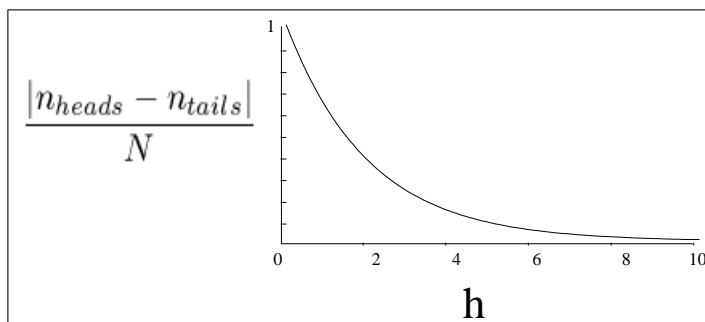


Figure 1.1: Bias of flips with height h

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:

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

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

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

Coin tosses appear random because we flip the coin to a much larger height than the correlation length ξ . Therefore 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 equivalence of the microscopic states (HEADS and TAILS). To stretch the argument a bit, we would call the preponderance of HEADS a “broken symmetry”; that is the macroscopic state (preponderance of HEADS) has a lower symmetry than the microscopic state (in which HEADS and TAILS are equally likely). This also happens in phase transitions, where $\xi/L = \mathcal{O}(1)$ because $\xi \approx L$ as $L \rightarrow \infty$, where L is the linear dimension 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 happened to Newton’s laws and the fact that one would expect to be able to calculate the trajectory of the coin exactly, given some initial condition? It is true that given an initial configuration, and a *precise* force, we can calculate the trajectory of the coin if we neglect inhomogeneities in the air. However, your thumb cannot apply a precise force for flipping to within one correlation length of height, and in practice computing the trajectory by using Newton’s laws is impossible. The analogous physical observation is that given a macroscopic state (say, at a fixed pressure or energy) it does not determine the underlying microscopic states. Many microscopic states are completely consistent with the specified macroscopic state.

These ideas, correlations on a microscopic scale, broken symmetries at a macroscopic scale, 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.

1.1 Random variables

In order to define a random variable x one needs to specify

1. The set of values that x can adopt (the sample space, or phase space later). An *event* is a subset of outcomes from the sample space. For example, in casting a die, the sample space is $\{1, 2, 3, 4, 5, 6\}$. A possible event E would be that the outcome be an even number $E = \{2, 4, 6\}$.

2. Each event (or single outcome) is assigned a probability $p(E)$ (or $p(x)$). For any event or value of x , we have,

$$p(x) \geq 0 \quad \int dx p(x) = 1.$$

and the additive property:

$$p(E_1 \text{ or } E_2) = p(E_1) + p(E_2) - p(E_1, E_2),$$

where $p(E_1, E_2)$ is the joint probability $p(E_1 \text{ and } E_2)$.

We interpret that $p(x)dx$ is the probability that the value of x lies between x and $x + dx$.

Of course, the integral above assumes that the sample space is continuous. In many cases, the values that x can adopt are discrete. In this case we could write,

$$p(x) = \sum_n p_n \delta(x - x_n),$$

so that

$$\int dx p(x) = \sum_n p_n = 1,$$

where p_n is the probability of occurrence of x_n .

Given a function of x , $f(x)$, we define its average as

$$\langle f(x) \rangle = \int f(x) p(x) dx.$$

In particular, we introduce the moments of x as

$$\mu_1 = \langle x \rangle = \int x p(x) dx$$

or the average of x ,

$$\mu_n = \langle x^n \rangle = \int x^n p(x) dx,$$

or the n^{th} moment of x . A special role is played by the variance

$$\sigma^2 = \langle (x - \langle x \rangle)^2 \rangle = \mu_2 - \mu_1^2.$$

The quantity σ is also known as the standard deviation.

Note, in passing, that not all probability distributions have variances or higher moments. See for example,

$$p(x) = \frac{1}{\pi} \frac{\gamma}{(x - a)^2 + \gamma^2}, \quad -\infty < x < \infty.$$

Although the distribution has a peak at $x = a$ and is symmetric around that value, not even $\langle x \rangle$ exists.

We next define the characteristic function $G(k)$ as the Fourier transform of the probability $p(x)$:

$$G(k) = \int dx e^{ikx} p(x) = \langle e^{ikx} \rangle. \quad (1.1)$$

Expanding the exponential in Taylor series leads to the following result:

$$G(k) = \sum_{m=0}^{\infty} \frac{(ik)^m}{m!} \mu_m. \quad (1.2)$$

It is also customary to define the cumulants of x through the following relation,

$$\ln G(k) = \sum_m \frac{(ik)^m}{m!} \langle\langle x^m \rangle\rangle,$$

where $\langle\langle x^m \rangle\rangle$ is the m^{th} cumulant of x . The lowest order cumulants are as follows,

$$\begin{aligned} \langle\langle x \rangle\rangle &= \mu_1 \\ \langle\langle x^2 \rangle\rangle &= \mu_2 - \mu_1^2 = \sigma^2 \\ \langle\langle x^3 \rangle\rangle &= \mu_3 - 3\mu_2\mu_1 + 2\mu_1^3 \\ \langle\langle x^4 \rangle\rangle &= \mu_4 - 4\mu_3\mu_1 - 3\mu_2^2 + 12\mu_2\mu_1^2 - 6\mu_1^4. \end{aligned}$$

1.1.1 Gaussian distribution

A very common distribution in what follows is the Gaussian. It is given by

$$p(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

With this definition, one has

$$\begin{aligned} \mu_{2n+1} &= 0, \\ \mu_{2n} &= (2n-1)!! = (2n-1)(2n-3)(2n-5)\dots 1, \\ \langle\langle x^m \rangle\rangle &= 0, \quad m > 2. \end{aligned}$$

1.2 Multivariate distributions

Most of what we will do in what follows, involves multidimensional (or infinite dimensional) spaces. Let x_1, x_2, \dots, x_r be random variables, one defines $p(x_1, x_2, \dots, x_r)$ as the joint probability distribution. This is the probability that all the variables, x_1, x_2, \dots, x_r have exactly a specified value. It is also useful to define the marginal probability as a partial integral,

$$p(x_1, \dots, x_s) = \int dx_{s+1} \dots dx_r p(x_1, x_2, \dots, x_r).$$

Of course, this is the probability that x_1, \dots, x_s have specified values, for *any* value of the remaining variables x_{s+1}, \dots, x_r .

A conditional probability can be defined $p(x_1, \dots, x_s | x_{s+1}, \dots, x_r)$ which is the probability that x_1, \dots, x_s have specified values, *given* also specified values of x_{s+1}, \dots, x_r . Bayes Theorem asserts that

$$p(x_1, \dots, x_r) = p(x_1, \dots, x_s | x_{s+1}, \dots, x_r) p(x_{s+1}, \dots, x_r).$$

We introduce now the concept of statistical independence. The sets $\{x_1, \dots, x_s\}$ and $\{x_{s+1}, \dots, x_r\}$ are said to be statistically independent if

$$p(x_1, \dots, x_r) = p(x_1, \dots, x_s) p(x_{s+1}, \dots, x_r),$$

that is, the joint probability factorizes. An equivalent statement is that

$$p(x_1, \dots, x_s | x_{s+1}, \dots, x_r) = p(x_1, \dots, x_s).$$

so that the distribution of x_1, \dots, x_s is not affected by prescribing different values of x_{s+1}, \dots, x_r .

Generalized moments are introduced now as

$$\langle x_1^m x_2^n \rangle = \int dx_1 \dots dx_r x_1^m x_2^n p(x_1, \dots, x_r)$$

Note that if x_1 and x_2 are independent, then

$$\langle x_1^m x_2^n \rangle = \langle x_1^m \rangle \langle x_2^n \rangle,$$

that is, multivariate moments factorize into products of the moments of the individual variables. One therefore defines correlations as deviations from statistical independence,

$$\langle \langle x_1 x_2 \rangle \rangle = \langle x_1 x_2 \rangle - \langle x_1 \rangle \langle x_2 \rangle.$$

The quantity $\langle \langle x_1 x_2 \rangle \rangle$, formally analogous to a cumulant, is identically zero if x_1 and x_2 are statistically independent. A nonzero cumulant indicates a statistical correlation between the two variables.

In general, if two variables x_1 and x_2 are independent, then

$$\langle \langle x_1^m x_2^n \rangle \rangle = 0 \quad m \neq 0 \quad \text{and} \quad n \neq 0.$$

1.2.1 Gaussian distribution

A multivariate Gaussian distribution is given by

$$p(x_1, \dots, x_r) = C e^{-\frac{1}{2} \sum A_{ij} x_i x_j - \sum B_i x_i},$$

where the matrix A has constant coefficients and is positive definite, and the vector B also has constant coefficients. The normalization constant is given by

$$C = (2\pi)^{-r/2} \sqrt{|A|} e^{-\frac{1}{2} B \cdot A^{-1} \cdot B}.$$

Averages are given by,

$$\langle x_i \rangle = - \sum (A^{-1})_{ij} B_j,$$

and second moments by,

$$\langle x_i x_j \rangle - \langle x_i \rangle \langle x_j \rangle = \langle \langle x_i x_j \rangle \rangle = (A^{-1})_{ij}.$$

Therefore, if the variables are uncorrelated, the matrix A^{-1} is diagonal.

An important result concerns the sum of random variables that are Gaussianly distributed. In general, if x_1, \dots, x_r are random variables, and we define $y = x_1 + \dots + x_r$, then

$$\langle \langle y \rangle \rangle = \sum_i \langle \langle x_i \rangle \rangle \quad \langle \langle y^2 \rangle \rangle = \sum_{ij} \langle \langle x_i x_j \rangle \rangle.$$

If, on the other hand, x_1, \dots, x_r are Gaussianly distributed, then $y = x_1 + \dots + x_r$ is also Gaussianly distributed. If, in addition, all the x_i are statistically independent, then

$$\langle y \rangle = \sum \langle x_i \rangle \quad \sigma_y^2 = \sum \sigma_{x_i}^2$$

Chapter 2

Large systems, and independence of parts

The systems that we will study are composed of N entities (particles, most often), with N very large (macroscopic). In order to accomplish a statistical description of such a system, we must first introduce the sample space. It is over this space (often called phase space or configuration space) that probabilities need to be defined, and averages calculated.

In order to put in perspective the difficulty of the task of just enumerating the states (sample points), and then computing averages, we recall a few results from Statistical Mechanics. In the canonical ensemble, the probability of being in state ν is

$$p_\nu = \frac{1}{Z} e^{-E_\nu/k_B T} \quad (2.1)$$

where the partition function is

$$Z = \sum_{\{\nu\}} e^{-E_\nu/k_B T}, \quad (2.2)$$

where the sum extends over all possible microscopic states of the system. The connection to thermodynamics is from the Helmholtz free energy $F = E - TS$ where S is the entropy, and

$$e^{-F/k_B T} = Z \quad (2.3)$$

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

$$Z = (e^{\frac{-f}{k_B T}})^N$$

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

$$\text{Number of states} = \sum_{\{\nu\}} \sim e^{\mathcal{O}(N)} \quad (2.4)$$

For example, for a set of discrete magnetic dipoles which each can have an orientation parallel anti parallel to a given direction, there are evidently

$$\text{dipole \#1} \times \text{dipole \#2} \dots \text{dipole \#N} = 2 \times 2 \times \dots 2 = 2^N = e^{N \ln 2} \quad (2.5)$$

total states.

Another example concerns the total number of states of a set of N particles in a volume V

$$\frac{V^N}{N!} \approx \frac{V^N}{N^N} = e^{N \ln \frac{V}{N}} \quad (2.6)$$

where the factor $N!$ accounts for indistinguishability.

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

$$\frac{10^{30}m}{10^{-10}m} \sim 10^{40} \quad (2.7)$$

This is in no way comparable to (!)

$$e^{10^{23}} \quad (2.8)$$

the number of states for a system of $N = 10^{23}$ particles. 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.

2.1 Independence of parts

In order to illustrate the concept of independence of parts (or subsystems) of a macroscopic system, it is worth thinking about experiments a bit. Say an experimentalist measures the speed of sound c in a sample, as shown. When doing the experiment, the researcher obtains some definite value c .

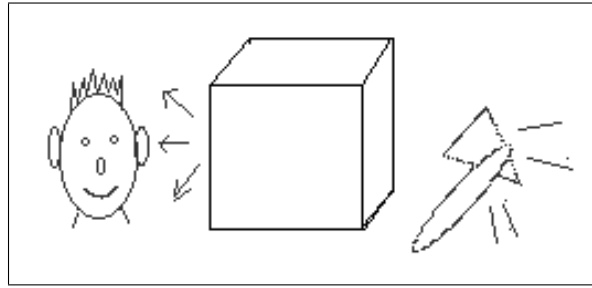


Figure 2.1: Experiment 1 on the whole system.

Afterwards, the experimentalist might cut up the sample into many parts so that his or her friends can repeat the same experiment. That is as shown in Fig. 2.2. Each experimenter will measure the

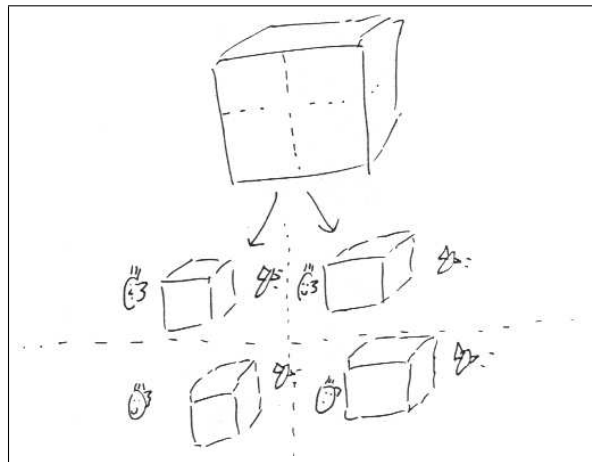


Figure 2.2: Experiment 2, on many parts of the same system as in experiment 1.

sound speed as shown, and they will all obtain slightly different values, but all close to c , the value obtained for the whole system. Why are the values different ? Which one is more accurate ?

It would seem that in experiment 2, since four independent values are obtained, averaging over them would give a more accurate determination of c because improved statistics. In fact, one can

continue breaking up the system like this, as shown in Fig. 2.3, and potentially keep improving the statistical accuracy in the determination of c for the material in question.

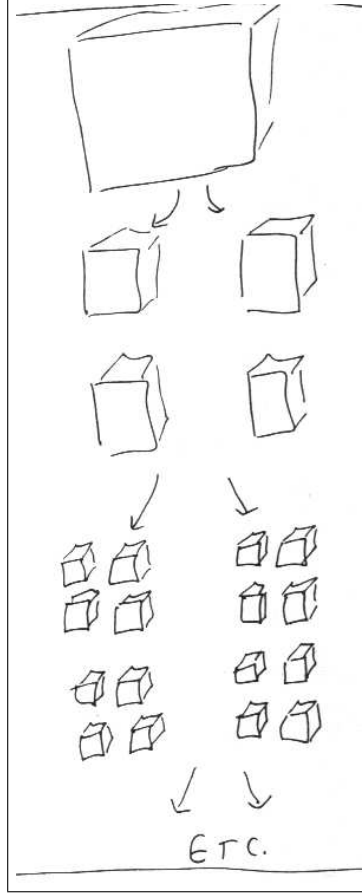


Figure 2.3: Experiment 3

Of course, one would expect that the expected value (and accuracy) of the determination of c in one experiment on the big system must be equivalent to many experiments conducted on individual parts or subsystems of it, as long as these individual parts are independent. By this it is meant that a system “self-averages” all its independent parts: the value obtained for the system as a whole is an average over the values obtained over the independent parts or subsystems. You can visualize this by imagining that in order to measure c a sound wave must traverse the entire system so that, in effect, the speed of propagation over the entire system will be an average over the speeds of propagation in all the subsystems along the path of the wave. To the extent that they are independent, then the speed measured over the whole system will be a simple average over the speeds in the various parts.

The number of independent parts N of a given system (to temporarily use the same notation as if they were particles) can be estimated to be

$$N = \mathcal{O}\left(\frac{L^3}{\xi^3}\right) \quad (2.9)$$

for a three dimensional system of linear scale L . The new quantity introduced here is the correlation length ξ , the length over which the motion of the microscopic constituents of our system are correlated, and hence cannot be taken as statistically independent. The correlation length must be at least as large as the size of an atom so $\xi \geq \mathcal{O}(\text{\AA})$, and indeed usually $\xi \approx \mathcal{O}\text{\AA}$. In short, if the microscopic motion of the system constituents is correlated over a volume of the order of ξ^3 , then our macroscopic system of volume L^3 can be approximately decomposed in V^3/ξ^3 independent subsystems.

2.2 Self averaging

In this section we explore the mathematical consequences of independence and self averaging, and in particular the consequences of the Central Limit Theorem of Probability Theory on Statistical Mechanics.

In order to do so, consider an extensive thermodynamic variable X . If a system consists of N independent parts as shown, then

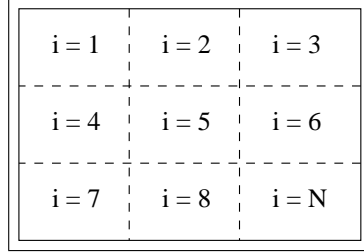


Figure 2.4: N independent parts

$$X = \sum_{i=1}^N X_i \quad (2.10)$$

where X_i is the value of X in part i . This definition applies to any microscopic configuration of the system. The statistical average of X is

$$\langle X \rangle = \sum_{i=1}^N \langle X_i \rangle \quad (2.11)$$

The average $\langle \cdot \rangle$ is understood over all possible microscopic configurations of the system, with probability distribution given by some appropriate statistical ensemble. Of course, the expectation is that in equilibrium, the statistical average $\langle X \rangle$ coincides with the spatial average of X over all the independent parts.

More precisely, the statistical average must equal

$$\langle X \rangle = \frac{\int_0^T dt X(t)}{T} = \frac{\int_0^T dt \sum_{i=1}^N X_i(t)}{T}, \quad (2.12)$$

where T is a microscopically very long time (although it could be macroscopically short). Of course,

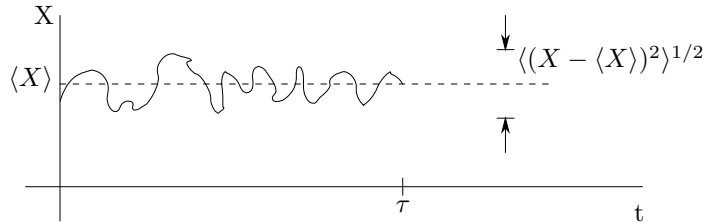


Figure 2.5: Time Average

$X(t)$ is a fluctuating quantity as the system evolves over phase space.

The concept of self averaging can now be stated as follows: The values of $X(t)$ over the entire system deviate negligibly from $\langle X \rangle$ when N is large and the parts are independent. If so, the statistical average equals the sum over the subsystems. In order to prove this statement, let us calculate the deviations of X from the average, $\langle (X - \langle X \rangle)^2 \rangle^{1/2}$ (see Fig. 2.5. First define $\Delta X =$

$X - \langle X \rangle$ so that $\langle \Delta X \rangle = \langle X \rangle - \langle X \rangle = 0$. The variance is

$$\begin{aligned} \langle (\Delta X)^2 \rangle &= \langle (X - \langle X \rangle)^2 \rangle \\ &= \langle X^2 - 2X\langle X \rangle + \langle X \rangle^2 \rangle \\ &= \langle X^2 \rangle - \langle X \rangle^2 \end{aligned} \quad (2.13)$$

Giving the condition, in passing,

$$\langle X^2 \rangle \geq \langle X \rangle^2 \quad (2.14)$$

since $\langle (\Delta X)^2 \rangle \geq 0$. By using the definition of X in terms of a sum over parts, we find,

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

which can be rewritten as,

$$\langle (\Delta X)^2 \rangle = \left\langle \sum_{\substack{i=1 \\ (i=j)}}^N (\Delta X_i)^2 \right\rangle + \left\langle \sum_{\substack{i=1 \\ j=1 \\ (i \neq j)}}^N (\Delta X_i)(\Delta X_j) \right\rangle \quad (2.15)$$

But part i is independent of part j , so that

$$\langle \Delta X_i \Delta X_j \rangle = \langle \Delta X_i \rangle \langle \Delta X_j \rangle = 0. \quad (2.16)$$

Therefore,

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

Since the standard deviation in each subsystem $\langle (\Delta X_i)^2 \rangle$ is finite and independent of N (say of $\mathcal{O}(1)$), then

$$\langle (\Delta X)^2 \rangle = \mathcal{O}(N), \quad (2.17)$$

or that the variance of X is proportional to the number of independent parts N . This implies that the size of the fluctuations relative to the average,

$$\frac{\langle (\Delta X)^2 \rangle}{\langle X \rangle^2} = \mathcal{O}\left(\frac{1}{N}\right), \quad (2.18)$$

and hence the relative size of fluctuations is very small if N is large. This is the self averaging result.

Note that, since $N = V/\xi^3$, we have

$$\frac{\langle (\Delta X)^2 \rangle}{\langle X \rangle^2} = \mathcal{O}\left(\frac{\xi^3}{V}\right) \quad (2.19)$$

so that fluctuations measure *microscopic* correlations. They would only be appreciable in the special limit in which $\frac{\xi^3}{V} \sim \mathcal{O}(1)$, that is, when the system cannot be decomposed in independent parts. This occurs in quantum mesoscopic systems (where V is small), and in phase transitions (where ξ is large).

2.3 Central Limit Theorem

In this section we show that the limiting distribution for X when the system is comprised of a large number of independent parts is a Gaussian distribution. This is the so called Central Limit theorem.

2.3.1 Probability of the sum of two independent variables

Consider x_1 and x_2 to be two independent variables, and define its sum $y = x_1 + x_2$. By definition,

$$p(y)\Delta y = \int_{y < x_1 + x_2 < y + \Delta y} dx_1 dx_2 p(x_1, x_2), \quad (2.20)$$

or

$$p(y) = \int dx_1 dx_2 p(x_1, x_2) \delta(y - x_1 - x_2). \quad (2.21)$$

If x_1 and x_2 are independent, then $p(x_1, x_2) = p_{x_1}(x_1)p_{x_2}(x_2)$. Hence

$$p(y) = \int dx_1 p_{x_1}(x_1) p_{x_2}(y - x_1). \quad (2.22)$$

The probability of the sum is the convolution of the individual probability distributions. By taking the Fourier transform, one has

$$G_y(k) = G_{x_1}(k) G_{x_2}(k). \quad (2.23)$$

For the particular case of Gaussian variables, we note that if $p(x)$ is a Gaussian distribution of mean μ and variance σ^2 , its characteristic function is

$$G_x(k) = e^{i\mu k - \frac{1}{2}\sigma^2 k^2}. \quad (2.24)$$

Therefore,

$$G_y(k) = e^{i(\mu_1 + \mu_2)k - \frac{1}{2}(\sigma_1^2 + \sigma_2^2)k^2}. \quad (2.25)$$

2.3.2 Central Limit Theorem

Consider a set of N independent variables X_i . We assume for simplicity that $\langle X_1 \rangle = 0$ (i.e., we are really considering the variable ΔX_i defined earlier). Otherwise, X_i is characterized by a *completely arbitrary distribution function*, not necessarily a Gaussian. We now define $y = \sum_{i=1}^N X_i$ and $z = y/\sqrt{N}$. Note that

$$G_z(k) = G_y\left(\frac{k}{\sqrt{N}}\right) = \left[G_{X_1}\left(\frac{k}{\sqrt{N}}\right)\right]^N.$$

The first equality follows from the definition of z , and the second from the fact that the variables X_i are independent. We have further assumed that all the distributions p_{X_i} are identical. This is the appropriate case for a macroscopic physical system. By definition,

$$G_z(k) = \int dz e^{ikz} p(z) = \int \frac{dy}{\sqrt{N}} e^{iky/\sqrt{N}} p(z) = \int dy e^{i(k/\sqrt{N}) \frac{y}{\sqrt{N}}} p(z).$$

Since $p(z)dz = p(y)dy$, we finally find $G_z(k) = G_y(k/\sqrt{N})$. This is the result used above.

For large N , we will only need the dependence of G_{X_1} for small values of its argument. We therefore expand in Taylor series,

$$\begin{aligned} G_{X_1}(k) &= \int e^{ikX_1} p(X_1) dX_1 = \int \left(1 + ikX_1 + \frac{(ik)^2}{2} X_1^2 + \dots\right) p(X_1) dX_1 \\ &= 1 + ik\langle X_1 \rangle + \frac{(ik)^2}{2} \langle X_1^2 \rangle + \dots \\ &= 1 - \frac{\sigma_{X_1}^2 k^2}{2} + \dots \end{aligned} \quad (2.26)$$

Therefore,

$$G_z(k) = \left[1 - \frac{\sigma_{X_1}^2 k^2}{2N} + \mathcal{O}\left(\frac{k}{\sqrt{N}}\right)^3\right]^N.$$

Now recall that

$$\lim_{N \rightarrow \infty} \left(1 + \frac{z}{N}\right)^N = e^z,$$

so that for large N ,

$$G_z(k) = e^{-\frac{1}{2}\sigma_{X_1}^2 k^2}.$$

In short, $z = (\sum X_i)/\sqrt{N}$ is a Gaussian variable with zero average and variance $\sigma_{X_1}^2$, *quite independently* of the form of $p(X_1)$.

In summary, the central limit theorem states that if X_1, \dots, X_N are independent, arbitrary but identically distributed random variables of zero mean, then

$$\langle (\sum X_i)^2 \rangle = N\sigma^2, \quad (2.27)$$

where σ^2 is the variance of each of the variables in $\{X_i\}$.

2.4 Correlations in Space

The results obtained so far apply to a large system, and perhaps it is not surprising that global fluctuations in a large system are small. The results for the global fluctuations have implications for the correlations of quantities, however. Consider the density of X , $x = X/N$. If we divide the system in elements of volume that are microscopically large, but macroscopically small, then $x = x(\vec{r})$ is a function of position. Consider now the correlation function

$$C(\vec{r}, \vec{r}') = \langle x(\vec{r})x(\vec{r}') \rangle - \langle x(\vec{r}) \rangle \langle x(\vec{r}') \rangle = \langle \Delta x(\vec{r}) \Delta x(\vec{r}') \rangle. \quad (2.28)$$

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

1. If the system is translationally invariant, then the origin of the coordinate system for \vec{r} and \vec{r}' is arbitrary. C can only depend on the separation between \vec{r} and \vec{r}' , not on their absolute value:

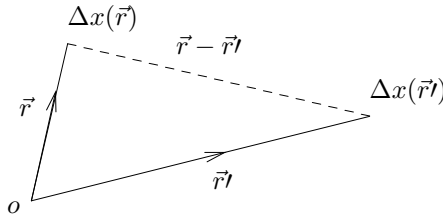


Figure 2.6: Translational invariance (only $\vec{r} - \vec{r}'$ is important)

$$C(\vec{r}, \vec{r}') = C(\vec{r} - \vec{r}') = C(\vec{r}' - \vec{r}) \quad (2.29)$$

and, of course,

$$\langle \Delta x(\vec{r}) \Delta x(\vec{r}') \rangle = \langle \Delta x(\vec{r} - \vec{r}') \Delta x(0) \rangle. \quad (2.30)$$

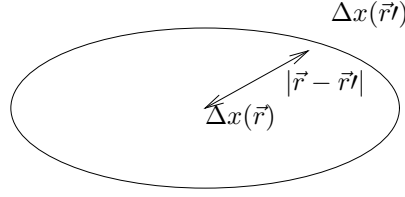
2. If the system is isotropic, then the direction of $\Delta x(\vec{r})$ from $\Delta x(\vec{r}')$ is unimportant, and only the magnitude of the distance $|\vec{r} - \vec{r}'|$ is physically relevant. Hence

$$C(\vec{r}, \vec{r}') = C(|\vec{r} - \vec{r}'|) \quad (2.31)$$

So what was a function of six variables, (\vec{r}, \vec{r}') , is now a function of only one: $|\vec{r} - \vec{r}'|$.

3. Usually correlations fall off with distance, so that points that are far apart are statistically independent. That is,

$$\langle \Delta x(r \rightarrow \infty) \Delta x(0) \rangle = \langle \Delta x(r \rightarrow \infty) \rangle \langle \Delta x(0) \rangle = 0, \quad (2.32)$$

Figure 2.7: Isotropy (only $|\vec{r} - \vec{r}'|$ is important)

or

$$C(r \rightarrow \infty) = 0. \quad (2.33)$$

The decay to zero will occur over distances of the order of the correlation length ξ precisely because of its definition.

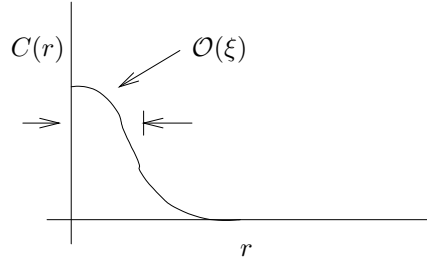


Figure 2.8: Typical correlation function

Finally, we can estimate in general the magnitude of the function C . The deviation of the density $x = X/N$ from its average can be written as,

$$\Delta x = \frac{\int d\vec{r} \Delta x(\vec{r})}{\int d\vec{r}} = \frac{1}{V} \int d\vec{r} \Delta x(\vec{r}). \quad (2.34)$$

Therefore

$$\frac{1}{V^2} \int d\vec{r} \int d\vec{r}' \langle \Delta x(\vec{r}) \Delta x(\vec{r}') \rangle = \langle (\Delta x)^2 \rangle \sim \mathcal{O}\left(\frac{1}{N}\right), \quad (2.35)$$

according to the Central Limit Theorem, or,

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

We now introduce the following coordinate transformation: $\vec{R} = \vec{r} - \vec{r}'$, $\vec{R}' = \frac{1}{2}(\vec{r} + \vec{r}')$. The Jacobian of the transformation is one, and we obtain

$$\int d\vec{R}' \int d\vec{R} C(\vec{R}) = V \int d\vec{R} C(\vec{R}) \sim \mathcal{O}\left(\frac{V^2}{N}\right) \quad (2.36)$$

or reintroducing the variable \vec{r} ,

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

Since $N = \frac{V}{\xi^3}$, we have

$$\int d\vec{r} C(\vec{r}) \sim \mathcal{O}(\xi^3). \quad (2.37)$$

This result provides the order of magnitude of the function C . Although many functions satisfy the constraint of Eq. (2.37) (see Fig. 2.8), it turns out that, usually,

$$C(r) \sim e^{-r/\xi} \quad (2.38)$$

for large r , a function that satisfies Eq. 2.37. For systems comprised of a large number of independent subsystems, the correlation function $C(r)$ usually decays exponentially, with a characteristic decay length equal to ξ .

In summary, we have focused on a system having $N \rightarrow \infty$ independent parts or subsystems. For $x = X/N$, where X is an extensive thermodynamic variable, and $N = V/\xi^3$ the number of independent parts in the volume V , with ξ is the correlation length, we have,

$$\langle x \rangle = \mathcal{O}(1) \quad (2.39)$$

$$\langle (\Delta x)^2 \rangle \sim \mathcal{O}\left(\frac{1}{N}\right) \quad (2.40)$$

$$p(x) \propto \frac{(x - \langle x \rangle)^2}{e^{2\langle (\Delta x)^2 \rangle}} \quad (2.41)$$

$$\int d\vec{r} \langle \Delta x(r) \Delta x(0) \rangle \sim \mathcal{O}(\xi^3) \quad (2.42)$$

This is the generic situation in classical statistical mechanics in which fluctuations are small. The purpose of this course, on the other hand, is to examine those cases in which $V/\xi^3 \sim \mathcal{O}(1)$ and hence these results break down!

2.5 (*) Central Limit Theorem through explicit calculation of moments

We present an alternative derivation of the central limit theorem that considers the explicit calculation of the moments of the variable X . 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 $p(X)$ is symmetric about the mean $\langle X \rangle$ as shown in Fig. 2.9. It turns out that one can show that the distribution function *is* symmetric as $(N \rightarrow \infty)$, but

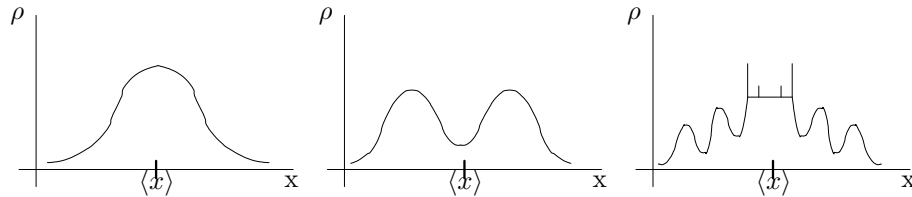


Figure 2.9: Symmetric distribution functions $\rho(x)$

lets just assume it so the derivation is simpler.

In order to obtain the distribution of X , $p(X)$ we could first calculate its characteristic function $G_X(k)$, which in turn can be obtained by calculating all the moments of X , as shown in Eq. (1.2). We therefore proceed to evaluate all the even moments

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

for positive integers n . We have already evaluated the case $n = 1$ (Eq. (2.17), so we start now by considering $n = 2$,

$$\begin{aligned} \langle (\Delta X)^4 \rangle &= \left\langle \sum_{i=1}^N \Delta X_i \sum_{j=1}^N \Delta X_j \sum_{k=1}^N \Delta X_k \sum_{l=1}^N \Delta X_l \right\rangle \\ &= \left\langle \sum_{\substack{i=1 \\ (i=j=k=l)}}^N (\Delta X_i)^4 \right\rangle + 3 \left\langle \sum_{\substack{i=1 \\ j=1 \\ (i \neq j)}}^N (\Delta X_i)^2 (\Delta X_j)^2 \right\rangle + \cancel{\mathcal{O}(\Delta X_i)^4} \quad (2.43) \\ &\sim \mathcal{O}(N) + \mathcal{O}(N^2) \end{aligned}$$

Therefore, as $N \rightarrow \infty$,

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

or

$$\langle (\Delta X)^4 \rangle = 3 \langle (\Delta X)^2 \rangle^2 \sim \mathcal{O}(N^2) \quad (2.44)$$

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

$$\langle (\Delta X)^{2n} \rangle \sim \langle (\Delta X)^2 \rangle^n \sim \mathcal{O}(N^n) \quad (2.45)$$

The constant of proportionality in Eq. 2.45 can be determined exactly. Note that in

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

there are $(2n - 1)$ ways to pair off the first two ΔX . After those two ΔX 's are considered, there remain $(2n - 3)$ ways to pair off the next two ΔX , and so on. Until finally, we get to the last two ΔX 's which have only way to be paired off. Hence

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

or more conventionally,

$$\langle (\Delta X)^{2n} \rangle = \frac{2n!}{2^n n!} \langle (\Delta X)^2 \rangle^n \quad (2.46)$$

In short, we have computed all the even moments of ΔX in terms of $\langle (\Delta X)^2 \rangle$. Of course, we also need the average $\langle X \rangle$ that enters the definition of $\langle \Delta X \rangle$. Given the definition of the characteristic function, we now have,

$$p(X) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{ikX} \sum_{n=0}^{\infty} \frac{(ik)^n}{n!} \langle X^n \rangle. \quad (2.47)$$

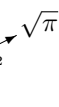
In our case, all odd moments vanish, so that,

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

But $\langle X^{2n} \rangle = \frac{(2n)!}{2^n n!} \langle X^2 \rangle^n$ (assuming $\langle X \rangle = 0$ for simplicity. The same derivation holds in the more general case). Hence,

$$p(X) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{ikX} \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{-k^2}{2} \langle X^2 \rangle \right)^n = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{ikX} e^{\frac{-k^2}{2} \langle X^2 \rangle} \quad (2.48)$$

This integral can be done by completing the square, that is adding and subtracting $\frac{-X^2}{2\langle X^2 \rangle}$, so that

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


The final result, allowing for a nonzero average $\langle X \rangle$ is,

$$p(X) = \frac{1}{\sqrt{2\pi \langle (\Delta X)^2 \rangle}} e^{\frac{-(X - \langle X \rangle)^2}{2 \langle (\Delta X)^2 \rangle}}, \quad (2.49)$$

a Gaussian distribution.

2.5. (*) CENTRAL LIMIT THEOREM THROUGH EXPLICIT CALCULATION OF MOMENTS 17

In passing, it is worth remembering the following important corollary,

$$\langle e^{ikX} \rangle = e^{\frac{-k^2}{2} \langle X^2 \rangle}, \quad (2.50)$$

for a Gaussian distribution.

This is again the central limit theorem for a system of many independent parts. Its consequences are most transparent if one deals with an intensive density $x = \frac{X}{N}$. Then,

$$\begin{aligned} \langle x \rangle &\sim \mathcal{O}(1) \\ \langle (\Delta x)^2 \rangle &\sim \mathcal{O}(1/N) \end{aligned} \quad (2.51)$$

Let

$$\langle (\Delta x)^2 \rangle \equiv \frac{a^2}{N} \quad (2.52)$$

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

$$p(x) \propto e^{-N \frac{(x - \langle x \rangle)^2}{2a^2}} \quad (2.53)$$

so the width of the Gaussian is exceedingly small. As stated earlier, fluctuations around the average

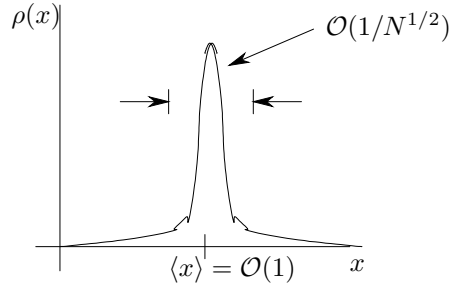


Figure 2.10: Gaussian (Width Exaggerated)

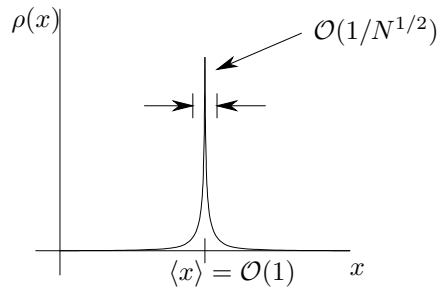


Figure 2.11: Gaussian (Width (sort of) to scale)

are small, and hence the average value is well defined for a system of many $N \rightarrow \infty$ independent parts. The distribution of x is Gaussian regardless of the distribution of the X_i in each of the parts or subsystems. This result applies to the equilibrium distribution of *any* thermodynamic variable, with the only assumption that the system be composed of many independent subsystems.

Chapter 3

Self similarity and Fractals

This chapter allows us to explore one situation in which a large object is not comprised of many independent parts. More precisely, by looking at objects that are self similar, we will uncover a class of systems in which spatial correlation functions do not decay exponentially with distance, the hallmark of statistical independence among the parts.

Fractals are mathematical constructs that seem to exist in fractional dimensions between 1 and 2, 2 and 3, etc.. They provide a useful starting point for thinking about systems in which $V/\xi^3 \sim \mathcal{O}(1)$, or if $V = L^3$, $\xi/L \sim \mathcal{O}(1)$. Figure 3.1 schematically shows a fractal object. The perimeter appears

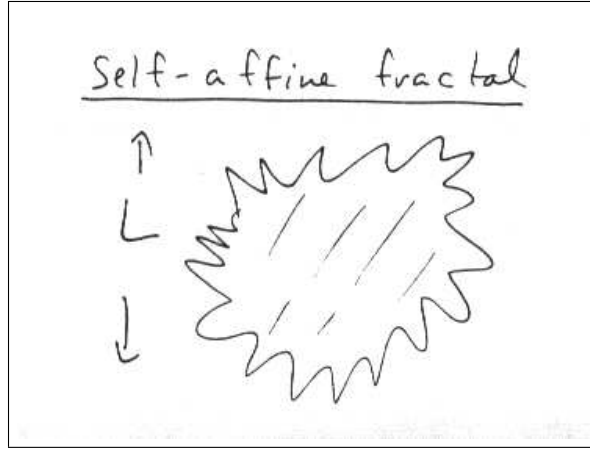


Figure 3.1: Self affine Fractal

very jagged, or as it is said in physics, *rough*. The self-affine fractal has the property that its perimeter length P satisfies,

$$P \sim L^{d_s} \quad \text{as } L \rightarrow \infty \quad (3.1)$$

where $d_s > 1$ is the so called self-affine fractal exponent. Note that for a circle of diameter L , $P_{\text{circle}} = \pi L$, while for a square of side L $P_{\text{square}} = 4L$. In fact, $P \sim L^1$ for any run-of-the-mill, non fractal, object in two spatial dimensions. More generally, the surface of a compact (non fractal) object in d spatial dimensions is

$$P_{\text{non fractal}} \sim L^{d-1} \quad (3.2)$$

It is usually the case that

$$d_s > d - 1 \quad (3.3)$$

Another example is given in Fig. 3.2. This object has a distribution of holes such that its mass M satisfies

$$M \sim L^{d_f} \quad \text{as } L \rightarrow \infty, \quad (3.4)$$

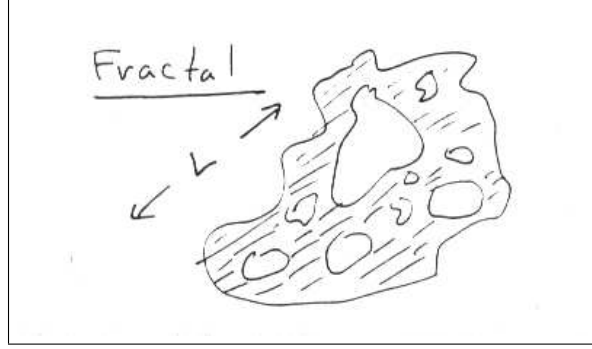


Figure 3.2: Volume Fractal

where d_f is the fractal dimension. Note again that for a circle $M = (\pi/4)\rho L^2$ if L is the diameter, while for a square $M = \rho L^2$. In fact, for run-of-the-mill compact objects

$$M_{\text{non fractal}} \sim L^d \quad (3.5)$$

in d dimensions. Usually

$$d_f < d \quad (3.6)$$

Since the mass density is normally defined as M/L^d in d dimensions, this would imply that

$$\rho_f \sim \frac{1}{L^{d-d_f}} \rightarrow 0 \quad (3.7)$$

as $L \rightarrow \infty$.

3.1 Von Koch Snowflake

We construct in this section a specific example of a fractal object, and compute its fractal dimension. In order to construct a von Koch fractal (or snowflake), follows the steps outlined in Fig. 3.3. In

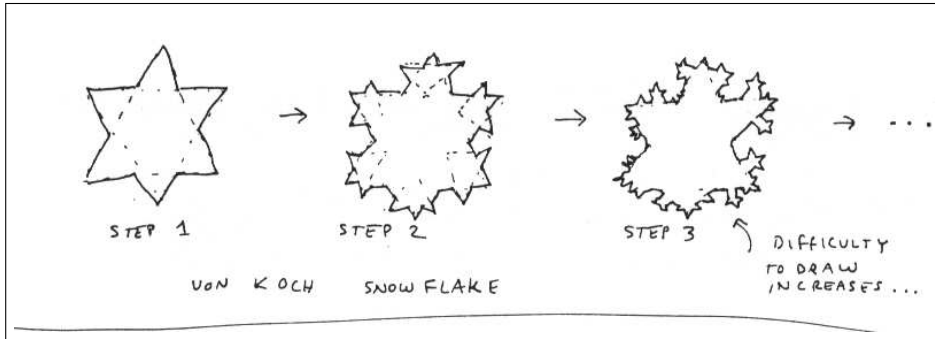


Figure 3.3: Von Koch Snowflake

order to find the self-affine or fractal exponent for the perimeter, imagine we had drawn one of these fractals recursively. Say the smallest straight element has length l . We calculate the perimeter P as we go through the recursion indicated in Fig. 3.4.

From the figure, we find that the perimeter $P_n = 4^n l$ after n iterations, whereas the end to end distance after n iterations is $L_n = 3^n l$. By taking the logarithm, we find $\ln(P/l) = n \ln 4$ and $\ln(L/l) = n \ln 3$. Hence,

$$\frac{\ln P/l}{\ln L/l} = \frac{\ln 4}{\ln 3} \quad \text{or} \quad P/l = (L/l)^{d_s}$$

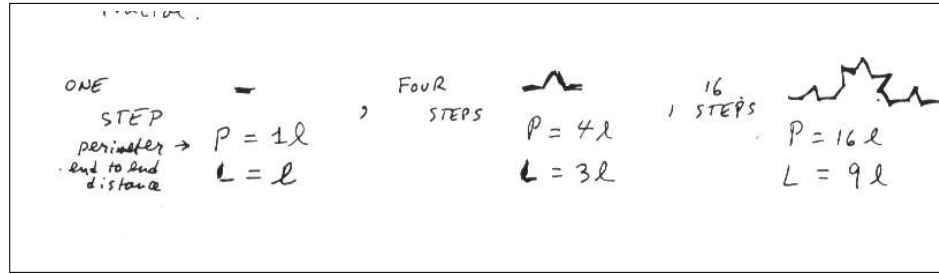


Figure 3.4:

with

$$d_s = \frac{\ln 4}{\ln 3} \approx 1.26 > d - 1 = 1. \quad (3.8)$$

3.2 Sierpinski Triangle

The second example that we discuss is the Sierpinski triangle. It can also be constructed recursively by following the steps outlined in Fig. 3.5. The iteration to construct this fractal is also shown in

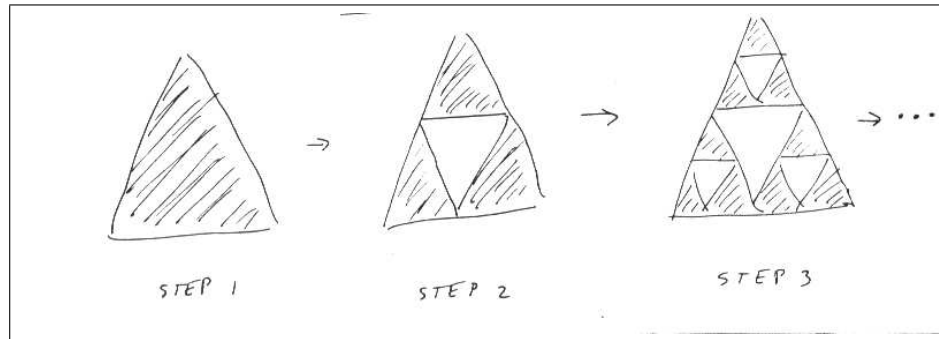


Figure 3.5: Sierpinski Triangle

Fig. 3.6. Define the mass of the smallest filled triangle as m and its volume as v . Successive masses are given by $M_n = 3^n m$ after n iterations, with a volume $V_n = 4^n v$. Since $V_n = \frac{1}{2} L_n^2$, and $v = \frac{1}{2} l^2$

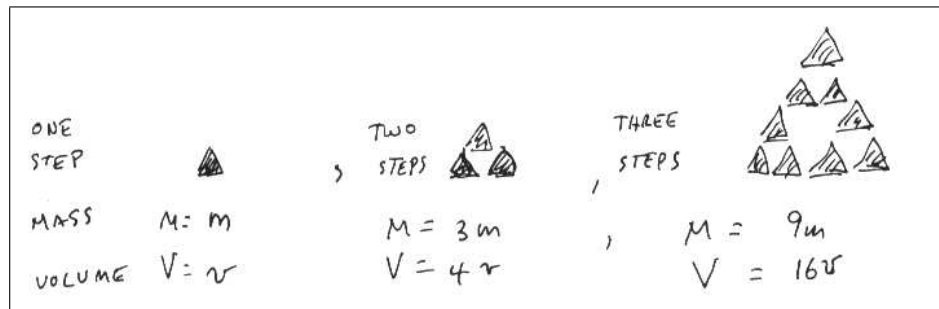


Figure 3.6:

we have $L_n = 2^n l$. Taking logarithms, we obtain $\ln(M/m) = n \ln 3$ and $\ln(L/l) = n \ln 2$, taking the ratio

$$\frac{\ln(M/m)}{\ln(L/l)} = \frac{\ln 3}{\ln 2} \quad \frac{M}{m} = \left(\frac{L}{l}\right)^{d_f},$$

with

$$d_f = \frac{\ln 3}{\ln 2} \approx 1.585 < d = 2. \quad (3.9)$$

3.3 Correlations in Self-Similar Objects

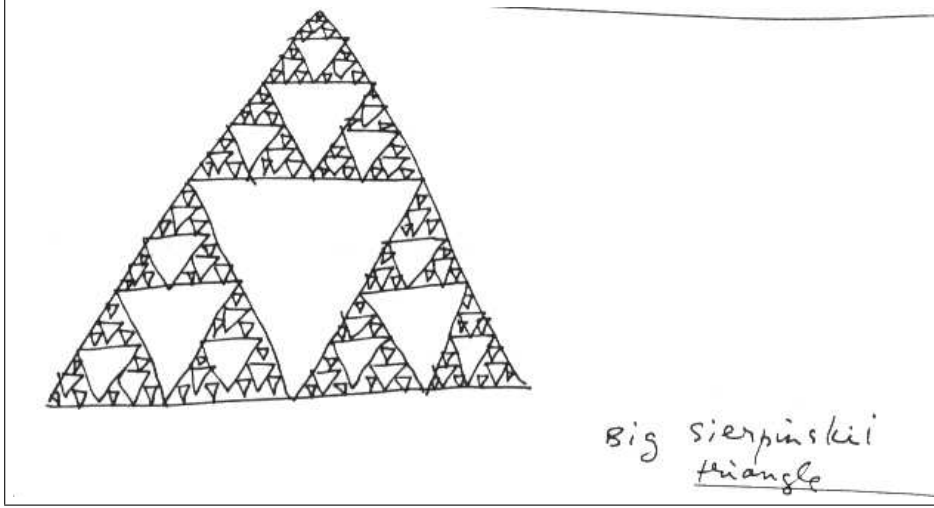


Figure 3.7:

The defining feature about a fractal is that it “looks the same” on different length scales: If you take the fractal in Fig. 3.7, select a small portion of it and magnify it, the magnified structure will be identical to the original fractal. This is the case irrespective of the size of the portion that is being magnified. Although this is true in this particular example, we will focus in what follows in a less restrictive definition: “statistical self similarity”. In this case, configurations and their magnified counterparts are geometrically similar only in a statistical sense. More specifically, any parameter of the structure that is invariant under uniform magnification has to be, on average, independent of scale.

This statement can be quantified by introducing the correlation function $C(r)$ of the density $n(r)$ in Fig. 3.7,

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

The average is, say, over all orientations in space for convenience. In the case of a physical system, the average would be the usual statistical average. If the structure is uniformly magnified by a factor λ : $r \rightarrow \lambda r$, then statistical self-similarity is defined as

$$C(\lambda r) = \lambda^{-p} C(r), \quad (3.10)$$

where the exponent p is a constant. Both the magnified $C(\lambda r)$ and the original $C(r)$ are functionally the same, except for a constant scale factor λ^p .

We first prove this relationship. The assumption of self similarity implies that $C(r) \propto C(r/b)$, or $C(r) = f(b) C(r/b)$ with $b = 1/\lambda$. If we take the derivative with respect to b , and let $r^* = r/b$ we find,

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

or,

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

Therefore,

$$\frac{d \ln f(b)}{d \ln b} = \frac{d \ln C(r^*)}{d \ln r^*} \quad (3.11)$$

Since the left hand side only depends on b , and the right hand side only on r^* , they both must equal a constant,

$$\frac{d \ln f(b)}{d \ln b} = \text{const.} = -p \quad f(b) \propto b^{-p} \quad (3.12)$$

and the recursion relation is

$$C(r) = b^{-p} C(r/b) \quad (3.13)$$

up to a multiplicative constant.

Equation (3.13) is sufficient to determine the functional form of $C(r)$. Since the magnification factor b is arbitrary, let us consider the special value

$$b = 1 + \epsilon, \quad \epsilon \ll 1 \quad (3.14)$$

Then, by expanding Eq. (3.13) in Taylor series, we have

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

At first order in ϵ we find,

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

The solution of which is

$$C \propto \frac{1}{r^p}. \quad (3.15)$$

In summary, self similarity implies that correlation functions are generalized homogeneous functions of their arguments,

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

relation that immediately leads to a power law dependent of the correlation function on distance,

$$C(r) \propto \frac{1}{r^p} \quad (3.16)$$

Usually, these relations apply in physical systems for large distances. At shorter distances (say the scale of an elemental triangle in the case at hand) self-similarity breaks down, and with it the mathematical relations that we have derived. In addition, for systems of finite size, self similarity also breaks down at scales of the order of the size of the system.

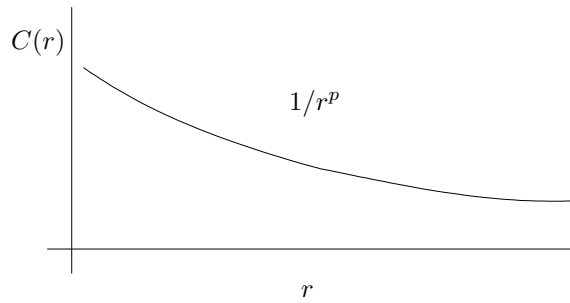


Figure 3.8: Decay of Correlations in a Self similar object

Note that the the dependence of $C(r)$ on r that we have just found is quite different from that which occurs in a system of many independent parts,

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

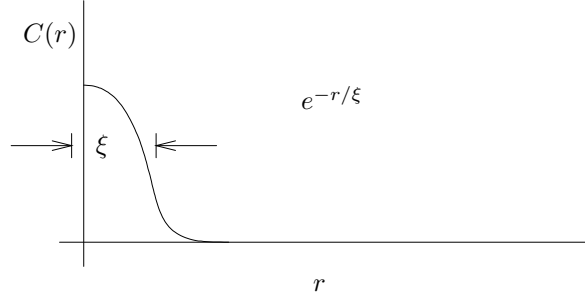


Figure 3.9: Decay of Correlations in a system of many independent parts

In particular, the relation

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

in d -dimensions that is satisfied for a system comprised of many independent parts is evidently *not* satisfied for a self similar system in which

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

if $d > p$. In this case, the integral diverges! (this formally implies a divergence in the variance of x):

$$\int d^d r C(r) \propto \int_a^L r^{d-1} dr \frac{1}{r^p} \sim L^{d-p}, \quad (3.17)$$

where a is a microscopic length (irrelevant here), and $L \rightarrow \infty$ the size of the system.

In summary, fractals (self similar or rough objects) do not satisfy the central limit theorem since the integral in Eq. 3.17 diverges. The study of rough surfaces is one of the focuses of the rest of the course.

Chapter 4

Review of Statistical Mechanics and Fluctuations

4.1 Thermodynamics

The thermodynamic state of a system in equilibrium is completely specified by a finite set of extensive variables called thermodynamic variables. A class of variables is related to conservation laws: e.g., energy E and number of particles N . Others to the system's extent: The volume V . This set also needs to be complete, a question that is easy to answer a posteriori for any given system, but that it is not trivial to address in newly discovered systems.

The thermodynamic description entails a huge reduction in the number of degrees of freedom needed to describe the state of the system in the so called thermodynamic limit ($N \rightarrow \infty$, $V \rightarrow \infty$, such that the number density $n = N/V$ is finite). Each classical particle, for example, has in three dimensions six degrees of freedom (position and momentum). So there are $\sim 10^{23}$ microscopic degrees of freedom, yet the equilibrium state of a fluid needs only three.

The first law of thermodynamics states that the energy of a closed system is only a function of the thermodynamic variables in equilibrium. It is extensive, and a function of state: every time the system is in the same macroscopic state, it has the same internal energy.

If the system is not closed, and can exchange work with the environment, the change in internal energy must equal the work exchanged, $dE = -dW$, where the work $dW > 0$ if done by the system. For a fluid system, mechanical work has the familiar form $dW = pdV$ where p is the (uniform) pressure of the fluid, and dV the change in volume. This latter relation only applies to a quasi-static process (very slow compared to microscopic relaxation times). If the process is not quasi static, the pressure in the fluid is not uniform. The work done still equals the change in energy, but one must use the equations of hydrodynamics to actually calculate it.

If the system is not closed, and can exchange more than work with the environment, then the first law adopts the form,

$$dQ = dE + dW,$$

where dQ is the heat exchanged between the system and the environment. By convention, $dQ > 0$ if it flows into the system. Note that whereas dE is the differential form of a function of state (E), neither dW and dQ are. Both depend on the specific thermodynamic path of the process.

The second law of thermodynamics introduces a new function of state, the entropy S . It is an extensive function of the thermodynamic variables, so that for a mono component fluid $S = S(E, N, V)$. The second law defines the state of thermodynamic equilibrium by saying that given the macroscopic constraints on a closed system (i.e., E, V, N), the state of equilibrium is the one that maximizes the entropy over the manifold of constraint states.

In particular, if the state A of the system includes some internal constraint (e.g., no particles in its left bottom corner), and the resulting entropy of this state is $S(A)$, then removing the internal constraint -thus leading to a new equilibrium state B - the fact that S needs to be maximized

immediately implies $S(B)/geS(A)$; i.e., the entropy can only increase in spontaneous processes in a closed system.

The second law has immediate consequences for the state of equilibrium. Consider a closed system, and imagine it decomposed in two subsystems, as indicated in Fig. 4.1. If the system in

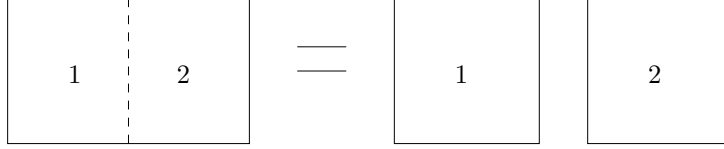


Figure 4.1: A system with two parts.

question is a fluid, then the three thermodynamic variables of the whole system (1+2): V, E, N satisfy,

$$\begin{aligned} V &= V_1 + V_2 \\ E &= E_1 + E_2 \\ N &= N_1 + N_2 \end{aligned}$$

We are now going to search the equilibrium state by requiring that S be maximized over any changes in V_1, E_1, N_1 , i.e. the way in which the three variables can be apportioned over the manifold of constrained states given by fixed (E, V, N) .

Since S is a function of the thermodynamic variables and it is extensive, we can write,

$$dS = dS_1 + dS_2 = \left(\frac{\partial S_1}{\partial E_1}\right) dE_1 + \left(\frac{\partial S_1}{\partial V_1}\right) dV_1 + \left(\frac{\partial S_1}{\partial N_1}\right) dN_1 + \left(\frac{\partial S_2}{\partial E_2}\right) dE_2 + \left(\frac{\partial S_2}{\partial V_2}\right) dV_2 + \left(\frac{\partial S_2}{\partial N_2}\right) dN_2.$$

Given the constraints, we have

$$\begin{aligned} dV_1 + dV_2 &= 0 \\ dE_1 + dE_2 &= 0 \\ dN_1 + dN_2 &= 0. \end{aligned}$$

Therefore,

$$dS = \left[\left(\frac{\partial S_1}{\partial V_1}\right) - \left(\frac{\partial S_2}{\partial V_2}\right)\right] dV_1 + \left[\left(\frac{\partial S_1}{\partial E_1}\right) - \left(\frac{\partial S_2}{\partial E_2}\right)\right] dE_1 + \left[\left(\frac{\partial S_1}{\partial N_1}\right) - \left(\frac{\partial S_2}{\partial N_2}\right)\right] dN_1.$$

Since E, V and N are independent variables, its variations are also independent. The condition of equilibrium is that S be maximized over any variation of internal variables, or $dS = 0$ in this equation. Therefore, sufficient conditions for equilibrium are,

$$\begin{aligned} \left(\frac{\partial S_1}{\partial V_1}\right) &= \left(\frac{\partial S_2}{\partial V_2}\right) = \frac{p}{T} \text{uniform} \\ \left(\frac{\partial S_1}{\partial E_1}\right) &= \left(\frac{\partial S_2}{\partial E_2}\right) = \frac{1}{T} \text{uniform} \\ \left(\frac{\partial S_1}{\partial N_1}\right) &= \left(\frac{\partial S_2}{\partial N_2}\right) = -\frac{\mu}{T} \text{uniform} \end{aligned}$$

In short, the condition of equilibrium leads to three intensive variables T, μ, p that must be spatially *uniform* in equilibrium. If T is not uniform, energy flows will occur. If p is not uniform, the relative volume of the two subsystems will change. If μ is not uniform, mass transport will take place. When the system regains equilibrium, all three intensive variables must be uniform.

Note that this derivation admits a straightforward generalization to more complex systems in which more than three thermodynamic variables are needed to specify the state of the system. In

this respect, given the number of necessary thermodynamic variables, the second law leads to as many variables that must be uniform in equilibrium, and they all correspond to the derivative of S with respect to each of the variables.

Further progress is achieved by the introduction the so called equations of state. These relate the values of the thermodynamic variables in the equilibrium state, and are system dependent. Most often they are determined experimentally. For example, the ideal gas equation of state is

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

supplemented with the condition $(\partial E/\partial V)_{NT} = 0$.

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

$$dp = \left(\frac{\partial p}{\partial V} \right)_{NT} dV + \left(\frac{\partial p}{\partial T} \right)_{NV} dT$$

and

$$dE = \left(\frac{\partial E}{\partial V} \right)_{NT} dV + \left(\frac{\partial E}{\partial T} \right)_{NV} dT$$

Clearly, if we know all the (T, V) dependence of these derivatives, such as $(\frac{\partial E}{\partial T})_{NV}$, we know the equations of state. In fact, most work focuses on determining 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., how to calculate k_T or C_V from first principles.

We finally mention that these quantities can have very striking behavior. Consider the phase diagram of a pure substance (fig. 4.2). The lines denote first order phase transition lines. The dot denotes a *continuous* phase transition, often called of second order. This particular point is known as the critical point. Imagine preparing a system at the pressure corresponding to the continuous transition, and increasing temperature from zero through that transition (see fig. 4.3). Figure 4.4 shows the expected behavior of k_T and C_V along that path. 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 are known to have power-law divergences. We will spend a lot of time on continuous, so called second order transitions in this course.

4.2 Statistical Mechanics

The summary just given about thermodynamics has a parallel in statistical mechanics. In principle, we believe that the observable value of a macroscopic variable G is an average over the temporal

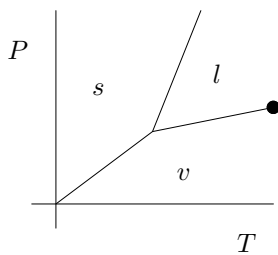


Figure 4.2: Solid, liquid, vapor phases of pure substance

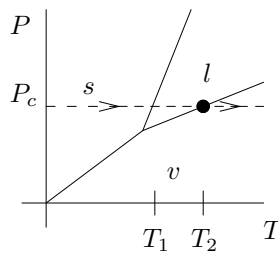


Figure 4.3:

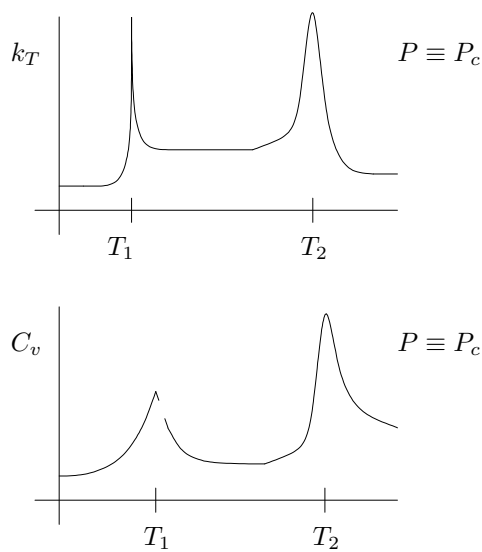


Figure 4.4:

evolution of the system over the manifold in phase space that corresponds to the thermodynamic state of equilibrium. Schematically,

$$G_{\text{obs}} = \frac{1}{\mathcal{N}} \sum_{i=1}^{\mathcal{N}} G(t_i),$$

where $G(t_i)$ is the value of G over a small interval around time t_i . The starting point is rewriting this equation as,

$$G_{\text{obs}} = \sum_{\nu} \left(\frac{1}{\mathcal{N}} \# \text{times in which the system is in state } \nu \right) G_{\nu}.$$

The quantity in parenthesis is the probability of occurrence of state ν while the system evolves in thermodynamic equilibrium, or p_{ν} . Therefore,

$$G_{\text{obs}} = \langle G \rangle = \sum_{\nu} p_{\nu} G_{\nu}.$$

The remarkable fact is that while the detailed trajectory over phase space is impossible to compute, the probability p_{ν} is known to great accuracy without ever solving the equations of motion (provided that G is a macroscopic variable). Several cases need to be distinguished.

4.2.1 Closed system

If $\Omega(N, V, E)$ is the total number of microscopic states for a closed system of fixed N, V, E , then,

$$p_{\nu} = \frac{1}{\Omega(N, V, E)}.$$

All states of the same energy occur with equal probability.

The connection with thermodynamics is through Boltzmann's formula,

$$S(N, V, E) = k_B \ln \Omega(N, V, E). \quad (4.1)$$

With this definition the entropy is extensive. If the system is decomposed into two subsystems, then $\Omega = \Omega_1 \Omega_2$, and hence $S = S_1 + S_2$.

Spontaneous processes lead to entropy increases as well. Removing an internal constraint in an otherwise closed system necessarily increases the number of accessible microscopic states Ω , and hence the entropy S .

4.2.2 System at constant temperature

In order to maintain a system at constant temperature it must be kept in equilibrium with a large heat bath with which it can exchange energy. When the combined system plus heat bath reach equilibrium (the whole is a closed system), the temperature of both must be the same.

Consider now a state of the system of energy E_{ν} . The probability of such a state in the system must be proportional to the number of states of the heat bath that are compatible with the system having an energy E_{ν} . If the combined energy of the system plus heat bath is E , then

$$p_{\nu} \propto \Omega(E_{\text{bath}}) = \Omega(E - E_{\nu}).$$

We also know that $E \gg E_{\nu}$ since we are considering a large heat bath. We now expand,

$$p_{\nu} \propto \exp [\ln \Omega(E - E_{\nu})] = \exp \left[\ln \Omega(E) - \frac{\partial \ln \Omega}{\partial E} E_{\nu} + \dots \right].$$

Recalling thermodynamics, $\beta = 1/k_B T = \frac{\partial \ln \Omega}{\partial E}$. Therefore, after appropriate normalization of p_ν by Z , the partition function, we find,

$$p_\nu = \frac{1}{Z} e^{-\beta E_\nu}. \quad (4.2)$$

Therefore, the states of an equilibrium system held at constant temperature are quite simply exponentially distributed according to their energy.

4.2.3 Other ensembles

Imagine now a system that can exchange both E and X with a bath. The first law of thermodynamics in this case is

$$TdS = dE + (k_B T \xi) dX, \quad (4.3)$$

so that

$$\left(T \frac{\partial S}{\partial X} \right)_E = k_B T \xi,$$

so that $k_B \xi$ is the conjugate variable to X .

Following the same argument given above we have now,

$$p_\nu \propto \Omega(E - E_\nu, X - X_\nu).$$

Expanding in Taylor series, we have

$$p_\nu \propto \exp [\ln \Omega(E - E_\nu, X - X_\nu)] = \exp \left[\ln \Omega(E, X) - \frac{\partial \Omega}{\partial E} E_\nu - \frac{\partial \Omega}{\partial X} X_\nu + \dots \right].$$

Therefore, given that $S = k_B \ln \Omega$ for the combined closed system, and introducing the partition function Z as normalization, we find

$$p_\nu = \frac{1}{Z} e^{-\beta E_\nu - \xi X_\nu} \quad Z(\beta, \xi) = \sum_\nu e^{-\beta E_\nu - \xi X_\nu}. \quad (4.4)$$

A particular example is the grand canonical distribution in which $X = N$, and hence $\xi = -\mu/k_B T$ and,

$$Z(\beta, \mu) = \sum_{N_\nu} \sum_{E_\nu} e^{-\beta E_\nu + \beta \mu N_\nu} \quad (4.5)$$

4.3 Fluctuations

Landau and Lifshitz are credited with having organized the foundations of the theory of fluctuations, although they attribute the theory to Einstein in their book on statistical physics. We follow their treatment.

Consider the density x of a thermodynamic variable X in a closed system. Let $\langle x \rangle$ be the equilibrium value, with $S(\langle x \rangle)$ its entropy. The value of $\langle x \rangle$ is such that it maximizes the entropy, or, equivalently, that corresponds to the largest number of microscopic states $\Omega(\langle x \rangle)$ given the macroscopic constraints in the entire systems. Imagine that the local density fluctuates to a different value x , which itself can correspond to $\Omega(x)$ different microscopic states that would yield the same local value of x . Obviously, $\Omega(x) < \Omega(\langle x \rangle)$, as $\langle x \rangle$ is the equilibrium value.

The key insight is to use Boltzmann's formula to assign the probability of occurrence of the fluctuation (*while the system remains in equilibrium*). First the probability of occurrence of x is defined:

$$p(x) = \frac{\Omega(x)}{\Omega(\langle x \rangle)},$$

as the system is closed, and hence, by using Boltzmann's formula,

$$p(x) = e^{\frac{1}{k_B} (S(x) - S(\langle x \rangle))}. \quad (4.6)$$

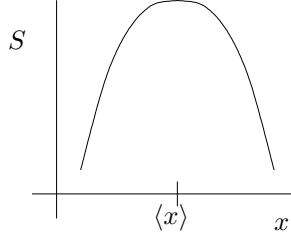


Figure 4.5:

This situation is schematically shown in Fig. 4.5.

Near the maximum of S we have

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

But $(\partial S / \partial x)_{x=\langle x \rangle} = 0$ as a condition for equilibrium, and $\partial^2 S / \partial x^2$ is negative since S is a maximum. Hence,

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

where we have defined

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

The negative sign is introduced anticipating that the second derivative is negative. Therefore probability of the fluctuation is

$$p(x) \propto e^{-(x - \langle x \rangle)^2 / 2\sigma^2}$$

or a *Gaussian distribution*. Since x is a local density (not microscopic) the distribution is sharply peaked at $x = \langle x \rangle$, and one normally allow integrals over $p(x)$ to vary from $x = -\infty$ to $x = \infty$. Then

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

i.e.,

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

This is called a fluctuation-dissipation relation of the first kind, and is a central result of fluctuation theory. It relates the variance of fluctuations of x to thermodynamic derivatives *taken in equilibrium*. They are also called correlation-response relations, as the second derivative of the entropy generically gives rise to response functions (heat capacities, compressibilities, etc.). In this way, e.g., the variance of fluctuations in energy is proportional to the equilibrium heat capacity.

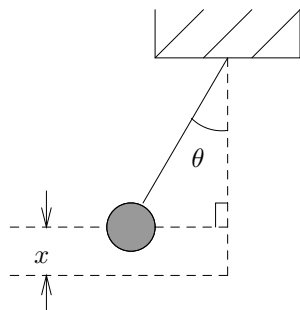
Note that since x is intensive we have

$$\langle (\Delta x)^2 \rangle \sim \mathcal{O}(1/N) = \mathcal{O}(\xi^3 / L^3).$$

Therefore

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

so that thermodynamic second derivatives are determined by microscopic correlations. This is why we need statistical mechanics to calculate thermodynamic second derivatives (or response functions). We need more degrees of freedom than simply x to do anything all that interesting.

Figure 4.6: Simple pendulum of mass m , length l , displaced an angle θ .

4.3.1 Simple mechanical example

Consider the simple pendulum shown in Fig. 4.6. We are going to study fluctuations in its height x induced presumably by random collisions with air molecules surrounding it. The amount of work done by the thermal background to raise the pendulum a height x is mgx . Hence $W = -mgx = -mgl\theta^2/l$ for small θ is the work done by the system. The quantity g is the gravitational acceleration, and we have used the fact that $x = l - l\cos\theta \approx l\frac{\theta^2}{2}$.

We assume next that, the pendulum being a mechanical system, nothing has been done to its internal degrees of freedom so as to change its internal energy. Therefore the work done, has to equal the change in entropy

$$\Delta S = -\frac{mgl\theta^2}{2T}$$

Since $\langle\theta\rangle = 0$ is the equilibrium state, we have

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

This is a very small fluctuation for a pendulum of macroscopic mass m . However, if the mass is sufficiently small, fluctuations are appreciable. This is the origin, for example, of fluctuations of small particles immersed in water, called Brownian motion, as studied by Einstein.

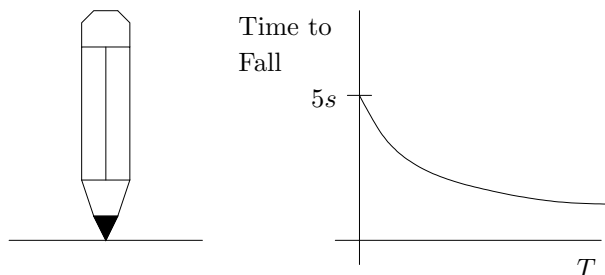


Figure 4.7:

There are many other similar situations described in the book by Landau and Lifshitz. A particularly intriguing one, not given in Landau's book, is the problem of a 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).

4.4 Thermodynamic Fluctuations

We now review the classical treatment of thermodynamic fluctuations. Imagine we have two subsystems making up a larger system. We imagine a fluctuation in which the two subsystem exchange energy and volume. Because of the changes in energy and volume of each subsystem, there is also a change in entropy in both, as well as a change in the total entropy of the system as a consequence of the fluctuation. Since the entropy is extensive, we write the entropy change due to the fluctuation

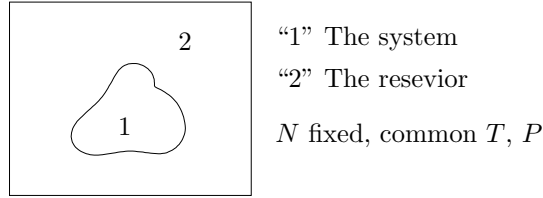


Figure 4.8:

as,

$$\begin{aligned}\Delta S &= \Delta S_1 + \Delta S_2 \\ &= \Delta S_1 + \underbrace{\frac{dE_2 + p dV_2}{T}}_{\text{second law for the reservoir}}\end{aligned}$$

When the reservoir increases its energy, system 1 decreases it by the same amount. The same is true of the volume. Because 1 and 2 are in thermal equilibrium, they both have the same pressure p and temperature T). Therefore we have

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

According to Eq. (4.6), the probability that such a fluctuation happens spontaneously is $e^{\Delta S/k_B}$, or

$$p \propto e^{-\frac{1}{k_B T}(\Delta E - T \Delta S + p \Delta V)} \quad (4.7)$$

where we have dropped the subscript “1” when referring to the system. In the case of a monocomponent fluid that we are considering, further reduction to only two independent variables is possible. Consider,

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

to second order in the fluctuation. Rearranging we have,

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

Recall that $(\frac{\partial E}{\partial S})_V = T$ and $(\frac{\partial E}{\partial V})_S = -p$. We then obtain,

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

Note the similarity with the differential form of the second law of thermodynamics. The extra term in brackets is of second order in the variations, and hence disappears from the differential form. However, it gives the lowest order correction to thermodynamics when considering fluctuations that are finite. Substituting back into Eq. (4.7) we obtain,

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

in which only second order terms in Δ remain. Further manipulation is possible by considering,

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

where we have used the definition of C_V and a Maxwell relation. Likewise

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

where κ_T is the isothermal compressibility. Putting these expressions together and simplifying yields

$$p(\Delta T, \Delta V) \propto e^{\left[\frac{-C_V}{2k_B T^2} (\Delta T)^2 - \frac{1}{2k_B T \kappa_T V} (\Delta V)^2\right]}.$$

As expected, the probability of the fluctuations is Gaussian. Furthermore, the probabilities of fluctuations in T and V are independent since

$$p(\Delta T, \Delta V) = p_T(\Delta T) \cdot p_V(\Delta V)$$

By inspection we have,

$$\langle(\Delta T)^2\rangle = \frac{k_B T^2}{C_V}, \quad \langle(\Delta V)^2\rangle = k_B T \kappa_T V, \quad \langle\Delta T \Delta V\rangle = 0. \quad (4.8)$$

The choice of T and V as independent variables is obviously arbitrary. Many other combinations are possible. For example, if we had chosen ΔS and Δp as our independent variables, we would have obtained

$$p(\Delta S, \Delta p) \propto e^{\left[\frac{-1}{2k_B C_p} (\Delta S)^2 - \frac{\kappa_S V}{2k_B T} (\Delta p)^2\right]}$$

so that

$$\langle(\Delta S)^2\rangle = k_B C_p, \quad \langle(\Delta p)^2\rangle = \frac{k_B T}{\kappa_S V}, \quad \langle\Delta S \Delta p\rangle = 0. \quad (4.9)$$

It is also possible to obtain higher order moments from the Gaussian distributions. So, for example,

$$\langle(\Delta S)^4\rangle = 3\langle(\Delta S)^2\rangle^2 = 3(k_B C_p)^2.$$

Note from all the relations derived that

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

for intensive variables, and

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

for extensive variables. This agrees with our general treatment of systems comprised of many independent parts, and is a general feature of thermodynamic fluctuations in any physical system. Also note the explicit relationship between fluctuations ($\langle(\Delta T)^2\rangle$), thermodynamic derivatives ($C_V = T(\partial S/\partial T)_V$), and microscopic correlations ($1/N = \xi^3/L^3$).

4.4.1 Correlation functions and response functions

Consider the result

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

and rewrite it in terms of the number density $n = N/V$ for fixed N . 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

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

We have on the other hand that Δn can be written as

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

Hence

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

If translational invariance of space is assumed, we can carry out one of the integrals,

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

But $\langle (\Delta n)^2 \rangle = n^2 k_B T \kappa_T / V$, so we obtain a so called thermodynamic “sum rule”:

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

The spatial integral of the correlation function equals a response function.

This relationship can be rewritten in Fourier space. If

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

we then have

$$\lim_{k \rightarrow 0} \hat{C}(k) = n^2 k_B T \kappa_T \quad (4.10)$$

The quantity $\hat{C}(k)$ is called the structure factor. It is directly observable by X Ray, neutron, or light scattering as we will see later. For a liquid, one typically obtains a function of the type shown in Fig. 4.9. The point at $k = 0$ is determined by the result Eq. (4.10).

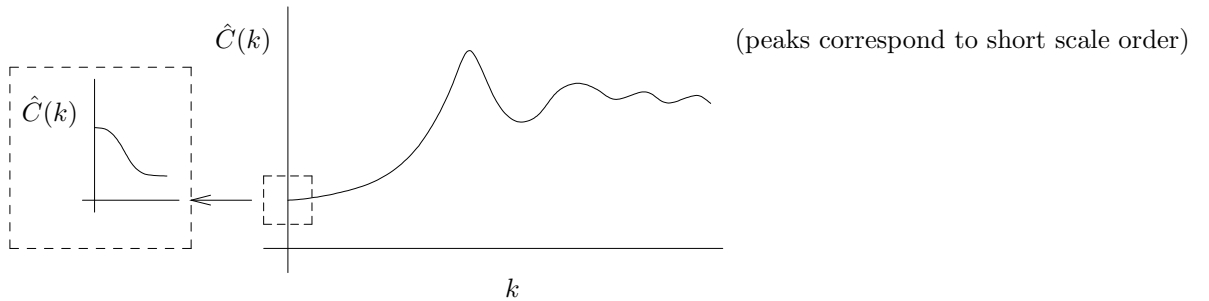


Figure 4.9: Typical structure factor for a liquid.

Chapter 5

Fluctuations of Surfaces

A striking fact is that, although most natural systems are inhomogeneous, with complicated form and structure, often the focus in thermodynamic studies is in uniform systems.

As an example, consider a mono component system which can be in the homogeneous phases of solid, liquid, or vapor. If the system is at constant temperature and pressure, the thermodynamic description is based on the Gibbs free energies $g_s(T, p)$, $g_l(T, p)$, or $g_v(T, p)$. In equilibrium, the state of the system is that which minimizes the Gibbs free energy at each point (T, P) . However, there is a possibility for inhomogeneous configurations 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-vapor coexistence.

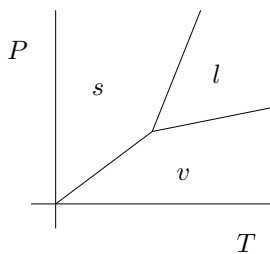


Figure 5.1:

These inhomogeneous states along the lines in the phase diagram are a set of measure zero relative to the area taken up by the p - T plane.

Much modern work in condensed matter physics concerns the study of inhomogeneous configurations. To understand their basic properties, we shall consider the most simple case, namely two equilibrium coexisting phases separated by a surface.

As every one knows, finite volume droplets are spherical, so that very large droplets are locally planar. We will consider first the case of a planar surface separating to coexisting phases. The phenomenology to be discussed is the same for liquid-vapor systems, oil-water systems, magnetic domain walls, as well as many other systems in nature. The reason for this is simple thermodynamics: A system with a surface has an extra free energy that is proportional to the area of that

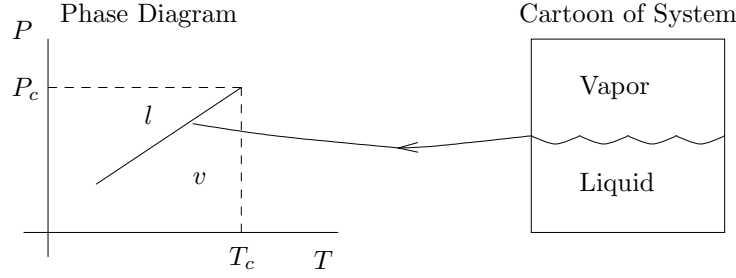


Figure 5.2:

surface. Using the Helmholtz free energy for a system at constant temperature and volume, we can write,

$$F = F_{\text{Bulk}} + \underbrace{\sigma \int d\vec{S}}_{\text{Surface Area}} \quad (5.1)$$

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

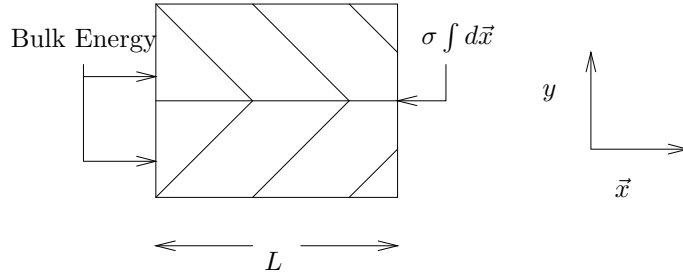


Figure 5.3:

due to the existence of a surface is minimized by minimizing the surface area. If constraints allow it, the surface will be planar. Note also that this surface energy is negligible in the thermodynamic limit. The bulk free energy grows as the volume of the system, whereas the surface excess energy only grows as the surface area. Hence, the second is negligible in a large system. Nevertheless, when fluctuations are allowed, fluctuations in the surface are important, and they will in fact be the subject matter of this chapter.

A coordinate system is introduced as sketched in Fig. (5.3). The coordinate y is perpendicular to the plane of the surface, while \vec{x} is a $(d-1)$ -dimensional vector parallel to it (in d spatial dimensions). Furthermore, the lateral dimension of the system is L , so its total volume is $V = L^d$, while the total surface area is $A = L^{d-1}$.

We shall limit our investigations to long length scale phenomena, namely surface configurations with characteristic wavelengths much larger than molecular dimensions.

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

where a is an approximate molecular size,

$$a \approx 5\text{\AA}. \quad (5.2)$$

Because of this constraint, there is the associated upper limit in wavenumbers,

$$k \leq \Lambda = \frac{2\pi}{a}.$$

A remarkable observation in statistical mechanics is that *many* microscopic systems can give rise to the same behavior on the long length scales just described. A familiar example is a Newtonian

fluid: all such fluids behave in the same way irrespective of their molecular composition. In fact, it is commonplace in statistical mechanics to go much much further than this, and simply introduce new microscopic models that although seemingly unrelated to the microscopic system in question, they share the same long wavelength properties.

5.1 Lattice models of fluctuating surfaces

As a first example of a simple microscopic model of a fluctuating surface, let us consider the model schematically shown in Fig. (5.4). We have discretized the \vec{x} and y axes. Instead of

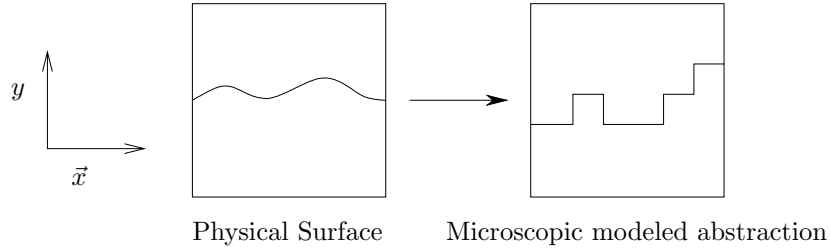


Figure 5.4:

having a continuous variable where $-L/2 \leq x \leq L/2$ and $-L/2 \leq y \leq L/2$ (in $d = 2$), we have discrete elements of surface at $i = -\frac{L}{2}, \dots, -3, -2, -1, 0, 1, 2, 3, \dots, \frac{L}{2}$, where L could be measured in units of a , the small length scale cutoff. We can likewise discretize the y directions as $j = -\frac{L}{2}, \dots, -3, -2, -1, 0, 1, 2, 3, \dots, \frac{L}{2}$. Actually, for greater convenience, let the index along the x axis vary as

$$i = 1, 2, 3, \dots, L \quad (5.3)$$

All the possible states of the system form a discrete set that can easily be enumerated. Any configuration of this fluctuating surface corresponds to a set of heights h_i , where h_i is an integer $-L/2 \leq h_i \leq L/2$ as shown in Fig. 5.5. Statistical mechanical sums over states $\{\nu\}$ can be written as, partition function is

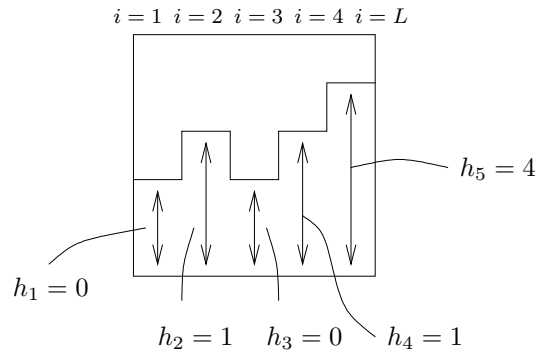


Figure 5.5:

$$\sum_{\{\nu\}} = \sum_{\{h_i\}} = \prod_{i=1}^L \sum_{h_i=-L/2}^{L/2} \quad (5.4)$$

This equation deserves some thought; its meaning is further illustrated in Fig. 5.5.

This model represents a striking simplification relative to a true microscopic description of a liquid-vapor interface, but as $L \rightarrow \infty$, it becomes perfectly accurate.

Second, we need to introduce the energy of each state of the system $E_\nu = E\{h_i\}$ in a way that is consistent with the free energy $\Delta F = \sigma S$ of the physical interface. Physically, we want to discourage surfaces with too much area, since the free energy is minimized by a flat surface. We

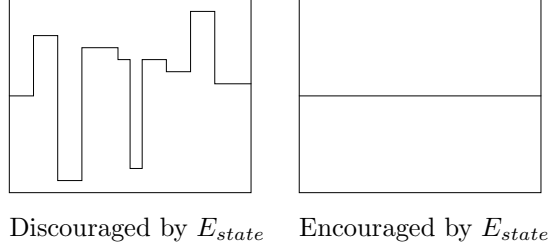


Figure 5.6:

therefore encourage local planarity by defining as excess energy of a surface,

$$E_\nu = J \sum_{i=1}^L (h_i - h_{i+1})^2 \quad (5.5)$$

where J is a positive constant. This particular choice is called the discrete Gaussian solid-on-solid model. With the choice of energy Eq. (5.5), the partition function is

$$Z = \prod_i \sum_{h_i} e^{-\frac{J}{k_B T} \sum_{i'=1}^L |h_{i'} - h_{i'+1}|^2} \quad (5.6)$$

This is for $d = 2$, but the generalization to $d=3$ is easy. Determination of this partition function produces the equilibrium macroscopic properties of the coexisting system with due allowance for all the possible configurations of the fluctuating surface that separates the two bulk phases. Note that this expression does not contain the bulk free energy of the phases as we have focused here on the surface part,

Although it is straightforward to solve for Z numerically, it is not possible to do so analytically. It turns out to be easier to solve for the continuum version of Eq. 5.6 which we shall consider next.

5.2 A continuum model of fluctuating surface

Let $h(\vec{x})$ be the height of the interface at any point \vec{x} as shown in Fig. (5.7). Note that this description

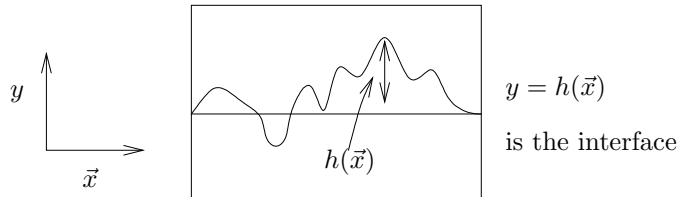


Figure 5.7:

is not completely general from the outset as it does not allow for overhangs or bubbles (Fig. 5.8). Neither one can be described with a single valued function $y = h(\vec{x})$. Of course, although it was not mentioned above, the “lattice model” (Fig. 5.5) also could not have overhangs nor bubbles.

We now have a continuum of states ν , each corresponding to a different configuration of the surface $h(\vec{x})$. Without being mathematically precise, let us mention that the continuum analogue of

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

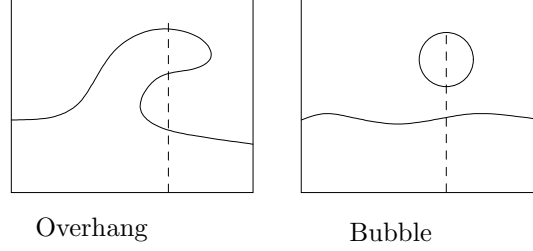


Figure 5.8:

Eq. (5.4) is

$$\begin{aligned} \sum_{\nu} &= \sum_{\{h_{\vec{x}}\}} = \prod_{\vec{x}} \int dh(\vec{x}) \\ &= \int \mathcal{D}h(\cdot) \end{aligned} \quad (5.7)$$

where the last step defines a functional integral. Things are a little subtle because \vec{x} is a continuous variable, and hence the integration is over all possible configurations (or fields) $h(\vec{x})$. In the sections that follow, we will discretize functional integrals before we actually compute them, so we do not need to go into the details of functional calculus here. Also note that whereas the discrete sum over states is dimensionless, $\prod_{\vec{x}} \int dh(\vec{x})$ has dimensions of $(\prod_{\vec{x}} L)$ (infinite dimensional). This constant factor of dimensionality does not affect the results, and we shall ignore it.

We next need to specify the energy of each state ν , or configuration. Since now each state is labeled by a function $h(\vec{x})$, the energy becomes a *functional* of $h(\vec{x})$. For the same reasons as outlines in the discrete model, we choose to assign each configuration an energy which is proportional to its surface $E\{h(\vec{x})\} = \sigma S$, where S is the $(d-1)$ dimensional surface area for a system in d dimensions, and σ is the -constant- interfacial tension.

In order to compute the surface area S of an arbitrary configuration $h(\vec{x})$, we first focus on the simple case of $d = 2$ (Fig. 5.9). Each surface element (perimeter in $d = 2$) is

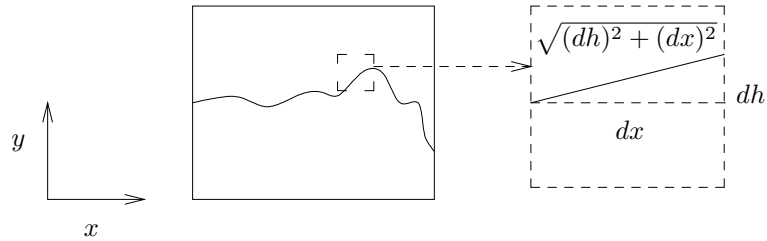


Figure 5.9:

$$\begin{aligned} dS &= \sqrt{(dh)^2 + (dx)^2} \\ &= dx \sqrt{1 + \left(\frac{dh}{dx}\right)^2} \end{aligned} \quad (5.8)$$

In d dimensions one has

$$d^{d-1}S = d^{d-1}x \sqrt{1 + \left(\frac{\partial h}{\partial \vec{x}}\right)^2} \quad (5.9)$$

so that the total surface area is

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

so that

$$E\{h(\vec{x})\} = \sigma \int d^{d-1}\vec{x} \sqrt{1 + \left(\frac{\partial h}{\partial \vec{x}}\right)^2} \quad (5.10)$$

It turns out that this model is not yet analytically tractable. One introduces the further restriction that

$$\left(\frac{\partial h}{\partial \vec{x}}\right)^2 \ll 1 \quad (5.11)$$

that is, that the surface fluctuations considered are only those of small amplitude. Expanding Eq. (5.10) in power series gives

$$E\{h(\vec{r})\} \simeq \sigma L^{d-1} + \underbrace{\frac{\sigma}{2} \int d^{d-1}\vec{x} \left(\frac{\partial h}{\partial \vec{x}}\right)^2}_{\text{Extra surface area due to fluctuation}} \quad (5.12)$$

Finally the partition function of this model is

$$Z = \sum_{\nu} e^{-E_{\nu}/k_B T} = e^{-\frac{\sigma L^{d-1}}{k_B T}} \left(\prod_{\vec{x}} \int dh(x) \exp \left[\frac{-\sigma}{2k_B T} \int d^{d-1}\vec{x} \left(\frac{\partial h}{\partial \vec{x}}\right)^2 \right] \right) \quad (5.13)$$

which is an infinite dimensional functional integral. Although it looks quite formidable, it is straightforward to evaluate. In particle physics a similar integral appears in the so called free-field theory.

5.2.1 Review of Fourier series

We will compute the functional integrals by Fourier transformation. We first a few basic expressions related to Fourier transforms. It is convenient to work with both continuous *and* discrete Fourier transforms. They are defined as

$$h(\vec{x}) = \begin{cases} \int \frac{d^{d-1}\vec{k}}{(2\pi)^{d-1}} e^{i\vec{k}\cdot\vec{x}} \hat{h}(\vec{k}) & \text{Continuoustransform} \\ \frac{1}{L^{d-1}} \sum_{\vec{k}} e^{i\vec{k}\cdot\vec{x}} \hat{h}_{\vec{k}} & \text{discrete series,} \end{cases} \quad (5.14)$$

where

$$\hat{h}(\vec{k}) = \int d^{d-1}\vec{x} e^{-i\vec{k}\cdot\vec{x}} h(\vec{x}) = \hat{h}_{\vec{k}} \quad (5.15)$$

For the particular case of the Dirac delta function, we have

$$\delta(\vec{x}) = \int \frac{d^{d-1}\vec{k}}{(2\pi)^{d-1}} e^{i\vec{k}\cdot\vec{x}}. \quad (5.16)$$

Consider now three dimensional space and a function $F = F(\vec{k})$ function of wavevector \vec{k} . Given a system of lateral dimension L , the corresponding discretization of wavenumber is $\Delta k = 2\pi/L$. Hence, element of volume in Fourier space (the inverse of the density of states) is $\Delta \vec{k} = (2\pi)^3/L^3$. In the limit $L \rightarrow \infty$ we have $\Delta k \rightarrow 0$ and the result,

$$\lim_{L \rightarrow \infty} \frac{1}{L^3} \sum_{\vec{k}} F(\vec{k}) = \int \frac{d\vec{k}}{(2\pi)^3} F(\vec{k}). \quad (5.17)$$

In the particular case that F is the Kronecker delta $F(\vec{k}) = \delta_{\vec{k} \vec{k}_0}$ we obtain

$$\delta_{\vec{k} \vec{k}_0} \rightarrow \frac{L^3}{(2\pi)^3} \delta(\vec{k} - \vec{k}_0). \quad (5.18)$$

We begin by using these results to simplify the form of $E\{h(\vec{x})\}$ in Eq. (5.12). 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 the integral over \vec{k}' . But $h(\vec{x})$ is real, hence $\hat{h}^*(\vec{k}) = \hat{h}(-\vec{k})$ (from Eq. (5.15)), and we have

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

or

$$\int d^{d-1}x \left(\frac{\partial h}{\partial \vec{x}} \right)^2 = \frac{1}{L^{d-1}} \sum_{\vec{k}} k^2 |\hat{h}_{\vec{k}}|^2 \quad (\text{discrete case}) \quad (5.19)$$

Hence the modes $\hat{h}_{\vec{k}}$ are *uncoupled* in Fourier space: the energy is a direct sum over \vec{k} where all the terms in the sum are independent of each other. Therefore we have for the energy of each configuration

$$E(\{\hat{h}_{\vec{k}}\}) = \sigma L^{d-1} + \frac{\sigma}{2L^{d-1}} \sum_{\vec{k}} k^2 |\hat{h}_{\vec{k}}|^2, \quad (5.20)$$

where the configuration is expressed in terms of the Fourier coefficients $\{\hat{h}_{\vec{k}}\}$.

In order to compute the partition function we need to replace the sum over all configurations $\{h(\vec{x})\}$ by a sum over the Fourier modes $\hat{h}_{\vec{k}}$. This is a little subtle because $h(\vec{x})$ is real while $\hat{h}_{\vec{k}}$ is complex, and we also have the relationship $\hat{h}_{\vec{k}}^* = \hat{h}_{-\vec{k}}$. This is standard in Fourier transform theory. Briefly,

$$\sum_{\nu} = \prod_{\vec{x}} \int dh(\vec{x}) = \prod'_{\vec{k}} \int d^2 \hat{h}_{\vec{k}} \quad (5.21)$$

where $d^2 \hat{h}_{\vec{k}} = d\Re(\hat{h}_{\vec{k}}) d\Im(\hat{h}_{\vec{k}})$, that is, the product of the real and imaginary parts, and $\prod'_{\vec{k}}$ means a restriction to only those modes that are independent; i.e., not related through $\hat{h}_{\vec{k}}^* = \hat{h}_{-\vec{k}}$. In two spatial dimensions, the sum extends over one half of the \vec{k} plane as shown in Fig. 7.12.

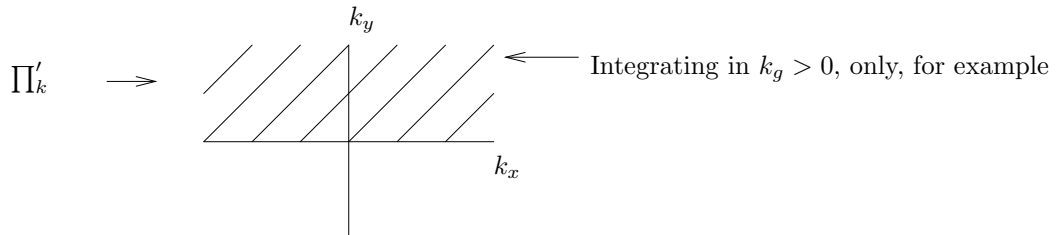


Figure 5.10:

5.2.2 Height-height correlation function

We will not attempt yet the calculation of the full partition function. As will be seen later, our approximation leads to Gaussian distributions and hence it is sufficient to compute second moments of the variable h . We address in this section the calculation of the correlation function

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

the correlation between the surface deflection at points \vec{x} and \vec{x}' . If we assume translational invariance on the plane of the interface, and for convenience we choose

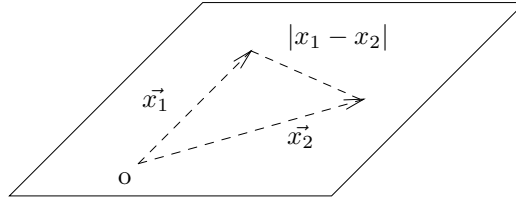


Figure 5.11: \vec{x} plane

$$\langle h \rangle = 0, \quad (5.22)$$

our only task is to evaluate the correlation function

$$G(x) = \langle h(\vec{x}) h(0) \rangle = \langle h(\vec{x} + \vec{x}') h(\vec{x}') \rangle \quad (5.23)$$

We first write,

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

Because of translational invariance, one of the integrals can be eliminated by letting,

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

The Jacobian of this transformation is one: $d\vec{x}d\vec{x}' = d\vec{y}d\vec{y}'$, and

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

The last integral is a delta function,

$$\langle \hat{h}(\vec{k}) \hat{h}(\vec{k}') \rangle = \int d\vec{y} e^{-i\vec{k}\cdot\vec{y}} G(y) (2\pi)^{d-1} \delta(\vec{k} + \vec{k}'). \quad (5.24)$$

Carrying out the integral over \vec{y} to define the Fourier transform of G , we obtain,

$$\langle \hat{h}(\vec{k}) \hat{h}(\vec{k}') \rangle = \hat{G}(\vec{k}) (2\pi)^{d-1} \delta(\vec{k} + \vec{k}'). \quad (5.25)$$

In short, the function $\langle \hat{h}(\vec{k}) \hat{h}(\vec{k}') \rangle$ is related to the Fourier transform of the correlation function $G(\vec{x})$ that we are trying to compute. Instead of calculating the function $G(\vec{x})$ directly we proceed now to calculate the function $\langle \hat{h}(\vec{k}) \hat{h}(\vec{k}') \rangle$ instead. And in order to do so, we turn to a discrete representation so that

$$\langle \hat{h}(\vec{k}) \hat{h}(\vec{k}') \rangle = \hat{G}_{\vec{k}} L^{d-1} \delta_{\vec{k} + \vec{k}', 0}. \quad (5.26)$$

Because of the delta function, all terms are zero except

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

It is this latter second moment of $\hat{h}_{\vec{k}}$ that we now proceed to calculate.

By definition of statistical average, we have (in the discrete representation)

$$\langle |\hat{h}_{\vec{k}}|^2 \rangle = \frac{\prod'_{\vec{k}'} \int d^2 \hat{h}_{\vec{k}'} |\hat{h}_{\vec{k}}|^2 e^{-\frac{\sigma}{2k_B T L^{d-1}} \sum_{\vec{k}''} k''^2 |\hat{h}_{\vec{k}''}|^2}}{\prod'_{\vec{k}'} \int d^2 \hat{h}_{\vec{k}'} e^{-\frac{\sigma}{2k_B T L^{d-1}} \sum_{\vec{k}''} k''^2 |\hat{h}_{\vec{k}''}|^2}} \quad (5.27)$$

Note that we have a multidimensional integral involving $d^2 \hat{h}_{\vec{k}'}$ as independent variables. Except for the integrals involving $\hat{h}_{\vec{k}}$, identical terms on the numerator and denominator cancel. In addition, we have to be careful because for any \vec{k} we have to consider two terms in the right hand side that do not cancel. This is so because

$$\sum_{\vec{k}''} k''^2 |\hat{h}_{\vec{k}''}|^2 = \sum_{\vec{k}''} k''^2 \hat{h}_{\vec{k}''} \hat{h}_{-\vec{k}''}.$$

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

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

giving a factor of 2 for each \vec{k} . With both considerations in mind, Eq. (5.27) reduces to,

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

(These steps are given in Goldenfeld's book in Section 6.3, p.174.) Now, since $\hat{h}_{\vec{k}}$ is a complex variable, let us write

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

(of course, both R and θ depend on \vec{k} , but this does not matter in this calculation). so that

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

as the integrals over θ in the numerator and denominator cancel (the integrand does not depend on θ). We have defined,

$$a = \frac{k_B T}{\sigma k^2} L^{d-1} \quad (5.28)$$

The Gaussian integral can be done exactly, to find our final result

$$\langle |\hat{h}_{\vec{k}}|^2 \rangle = \frac{k_B T}{\sigma k^2} L^{d-1} \quad (5.29)$$

Therefore, from Eq. 5.26 we have

$$\hat{G}(\vec{k}) = \frac{k_B T}{\sigma} \frac{1}{k^2}. \quad (5.30)$$

The spatial correlation function

$$G(\vec{x}) = \langle h(\vec{x}) h(0) \rangle \quad (5.31)$$

can be found by inverse Fourier transform of Eq. (5.30).

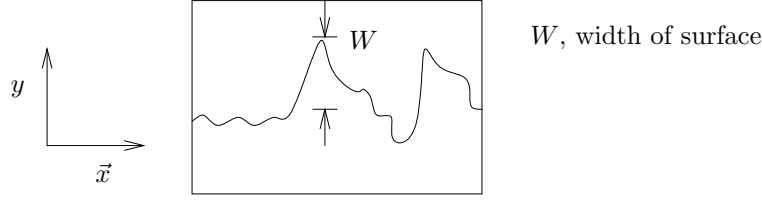


Figure 5.12: Schematic representation of a fluctuating surface and the definition of its width.

Surface width and roughening

Prior to obtaining $G(\vec{x})$, we obtain the equilibrium width of the fluctuating surface. We define the width w as the root mean square of the fluctuations of h averaged over the \vec{x} plane,

$$w^2 = \frac{1}{L^{d-1}} \int d^{d-1}\vec{x} \langle (h(\vec{x}) - \langle h(\vec{x}) \rangle)^2 \rangle \quad (5.32)$$

We set the reference surface at a point such that $\langle h(\vec{x}) \rangle = 0$. Note that since $\langle h(\vec{x})h(\vec{x}) \rangle = G(\vec{x} = 0)$, it follows that $w^2 = G(0)$ with, by definition of the Fourier transform,

$$G(0) = \int \frac{d^{d-1}\vec{k}}{(2\pi)^{d-1}} \frac{k_B T}{\sigma k^2}. \quad (5.33)$$

Before we carry out the integral we note that it potentially has divergences as $\vec{k} \rightarrow 0$, or $\vec{k} \rightarrow \infty$. In order to avoid them, we will always restrict our domain of integration to $\frac{2\pi}{L} \leq |\vec{k}| \leq \Lambda = \frac{\pi}{a}$. The lattice constant a provides a natural cut off at large k (or short length scales), as our calculations and models are not really meaningful below those scales. At the lower end, we introduce the inverse lateral dimension of the system L as the cut off. We will carry out the integrals with this lower limit, and then explore the consequence of taking the thermodynamic limit $L \rightarrow \infty$.

With these two cut offs in mind, it is easy to calculate the integral Eq. 5.33. First in $d = 2$ (so that the interface is one dimensional) we find,

$$G(0) = \frac{k_B T}{\sigma} \frac{1}{2\pi} \int_{-\Lambda}^{\Lambda} dk \frac{1}{k^2} = \frac{k_B T}{\pi \sigma} \int_{\frac{2\pi}{L}}^{\Lambda} dk \frac{1}{k^2} \simeq \frac{k_B T}{2\pi^2 \sigma} L \quad (5.34)$$

in the limit of large L and finite Λ . Therefore in $d = 2$,

$$w = \sqrt{\frac{k_B T}{2\pi^2 \sigma}} L^{1/2} \quad (5.35)$$

Note the perhaps surprising result that the width w diverges in the thermodynamic limit of $L \rightarrow \infty$. This result is a direct consequence of the spectrum of surface fluctuations scaling as $\langle |\hat{h}_{\vec{k}}|^2 \rangle \sim 1/k^2$, which in turn is a direct consequence of the fact that the energy spectrum of surface fluctuations is proportional to $(\partial h / \partial \vec{x})^2$. In physical terms, surface fluctuations (i.e., deviations from planarity) increase the energy of the configuration by an amount proportional to its gradient squared. This increase is small enough for small gradients so that the presence of these distortions at finite temperature dominates the averages, and leads to a divergent variance of surface fluctuations. We will return to this issue several times in this course, especially when we discuss the consequences of broken symmetries.

Before proceeding, it is instructive to point out that the surface width has a strong dependence on the dimensionality of space. Indeed, in order to compute w in $d = 3$, we first note that

$$G(0) = \frac{k_B T}{\sigma} \frac{1}{4\pi^2} 2\pi \int_{\frac{2\pi}{L}}^{\Lambda} \frac{k dk}{k^2} = \frac{k_B T}{2\pi \sigma} \ln(L\Lambda/2\pi). \quad (5.36)$$

In $d = 3$, a two dimensional surface has a width

$$w = \sqrt{\frac{k_B T}{2\pi\sigma}} \left(\ln \frac{L\Lambda}{2\pi} \right)^{1/2} \quad (5.37)$$

We further illustrate the dependence of w on dimensionality by quoting results that have been obtained when d is considered to be a continuous variables (we will not give the details here). It has been found that,

$$w = \begin{cases} \text{undefined,} & d \leq 1 \\ L^x, & 1 < d < 3 \\ (\ln L)^{1/2}, & d = 3 \\ \text{constant,} & d > 3 \end{cases} \quad (5.38)$$

where the exponent x is given by,

$$x = \frac{3-d}{2} \quad (5.39)$$

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

The divergence of w with L for $d < 3$ shown in Eq. 5.37 is called roughening; the surface is macroscopically rough. Note that although the width for $1 < d < 3$ diverges, it is in fact not large

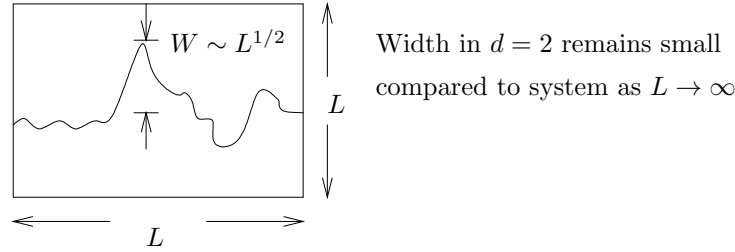


Figure 5.13:

compared to the lateral dimension of the system itself as $L \rightarrow \infty$:

$$\frac{w}{L} = \frac{1}{L^{1-x}} = \frac{1}{L^{(d-1)/2}} \quad (5.40)$$

Hence the width is small for $d > 1$. In $d = 1$, it would appear that the width is comparable to the system size. This means the interface, and indeed phase coexistence itself, does not exist. This is one signature of what is called *lower critical dimension*, at and below which phase coexistence cannot occur at any nonzero temperature. Fluctuations are strong enough to prevent the two phases from being separated by a well defined surface.

In more general cases (not just the one considered), a roughening exponent x is defined by the relation $w \sim L^x$ as $L \rightarrow \infty$. This can be shown to be equivalent to a definition of the surface correlation exponent η_s as

$$\hat{G}(k) \sim \frac{1}{k^{2-\eta_s}} \quad (5.41)$$

as $k \rightarrow 0$. (In the case discussed above $\eta_s = 0$, of course.) The two exponents are then seen to be related by the formula

$$x = \frac{3-d}{2} - \frac{\eta_s}{2} \quad (5.42)$$

The surface turns out to be self-affine in exactly the same way that we discussed earlier for fractal structures (the Von Koch snowflake) since correlation functions like $\hat{G}(\vec{k})$ obey power laws in k . The self-affine or fractal dimension of the rough surface can be shown to be

$$d_s = x + d - 1 \quad (5.43)$$

Spatial correlation function

The spatial correlation function $G(x)$ can be obtained by inverse Fourier transformation of Eq. (5.30). It is not necessary to do the calculation, however, as one can rely on known results of electrostatics. The Poisson equation for the electrostatic potential G due to a point charge at \vec{r}_2 is

$$\nabla^2 G(\vec{r}_1, \vec{r}_2) = -\delta(\vec{r}_1 - \vec{r}_2).$$

The Fourier transform of this equation yields,

$$-k^2 \hat{G}(\vec{k}) = -1, \quad \hat{G}(\vec{k}) = \frac{1}{k^2}.$$

On the other hand, the electrostatic potential is known,

$$G(\vec{r}_1, \vec{r}_2) = \frac{1}{4\pi} \frac{1}{|\vec{r}_1 - \vec{r}_2|} \quad d = 3 \quad (5.44)$$

$$G(\vec{r}_1, \vec{r}_2) = -\frac{1}{2\pi} \ln |\vec{r}_1 - \vec{r}_2| \quad d = 2. \quad (5.45)$$

Therefore in $d = 3$ or a two dimensional surface, one has

$$G(\vec{x}_1, \vec{x}_2) = -\frac{k_B T}{2\pi\sigma} \ln |\vec{x}_1 - \vec{x}_2| \quad (5.46)$$

In the case of $d = 2$, or a one dimensional interface, the correlation function is

$$G(x_1, x_2) \sim |x_1 - x_2|,$$

that is, correlations *grow* with distance.

It is instructive to redo this calculation explicitly, albeit for a slightly different correlation function,

$$g(\vec{x}) = \langle (h(\vec{x}) - h(0))^2 \rangle \quad (5.47)$$

which reduces to

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

The interpretation of $g(\vec{x})$ from Eq. (5.47) is simple: It is the (squared) difference in the heights between two points separated by \vec{x} , as shown in Fig. 5.14. Of course, $g(x \rightarrow L) = 2w^2$ because

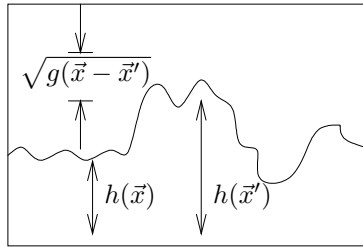


Figure 5.14:

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

Substituting explicitly, we find

$$\begin{aligned} g(\vec{x}) &= 2 \frac{k_B T}{\sigma} \frac{1}{(2\pi)^{d-1}} \int d^{d-1} \vec{k} \left(\frac{1 - e^{i\vec{k} \cdot \vec{x}}}{k^2} \right) \\ &= 2 \frac{k_B T}{\sigma} \frac{1}{(2\pi)^{d-1}} \int d^{d-1} \vec{k} \left(\frac{1 - \cos \vec{k} \cdot \vec{x}}{k^2} \right) \end{aligned} \quad (5.48)$$

where the equality follows from the fact that the integrand (except the exponential term) is even in k , so that the sin part of the Fourier transform vanishes.

Consider first the case $d = 2$, or a one dimensional surface.

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

Let $u = kx$,

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

Now let $\Lambda = \infty$ as there is no divergence at short distances,

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

where we have used the result

$$\int_0^{\infty} du \frac{1 - \cos u}{u^2} = \frac{\pi}{2}.$$

Therefore, we can write for $d = 2$

$$g(x, L) = \frac{k_B T}{\sigma} x f(x/L) \quad (5.49)$$

where we have defined

$$f(y) = 1 - \frac{2}{\pi} \int_0^{2\pi y} du \frac{1 - \cos u}{u^2} \quad (5.50)$$

Equation (5.49) is our final result. It shows that for a one dimensional interface, the correlation function grows linearly with x , with a factor that is a function of x/L . It is instructive to consider the limits of the function f explicitly:

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

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

This function is called a scaling function and is schematically shown in Fig. 5.15. At short distances

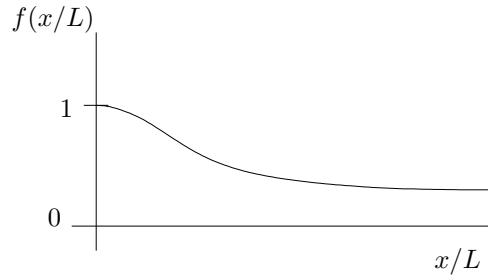


Figure 5.15:

compared with the size of the system ($x/L \rightarrow 0$) $f \rightarrow 1$ and correlations indeed grow linearly.

A similar calculation can be carried out in $d = 3$, to find

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

Evaluating this integral is a bit more complicated, and we just quote the answer

$$g(r) = \frac{k_B T}{4\pi \sigma} \ln r \Lambda \quad (5.51)$$

The calculation can be carried out for a continuous dimension d , and it can be shown that the result is

$$g(r) = \begin{cases} \text{undefined}, & d = 1 \\ r^{2x}, & 1 < d < 3 \\ \ln r, & d = 3 \\ \text{constant}, & d > 3 \end{cases} \quad (5.52)$$

where, again $x = (3 - d)/2$. Correlations are power law functions.

5.2.3 Partition function, free energy, and surface tension

Since the probability distribution function of $\{\hat{h}_{\vec{k}}\}$ is a Gaussian, one can calculate exactly in this case the partition function and the free energy. We do this next. Recall Eq. (5.13),

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

Switching to Fourier space, and by using Eqs. (5.19) and (5.21) we find,

$$e^{-F/k_B T} = Z = e^{-\frac{\sigma L^{d-1}}{k_B T}} \left(\int \prod_{\vec{k}} d^2 \hat{h}_{\vec{k}} \right) \exp \left[\frac{-\sigma}{2k_B T} \frac{1}{L^{d-1}} \sum_k k^2 |\hat{h}_k|^2 \right] \quad (5.53)$$

where we have used the definition of the free energy F in terms of the partition function Z . Equation Eq. (5.53) can be rewritten as

$$Z = e^{-\frac{\sigma L^{d-1}}{k_B T}} \prod_{\vec{k}}' \int d^2 \hat{h}_{\vec{k}} e^{-\frac{\sigma}{2k_B T} \frac{k^2}{L^{d-1}} |\hat{h}_k|^2}$$

since all modes are independent. The Fourier amplitudes $\hat{h}_{\vec{k}}$ are complex so that we write

$$\int d^2 \hat{h}_{\vec{k}} = \int_{-\infty}^{\infty} d\Re(\hat{h}_{\vec{k}}) \int_{-\infty}^{\infty} d\Im(\hat{h}_{\vec{k}}) = \int_0^{2\pi} d\theta \int_0^{\infty} R dR, \quad (5.54)$$

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

$$Z = e^{-\frac{\sigma L^{d-1}}{k_B T}} \prod_{\vec{k}}' \int_0^{2\pi} d\theta \int_0^{\infty} R dR e^{-\left(\frac{\sigma}{2k_B T} \frac{k^2}{L^{d-1}}\right) R^2}$$

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

$$\begin{aligned} Z &= e^{-\frac{\sigma L^{d-1}}{k_B T}} \prod_{\vec{k}}' 2\pi \frac{1}{2\left(\frac{\sigma}{2k_B T} \frac{k^2}{L^{d-1}}\right)} \int_0^{\infty} du e^{-u} \xrightarrow{0} \\ &= e^{-\frac{\sigma L^{d-1}}{k_B T}} \prod_{\vec{k}}' \frac{2\pi k_B T L^{d-1}}{\sigma k^2} \\ &= e^{-\frac{\sigma L^{d-1}}{k_B T}} e^{\ln \prod_{\vec{k}}' \frac{2\pi k_B T L^{d-1}}{\sigma k^2}} \\ &= \exp \left(-\frac{\sigma L^{d-1}}{k_B T} + \sum_{\vec{k}}' \ln \frac{2\pi k_B T L^{d-1}}{\sigma k^2} \right) \end{aligned}$$

This is our final result (eliminating the restriction on the sum over \vec{k}),

$$Z = \exp \left(-\frac{\sigma L^{d-1}}{k_B T} + \frac{1}{2} \sum_{\vec{k}} \ln \frac{2\pi k_B T L^{d-1}}{\sigma k^2} \right) \quad (5.55)$$

and hence the free energy is

$$F = \sigma L^{d-1} - \frac{k_B T}{2} \sum_{\vec{k}} \ln \frac{2\pi k_B T L^{d-1}}{\sigma k^2}. \quad (5.56)$$

This is a very interesting result. The free energy consists of two parts. The first one is the energy due to the planar surface, and it is proportional to the interfacial tension σ . This is the zero temperature limit of Eq. (5.56) in which only the energy (and not the entropy) contribute to the free energy. At $T \neq 0$, fluctuations contribute to the entropy as the system explores a larger region of configuration space, leading to the second term in the equation. Note how the contribution from each mode k appears weighted by $1/k^2$.

The thermodynamic surface tension σ^* is defined as the change in free energy relative to the change in surface,

$$\sigma^* = \left(\frac{\partial F}{\partial S} \right)_{T,V,N} = \left(\frac{\partial F}{\partial L^{d-1}} \right)_{T,V,N} \quad (5.57)$$

which includes, as it should, the effect of fluctuations. In computing the derivative with respect to L it is important to note that \vec{k} do implicitly depend on the dimension L (see Fig. 5.16). The dependence of k on L is simple, and it follows that the combination (kL) is precisely independent of L . We therefore rewrite Eq. (5.56) as,

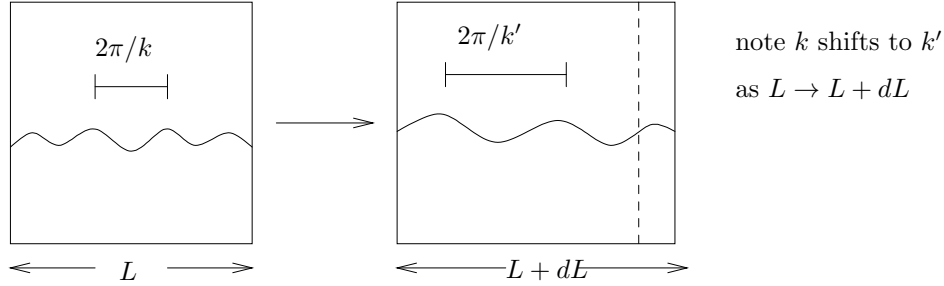


Figure 5.16:

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

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,

$$F = \sigma L^{d-1} - \left[\frac{k_B T}{2} \left(\frac{d+1}{d-1} \right) (\ln L^{d-1}) \sum_{\vec{k}} 1 \right] - \frac{k_B T}{2} \sum_{\vec{k}} \ln \frac{2\pi k_B T}{\sigma (kL)^2}. \quad (5.58)$$

The second term in the right hand side is independent of \vec{k} and hence it has all been removed from the sum, leaving only the term $\sum_{\vec{k}} 1$ which equals the total number of states. Also note that the

third term in the right hand side is independent of L as only the combination (kL) appears here. Therefore,

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

in the discrete representation that we are using. In order to compute the total number of states, it is better to use a continuum representation,

$$\sigma^* = \sigma - \frac{k_B T}{2} \frac{d+1}{d-1} \int \frac{d^{d-1} \vec{k}}{(2\pi)^{d-1}}, \quad (5.59)$$

which yields,

$$\int \frac{d^{d-1} \vec{k}}{(2\pi)^{d-1}} = \frac{1}{2\pi} \int_{-\Lambda}^{\Lambda} dk = \frac{\Lambda}{\pi} \quad d = 2 \quad (5.60)$$

$$\int \frac{d^{d-1} \vec{k}}{(2\pi)^{d-1}} = \frac{1}{4\pi^2} \int_0^{2\pi} d\theta \int_0^{\Lambda} k dk = \frac{1}{4\pi^2} \pi \Lambda^2 = \frac{\Lambda^2}{4\pi} \quad d = 3 \quad (5.61)$$

$$(5.62)$$

where $\Lambda = \pi/a$. We find for the surface tension,

$$\begin{aligned} \sigma^* &= \sigma - \frac{3k_B T}{2\pi} \Lambda \quad d = 2, \\ \sigma^* &= \sigma - \frac{k_B T}{4\pi} \Lambda^2 \quad d = 3. \end{aligned} \quad (5.63)$$

Note that the surface tension decreases with temperature. At T increases, fluctuations away from planarity become more important, and therefore the surface appears, on average, to be more floppy. This is of course a measurable effect, one that follows directly from the effect of thermal fluctuations on macroscopic quantities. Note also that by increasing the temperature, the surface tension σ^* becomes zero. At this point, fluctuations are so large that there is effectively no restoring force on the surface to bring it back to planarity. In effect, it disappears, and with it, the coexistence of two separate phases.

5.3 Impossibility of Phase Coexistence in d=1

The surface tension σ^* is not defined for $d = 1$. In our model, σ^* actually becomes $-\infty$ at $d = 1$. This anomaly signals the lack of two phase coexistence in one dimension. In fact, this is a theorem due to Landau and Peierls, which is worth quickly proving. (See Goldenfeld's book, p.46.)

A homogeneous phase has a bulk energy $E_B \propto L^d$. Adding a surface to the system adds an extra energy $E_S \propto L^{d-1}$. Therefore, if one adds an interface to an initially uniform phase, there is a net increase in the energy of the system by $\Delta E \propto L^{d-1}$. Since the energy increases by

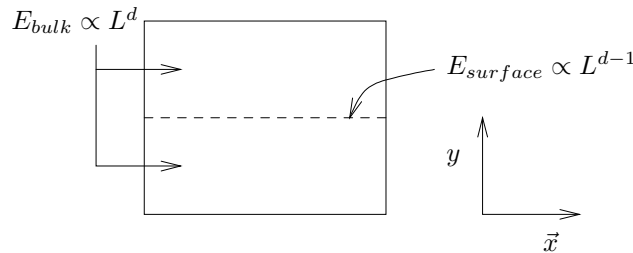


Figure 5.17:

adding interfaces, it looks like the system wants to have the smallest possible number of them. In conventional coexistence, this would mean two phases separated by a single surface.

However, at finite temperature the equilibrium state of the system is the one that minimizes the free energy $F = E - TS$, not just the energy E . In order to estimate the entropy change due to introducing an interface in the system, we need to count the number of microscopic states that are consistent with one single macroscopic interface (see Fig. 5.18). The total number of such states

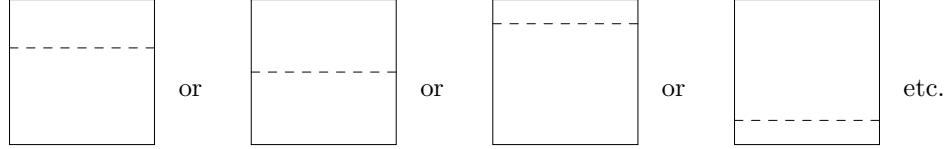


Figure 5.18:

is L along the y axis as shown, as the interface could be anywhere along that axis. Of course the interface could be in any location along the remaining axes, and hence there are $\sim L^d$ ways to place a single interface in our d dimensional system. The entropy gained by introducing a single surface is therefore

$$\Delta S \sim k_B \ln L^d = k_B d \ln L. \quad (5.64)$$

Hence, the total change in the free energy due to introducing a single surface is

$$\Delta F = \mathcal{O}(L^{d-1}) - k_B T d \mathcal{O}(\ln L) \quad (5.65)$$

For $d > 1$ and $L \rightarrow \infty$, the free energy increases if a surface is added to the system. Under normal circumstances, a system with no interfaces will be the equilibrium state. But for $d = 1$ (and $T > 0$) the energy loss due to introducing an interface is finite, whereas the entropy gain increases as $\ln L$. Therefore for sufficiently large L , the system can always reduce its free energy by adding one interface.

Of course, once an interface is added, the same argument can be repeated for any of the two bulk phases resulting in the conclusion that it would be favorable to add a second interface, and a third, and so on. This means that coexisting phases are not stable for $d = 1$. Since adding one interface decreases F , we can add another and another and another, and keep on decreasing F until there are no longer bulk coexisting phases, just a mishmash of tiny regions of phase fluctuations. Hence, for systems with short-ranged interactions (consistent with our assumption that the energy of an interface is simply $\sim L^{d-1}$), and $T > 0$ (so that entropy matters), there can be no phase coexistence in $d = 1$.

5.4 Order of magnitude calculations in $d=3$

Some of our earlier consideration may seem a little academic, including the fact that they are done in arbitrary dimension d . We mention, for example, that our results for $d = 2$ apply to films absorbed on liquid surfaces and to ledges in crystal surfaces, as shown in Fig. 5.19. Experiments have 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, Eqs. (5.37) and (5.63),

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

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

where we have defined

$$T^* = \frac{\sigma}{4\pi k_B (\Lambda/2\pi)^2} \quad (5.66)$$

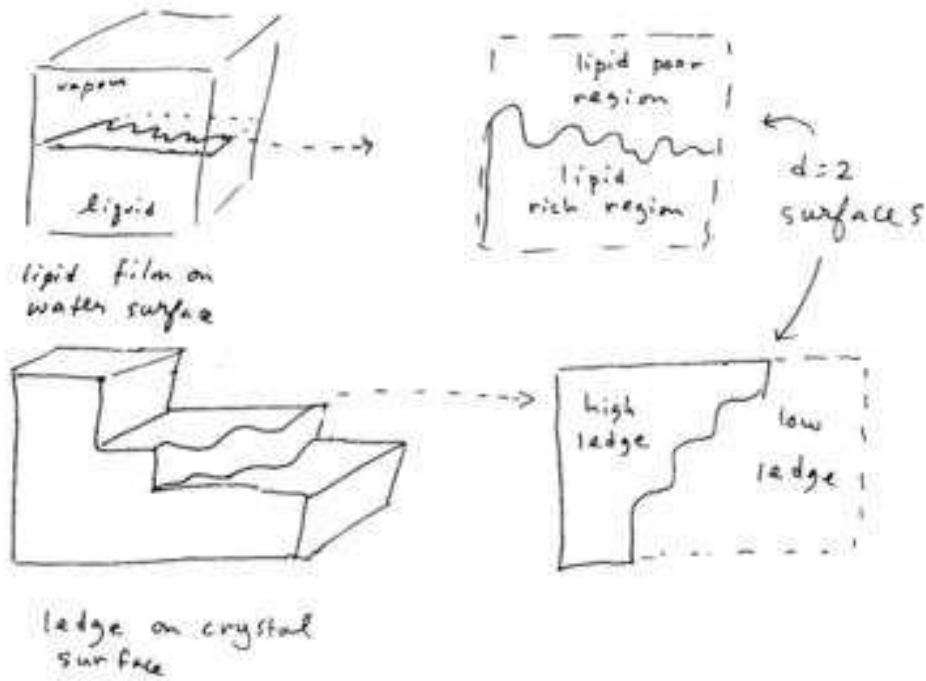


Figure 5.19:

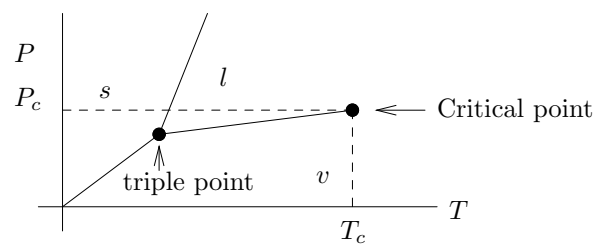


Figure 5.20: Phase Diagram

Consider now the phase diagram of a pure substance (Fig. 5.20). Our analysis describes the behavior along the liquid-vapor coexistence line where an interface separates the two bulk equilibrium phases. Near the triple point of a typical fluid, the surface tension is approximately $\sqrt{\frac{k_B T_t}{2\pi\sigma}} \approx 1\text{\AA}$, where T_t is the triple point temperature. Similarly, one roughly has $\frac{\Lambda}{2\pi} \approx 10\text{\AA}$. Therefore,

$$w = 1\text{\AA} \left(\ln \frac{L}{10\text{\AA}} \right)^{1/2} \quad (5.67)$$

In real systems, it turns out that the divergence is *very* weak:

$$w = \begin{cases} 2.6\text{\AA} & L = 10000\text{\AA} \text{ (light wavelength)} \\ 4.3\text{\AA} & L = 10 \text{ cm (coffee cup)} \\ 5.9\text{\AA} & L = 1000 \text{ km (ocean)} \\ 9.5\text{\AA} & L = 10^{30} \text{ m (universe)} \end{cases}$$

Of course, in a liquid-vapor system on Earth, gravity plays a role. The extra energy due to gravity is

$$E_G = \frac{1}{2} \int d^{d-1}\vec{x} (\rho_l - \rho_v) g h^2(x)$$

where $\rho_{l,v}$ is the mass density of the liquid, vapor, and g is the intensity of the gravitational field. This extra energy changes our results in a way you should work out for yourself.

As far as the surface tension is concerned, note that σ^* vanishes at T^* (assuming $\sigma \neq \sigma(T)$ is a constant). It is tempting, although a little too rough, to interpret $T^* = T_c$ to be the critical point. With numbers, one gets $T^* \approx 1.4T_t$, which is the right order of magnitude for T_c . One indication that our model is only approximate is that fact that is known that

$$\sigma^* \sim \left(1 - \frac{T}{T_c} \right)^\mu$$

for $T \rightarrow T_c$, where $\mu \approx 1.24$ in $d = 3$. So our theory and interpretation of T^* as T_c gives $\mu = 1$ instead. This is not too bad, considering we only went to lowest order in the expansion of $E_\nu = \int d\vec{x} \sqrt{1 + (\partial h / \partial \vec{x})^2}$ which clearly underestimates the fluctuations of $h(x)$ near the critical point, and did not consider bulk properties at all.

5.5 (*) A different view of gradient terms

Quite obviously, terms proportional to gradients of $h(\vec{x})$ in the energy E increase the energy of non uniform configuration, and hence favor a planar surface. We briefly discuss here a different interpretation of these terms, to show that they can be understood as the nonlocal interaction between neighboring points $h(\vec{x})$ and $h(\vec{x}')$. Consider

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

where we have used 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. We find

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

We have repeatedly integrated by parts recalling that terms involving surfaces at ∞ vanish. Hence,

$$\int d^{d-1} \vec{x} \left(\frac{\partial h}{\partial \vec{x}} \right)^2 = \int d^{d-1} \vec{x} \int d^{d-1} \vec{x}' h(\vec{x}) M(\vec{x} - \vec{x}') h(\vec{x}') \quad (5.68)$$

where the kernel M is,

$$M(\vec{x}) = -\frac{\partial^2}{\partial \vec{x}^2} \delta(\vec{x}) \quad (5.69)$$

If the delta function is approximated by a continuous function of small but finite width, the M has the form shown in Fig. 5.21. This type of interaction matrix appears in many different contexts. It

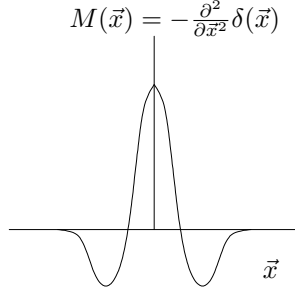


Figure 5.21:

causes $h(\vec{x})$ to interact with $h(\vec{x}')$ in their contribution to the energy E .

Chapter 6

Broken Symmetry and Correlation Functions

A major concept in modern Physics is symmetry. The results obtained for fluctuating surfaces are but examples of more general properties of correlations in systems with broken symmetries. We explore such general properties in this chapter.

First, let us elaborate on what is a spontaneously broken symmetry. In the case of a fluctuating surface, note that in the partition function

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

the energy of a state has the following “symmetry” or invariance

$$E_{\nu}(h + \text{constant}) = E_{\nu}(h), \quad (6.1)$$

because

$$E_{\nu} = \sigma L^{d-1} + \frac{\sigma}{2} \int d^{d-1} \vec{x} \left(\frac{dh}{d\vec{x}} \right)^2$$

so that it depends on derivatives of h , but not on the value of h itself. In other words, the energy of an interface is invariant under the uniform translation of the entire interface by an arbitrary constant, and so the partition function also has the same invariance. Therefore we would expect this symmetry to be respected by any quantity that is obtained from the partition function Z . However in a given experiment that average location of the physical interface $\langle h \rangle$ has one and only one value. In other words, even though by symmetry the interface could be on average at any location in the system, in a physical experiment it picks only one specific value. This is referred to as spontaneously broken symmetry, as in one physical realization the configuration of the interface breaks the symmetry of the system (under translation or rotation in this case). Correlations and averages do not respect that full symmetry of the energy E_{ν} or the partition function Z .

Homogeneous systems, whether liquid or vapor, are translationally and rotationally invariant. However, a system with a surface, as drawn in Fig. 6.1, only has translational invariance along the plane of the interface \vec{x} but not along the direction normal to the interface y . Hence the system with a surface has a lower symmetry compared to the bulk liquid or vapor systems, so that some symmetry must have been broken. The perfect symmetry of the homogeneous system is lost precisely because h , the broken symmetry variable, has a definite value in any given experimental realization, and hence the state breaks the symmetry of the underlying system’s energy.

It is instructive to qualitatively argue that the facts that Z and E_{ν} have a particular symmetry has implications on the form of the correlation functions involving the broken symmetry variable h . Consider the case of $d = 2$. Since a surface can be anywhere with equal probability, it will be at different locations in different systems or in different experiments. This is illustrated in Fig. 6.2. Imagine now a very large system thought of being composed of the parts shown in Fig. 6.2. In a part

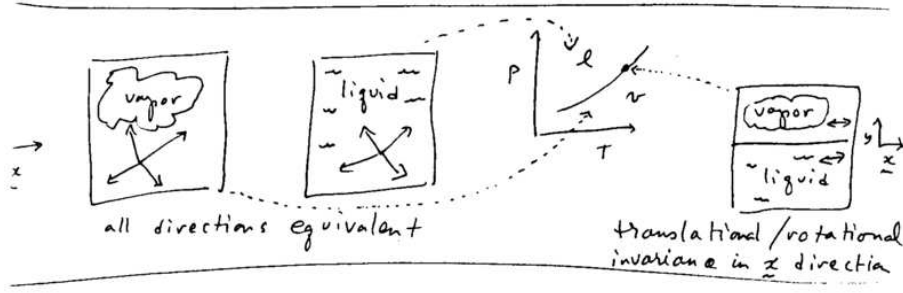


Figure 6.1:

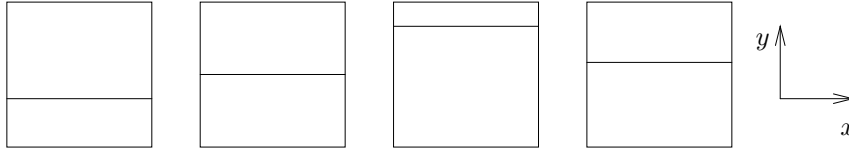


Figure 6.2:

of the system $\langle h \rangle$ might have one particular value, while far away along the x axis, it might want to have another value (both of the same energy, and hence should appear with equal probability). This hypothetical situation is depicted schematically in Fig. 6.3. Imagine we now approximate the interface in the figure by a sinusoidal function of wavenumber k , with $k \ll 1$. Since we know that the energy cost of producing this distortion is

$$\Delta E(k) = \frac{\sigma}{2} k^2 |\hat{h}_k|^2,$$

we find that $\Delta E(k) \rightarrow 0$ as $k \rightarrow 0$. Indeed, if one has a very large system encompassing the four parts shown in Fig. 6.2, with the four subsystems behaving like independent subsystems because of

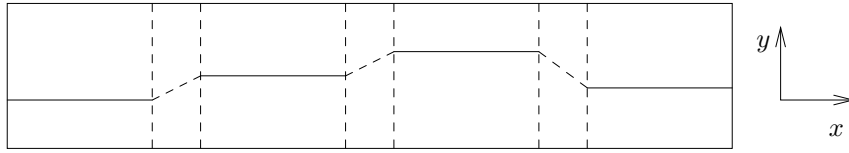


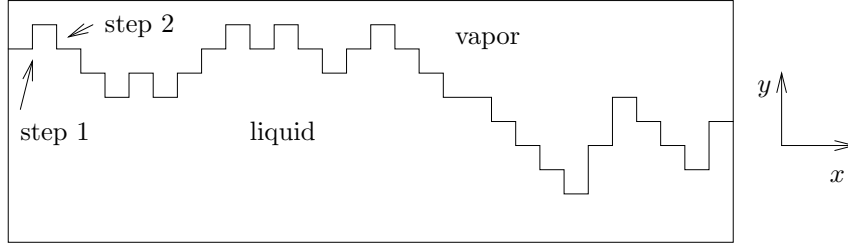
Figure 6.3: Surface of a very big system

the symmetry, the energy cost of such a fluctuations is infinitesimal as $k \rightarrow 0$. If we now use the equipartition of energy, one would find,

$$\frac{k_B T}{2} = \frac{1}{(L^{d-1})^2} \frac{\sigma}{2} k^2 \langle |\hat{h}_k|^2 \rangle, \quad \langle |\hat{h}_k|^2 \rangle \sim \frac{k_B T}{\sigma} \frac{1}{k^2}.$$

As we know, the Fourier transform of the correlation function of the broken symmetry variable h diverges as $1/k^2$ as $k \rightarrow 0$. However one needs not know the form of E_ν to obtain the result $\langle |\hat{h}_k|^2 \rangle \sim 1/k^2$. Rather, it is sufficient to know that the interface breaks a symmetry of Z and E_ν . This result is the main topic of this chapter, and it is known as Goldstone Theorem.

Let us recast this argument in a way that does not require the use of the equipartition theorem. As shown in Fig. 6.3, fluctuations of size, say, $\pm \Delta y$ occur over a long distance Δx . The actual values of Δy (which could be $\sim \text{\AA}$) and of Δx (which could be $\sim \text{meters}$) are not important for the argument. Since h is a broken symmetry variable, if one considers a displacement on the interface of the order of Δx to the right, then the interface would be distorted up or down *at random* by

Figure 6.4: Interface (in units of Δx and Δy)

an amount Δy . This is shown schematically in Fig. 6.4. Of course, the sequence of displacements constitutes a one dimensional random walk along the y axis with x playing the role of time. For a random walk, the root mean square deviation of y from its initial value is

$$y_{\text{RMS}} \sim x^{1/2}$$

for large “time” x . The deviation y_{RMS} is related to the correlation function $(g(x))^{1/2}$ from Eq. 5.47, so we have $g(x) \sim x$. Fourier transforming gives $\hat{G}(k) \sim 1/k^2$ for small k , and $\langle |\hat{h}_k|^2 \rangle \sim 1/k^2$. (The Fourier transform is not simple; $g(x)$ is a divergent function of x , leading to the singularity of $\hat{G}(k)$ at small k). This argument is only in $d = 2$, but it can be generalized to higher dimensions.

It is the fact that h is a broken symmetry variable (and hence the random walk of the variable y) which gives rise to the $1/k^2$ dependence; no specific details about the interaction energies are necessary. Note also that fluctuations in a way react to a symmetry breaking by “shaking it” in such a way as to try and restore the broken symmetry. The rough width of the surface is this attempt to remove it from the system and so restore the full translational invariance in the y direction, as

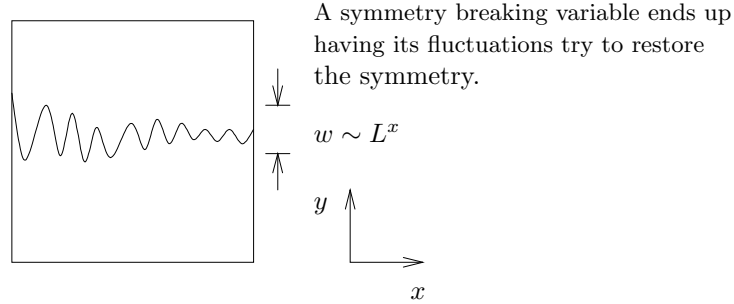


Figure 6.5:

shown in Fig. 6.5. In fact for $d = 1$, the fluctuations are so strong that the full symmetry *is* in fact restored, and one has an equilibrium state with the full translational invariance in the y direction (since coexisting phases cannot exist). In $1 < d \leq 3$, 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 our cartoons of rough surfaces attempt to show them as self similar and self affine objects much in the same sense as the von-Koch snowflake discussed earlier. Note an important new ingredient: *broken symmetry*. Spontaneously broken symmetry leads to power law correlations and to self similar structures.

6.1 Goldstone's Theorem

Let $B(\vec{r})$ be a general broken continuous symmetry variable, with \vec{r} a field point in d dimensions). For a system with translational invariance,

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

The theorem states that

$$\lim_{k \rightarrow 0} \hat{G}(k) \rightarrow \infty$$

with a divergence no weaker than

$$\hat{G}(k) \sim 1/k^2$$

We will not give a proof of theorem, rather present its physical underpinnings. For simplicity, we will consider $\langle B \rangle = 0$. If this is not the case, one can define $B' = B - \langle B \rangle$. We start by expanding the free energy as a Taylor series in B around the equilibrium state

$$F(B) = F_0 + \left(\frac{\partial F}{\partial B} \right)_0 B + \frac{1}{2} \left(\frac{\partial^2 F}{\partial B^2} \right)_0 B^2 + \dots$$

The first term in the right hand side F_0 is a constant, and $\left(\frac{\partial F}{\partial B} \right)_0 = 0$ if the system is in equilibrium. Hence the Taylor series is just,

$$F(B) = \frac{1}{2} \left(\frac{\partial^2 F}{\partial B^2} \right)_0 B^2 + \dots \quad (6.2)$$

We have to be a bit more careful with the Taylor expansion given that B is really a function of \vec{r} . Treated as a function of many variables, the correct expansion reads

$$F(B) = \frac{1}{2} \int d^d \vec{r} \int d^d \vec{r}' \left(\frac{\partial^2 F}{\partial B(\vec{r}) \partial B(\vec{r}')} \right)_0 B(\vec{r}) B(\vec{r}') \quad (6.3)$$

Since $\frac{\partial^2 F}{\partial B(\vec{r}) \partial B(\vec{r}'')}$ is a thermodynamic derivative, in a system which is translationally invariant, we write

$$\frac{\partial^2 F}{\partial B(\vec{r}) \partial B(\vec{r}'')} = C(\vec{r}, \vec{r}') = C(\vec{r} - \vec{r}') \quad (6.4)$$

By using Parseval's theorem in Fourier transforms, one has

$$F(B) = \frac{1}{2} \int \frac{d^d \vec{k}}{(2\pi)^d} \int \frac{d^d \vec{k}'}{(2\pi)^d} \left(\frac{\partial^2 F}{\partial \hat{B}(\vec{k}) \partial \hat{B}(\vec{k}')} \right)_0 \hat{B}(\vec{k}) \hat{B}(\vec{k}'). \quad (6.5)$$

Translational invariance implies

$$\frac{\partial^2 F}{\partial \hat{B}(\vec{k}) \partial \hat{B}(\vec{k}')} = (2\pi)^d \delta(\vec{k} + \vec{k}') \hat{C}(k) \quad (6.6)$$

If $B(\vec{x})$ is real, then $\hat{B}^*(-k) = \hat{B}(k)$. We finally obtain

$$F(B) = \frac{1}{2} \int \frac{d^d k}{(2\pi)^d} \hat{C}(k) |\hat{B}(k)|^2. \quad (6.7)$$

If we consider a discrete system, we would obtain instead,

$$F(B) = \frac{1}{2L^d} \sum_{\vec{k}} \hat{C}_{\vec{k}} |\hat{B}_{\vec{k}}|^2 \quad (6.8)$$

So far this is straightforward, and we have not used the knowledge that B is a broken symmetry variable. This has important implications on the dependence of the coefficient $\hat{C}(\vec{k})$. Anticipating the singular nature of the response near $k = 0$, let us expand

$$\hat{C}(\vec{k}) = \text{constant} + \text{constant } k^2 + \text{constant } k^4 + \dots \quad (6.9)$$

where we have assumed that \hat{C} is analytic near $\vec{k} = 0$. Analyticity of this function of many variables allows terms in the expansion of the form $\vec{k} \cdot \vec{k} = k^2$ and $(\vec{k} \cdot \vec{k})^2 = k^4$, but *not* non analytic terms such as $(\vec{k} \cdot \vec{k})|\vec{k}| = k^3$ (i.e., that have an absolute value sign).

We now use the fact that B *breaks* a symmetry of F . Hence F can have *no* dependence on B in the limit $\vec{k} = 0$. Otherwise, the free energy would depend on the spatial average of B , which would contradict the assumption B representing a symmetry of the system. Therefore

$$\lim_{k \rightarrow 0} \hat{C}(\vec{k}) = 0 \quad (6.10)$$

and, if we have analyticity,

$$\hat{C}(k) \sim k^2 \text{ as } k \rightarrow 0 \quad (6.11)$$

This gives

$$F(B) \propto \sum_{\vec{k}} k^2 |\vec{B}_{\vec{k}}|^2 \quad (6.12)$$

This is the same as Eq. (5.20) for the case of a fluctuating surface, so we have

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

where

$$\hat{G}(\vec{k}) \propto \frac{1}{k^2}, \quad \vec{k} \rightarrow 0.$$

If $\hat{C}(\vec{k})$ is not analytic as assumed in Eq. (6.9), we would get the weaker condition

$$\lim_{k \rightarrow 0} \hat{G}(\vec{k}) \rightarrow \infty$$

Finally, it is interesting to note that if B is *not* a broken symmetry variable, from Eq. (6.9) we have

$$F(B) = \text{constant} \sum_{\vec{k}} (k^2 + \xi^{-2}) |\hat{B}_{\vec{k}}|^2$$

where ξ^{-2} is a constant. Taken back to $\langle \hat{B}(\vec{k}) \hat{B}(\vec{k}') \rangle = (2\pi)^d \delta(\vec{k} + \vec{k}') \hat{G}(\vec{k})$ it leads to

$$\hat{G}(\vec{k}) \propto \frac{1}{k^2 + \xi^{-2}} \quad (6.13)$$

as $k \rightarrow 0$. Inverse Fourier transformation gives

$$\langle B(\vec{r}) B(0) \rangle = G(\vec{r}) \propto e^{-r/\xi}$$

as $r \rightarrow \infty$. In short, absence of the broken symmetry removes power law correlations, and brings instead an exponential decay of the correlation function. This was the signature of a system comprised of many independent sub systems. The quantity ξ can be identified as the correlation length discussed earlier.

Chapter 7

Equilibrium correlation functions and Scattering

Equilibrium fluctuations can be understood as local deviations in a subsystem that is part of a larger system in equilibrium. We study in this chapter how to describe this equilibrium fluctuations.

Consider as an example a system of N classical and indistinguishable particles at constant temperature. The partition function is

$$Z(N, V, T) = \frac{1}{N!} \int \frac{d^N \vec{r} d^N \vec{p}}{h^{3N}} e^{-\beta E(\vec{r}^N, \vec{p}^N)}.$$

It is useful to divide the energy into kinetic and potential, and the joint probability of space and momentum,

$$p(\vec{r}^N, \vec{p}^N) = \phi(\vec{p}^N) P(\vec{r}^N),$$

with

$$P(\vec{r}^N) = \frac{e^{-\beta V(\vec{r}^N)}}{\int d^N \vec{r} e^{-\beta V(\vec{r}^N)}} \quad (7.1)$$

where V is the potential energy function. By definition, P is the probability of observing the system in any specific spatial configuration while it evolves in configuration space in equilibrium. For an interacting system, P does not factorize in simpler distribution functions involving a single particle, and hence one often resorts to defining marginal probability distribution functions. The function P itself has too much information to be of any use if it were possible to compute it.

The first function is

$$\rho_1(\vec{r}) = N \int d\vec{r}_2 \dots d\vec{r}_N P(\vec{r}^N). \quad (7.2)$$

The integral is the probability that particle 1 is at \vec{r} . Since all particles are equivalent, the prefactor N leads to ρ_1 being the probability that *any* particle is at \vec{r} . The definition is such that if the system is uniform $P(\vec{r}^N) = 1/V^N$, and $\rho_1 = N/V = \rho$, the equilibrium density.

Analogously, one defines

$$\rho_2(\vec{r}_1, \vec{r}_2) = N(N-1) \int d\vec{r}_3 \dots d\vec{r}_N P(\vec{r}^N), \quad (7.3)$$

which is the joint probability of finding any one particle at \vec{r}_1 and any other particle at \vec{r}_2 . If the system is uniform, then $\rho_2 = N(N-1)/V^2 \simeq \rho^2$. In the case of the density ρ , it is customary to define the correlation function g as

$$g(\vec{r}_1, \vec{r}_2) = \frac{\rho_2(\vec{r}_1, \vec{r}_2)}{\rho^2}. \quad (7.4)$$

This function has an easy interpretation from probability theory. Recall that $P(A, B) = P(A|B)P(B)$. The joint probability distribution (for a system with translational invariance) would be $P(A, B) = \rho_2(0, \vec{r})$, and since $P(B) = \rho$, then

$$P(A|B) = \frac{P(A, B)}{P(B)} = \rho g(\vec{r}). \quad (7.5)$$

That is, $\rho g(\vec{r})$ is the conditional probability of finding a particle at \vec{r} *given that* there is a particle at the origin.

7.1 Measurement of $g(\vec{r})$ by diffraction

Consider a radiation scattering experiment as indicated in Fig. 7.1. Scattering experiments involve

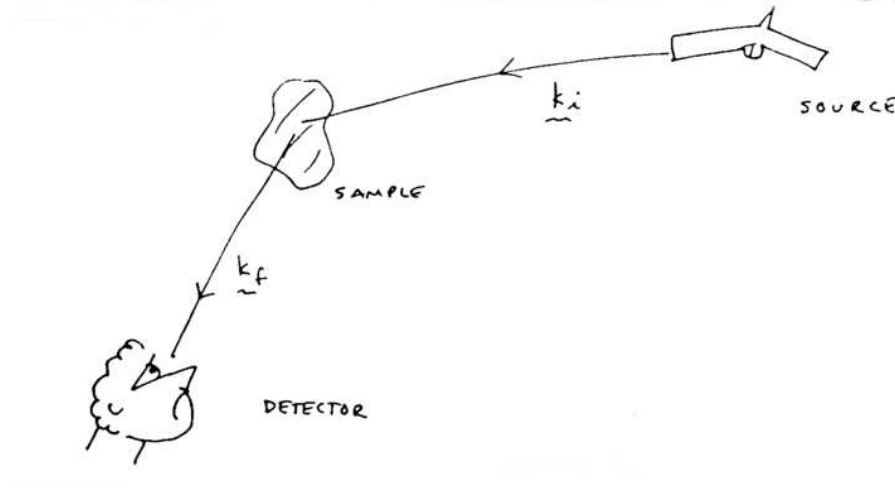


Figure 7.1:

the change in momentum of some incident particle or radiation by a sample. The radiation of choice depends on the scale of the fluctuations that one wishes to probe. In general, the wavelength of the radiation must be of the same order of magnitude as the spatial scale of the fluctuations.

The scattering wave following a scattering event in the sample is a radial outgoing wave. Its electric field is

$$E \sim E_0 \frac{e^{i\vec{k}_f \cdot \vec{r}}}{r}.$$

where \vec{k}_f is the outgoing wavevector, and ω the angular frequency of the wave. We consider here only elastic scattering in which the frequency of the outgoing wave is the same as the frequency of the incident wave (inelastic scattering is also used as a diagnostic tool, but we will not address it here). The wave that arrives at the detector after scattering off an element of volume in the sample at \vec{r}_j is,

$$E \sim E_0 \frac{1}{|\vec{R}_D - \vec{r}_j|} e^{i\vec{k}_i \cdot \vec{r}_j + i\vec{k}_f \cdot (\vec{R}_D - \vec{r}_j)} e^{-i\omega t}. \quad (7.6)$$

considering the path that the radiation has traveled from the source, to the sample, and to the detector. We now define the scattering wavevector as $\vec{k} = \vec{k}_f - \vec{k}_i$ and rewrite,

$$E \sim E_0 \frac{1}{|\vec{R}_D - \vec{R}_S|} e^{i\vec{k}_f \cdot \vec{R}_D} e^{-i\vec{k} \cdot \vec{r}_j} e^{-i\omega t},$$

where we have assumed that the detector is far from the sample and hence $|\vec{R}_D - \vec{r}_j| \approx |\vec{R}_D - \vec{R}_S|$, where \vec{R}_S is the location of the sample.

If one now adds the scattering from all elements of volume in the sample, the radiation that arrives at the detector is,

$$E \sim E_0 \frac{1}{|\vec{R}_D - \vec{R}_S|} e^{i\vec{k}_f \cdot \vec{R}_D} \left(\sum_j e^{-i\vec{k} \cdot \vec{r}_j} \right) e^{-i\omega t}.$$

We now compute the intensity of the radiation $I(\vec{k}) = E^*(\vec{k})E(\vec{k})$ and compute the statistical average, to find,

$$I(\vec{k}_f) \sim \frac{1}{|\vec{R}_D - \vec{R}_S|^2} N S(\vec{k}), \quad (7.7)$$

where we have defined the so called structure factor as,

$$S(\vec{k}) = \frac{1}{N} \left\langle \sum_{j=1}^N \sum_{m=1}^N e^{i\vec{k} \cdot (\vec{r}_m - \vec{r}_j)} \right\rangle. \quad (7.8)$$

The scattering intensity is a function of \vec{k} , the scattering wavevector.

We now show how to relate the structure factor to the correlation function. Write,

$$S(\vec{k}) = \frac{1}{N} \left\langle \sum_{l=1, l=j}^N 1 + \sum_{l \neq j}^N e^{i\vec{k} \cdot (\vec{r}_l - \vec{r}_j)} \right\rangle = \frac{1}{N} N + \frac{1}{N} N(N-1) \langle e^{i\vec{k} \cdot (\vec{r}_l - \vec{r}_j)} \rangle.$$

By writing the thermal average explicitly, we find,

$$S(\vec{k}) = 1 + \frac{N(N-1)}{N} \frac{\int d^N \vec{r} e^{-\beta V(\vec{r}^N)} e^{i\vec{k} \cdot (\vec{r}_1 - \vec{r}_2)}}{\int d^N \vec{r} e^{-\beta V(\vec{r}^N)}} = 1 + \frac{N(N-1)}{N} \frac{\int d\vec{r}_1 \int d\vec{r}_2 e^{i\vec{k} \cdot (\vec{r}_1 - \vec{r}_2)} \int d\vec{r}_3 \dots d\vec{r}_N e^{-\beta V(\vec{r}^N)}}{\int d^N \vec{r} e^{-\beta V(\vec{r}^N)}}.$$

The last integral in the right hand side and the numerator allow us to introduce the correlation function ρ_2 ,

$$S(\vec{k}) = 1 + \frac{1}{N} \int d\vec{r}_1 d\vec{r}_2 \rho_2(\vec{r}_1, \vec{r}_2) e^{i\vec{k} \cdot (\vec{r}_1 - \vec{r}_2)}.$$

Change $\vec{r}_2 = \vec{r}_1 + \vec{r}$, assume translational invariance, and integrate over \vec{r}_1 to obtain our final result

$$S(\vec{k}) = 1 + \frac{1}{N} \int d\vec{r}_1 \int d\vec{r} \rho^2 g(\vec{r}) e^{-i\vec{k} \cdot \vec{r}} = 1 + \rho \int d\vec{r} g(\vec{r}) e^{-i\vec{k} \cdot \vec{r}}. \quad (7.9)$$

In short, the scattered intensity is proportional to the structure factor as shown in Eq. (7.7), and the structure factor is the Fourier transform of the pair correlation function, Eq. (7.9). Therefore the scattered intensity measured in a diffraction experiment is directly the Fourier transform of the pair correlation function.

In practice, experiments measure the diffraction intensity as a function of the scattering angle θ , the angle between the incident wavevector \vec{k}_i and the outgoing wavevector \vec{k}_f . Since the scattering wavevector $\vec{k} = \vec{k}_f - \vec{k}_i$ and since we focus on elastic scattering $|\vec{k}_i| = |\vec{k}_f|$, we then have $k = k_i + \sin(\theta/2) + k_f \sin(\theta/2)$ or

$$k = \frac{4\pi}{\lambda_i} \sin \frac{\theta}{2}.$$

Hence changing the location of the detector relative to the sample (and hence changing θ produces the scattered intensity for different wavevectors \vec{k} .

One can distinguish two cases for $\hat{\rho}(q)$: (i) it corresponds to a local density or local concentration that is uniform in equilibrium, and (ii), it corresponds to a more general order parameter that may not be uniform in equilibrium (e.g., periodic). In the first case, Fourier transformation of a uniform density gives

$$S(\vec{k}) \propto \delta(\vec{k}). \quad (7.10)$$

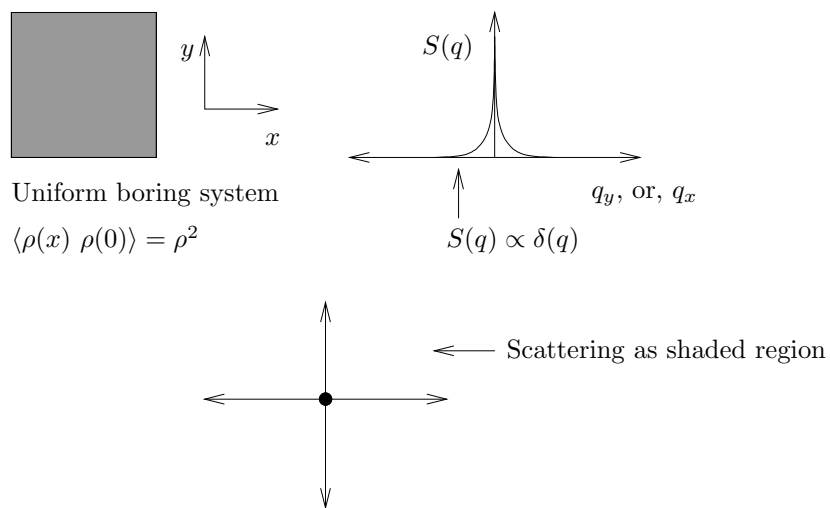


Figure 7.2:

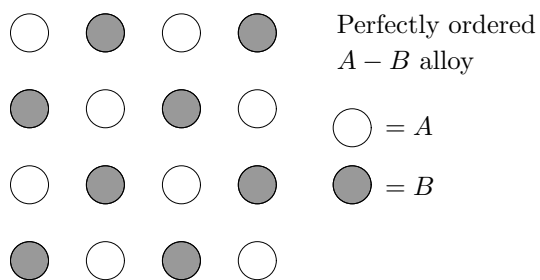


Figure 7.3:

The cartoon in Fig. 7.2 may make this clearer. The second case is more subtle. An example is the sub lattice concentration of a binary alloy. Say atoms of the two species are A and B, and assume that locally A wants to be neighbors with B but not with another A. This makes the equilibrium state look like a checker-board as shown in Fig. 7.3.

Assuming that the scattered wave is different from atoms A or B atoms this equilibrium structure can be seen by analyzing the structure factor $S(k)$. Note that the A atoms occupy a sub lattice, which has twice the lattice spacing of the original lattice as shown in Fig. 7.4. Hence the density

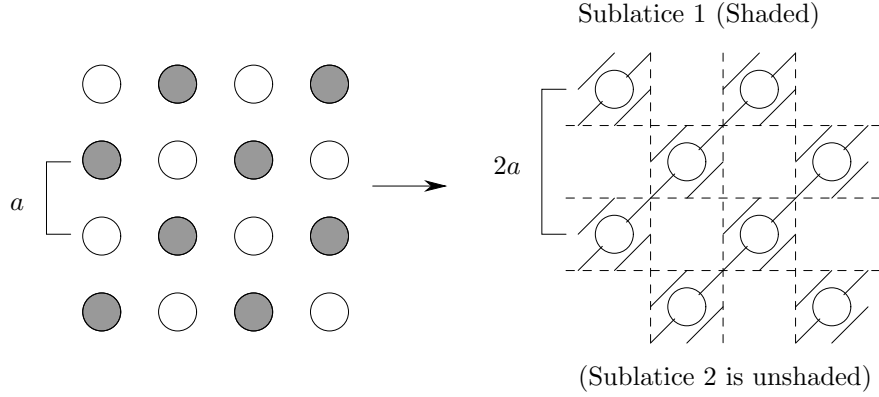


Figure 7.4:

of A atoms is uniform on this sub lattice with twice the lattice constant of the original system, and the scattering will show peaks *not* at $\vec{k} = 0$ but at $\vec{k} = \vec{k}_0$ corresponding to that structure,

$$S(\vec{k}) \propto \sum_{\vec{k}_0} \delta(\vec{k} - \vec{k}_0) \quad (7.11)$$

where $k_0 = 2\pi/(2a)$, and a is the original lattice spacing. The same scattered intensity is observed

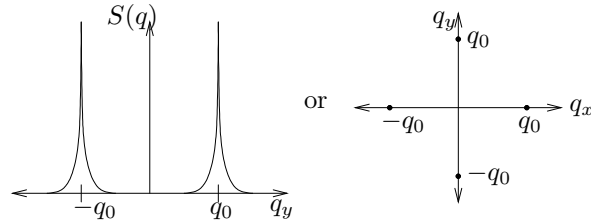


Figure 7.5:

for a crystal in which Bragg peaks form at specific positions in \vec{k} 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 a single crystal perfectly aligned with the incident radiation so that one obtains spots on the k_x and k_y axes as shown. Instead they appear at some random orientation (Fig. 7.6). This might seem trivial (one could just realign the sample), except that often a real crystal is a poly crystal, that is, it is comprised of many single crystal domains of different orientations that scatter simultaneously. In this case, all the orientations indicated above are smeared into a ring of radius k_0 .

7.2 Scattering off a flat interface

The two examples just given concern systems that are single phase. Consider now the case of a flat interface -without roughening- separating two uniform phases as shown in Fig. 7.8. For simplicity

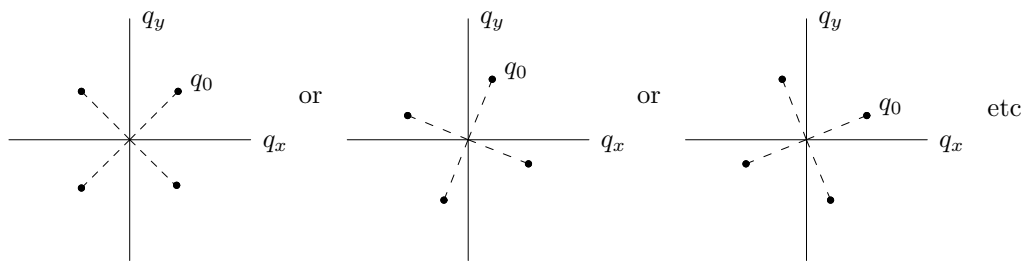


Figure 7.6:

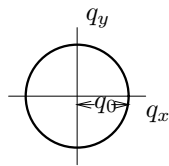


Figure 7.7: Average of Many Crystallites

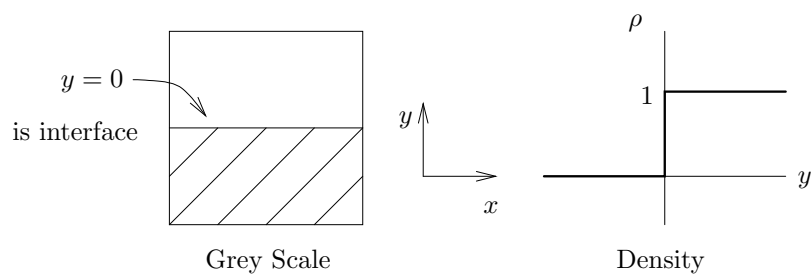


Figure 7.8:

let us focus on an interface embedded in dimension $d = 2$. If the density is uniform in the bulk phases, we can write, $\rho(x, y) = A \theta(y) + B$, 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, $\rho(x, y) = \theta(y)$

We now proceed to compute the scattering intensity that follows from this configuration. In order to do so, we first need to determine the Fourier transform of the Heaviside function θ . We first note that $\partial_y \rho(y) = \delta(y)$. Hence $ik_y \hat{\rho}(k_y) = 1$. We therefore find $\hat{\rho}(k_y) = 1/ik_y$. This result is correct up to a constant: $\partial_y(\rho(y) + C) = \delta(y)$ also satisfies the equation for any arbitrary constant C . Fourier transforming, we find $ik_y (\hat{\rho}(k_y) + C\delta(y)) = 1$. In order to determine the constant C we note that

$$\theta(x) + \theta(-x) = 1$$

Let us compute,

$$\hat{\theta}(k_y) = \int_{-\infty}^{\infty} dy \theta(y) e^{-ik_y y} = \int_{-\infty}^{\infty} (-dy) \theta(-y) e^{-ik_y(-y)} = - \int_{\infty}^{-\infty} dy (1 - \theta(y)) e^{-ik_y(-y)},$$

and hence,

$$\hat{\theta}(k_y) = \int_{-\infty}^{\infty} e^{-i(-k_y)y} - \int_{-\infty}^{\infty} dy \theta(y) e^{-i(-k_y)y} = \delta(k_y) - \hat{\theta}(-k_y).$$

Finally then $\hat{\theta}(k_y) + \hat{\theta}(-k_y) = \delta(k_y)$. This relation applied to ρ gives,

$$\hat{\rho}(k_y) + \hat{\rho}(-k_y) = \delta(k_y).$$

Substituting the expression for $\hat{\rho}$ above, we have,

$$\frac{1}{ik} - C\delta(k_y) + \frac{1}{-ik_y} - C\delta(k_y) = \delta(k_y), \quad \text{or} \quad C = -\frac{1}{2}.$$

The final result for the Fourier transform of the configuration with an interface is

$$\hat{\rho}(k_y) = \frac{1}{ik_y} + \frac{1}{2}\delta(k_y). \quad (7.12)$$

More properly in two dimensions, we should write

$$\hat{\rho}(\vec{k}) = \frac{1}{ik_y} \delta(k_x) + \frac{1}{2} \delta(\vec{k}), \quad (7.13)$$

as the system is uniform in the x direction. Therefore

$$S(k) = |\hat{\rho}(k)|^2 \propto \frac{1}{k_y^2} \delta(k_x) \quad (7.14)$$

where we have omitted the delta function from bulk scattering.

Assume now more generally that the unit vector normal to the interface is \hat{n} , and that the unit tangential vector is \hat{n}_\perp , (Fig. 7.9), then

$$S(k) \propto \frac{1}{(\vec{k} \cdot \hat{n})^2} \delta(\vec{k} \cdot \hat{n}_\perp) \quad (7.15)$$

Then the scattered intensity would appear as shown in Fig. 7.10: a streak pointed in the direction of \hat{n} (\hat{y} in Fig. 7.10).

7.3 Scattering from a line defect

Assume an idealized model of a line defect (density inhomogeneity) given by,

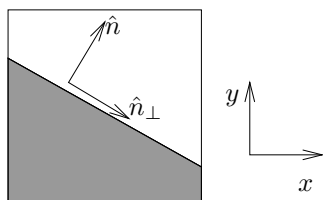


Figure 7.9:

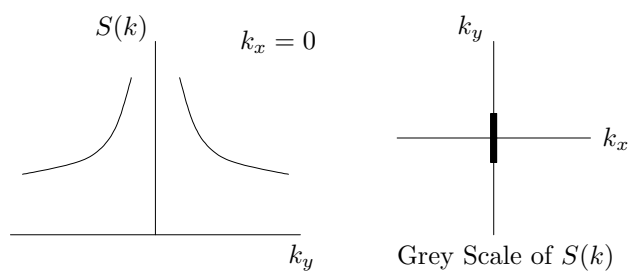
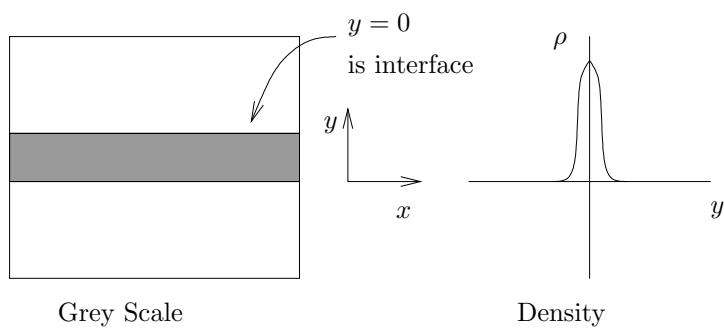
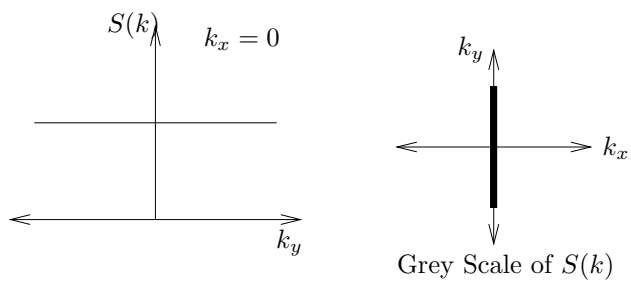
Figure 7.10: Scattering from interface $y = 0$ 

Figure 7.11:

Figure 7.12: Scattering from a line defect at $y = 0$

$$\rho(x, y) = \delta(y). \quad (7.16)$$

Then $\hat{\rho} = \delta(k_x)$ and $S(k) \propto \delta(k_x)$. If the unit vector normal to the surface of the defect is \hat{n} , and the tangential vector is \hat{n}_\perp , the proper expression is $S(k) \propto \delta(\vec{k} \cdot \hat{n}_\perp)$ for a line defect.

For a defect of small (but finite) extent, we can imagine a situation in which the density within the defect region ρ is not uniform, but it rather corresponds to a modulated phase of the type discussed above for a binary alloy. We then have to incorporate the wavenumber \vec{k}_0 corresponding to the modulated phase. One finds,

$$S(k) \propto \delta((\vec{k} - \vec{k}_0) \cdot \hat{n}_\perp) \quad (7.17)$$

A similar argument can be advanced for the case of an interface now separating modulated phases. One would find,

$$S(k) \propto \frac{1}{|(\vec{k} - \vec{k}_0) \cdot \hat{n}|^2} \delta((\vec{k} - \vec{k}_0) \cdot \hat{n}_\perp). \quad (7.18)$$

7.4 Scattering from rough interfaces

So far we have focused on idealized situations in which the interface is simply a line of discontinuity, and ignored any possible structure.

Before addressing the case of a rough interface, let us consider a diffuse interface in which the density profile is smooth instead of discontinuous, as shown in Fig. 7.13. Although the precise

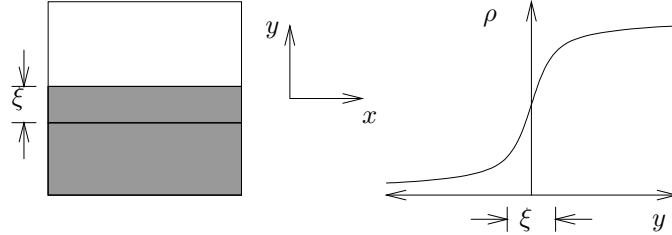


Figure 7.13: A diffuse interface, with no roughness

functional dependence does not matter, let us consider the following case:

$$\rho(x, y) \sim \tanh(y/\xi), \quad (7.19)$$

with ξ assumed small determining the width of the interface. In this case, one can show that

$$S(k) = |\hat{\rho}(k)|^2 \sim \frac{\delta(k_x)}{k_y^2} (1 + \mathcal{O}(\xi k_y)^2) \quad (7.20)$$

for an interface at $y = 0$. Diffuseness only effects scattering at $k_x = 0$ and for large k_y of order $1/\xi$

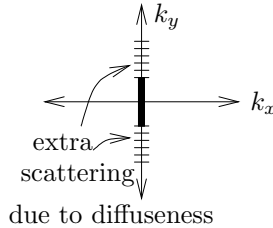


Figure 7.14:

(again, assuming that ξ is small so that ξk_y is comparable to 1). The resulting scattering is sketched in Fig. 7.14. This result is general, and does not depend on our choice of \tanh for ρ .

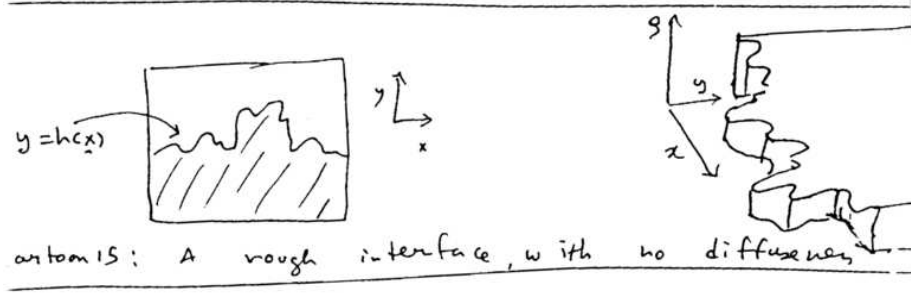


Figure 7.15: A very rough interface, with no diffuseness

A rough interface is not planar, but rather described by some profile $y = h(x)$ in, say $d = 2$. If the rough interface is not diffuse, that is, if it can be locally approximated by a surface of discontinuity, the density satisfies,

$$\rho(x, y) = \theta(y - h(x)). \quad (7.21)$$

If $h(x)$ involves only small distortions, then by expanding in Taylor series we can write,

$$\rho \approx \theta(y) - h(x)\delta(y) + \dots \quad (7.22)$$

since $d\theta/dy = \delta(y)$. Taking the Fourier transform and squaring gives

$$S(k) = \langle |\hat{\rho}(k)|^2 \rangle \sim \frac{\delta(k_x)}{k_y^2} + \langle |\hat{h}(k_x)|^2 \rangle, \quad (7.23)$$

where we have omitted unimportant delta functions at the origin. We obtain the rather remarkable result that the intensity of the scattered radiation depends on to the spectrum of surface fluctuations $\langle |\hat{h}(k_x)|^2 \rangle$. (In earlier cases in this chapter we did not consider thermal averages as the configuration of the interface was unique and fixed. For a rough surface we can consider all possible equilibrium configurations that the system adopts during the duration of the scattering experiment).

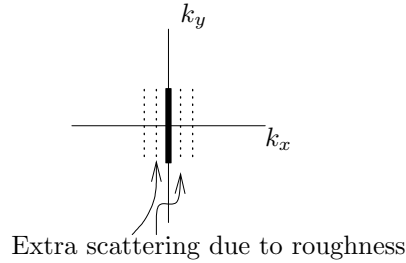


Figure 7.16:

7.5 (*) Scattering from many planar and randomly oriented surfaces

Evidently the scattering from many randomly oriented surfaces gives a pinwheel like cluster of streaks as shown in Fig. 7.17. Angularly averaging over all orientations of the interface gives

$$S = \frac{\int d\hat{n} |\hat{\rho}_k|^2}{\int d\hat{n}} \quad (7.24)$$

where \hat{n} is the unit vector normal to the surface. We write

$$\hat{n} = -\sin\theta\hat{x} + \cos\theta\hat{y} \quad (7.25)$$

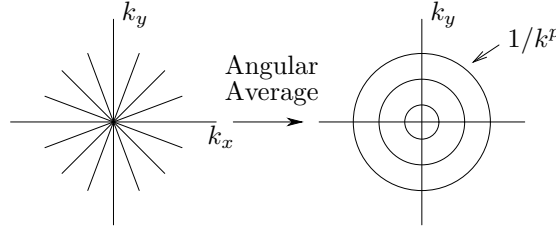


Figure 7.17:

and

$$\hat{n}_\perp = \cos \theta \hat{x} + \sin \theta \hat{y} \quad (7.26)$$

with

$$\vec{k} = k \cos \beta \hat{x} + k \sin \beta \hat{y} \quad (7.27)$$

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 $d = 2$ and assuming a uniform distribution of orientations θ)

$$\begin{aligned} S &= \frac{1}{2\pi} \int_0^{2\pi} d\theta \frac{\delta[k \cos \beta \cos \theta + k \sin \beta \sin \theta]}{|-k \cos \beta \sin \theta + k \sin \beta \cos \theta|^2} \\ &= \frac{1}{2\pi} \frac{1}{k^3} \int_0^{2\pi} d\theta \frac{\delta[\cos(\theta - \beta)]}{|\sin(\theta - \beta)|^2} \\ &= \frac{1}{2\pi} \frac{1}{k^3} \int_0^{2\pi} d\theta \frac{\delta[\cos \theta]}{|\sin \theta|^2} \end{aligned}$$

But $\cos \theta = 0$ at $\theta = \pi/2, 3\pi/2$ where $|\sin \theta|^2 = 1$, so that

$$S(k) = \frac{1}{\pi} \frac{1}{k^3} \quad (7.28)$$

Similarly, in arbitrary dimension one has

$$S(k) \propto \frac{\delta(\vec{k} \cdot \hat{n}_\perp)}{|\vec{k} \cdot \hat{n}|^2} \sim \frac{\delta k^{d-1} n_\perp}{k^2} \sim \frac{1}{k^{2+d-1}} = \frac{1}{k^{d+1}}. \quad (7.29)$$

We have used the facts that the tangent space to the surface is $d - 1$ dimensional, and that $\delta(ax) = \delta(x)/a$. for scattering from a surface.

Scattering from a line defect which has $|\hat{\rho}(k)|^2 \propto \delta(k_x)$, one similarly obtains

$$S(k) \propto \frac{1}{k^{d-1}} \quad (7.30)$$

We shall call these results, Porod's Law. They involve *only* geometrical considerations about the interface.

The case where ρ is not the order parameter, but a modulated structure exists instead is more subtle. If the wavenumber of the modulated phase is \vec{k}_0 then this introduces a new angle via

$$\vec{k}_0 = k_0 \cos \alpha \hat{x} + k_0 \sin \alpha \hat{y} \quad (7.31)$$

Now there are two possible averages

1. α or β , average over crystallite orientation, or, angles in a non angularly resolved detector. These are equivalent (look at the figure for a few seconds), and so we only need to do one of these two averages.
2. θ , averages over surface orientation.

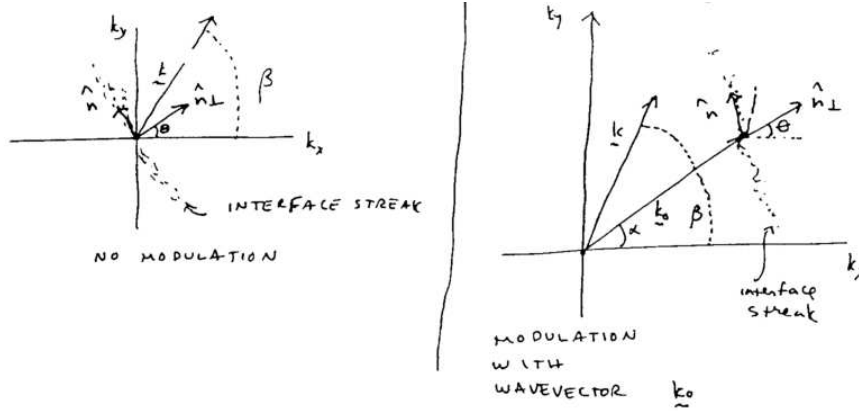


Figure 7.18:

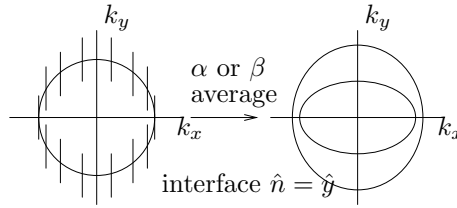


Figure 7.19:

If we average over α or β , for a fixed \hat{n} , it is easy to anticipate the answer as shown in fig. 7.19, where in the cartoon the regions around $(k_x = k_0, k_y = 0)$ retain the singularity of the original $\delta(k_x - k_0)/(k_y - k_0)^2$.

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 $\vec{k} = 0$

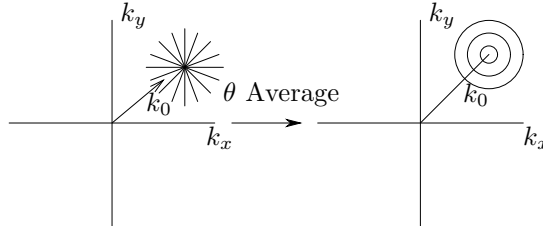


Figure 7.20:

to $\vec{k} = \vec{k}_0$, that is

$$S_{\text{surface}}(k) \propto \frac{1}{|\vec{k} - \vec{k}_0|^{d+1}} \quad (7.32)$$

$$S_{\text{line}}(k) \propto \frac{1}{|\vec{k} - \vec{k}_0|^{d-1}}$$

If a further average is done over crystallite orientation, we can take these expressions and conduct an angular average over the angle between \vec{k} and \vec{k}_0 . Clearly after averaging we will have

$$\bar{S} = S\left(\frac{|k| - |k_0|}{|k_0|}\right) \equiv S(\Delta k) \quad (7.33)$$

First, let

$$\phi = \beta - \alpha \quad (7.34)$$

so that

$$(\vec{k} - \vec{k}_0)^2 = k^2 - 2kk_0 \cos \phi + k_0^2$$

or, on using

$$k \equiv k_0(\Delta k + 1) \quad (7.35)$$

we have

$$(\vec{k} - \vec{k}_0)^2 = k_0^2 \{ (\Delta k)^2 + 2(1 - \cos \phi)(\Delta k + 1) \} \quad (7.36)$$

Hence in two dimensions, for a surface, we have

$$\bar{S}(\Delta k) = \frac{1}{2\pi k_0^3} \int_0^{2\pi} d\phi \frac{1}{|(\Delta k)^2 + 2(1 - \cos \phi)(\Delta k + 1)|^{3/2}} \quad (7.37)$$

which is a complex integral. We will work out its asymptotic limits $\Delta k \gg 1$ and $\Delta k \ll 1$.

First, 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)

$$\bar{S}(\Delta k) \sim \frac{1}{(\Delta k)^{d+1}}, \quad \Delta k \gg 1 \quad (7.38)$$

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. So we keep a small 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,

$$\bar{S}(\Delta k) \sim \frac{1}{(\Delta k)^2} \text{ as } \Delta k \rightarrow 0 \quad (7.39)$$

for *all* dimensions d .

If one does this in d dimensions with

$$S = \frac{1}{|\vec{k} - \vec{k}_0|^\gamma} \quad (7.40)$$

then

$$\bar{S} = \begin{cases} \frac{1}{(k-k_0)^{\gamma-(d-1)}}, & \frac{k-k_0}{k_0} \ll 1 \\ \frac{1}{(k-k_0)^\gamma}, & \frac{k-k_0}{k_0} \gg 1 \end{cases} \quad (7.41)$$

If $\gamma - (d-1) > 0$. For the case $\gamma = d-1$, one obtains

$$\bar{S} \sim \begin{cases} -\ln(k - k_0), & \frac{k-k_0}{k_0} \ll 1 \\ \frac{1}{(k-k_0)^{d-1}}, & \frac{k-k_0}{k_0} \gg 1 \end{cases} \quad (7.42)$$

It should be noted that the actual scattering is given by a difficult integral like Eq. 7.37, which is

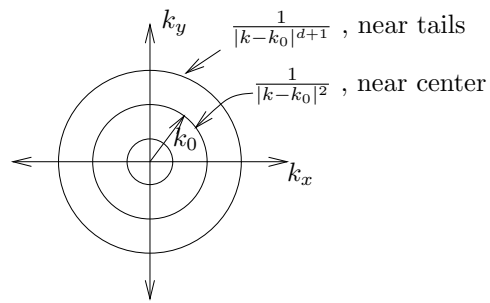


Figure 7.21:

pretty close to an elliptical integral.

Chapter 8

Fluctuations and Crystalline Order

We explore in this chapter an important example of a spontaneously broken symmetry: the formation of a solid crystalline lattice. We will pay special attention to the role of thermal fluctuations on lattice structure, and how they can be measured through diffraction experiments.

Consider a regular crystalline lattice such as the one shown in Fig. 8.1. The atomic density ρ

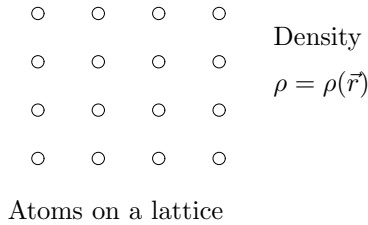


Figure 8.1:

is a function of \vec{r} as shown. A crystalline lattice, although it looks quite symmetric in a common sense way, it actually has very little symmetry compared to a liquid or a vapor. Both liquid or vapor are translationally invariant and isotropic. A crystalline solid on the other hand, has a very restricted symmetry. As you can see from Fig. 8.1, the lattice is invariant only under a discrete set of translations:

$$\vec{r} \rightarrow \vec{r} + r_0(n\hat{x} + m\hat{y}) \quad (8.1)$$

where n and m are integers (positive or negative), and r_0 is the lattice constant. Clearly the crystal

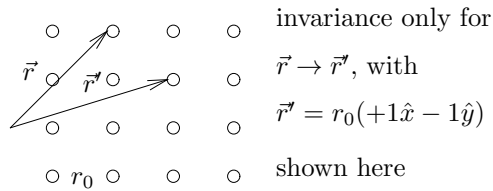


Figure 8.2:

has less symmetry than both a liquid and a vapor, so that the formation of a crystalline solid from the liquid phase necessarily must break the translational and rotational invariance of the liquid. As a consequence, there must be a broken symmetry variable that satisfies Goldstone's theorem.

In order to pursue this further, one introduces the displacement field $\vec{u}(\vec{r})$. Let \vec{r}_0 be the equilibrium position of a lattice point, and assume that under the influence of thermal fluctuation the particle that is located at \vec{r}_0 in equilibrium is displaced to \vec{r} . The displacement is then simply $\vec{u} = \vec{r} - \vec{r}_0$. This same quantity can be defined for each lattice point, and in the continuum limit, the displacement becomes a continuum field $\vec{u}(\vec{r})$ that describes the local lattice distortion with

reference to its equilibrium configuration. One such distorted configuration and the associated displacements is shown in Fig. 8.3. Therefore, the instantaneous density of any given configuration can be described as $\rho(\vec{r} + \vec{u}(\vec{r}))$ where $\vec{u}(\vec{r})$ is the local displacement vector, as shown in Fig. 8.3.

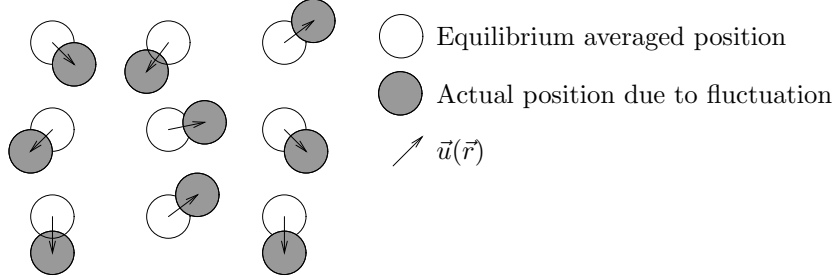


Figure 8.3: Distorted configuration due to thermal fluctuations

We are now in a position to understand what symmetry is broken spontaneously by the formation of a crystalline lattice. Consider a uniform displacement \vec{u} . Figure 8.4 shows two such cases. If all three configurations have the same lattice constant, they must have the same energy. In other words, the energy must be invariant under displacements of the entire lattice by a constant \vec{u} . However,

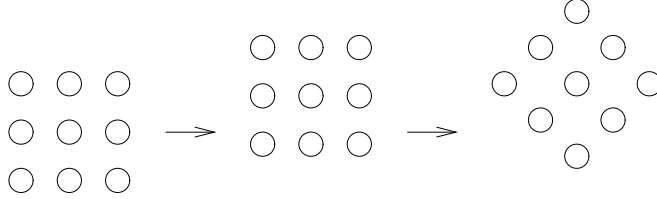


Figure 8.4:

in any given experiment a definite value for $\langle u(\vec{r}) \rangle = 0$ is obtained corresponding to a particular location of the lattice. Therefore, the actual lattice breaks the translational and rotational invariance of the energy of the system. In all respects, the variable \vec{u} plays the same role in crystalline solids as the variable $h(\vec{x})$ played in our earlier analysis of interface fluctuations.

By appealing now to Goldstone's Theorem, we may anticipate that in Fourier space, if all directions of fluctuation are equally likely, then

$$\langle \hat{u}(\vec{k}) \hat{u}(\vec{k}') \rangle \propto (2\pi)^d \delta(\vec{k} + \vec{k}') \left(\frac{1}{k^2} \right) \quad (8.2)$$

In order to pursue this analogy further, consider for a moment a lattice modeled as a one dimensional chain of particles connected by harmonic springs. The argument can be easily generalized to higher dimensions and to more realistic inter atomic potentials. Let u_n be the instantaneous displacement of the n particle in the chain. The energy of any configuration is given by,

$$E = \frac{C}{2} \sum_n (u_{n+1} - u_n)^2, \quad (8.3)$$

where we have assumed that C is the spring constant, and also neglected chain end effects by, for example, assuming periodic boundary conditions. Let a be the equilibrium separation between particles, or lattice constant. Then,

$$E = \frac{Ca^2}{2} \sum_n \frac{(u_{n+1} - u_n)^2}{a^2}.$$

We now take the continuum limit $a \rightarrow 0$ and note that in this limit $(u_{n+1} - u_n)/a \rightarrow du/dx$ where x is the coordinate along the chain, and further define $K = Ca^2$ finite. In this limit,

$$E = \frac{K}{2} \int dx \left(\frac{du}{dx} \right)^2.$$

This energy is formally the same as the energy of an interface fluctuation in the limit of small distortions. The important observation is that the energy does not depend on u directly, but rather on its derivative du/dx . Therefore, any two configurations that differ by a constant displacement, have the same energy. The energy is invariant under uniform displacement.

Given Eq. (8.2) we immediately find,

$$\langle u^2 \rangle = \int_{2\pi/L}^{\Lambda} \frac{d^d k}{(2\pi)^d} \frac{1}{k^2},$$

so that $\langle u^2 \rangle$ is finite in $d = 3$, whereas $\langle u^2 \rangle \sim \ln L$ in $d = 2$ and $\langle u^2 \rangle \sim L$ in $d = 1$. Due to the broken symmetry, fluctuations in displacement diverge with the system size below $d = 3$, and are finite in $d = 3$. We will examine below the consequences of this behavior.

8.1 Calculation of $\langle \rho \rangle$ and $\langle \rho(0)\rho(\vec{r}) \rangle$

We consider in this section the consequences of Eq. (8.2) on $\langle \rho(\vec{r}) \rangle$ and on the results of scattering experiments. First note that the density of a perfectly periodic crystalline lattice can be written as

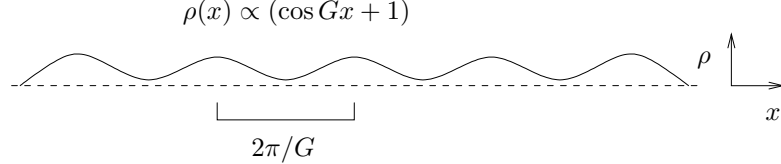


Figure 8.5: Rough depiction of a one dimensional periodic variation of density

$$\rho(x) = \sum_G e^{iGx} \hat{a}_G \quad (8.4)$$

following Fourier's decomposition theorem. The wavenumbers G are called reciprocal lattice vectors in $d = 3$, and \hat{a}_G are the corresponding Fourier coefficients. Each particular crystalline structure (e.g., cubic, tetragonal, etc.) is characterized by a definite set of vectors $\{\vec{G}\}$, and the details of the molecular structure of the solid are contained in $\hat{a}_{\vec{G}}$.

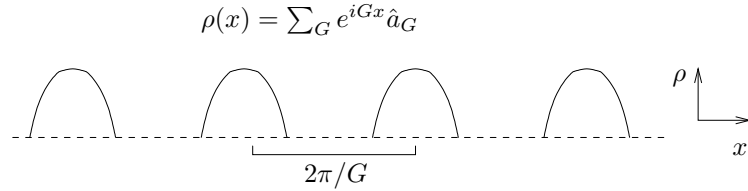


Figure 8.6: A more accurate depiction of a one dimensional variation in density.

The instantaneous density of any distorted configuration can be written in terms of the displacement field,

$$\rho(\vec{r} + \vec{u}(\vec{r})) = \sum_{\vec{G}} e^{i\vec{G} \cdot (\vec{r} + \vec{u}(\vec{r}))} \hat{a}_{\vec{G}} \quad (8.5)$$

where both sets $\{\vec{G}\}$ and $\{\hat{a}_{\vec{G}}\}$ correspond to the undistorted structure. The statistical average is simply,

$$\langle \rho(\vec{r}) \rangle = \sum_{\vec{G}} e^{i\vec{G} \cdot \vec{r}} \hat{a}_{\vec{G}} \langle e^{i\vec{G} \cdot \vec{u}(\vec{r})} \rangle \quad (8.6)$$

In order to calculate the average we assume that $\vec{u}(\vec{r})$ is a Gaussian variable. We mention a mathematical property of Gaussianly distributed variables. If x is Gaussianly distributed of zero mean, then

$$\langle e^{\xi x} \rangle = e^{\frac{\xi^2}{2} \langle x^2 \rangle} \quad (8.7)$$

for any constant ξ . This result can be generalized to a multivariate Gaussian distribution as follows

$$\langle e^{\sum \xi_n x_n} \rangle = e^{\frac{1}{2} \sum_{m,n} \xi_m \xi_n \langle x_m x_n \rangle}.$$

With this property and the assumption that \vec{u} is Gaussian, one has,

$$\langle e^{i\vec{G} \cdot \vec{u}} \rangle = e^{-\frac{1}{2} \vec{G} \vec{G} : \langle \vec{u} \vec{u} \rangle} \quad (8.8)$$

If fluctuations of \vec{u} along different directions are uncorrelated $\langle u_\alpha u_\beta \rangle \propto \delta_{\alpha\beta}$, then

$$\langle e^{-i\vec{G} \cdot \vec{u}} \rangle \approx e^{-\frac{1}{2} G^2 \langle u^2 \rangle}$$

The variance of the displacement field has been given earlier from our result Eq. (8.2)

$$\langle u^2 \rangle = \int_{2\pi/L}^{\Lambda} \frac{d^d k}{(2\pi)^d} \frac{1}{k^2} \propto \begin{cases} \text{constant} & d = 3 \\ \ln L & d = 2 \\ L & d = 1 \end{cases} \quad (8.9)$$

Substitution of this result into Eq. (8.6) leads to the desired result,

$$\langle \rho \rangle = \sum_{\vec{G}} \exp -\frac{1}{2} G^2 \langle u^2 \rangle e^{i\vec{G} \cdot \vec{r}} \hat{a}_{\vec{G}} \quad (8.10)$$

where $\exp -\frac{1}{2} G^2 \langle u^2 \rangle$ is the so called Debye-Waller factor. This factor is a constant for each G in $d = 3$ and hence originates relatively small changes in the average density. In $d = 2$ one has instead

$$\langle \rho \rangle = \sum_{\vec{G}} e^{i\vec{G} \cdot \vec{r}} \hat{a}_{\vec{G}} e^{-\frac{1}{2} G^2 \ln L} \quad (8.11)$$

except for unimportant constants. Note that in the strict limit $L \rightarrow \infty$ the exponential term vanishes, and hence $\langle \rho \rangle = \text{constant}$. The periodicity of the lattice is completely destroyed by fluctuations in the displacement field. Since the divergence with L is only logarithmic, this effect is in practice weak. It is more pronounced, however, in $d = 1$

$$\langle \rho \rangle = \sum_{\vec{G}} e^{i\vec{G} \cdot \vec{r}} \hat{a}_{\vec{G}} e^{-\frac{1}{2} G^2 L} \quad (8.12)$$

which is a large effect, much greater than the periodicity as $L \rightarrow \infty$. This is also related to the 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. 8.12 is $G = 0$ as $L \rightarrow \infty$ which gives $\rho(x) \rightarrow \text{constant}$, due to fluctuations in $\vec{u}(\vec{r})$.

8.1.1 Scattering intensity

We now turn to the calculation of the scattering intensity $I(k)$, where

$$I(k) \propto |\hat{\rho}_k|^2 \quad (8.13)$$

as indicated earlier. As a reference, we consider first a perfect lattice (in the absence of fluctuations or displacements). Then from Eq. (8.4)

$$\begin{aligned} \hat{\rho}_k &= \int d^d r \, e^{-i\vec{k}\cdot\vec{r}} \left\{ \sum_{\vec{G}} e^{i\vec{G}\cdot\vec{r}} \hat{a}_{\vec{G}} \right\} \\ &= \sum_{\vec{G}} \hat{a}_{\vec{G}} \int d^d r \, e^{-i(\vec{k}-\vec{G})\cdot\vec{r}} = \sum_{\vec{G}} \hat{a}_{\vec{G}} \delta(\vec{k}-\vec{G}). \end{aligned}$$

Therefore,

$$|\hat{\rho}_k|^2 = \sum_{\vec{G}, \vec{G}'} \hat{a}_{\vec{G}} \hat{a}_{\vec{G}'}^* \delta(\vec{k}-\vec{G}) \delta(\vec{k}-\vec{G}') \propto \sum_{\vec{G}} |\hat{a}_{\vec{G}}|^2 \delta(\vec{k}-\vec{G}), \quad (8.14)$$

since $\vec{k} = \vec{G}$ and $\vec{k} = \vec{G}'$, hence $\vec{G}' = \vec{G}$. The quantity $|\hat{a}_{\vec{G}}|^2$ is called the atomic form factor and describes the molecular arrangements within a unit cell of the crystal. The scattering intensity that corresponds to Eq. (8.14) displays well defined peaks at $\vec{k} = \vec{G}$, the scattering vectors that correspond to the reciprocal lattice vector. This is schematically shown in Fig. 8.7.

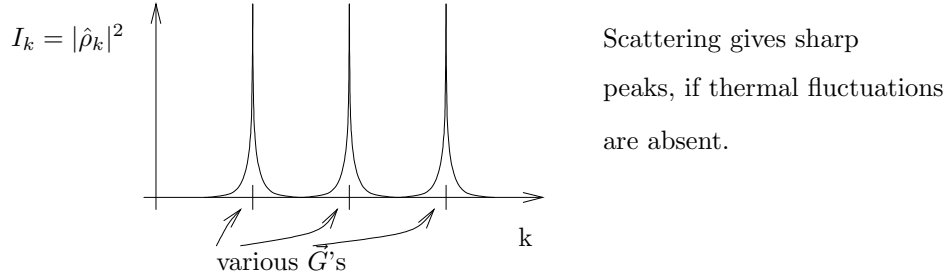


Figure 8.7:

We next consider the general case of a fluctuating crystal in which the displacement field is not zero. Then,

$$\hat{\rho}_k = \int d^d \vec{r} \, e^{-i\vec{k}\cdot\vec{r}} \left(\sum_{\vec{G}} e^{i\vec{G}\cdot\vec{r}} \hat{a}_{\vec{G}} e^{i\vec{G}\cdot\vec{u}(\vec{r})} \right)$$

or,

$$\hat{\rho}_k = \sum_{\vec{G}} \hat{a}_{\vec{G}} \int d^d \vec{r} \, e^{-i(\vec{k}-\vec{G})\cdot\vec{r}} e^{i\vec{G}\cdot\vec{u}(\vec{r})} \quad (8.15)$$

so that

$$\hat{\rho}_k^* = \sum_{\vec{G}'} \hat{a}_{\vec{G}'}^* \int d^d \vec{r}' \, e^{i(\vec{k}-\vec{G}')\cdot\vec{r}'} e^{-i\vec{G}'\cdot\vec{u}(\vec{r}')}$$

Hence the structure factor is

$$\langle |\hat{\rho}_k|^2 \rangle = \sum_{\vec{G}, \vec{G}'} \hat{a}_{\vec{G}} \hat{a}_{\vec{G}'}^* \int d^d \vec{r} \int d^d \vec{r}' \, e^{-i(\vec{k}-\vec{G})\cdot\vec{r}} e^{i(\vec{k}-\vec{G}')\cdot\vec{r}'} \langle e^{i\vec{G}\cdot\vec{u}(\vec{r}) - i\vec{G}'\cdot\vec{u}(\vec{r}')} \rangle \quad (8.16)$$

If the system overall is invariant under translation we have that,

$$\langle e^{i\vec{G}\cdot\vec{u}(\vec{r}) - i\vec{G}'\cdot\vec{u}(\vec{r}')} \rangle = f(\vec{r} - \vec{r}') \quad (8.17)$$

so that one of the integrals can be eliminated,

$$\langle |\hat{\rho}_k|^2 \rangle = \sum_{\vec{G}, \vec{G}'} \hat{a}_G \hat{a}_G^* \int d^d \vec{r} e^{-i(\vec{k}-\vec{G}) \cdot \vec{r}} (2\pi)^d \delta(\vec{G} - \vec{G}') f(\vec{r}) \quad (8.18)$$

where

$$f(\vec{r}) = \langle e^{i\vec{G} \cdot \vec{u}(r) - i\vec{G} \cdot u(0)} \rangle.$$

Evaluating the square, and by assuming as done earlier that the displacement field is a Gaussian variable, and that different components of \vec{u} are uncorrelated, we find,

$$\begin{aligned} f(\vec{r}) &= e^{-\frac{G^2}{2} \langle (u(\vec{r}) - u(0))^2 \rangle} \\ &= e^{-G^2 \langle u^2 \rangle} e^{G^2 \langle u(r)u(0) \rangle} \end{aligned} \quad (8.19)$$

and therefore,

$$\langle |\hat{\rho}_k|^2 \rangle = \sum_{\vec{G}} |\hat{a}_G|^2 e^{-G^2 \langle u^2 \rangle} \int d^d r e^{-i(\vec{k}-\vec{G}) \cdot \vec{r}} e^{G^2 \langle u(\vec{r})u(0) \rangle} \quad (8.20)$$

The correlation function appearing in the exponent $\langle u(\vec{r})u(0) \rangle$ can be calculated by inverse Fourier transformation of $\langle |\hat{u}(k)|^2 \rangle \sim \frac{1}{k^2}$:

$$\langle u(\vec{r})u(0) \rangle \propto \begin{cases} \frac{1}{r} & d = 3 \\ -\ln r, & d = 2 \\ -r, & d = 1 \end{cases} \quad (8.21)$$

In $d = 3$ it can be shown that the line widths are broadened by the Debye-Waller factor but do not change location. There is no effect from the decaying correlation function. the exponent. with

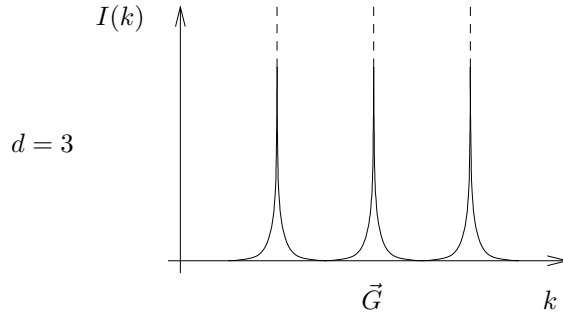


Figure 8.8: Delta Function Singularities

the same delta function peaks. In $d = 2$, the situation is more interesting:

$$\begin{aligned} \langle |\hat{\rho}_k|^2 \rangle &= \sum_{\vec{G}} e^{-G^2 \langle u^2 \rangle} |\hat{a}_G|^2 \int d^d r e^{-i(\vec{k}-\vec{G}) \cdot \vec{r}} e^{-c \ln r} \\ &= \sum_{\vec{G}} e^{-G^2 \langle u^2 \rangle} |\hat{a}_G|^2 \underbrace{\int d^d r \frac{e^{-i(\vec{k}-\vec{G}) \cdot \vec{r}}}{r^c}}_{\text{this is the Fourier transform of a power law}} \\ &= \sum_{\vec{G}} e^{-G^2 \langle u^2 \rangle} |\hat{a}_G|^2 \frac{1}{|\vec{k} - \vec{G}|^{d-c}}, \end{aligned}$$

where c is a constant. It is customary to set $(d - c) = 2 - \eta$, and write

$$\langle |\hat{\rho}_k|^2 \rangle = \sum_{\vec{G}} e^{-G^2 \langle u^2 \rangle} |\hat{a}_G|^2 \frac{1}{|\vec{k} - \vec{G}|^{2-\eta}}. \quad (8.22)$$

Thermal fluctuations in $d = 2$ change the delta function singularities at Bragg peak locations that are characteristic of long ranged crystalline order into power law singularities, still at wavevectors equal the reciprocal lattice vectors (Fig. 8.9). This broadening of the spectral lines reflects the

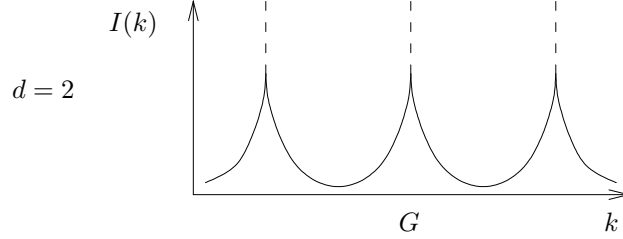


Figure 8.9: Power Law Singularities

enhanced effect of thermal fluctuations in $d = 2$.

The role of fluctuations is even more dramatic in $d = 1$. We first recall that

$$\int_{-\infty}^{\infty} dx e^{-b|x|} e^{-ikx} = \frac{2b}{b^2 + k^2},$$

so that

$$\langle |\hat{\rho}_k|^2 \rangle \propto \sum_{\vec{G}} |\hat{a}_G|^2 \underbrace{\int d^d r e^{-i(\vec{k}-\vec{G}) \cdot \vec{r}} e^{-cr}}_{\text{this is the Fourier transform of an exponential. But}} \\ \int d^d r e^{i\vec{k}' \cdot \vec{r}} e^{-r/\xi} \propto \frac{1}{(k')^2 + \xi^{-2}} \text{ for small } k'$$

or,

$$\langle |\hat{\rho}_k|^2 \rangle \propto \sum_G |\hat{a}_G|^2 \frac{1}{(k - G)^2 + \xi^{-2}} \quad (8.23)$$

where ξ^{-2} is a constant. The resulting scattering intensity is sketched in Fig. 8.10. Thermal fluctuations have removed any traces of any singularity at $k = G$, which indicates the lack of any long range order (any periodicity) in a one dimensional crystal.

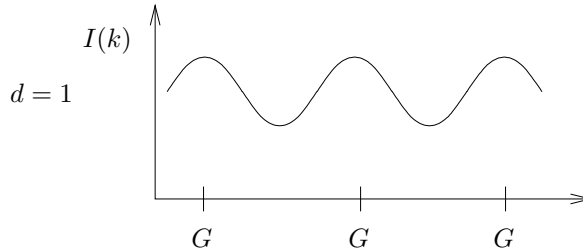


Figure 8.10: No Singularities

There are two other important consequences associated with spontaneously broken symmetries which we will not discuss here, but that you can look up for yourself. The first one is that continuous broken symmetries of the type that we have encountered are usually associated with so called soft modes. They are normal modes that correspond to long wave excitations or fluctuations that cost vanishingly low energy. For fluctuating surfaces, these modes are called capillary waves; for crystalline solids, they are known as phonons. The second important concept is that of “rigidity”. For a crystalline solid, this is the simple concept of rigidity: the solid appears rigid because any

distortion requires that many atoms move in concert. This is a direct consequence of the broken symmetry. The same concept applies to other broken symmetries.

Chapter 9

Ising Model

This is the first in a sequence of chapters in which we will explore the consequences of a spontaneously broken symmetry as it appears at a phase transition. We begin in this chapter by describing the Ising model of ferromagnetism which we will use to illustrate the general concepts.

There are many phenomena in nature in which a large number of constituents interact to give some global behavior. A magnet, for example, is composed of many interacting magnetic dipoles of molecular origin. A liquid is made up of many interacting molecules. A computer program has many interacting commands. The economy, the psychology of road rage, plate tectonics, and many other phenomena share the fact that all involve global behavior due to local interactions between a large number of entities.

The Ising model was introduced as a simple model of ferromagnetism in an attempt to describe the collective behavior that emerges at the Curie point of magnets: at this point the material turns from a paramagnet into a permanent magnet (or ferromagnet). It is defined on a regular lattice, each site of which contains a magnetic dipole or spin. In the simplest version, each spin can only appear in two microscopic states $+1$ (for up, in some direction) or -1 (down) (Fig. 9.1). The value of the spin variable at any lattice site is determined by the spin's (usually short range) interaction with the other neighboring spins. A sketch of the model in two dimensions on a square lattice is shown in Fig. 9.2.

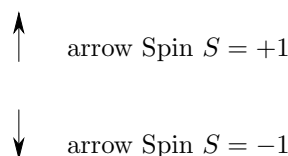


Figure 9.1:

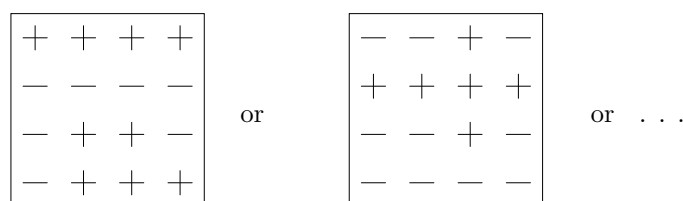


Figure 9.2: Possible configurations for a $N = 4^2 = 16$ system.

In order to calculate the partition function of this model, we would express the sum over states

as

$$\sum_{\nu} = \sum_{s_1=-1}^{+1} \sum_{s_2=-1}^{+1} \dots \sum_{s_N=-1}^{+1} = \prod_{i=1}^N \sum_{s_i=-1}^{+1}$$

where N is the total number of spins in a lattice of arbitrary dimensionality. The energy of any configuration $e^{E_{\nu}/k_B T}$ must now be prescribed. Since we want to model a ferromagnet, where all the spins are aligned either $+1$ or -1 , we make the spin of site i to want to oriented parallel to its neighbors. That is, for the spin on site i we define the following interaction energy,

$$E_i = \sum_{\text{neighbors of } i} (S_i - S_{\text{neighbors}})^2 = S_i^2 - 2 \sum_{\text{neighbors of } i} S_i S_{\text{neigh.}} + S_{\text{neigh.}}^2 \quad (9.1)$$

But $S_i^2 = 1$ always since $S_i = \pm 1$, so up to a constant (additive and multiplicative)

$$E_i = - \sum_{\text{neighbors of } i} S_i S_{\text{neighbors of } i}$$

On a square lattice, the neighbors of i are the spin surrounding it (see Fig. 9.3 for a two dimensional case).

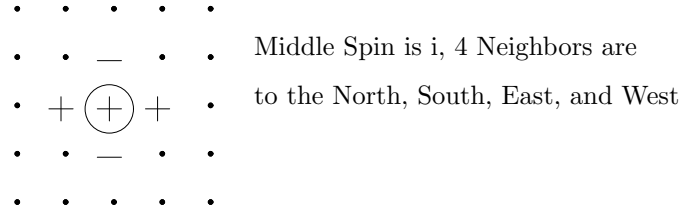


Figure 9.3: Four nearest neighbors on a square lattice.

This choice of interaction energy reflects the assumption that interactions at the microscopic level are local in nature. The choice of the 4 spins *above*, *below* and to the *right* and *left* is special (restricted to “nearest neighbors”), and other choices exist in the literature. For example, one can include interactions with spins on the diagonals (NW, NE, SE, SW) that are “next nearest neighbors”. It doesn’t actually matter for our purpose here how many are included, so long as the interaction is local. The total energy of any microscopic configuration is the sum over all spins of the interactions just described. We write,

$$E_{\nu} = -J \sum_{\langle ij \rangle} S_i S_j$$

where J is a positive interaction constant that sets the energy scale, and $\langle ij \rangle$ denotes a sum over all i and $j = 1, 2, \dots, N$, provided i and j are nearest neighbors, with no double counting. Note that the sum over spins has

$$\sum_{\langle ij \rangle} = \underbrace{\sum_{i=1}^N \sum_{j=1}^N}_{\text{neighbors, no double counting}} = N \times \frac{4}{2}$$

terms. If there were “ q ” nearest neighbors instead, corresponding to a different interaction pattern, to a different lattice (say hexagonal), or to a different dimension of space

$$\sum_{\langle ij \rangle} = \frac{q}{2} N$$

The value of q nearest neighbors is $q = 2$ for the one dimensional Ising model, $q = 4$ for a two dimensional square lattice, and $q = 6$ for a simple cubic lattice in three dimensions. The lowest energy is attained when all spins are aligned along the same orientation, in which case $E = -\frac{qJ}{2} N$.

One can also include in this model the effect of an externally imposed magnetic field H , which tends to align spins parallel to the field. The energy of a configuration inside a magnetic field is

$$E_\nu = -J \sum_{\langle ij \rangle} S_i S_j - H \sum_i S_i. \quad (9.2)$$

The corresponding partition function is

$$Z(N, \frac{J}{k_B T}, \frac{H}{k_B T}) = \prod_{i=1}^N \sum_{S_i=-1}^{+1} e^{(\frac{J}{k_B T}) \sum_{\langle jk \rangle} S_j S_k + (\frac{H}{k_B T}) \sum_i S_i} \quad (9.3)$$

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 yet been solved.

The most interesting consequence of the Ising model is that it exhibits a broken symmetry (for $H = 0$). In the absence of a field, the energy of a state and hence the partition function, have an exact symmetry $S_i \leftrightarrow -S_i$ (the group associated with this symmetry is Z_2 .) Hence $\langle S_i \rangle \leftrightarrow -\langle S_i \rangle$ and therefore

$$\langle S_i \rangle = 0$$

In fact, for dimension $d = 2$ and 3 this is not true. An experimental realization spontaneously breaks this symmetry exhibiting either positive or negative average magnetization $\langle S_i \rangle$. The exact result for the magnetization in $d = 2$ is surprisingly simple:

$$m = \langle S_i \rangle = \begin{cases} \pm \left(1 - \operatorname{cosech}^2 \frac{2J}{k_B T}\right)^{1/8}, & T < T_c \\ 0, & T > T_c \end{cases}$$

The critical temperature $T_c \approx 2.269J/k_B$ is where $(1 - \operatorname{cosech}^2 \frac{2J}{k_B T})^{1/8}$ vanishes.

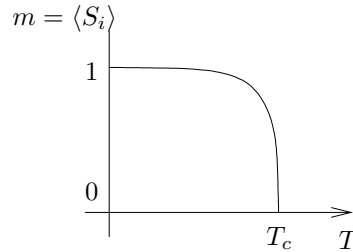


Figure 9.4: Two-dimensional square lattice Ising model magnetization per spin

A qualitative interpretation of this transition from zero to nonzero average magnetization at T_c is easy to give. Consider the 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 dominates the minimization of F . Clearly entropy is maximized by having the spins helter-skelter at random: $\langle S_i \rangle = 0$, since plus and minus one would appear with the same probability.

The transition temperature between $\langle S_i \rangle = 1$ and $\langle S_i \rangle = 0$ should be when the thermal energy is comparable to the energy per spin $k_B T = \mathcal{O}(\frac{q}{2}J)$, which is in the right ballpark. However, there is no simple explanation for the abrupt increase of the magnetization per spin $m = \langle S_i \rangle$ at T_c . This abruptness is in contrast with the case of a paramagnet that shows a gradual decrease in magnetization with increasing temperature (and in the presence of an external field H). In the ferromagnet, the fundamental up-down symmetry is broken at exactly T_c .

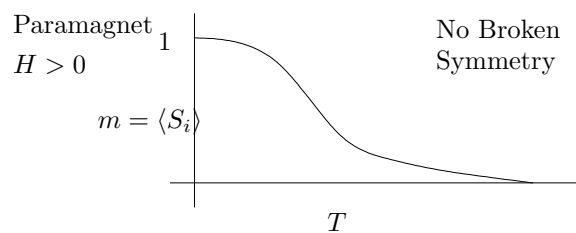


Figure 9.5:

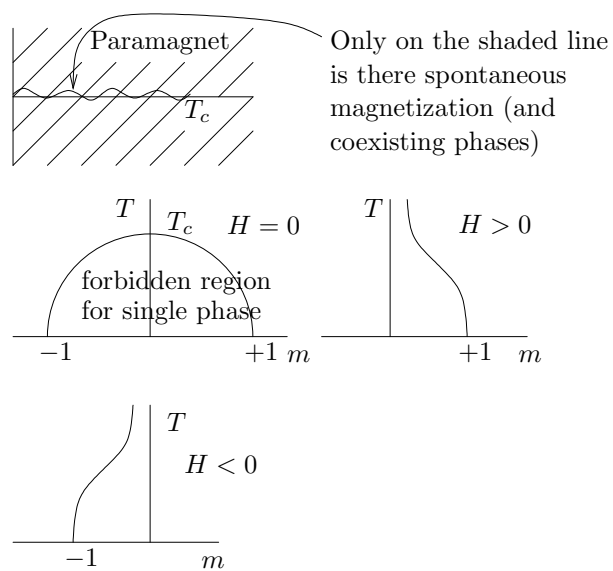


Figure 9.6:

The phase diagram for the Ising model (or indeed for any ferromagnet) is sketched in Fig. 9.6. The true breaking of the up-down symmetry only appears for $H = 0$. Otherwise, there is always a finite, residual magnetization at any temperature.

9.1 (*) Lattice Gas Model

In this model, one still has a lattice with sites labeled $i = 1, 2, \dots, N$, but now there is a concentration variable $c_i = 0$ or 1 defined at each site (i.e., the site is either empty or occupied). The most simple model of interaction is the lattice gas model defined as

$$E_\nu^{LG} = -\vartheta \sum_{\langle ij \rangle} c_i c_j - \mu \sum_i c_i$$

where ϑ is an interaction energy, 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 chemical species), or as a liquid vapor mixture (where c_i is the local density; e.g. zero for the vapor). This model can be mapped onto the Ising model by introducing the transformation,

$$S_i = 2c_i - 1.$$

Then

$$\begin{aligned} E_\nu^{Ising} &= -J \sum_{\langle ij \rangle} (2c_i - 1)(2c_j - 1) - H \sum_i (2c_i - 1) \\ &= -4J \sum_{\langle ij \rangle} c_i c_j + 4J \sum_{\langle ij \rangle} c_i + \text{constant} - 2H \sum_i c_i + \text{constant} \\ &= -4J \sum_{\langle ij \rangle} c_i c_j + 4J \frac{q}{2} \sum_i c_i - 2H \sum_i c_i + \text{constant} \end{aligned}$$

and

$$E_\nu^{Ising} = -4J \sum_{\langle ij \rangle} c_i c_j - 2(H - Jq) \sum_i c_i + \text{constant}.$$

Since the additive constant is unimportant, this is the same interaction energy as E_ν^{LG} , provided that

$$\vartheta = 4J \quad \mu = 2(H - Jq).$$

The phase transition occurs at a specific value of the chemical potential $\mu_c = -2Jq$. It is sometimes said that the liquid-vapor transition does not involve a change of symmetry (associated with the group Z_2) as opposed to the Ising model. This is not correct. However, the external chemical potential must have a carefully tuned value to get that possibility.

9.2 One dimensional Ising Model. Transfer matrix solution

It is easy to exactly solve the one dimensional Ising model. Of course, we already know that it cannot exhibit long range order, and hence have a true phase transition. Hence we expect

$$\langle S_i \rangle = m = 0$$

for all $T > 0$.

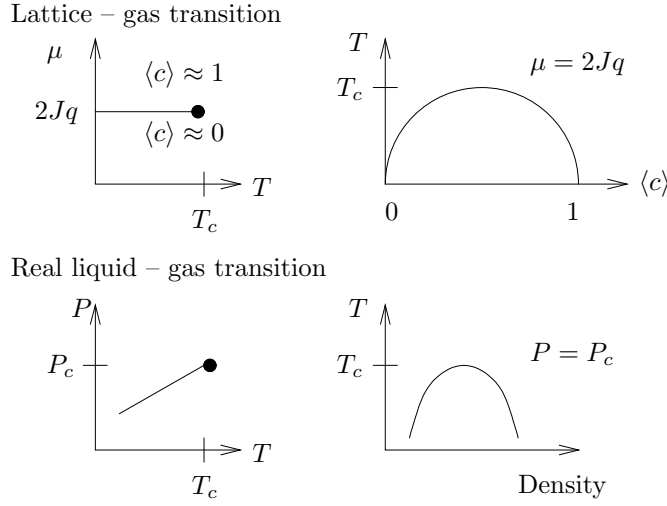


Figure 9.7:

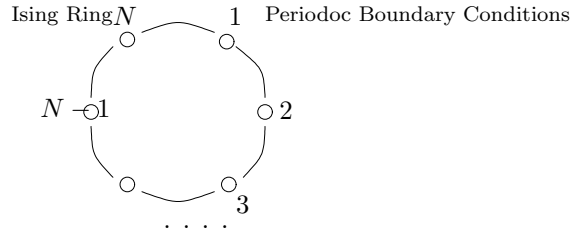


Figure 9.8:

9.2.1 No magnetic field

Consider a one dimensional chain of spin variables $S_i = \pm 1$, $i = 1, \dots, N$. We also consider the so called periodic boundary conditions, such that $S_{N+1} = S_1$. This system is shown in Fig. 9.8. The partition function is,

$$Z = \sum_{\{S_i\}} e^{\beta J \sum_{i=1}^N S_i S_{i+1}} = \sum_{S_1=\pm 1} \sum_{S_2=\pm 1} \dots \sum_{S_N=\pm 1} e^{\beta J \sum_{i=1}^N S_i S_{i+1}}.$$

The exponential of the sum can be written as a product of exponentials:

$$Z = \sum_{S_1=\pm 1} \sum_{S_2=\pm 1} \dots \sum_{S_N=\pm 1} [e^{\beta J S_1 S_2}] [e^{\beta J S_2 S_3}] \dots [e^{\beta J S_N S_1}] \quad (9.4)$$

The transfer matrix method involves representing each term in the product as a matrix:

$$T_{\sigma_i \sigma_{i+1}} = e^{\beta J S_i S_{i+1}} \quad (9.5)$$

In this representation, the elements of the matrix are labeled by the variable S (instead of just the conventional positive integers):

$$T_{S_i S_{i+1}} = \begin{pmatrix} T_{11} & T_{1-1} \\ T_{-11} & T_{-1-1} \end{pmatrix} = \begin{pmatrix} e^{\beta J} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta J} \end{pmatrix}$$

With the definition given in Eq. (9.5), the partition function can now be written as,

$$Z = \sum_{S_1=\pm 1} \sum_{S_2=\pm 1} \dots \sum_{S_N=\pm 1} T_{S_1 S_2} T_{S_2 S_3} \dots T_{S_N S_1}$$

Next consider the usual rules of matrix multiplication (A, B and C are matrices):

$$A = B \cdot C \quad \text{is explicitly} \quad A_{ik} = \sum_j B_{ij} C_{jk} \quad (9.6)$$

Therefore, in our notation,

$$\sum_{S_3=\pm 1} T_{S_2 S_3} T_{S_3 S_4},$$

can be understood as the product of the two transfer matrices in question, exactly as in Eq. (9.6) with the spin variables being the matrix indices. With this in mind, Eq. (9.4) can be written as,

$$Z = \sum_{S_1=\pm 1} T_{S_1 S_1}^N$$

since all the matrices $T_{S_i S_{i+1}}$ are identical, and after having performed the N matrix products implied by the sums over the N spin variables, and taking into account that $S_{N+1} = S_1$. All that is left, is the final sum over S_1 .

Given the definition of trace of a matrix,

$$\text{Tr}(A) = \sum_i A_{ii},$$

then,

$$Z = \text{Tr}(T^N).$$

The trace of T^N can be calculated by diagonalization. The matrix T is symmetric and hence can be written in diagonal form with real eigenvalues. Let D be the diagonal matrix, then

$$D = B^{-1} T B,$$

where B is the matrix implementing the basis set change to diagonalize T . Given the cyclic property of the trace,

$$\text{Tr}(ABC) = \text{Tr}(CAB),$$

one has that

$$\text{Tr}(D^N) = \text{Tr}(B^1 T B \ B^1 T B \ \dots) = \text{Tr}(B^{-1} T^N B) = \text{Tr}(B B^{-1} T^N) = \text{Tr}(T^N).$$

All we have to do now is to diagonalize T , so that

$$Z = \text{Tr}(D^N) = \lambda_1^N + \lambda_2^N,$$

where λ_1 and λ_2 are the two eigenvalues of T .

In order to find the eigenvalues, we compute

$$\begin{vmatrix} e^{\beta J} - \lambda & e^{-\beta J} \\ e^{-\beta J} & e^{\beta J} - \lambda \end{vmatrix} = 0.$$

One finds,

$$\lambda_{1,2} = e^{\beta J} \pm e^{-\beta J},$$

so that the free energy,

$$F = -k_B T \ln Z = -k_B T \ln [\lambda_1^N + \lambda_2^N] = -k_B T \left[\ln \lambda_1^N + \ln \left(1 + \left(\frac{\lambda_1}{\lambda_2} \right)^N \right) \right]$$

Assume that $\lambda_1 > \lambda_2$ and that N is very large (thermodynamic limit), then,

$$F = -k_B T \ln \lambda_1^N = -k_B T N \ln (\cosh \beta J). \quad (9.7)$$

This is the exact partition function for the Ising model in one dimension.

9.2.2 One dimensional Ising model in a magnetic field

The partition function in this case is,

$$Z = \sum_{\{S_i\}} e^{\beta(J \sum_{i=1}^N S_i S_{i+1} + H \sum_{i=1}^M S_i)} = \sum_{\{S_i\}} \prod_i e^{\beta J S_i S_{i+1} + \frac{\beta H}{2} (S_i + S_{i+1})}$$

The product can be expressed again in terms of a transfer matrix, now given by

$$T_{S_i S_{i+1}} = \begin{pmatrix} e^{\beta(J+H)} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta(J-H)} \end{pmatrix}$$

Diagonalization of this matrix leads to the following two eigenvalues,

$$\lambda_{1,2} = e^{\beta J} \cosh(\beta H) \pm \sqrt{e^{2\beta J} \sinh^2(\beta H) + e^{-2\beta J}},$$

with the free energy in this case given by

$$F = -k_B T N \ln \lambda_1,$$

which reduces to Eq. (9.7) in the limit of $H = 0$.

The average magnetization is given by,

$$\langle S \rangle = \frac{k_B T}{N} \frac{\partial \ln Z}{\partial H} = \frac{k_B T}{\lambda_1} \frac{\partial \lambda_1}{\partial H} = \frac{\sinh \beta h}{\sqrt{\sinh^2 \beta h + e^{-4\beta J}}}.$$

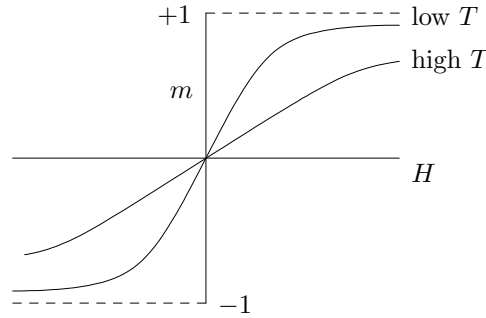


Figure 9.9: Magnetization of the $d = 1$ Ising model

The exact solution shows that as the field vanishes $H \rightarrow 0$, the order parameter vanishes $\langle S \rangle = 0$, for all temperatures. This model in one dimension does *not* exhibit a phase transition. Note, however, that something peculiar happens at $T = 0$. The limits $(T \rightarrow 0, h \rightarrow 0)$ and $(H \rightarrow 0, T \rightarrow 0)$ are different (i.e., the limit is singular). In particular, for $\beta \rightarrow \infty$, $\langle S \rangle = 1$ for any field H .

In fact the system of spins is perfectly ordered at $T = 0$ and $H = 0$. The exact free energy is at $T = 0$

$$F = -k_B T N \ln(e^{\beta J}) = -NJ,$$

which corresponds to all spins being parallel to each other (a perfectly ordered state). Therefore it is commonly said that the Ising model in one dimension has a phase transition only at $T = 0$, but that the system is disordered at any finite temperature.

9.2.3 Pair correlation function. Exact solution

It is also possible to compute exactly the pair correlation function of the Ising model in one dimension. It is useful to examine its behavior with spin separation, especially in the limit $T \rightarrow 0$. We define the spin pair correlation function as,

$$g(j) = \langle S_i S_{i+j} \rangle - \langle S_i \rangle \langle S_{i+j} \rangle,$$

the analog of the pair distribution function $h(r)$ defined for fluids. The calculation is greatly simplified by assuming that the interaction constant J changes from site to site $J = J_i$, and taking all $J_i = J$ at the end of the calculation. In this case, we have (for $H = 0$)

$$\langle S_i S_{i+j} \rangle = \frac{1}{Z} \sum_{\{S_i\}} S_i S_{i+j} e^{\beta \sum_l J_l S_l S_{l+1}} = \frac{1}{Z} \sum_{\{S_i\}} (S_i S_{i+1}) (S_{i+1} S_{i+2}) \dots (S_{i+j-1} S_{i+j}) e^{\beta \sum_l J_l S_l S_{l+1}}$$

by taking advantage of the fact that since $S_i = \pm 1$, then $S_i^2 = 1$.

Now note that each pair in the product can be written as,

$$(S_i S_{i+1}) e^{\beta \sum_l J_l S_l S_{l+1}} = \frac{\partial}{\partial (\beta J_i)} e^{\beta \sum_l J_l S_l S_{l+1}},$$

and therefore

$$\langle S_i S_{i+j} \rangle = \frac{1}{Z} \frac{1}{\beta^j} \left(\frac{\partial^j Z}{\partial J_i \partial J_{i+1} \dots \partial J_{i+j-1}} \right)_{J_l=J}.$$

Since the exact partition function for this case in which J is a function of the lattice site l is

$$Z = \prod_{l=1}^N (2 \cosh \beta J_l),$$

one has,

$$\frac{\partial Z}{\partial J_i} = \prod_{l \neq i} (2 \cosh \beta J_l) 2\beta \sinh \beta J_i,$$

and

$$\frac{\partial^2 Z}{\partial J_i \partial J_{i+1}} = \prod_{l \neq i, l \neq i+1} (2 \cosh \beta J_l) (2\beta \sinh \beta J_i) (2\beta \sinh \beta J_{i+1})$$

and repeating the process for the j derivatives

$$\langle S_i S_{i+j} \rangle = \frac{1}{\prod_l 2 \cosh \beta J_l} \frac{1}{\beta^j} \prod_{l \neq i, \dots, l \neq i+j-1} (2 \cosh \beta J_l) (2\beta)^j (\sinh \beta J)^j,$$

where in the last term of the right hand side we have already taken into account the fact that at the end of the calculation all the $J_l = J$. Taking this into account in the remaining terms, one has,

$$\langle S_i S_{i+j} \rangle = \frac{1}{(2 \cosh \beta J)^N} \frac{1}{\beta^N} (2 \cosh \beta J)^{N-j} (2\beta)^j (\sinh \beta J)^j = (\tanh \beta J)^j.$$

Finally, one defines a correlation length,

$$\xi = -\frac{1}{\ln \tanh \beta J},$$

so that the final result for the correlation function is,

$$\langle S_i S_{i+j} \rangle = e^{-j/\xi}$$

namely, correlations decay exponentially as a function of the separation between the spins, with a characteristic decay length ξ . There are no power law correlations as there is no phase transition at any finite temperature.

Consider now the magnetic susceptibility given by

$$\chi = \beta \sum_j \langle S_i S_{i+j} \rangle$$

and given the exponential decay of the correlation, the susceptibility is finite. Of course, this is in agreement with the observation that there is no phase transition at finite temperature in the one dimensional Ising model. However, note that ξ behaves anomalously as one approaches $T = 0$, also consistent with long ranged order appearing exactly at $T = 0$. In particular, as $\beta \rightarrow \infty$, $\tanh \beta J \rightarrow 1$ so that the correlation function diverges logarithmically as $T = 0$ is approached, reflecting the emergence of long ranged correlations and the eventual ordered state at $T = 0$.

9.3 Mean Field Theory

The one dimensional Ising model can be solved exactly, but its usefulness is limited as it does not show the existence of a finite temperature phase transition. Short of an exact solution in three dimensions, there is an approximate scheme, mean field theory, that yields quite useful results, although its predictions disagree with experiments in several important aspects.

The energy of each spin in the lattice depends on the local field created by its neighbors. The energy of the neighbors depends on the spin in question and that of their own neighbors, thus leading to a complex coupled system. However, when the ferromagnet is globally magnetized or demagnetized, the local field on any spin does not fluctuate very much from spin to spin. The mean field approximation is based on replacing the local field on any given spin by an average spin as determined over the entire system.

We first calculate the partition function of the Ising model by using the mean field approximation. The partition function is given by,

$$Z = \sum_{\{S_i\}} e^{\beta(J \sum S_i S_j + H \sum S_i)}$$

The difficulty in the evaluation function resides in the coupling term $S_i S_j$. The approximation involves replacing S_j by the (yet unknown) average $\langle S \rangle$,

$$Z \approx \sum_{\{S_i\}} e^{\beta(J \sum S_i \langle S \rangle + H \sum S_i)} = \sum_{\{S_i\}} e^{\beta(Jq \langle S \rangle + H) \sum S_i}$$

The sum in the exponent can be turned into a product of uncoupled exponentials. The product can be further interchanged with the sum over states

$$Z = \sum_{\{S_i\}} \prod_i e^{\beta(Jq \langle S \rangle + H) \sum S_i} = \prod_i \sum_{S_i = \pm 1} e^{\beta(Jq \langle S \rangle + H) \sum S_i}$$

Finally,

$$Z = \left[e^{\beta(Jq \langle S \rangle + H)} + e^{-\beta(Jq \langle S \rangle + H)} \right]^N. \quad (9.8)$$

This completes our calculation of the partition function. The solution has not yet been found as Z still depends on the unknown average $\langle S \rangle$. We obtain the latter quantity self consistently from the partition function just obtained,

$$\langle S \rangle = \frac{1}{N} \sum_i S_i = \frac{k_B T}{N} \frac{\partial \ln Z}{\partial H}. \quad (9.9)$$

By combining Eqs. (9.9) and (9.8) one finds,

$$\langle S \rangle = \tanh [\beta (Jq \langle S \rangle + H)], \quad (9.10)$$

which is an implicit equation for the average magnetization $\langle S \rangle$.

Equation (9.10) does lead to a phase transition point at which the average magnetization $\langle S \rangle$ abruptly deviates from zero. Let us expand $\langle S \rangle$ around zero in the vicinity of this point by recalling that $\tanh x \simeq x - x^3/3 + \dots$ for small x ,

$$\langle S \rangle \simeq \beta Jq \langle S \rangle - \frac{1}{3} (\beta Jq)^3 \langle S \rangle^3 + \dots$$

This equation has a solution of zero magnetization $\langle S \rangle = 0$, but also a pair of solutions of nonzero magnetization as long as $\beta Jq > 1$. Therefore, the mean field approximation leads to a transition temperature at

$$T_c = \frac{Jq}{k_B}$$

These other two solutions are,

$$\langle S \rangle^2 = \frac{3(\beta Jq - 1)}{(\beta Jq)^3}$$

It is instructive to examine their temperature dependence near T_c . By expanding $\beta = \beta_c + \delta\beta$ and by noting that $\delta\beta = -\delta T/(k_B T^2)$, we find,

$$\langle S \rangle \sim \pm \sqrt{-\delta T} = (T_c - T)^{1/2}.$$

In summary, the mean field approximation not only predicts a critical temperature separating a high temperature phase of zero magnetization and a low temperature phase with two -degenerate- phases of nonzero magnetization, but it also predicts the dependence of the magnetization on temperature near T_c . Specifically, the magnetization vanishes as T_c is approached from below as a power law with an exponent of $1/2$. In what follows, we introduce a magnetization exponent β by

$$m \propto (1 - T/T_c)^\beta$$

where the critical exponent in the mean field approximation is $\beta = 1/2$. The parabolic dependence

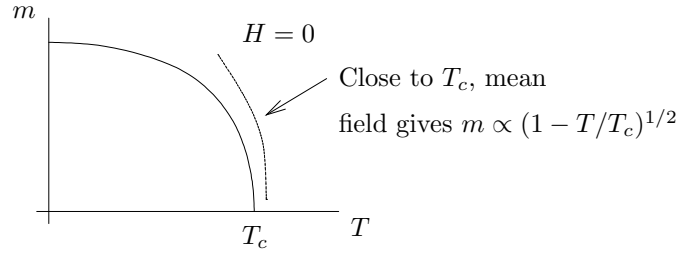


Figure 9.10:

of m versus temperature near T_c is shown in Fig. 9.11.

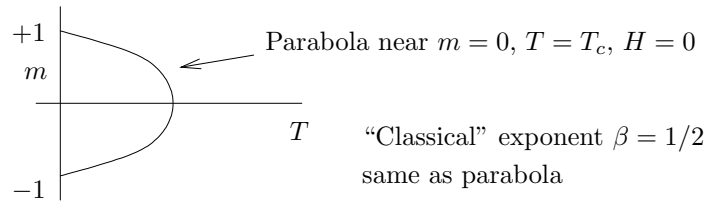


Figure 9.11:

The mean field results just derived depend on the dimensionality of the system, but only through the number of neighbors q . We find,

$$k_B T_c = qJ = \begin{cases} 2J & d = 1 \\ 4J & d = 2 \\ 6J & d = 3 \end{cases}$$

The result for $d = 1$ is clearly incorrect since $T_c = 0$. In $d = 2$, Onsager's exact solution yields $k_B T_c \approx 2.27J$, so the mean field estimate is rather poor. Note also that Onsager's solution yields $\beta = 1/8$ in $d = 2$, quite different from the mean field result $\beta = 1/2$. In three dimensions numerical results yield $k_B T_c \sim 4.5J$ and $\beta = 0.313\dots$, closer to the mean field estimates, but still quantitatively incorrect. Strangely enough, numerical studies conducted for lattices in high dimensions ($d \geq 4$) do find $\beta = 1/2$ (although there are logarithmic corrections in $d = 4$).

There are a number of other predictions made by the mean field approximation. For example, for $T = T_c$, 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 around $m = 0$ we find

$$H = k_B T_c \left(m + \frac{1}{3}m^3 + \dots - m\right)$$

or

$$H = \frac{k_B T_c}{3} m^3.$$

This dependence defines another critical exponent

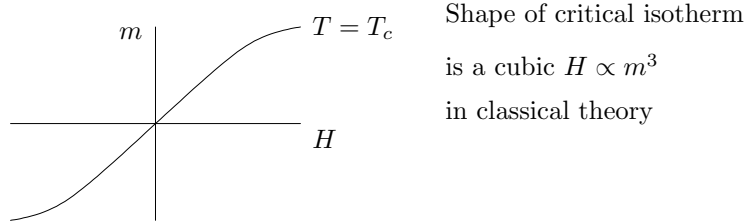


Figure 9.12:

$$H \propto m^\delta \quad \text{at} \quad T = T_c$$

where $\delta = 3$ in the mean field approximation.

A second example concerns the dependence of the magnetic susceptibility

$$\chi_T = \left(\frac{\partial m}{\partial H}\right)_T.$$

The susceptibility is the analog of the compressibility in liquid gas systems

$$\kappa_T = -\frac{1}{V} \left(\frac{\partial V}{\partial P}\right)_T$$

The magnetic susceptibility relates the change in magnetization that is brought about by a change in magnetic field. A large χ_T means the material is very “susceptible” to being magnetized. Given the mean field solution,

$$m = \tanh \frac{k_B T_c m + H}{k_B T}$$

we have

$$H = k_B T \cdot \tanh^{-1} m - k_B T_c m$$

so that

$$\left(\frac{\partial H}{\partial m}\right)_T = k_B T \frac{1}{1 - m^2} - k_B T_c.$$

Therefore

$$\chi_T = \frac{1}{(T - T_c)^\gamma}$$

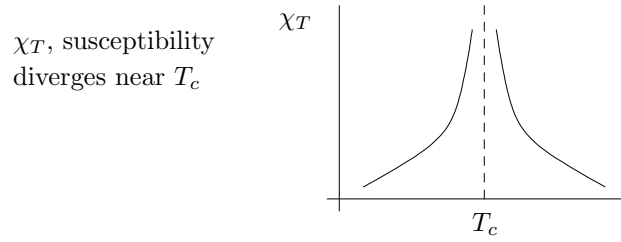


Figure 9.13:

close to T_c with $\gamma = 1$ in the mean field approximation. The divergence of the susceptibility near T_c is shown in Fig. 9.13.

It is interesting to dwell some more on the origin of the divergence of the susceptibility near the transition point. From its definition,

$$\chi - \frac{\partial \langle S \rangle}{\partial H} = \frac{\partial}{\partial H} \frac{1}{Z} \sum_{\{S_i\}} S_0 e^{-\beta(-J \sum S_i S_j - H \sum S_i)}.$$

All spins are equivalent, so we have picked the average of one of them (S_0) as the average spin in the system. By taking the derivative with respect to H , we find,

$$\chi = -\frac{1}{Z^2} \frac{\partial Z}{\partial H} \langle S \rangle + \frac{1}{Z} \sum_{\{S_i\}} S_0 e^{\beta(J \sum S_i S_j + H \sum S_i)} \left(\beta \sum_i S_i \right).$$

Recalling that $\langle S \rangle = (k_B T / N) \partial \ln Z / \partial H$, we find,

$$\chi = -\frac{N}{k_B T} \langle S \rangle^2 + \frac{\beta}{Z} \sum_{\{S_i\}} S_0 \left(\sum S_i \right) e^{\beta(-J \sum S_i S_j + H \sum S_i)} = -\frac{N}{k_B T} \langle S \rangle^2 + \beta \sum \langle S_0 S_i \rangle.$$

Rearranging we find,

$$\chi = \beta \sum_i (\langle S_0 S_i \rangle - \langle S_0 \rangle \langle S_i \rangle). \quad (9.11)$$

Equation (9.11) is the analog of the sum rule found for liquids

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

These are but two examples of equations that relate thermodynamic response functions to integrals of the corresponding correlation functions.

We have just argued that the susceptibility diverges near T_c , yet given that $S_i = \pm 1$, all terms in Eq. (9.11) are smaller than one. The origin of the divergence must be in the number of terms that contribute to the sum, or equivalently, in very slowly decaying correlations in S . Long range correlations among the spins near T_c must be responsible for the divergence of χ . In three dimensions, for example, the correlation function must decay as

$$\langle S_0 S_i \rangle \sim \frac{1}{|\vec{r}_0 - \vec{r}_i|^3}$$

or slower for the susceptibility to diverge. Certainly, exponentially decaying correlations are inconsistent with a divergent susceptibility.

As a third example we consider the energy and its derivative, the heat capacity,

$$\langle E \rangle = \left\langle -J \sum_{\langle ij \rangle} S_i S_j - H \sum_i S_i \right\rangle$$

In the mean field approximation all the spins are independent, 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 per spin is defined to be

$$c = \frac{1}{N} \left(\frac{\partial E}{\partial T} \right)_N.$$

Since near T_c and $H = 0$,

$$m^2 \propto \begin{cases} (T_c - T) & T < T_c \\ 0 & T > T_c \end{cases}$$

we have

$$\langle E \rangle = \begin{cases} -\text{constant}(T_c - T), & T < T_c \\ 0, & T > T_c \end{cases}$$

and hence

$$c = \begin{cases} \text{constant} & T < T_c \\ 0 & T > T_c \end{cases}$$

If we had included the full temperature dependence away from T_c , the specific heat looks more like a saw tooth, as drawn in Fig. 9.14. More generally, the specific heat diverges near the critical

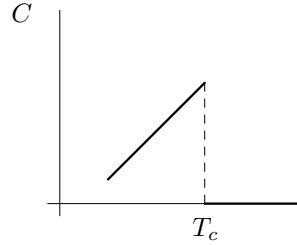


Figure 9.14: Discontinuity in C in mean field theory.

temperature, and the divergence is used to defined another critical exponent through

$$c \propto |T - T_c|^{-\alpha}$$

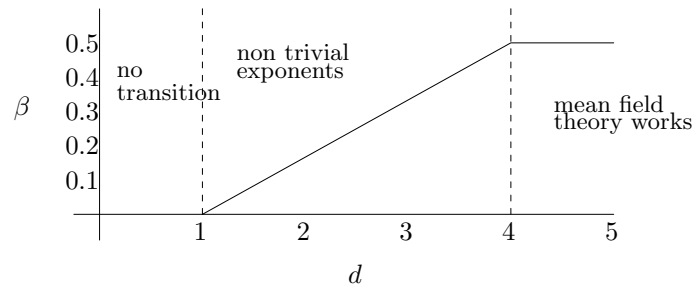
This discontinuity in c , a second derivative of the free energy, is one reason continuous phase transitions were called “second order”. Now it is known that this is an artifact of mean field theory.

To summarize, near the critical point we have,

$$\begin{aligned} m &\sim (T_c - T)^\beta, & T < T_c, H = 0 \\ H &\sim m^\delta, & T = T_c \\ X_T &\sim |T - T_c|^{-\gamma}, & H = 0 \\ c &\sim |T - T_c|^{-\alpha}, & H = 0 \end{aligned}$$

We summarize in the following table the predictions from mean field theory, the results of Onsager’s exact solution in $d = 2$, and approximate numerical results which have been obtained in $d = 3$.

	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

Figure 9.15: Variation of critical exponent β with spatial dimension

We will address later why mean field theory works for $d \geq 4$. Figure 9.15 sketches for example the dependence of the critical exponent β as a function of a continuous dimension variable d . For now we will leave the Ising model behind.

Chapter 10

Continuum theories of the Ising model

Nowadays, it is well established that, near the critical point, one can either solve the Ising model in its original discrete form

$$E = -J \sum_{\langle ij \rangle} S_i S_j - H \sum_i S_i$$

with

$$Z = \sum_{\{S_i\}} e^{-E/k_B T}$$

or the so called 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

$$Z = \int \mathcal{D}\psi e^{-\beta F}.$$

The continuum approximation here is in the same spirit as that introduced when we used a Gaussian, continuum model for interface fluctuations instead of a discrete version based on interface steps. Although it may not appear that way at first sight, the field theory description of the Ising model has proved to be a very useful simplification near the critical point, yielding the same quantitative behavior as the original, discrete model.

It is possible to obtain one from the other in an approximate way as we now show. First, define $\psi(\vec{r})$ to be the spin density over a small volume V ,

$$\frac{1}{N_V} \approx \frac{1}{V} \int_V d\vec{r} \psi(\vec{r}).$$

With this in mind, we write,

$$H \sum_i S_i = \frac{HN}{V} \int d^d \vec{x} \psi(\vec{x}).$$

Next, note that $(S_i - S_j)^2 = 2(1 - S_i S_j)$ since $S_i^2 = S_j^2 = 1$. Therefore

$$-J \sum_{\langle i,j \rangle} S_i S_j = -\frac{J}{2} \sum_{i,j} S_i S_j = \frac{J}{2} \sum_{i,j} \left[\frac{1}{2} (S_i - S_j)^2 - 1 \right] = \frac{J}{4} \sum_{i,j} (S_i - S_j)^2 - \frac{JN^2}{2}.$$

We next approximate,

$$\frac{(S_i - S_j)^2}{\Delta x^2} \approx \left(\frac{\partial \psi(x)}{\partial x} \right)^2.$$

This is clear in one dimension, but it also holds in more than one dimension with the nearest neighbor Ising model. We then find,

$$-J \sum_{\langle i,j \rangle} S_i S_j \approx \frac{J \Delta x^2 N}{2V} \int d^d \vec{x} (\nabla \psi)^2 - \frac{J N^2}{2}.$$

We neglect the second term in the right hand side as it is only an additive constant. We also define the constant $K = \frac{J \Delta x^2 N}{4V}$. Therefore in the continuum limit, the energy of any spin configuration is given by,

$$E = \frac{K}{2} \int d^d \vec{x} (\nabla \psi)^2 - H \int d^d \vec{x} \psi(\vec{x}).$$

There is, however, an additional complication in the calculation of the partition function. In the original model, the discrete spins are constrained to be ± 1 , whereas ψ is a continuum variable. The standard procedure involved allowing arbitrary values of ψ in the configurations, but to penalize -via an increased energy- those configurations that locally deviate from $\psi = \pm 1$. Schematically,

$$\sum_{\{S_i = \pm 1\}} = \prod_i \sum_{S_i = -\infty}^{S_i = \infty} \delta_{S_i^2, 1} \propto \prod_{\vec{x}} \int d\psi(\vec{x}) \delta(\psi^2(\vec{x}) - 1).$$

Where we have extended the sum but introduced a Kronecker delta, followed by the continuum approximation. In order to incorporate the constraint given by the delta function into the energy, we introduce an exponential representation of the delta function,

$$\delta(x) = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon \sqrt{\pi}} e^{-x^2/\epsilon^2} :$$

leading to

$$\prod_{\vec{x}} \int d\psi(\vec{x}) \frac{1}{\epsilon \sqrt{\pi}} e^{-\frac{1}{\epsilon^2} (\psi^2 - 1)^2}.$$

Now, by expanding the square in the exponent, redefining the constants appropriately, and removing constant terms, we arrive at our field theory model,

$$E = \int d^d \vec{x} \left[\frac{K}{2} (\nabla \psi)^2 + f(\psi) \right], \quad f(\psi) = \frac{r}{2} \psi^2 + \frac{u}{4} \psi^4 - H \psi. \quad (10.1)$$

This is the so called ψ^4 field theory. The critical point of this model is at $r = 0$ so that $r > 0$, $T > T_c$ and $r < 0$ for $T < T_c$. With this choice, the function $f(\psi)$ is shown in Fig. 10.1. For $r > 0$, there

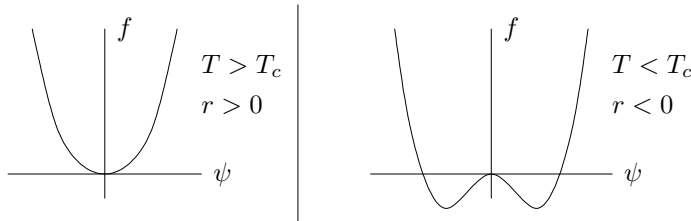


Figure 10.1: Uniform part of the energy E in the absence of magnetic field.

is a single minimum of the energy that corresponds to an average value of $\psi = 0$. Below T_c there are two minima in the energy that correspond to states of finite magnetization. It is customary to

assume $r \propto (T - T_c)$, $u > 0$ so that $\psi \sim \sqrt{T_c - T}$ in the absence of fluctuations, as given by mean field theory. We will elaborate on these issues below when we describe Landau's theory.

The ultimate goal of the theory is the determination of the critical exponents α , β , γ , δ , η , and ν defined earlier, and to understand the observed experimental deviations from mean field theory. For reference, we list here the definition of the various exponents.

$$F(T) \sim |T - T_c|^{2-\alpha}, \quad H = 0 \quad C = \frac{\partial^2 F}{\partial T^2} \sim |T - T_c|^{-\alpha}.$$

$$\psi \sim (T_c - T)^\beta, \quad H = 0, T < T_c$$

$$\chi_T = \left(\frac{\partial \psi}{\partial H}\right)_T \sim |T - T_c|^{-\gamma}, \quad H = 0$$

$$\psi \sim H^{1/\delta}, \quad T = T_c$$

$$\langle \psi(\vec{x}) \psi(0) \rangle \sim \frac{1}{x^{d-2+\eta}}, \quad T = T_c, H = 0 \quad \langle \psi(\vec{x}) \psi(0) \rangle \sim \begin{cases} e^{-x/\xi}, & T > T_c \\ \langle \psi \rangle^2 + \mathcal{O}(e^{-x/\xi}), & T < T_c \end{cases}$$

and

$$\xi \sim |T - T_c|^{-\nu}, \quad H = 0$$

This is the first time we have mentioned η and ν in this context.

10.1 Landau theory

Landau introduced the ψ^4 theory in a different way. He asserted that the continuous phase transition at the critical point involves going from a disordered phase with high symmetry to an ordered phase with lower symmetry. To describe the transition and its symmetry reduction he introduced an order parameter, nonzero in the ordered phase only. For example, the order parameter in a ferromagnet is the magnetization. In general, identifying the order parameter and the symmetry broken at T_c can be quite subtle. For the time being we will restrict ourselves to a scalar order parameter, where the symmetry broken is Z_2 , that is, when $\psi \leftrightarrow -\psi$ symmetry is spontaneously broken. A slightly more complicated example would involve a magnetic dipole or spin confined to a plane (dimension $d = 2$). In this case all spin orientations would be equivalent, so that the order parameter is now a vector in the plane $\vec{\psi} = \psi_x \hat{x} + \psi_y \hat{y}$. The group corresponding to this rotation symmetry is $O(2)$. A spin in $d = 3$ would have $\vec{\psi} = \psi_x \hat{x} + \psi_y \hat{y} + \psi_z \hat{z}$, and if all directions are equivalent the symmetry group is $O(3)$.

Landau argued that near the point at which the symmetry is spontaneously broken, the free energy must be a function of the order parameter ψ . Hence

$$F = \int d^d x f(\psi)$$

and he further assumed that near the transition (where ψ abruptly deviates from zero), the free energy must have an expansion in powers of ψ ,

$$f(\psi) = \frac{r}{2} \psi^2 + \frac{u}{4} \psi^4 + \dots$$

In this case, the expansion has been chosen to preserve the symmetry $\psi \leftrightarrow -\psi$ so that no odd terms in ψ would appear. More complicated order parameters and symmetries would lead to different expansions of f . In the event that an external magnetic field breaks the symmetry, another term is added to the free energy,

$$-H \int d^d x \psi$$

Since at high temperatures, in the disordered phase, $\psi = 0$, one chooses $r > 0$ for $T > T_c$. However, at low temperatures, we have an ordered phase with $\psi \neq 0$, and hence $r < 0$. To make

sure that the free energy cannot be reduced by further increasing ψ he set $u > 0$. There is no need for u to have any special dependence in T near T_c , so it is taken to be a constant. It is customary to expand r in Taylor series,

$$r = r_0(T - T_c) + \dots$$

and keep only the lowest order term, with $r_0 > 0$.

Finally, spatial fluctuations in order parameter cost energy and would be suppressed. A way to incorporate this into F is to assume that it also depends on gradients of ψ . This dependence can in turn be expanded in powers of $\vec{\nabla}$ and ψ , so that the lowest order, nontrivial term is

$$\frac{1}{2}K \int d^d \vec{r} |\nabla \psi|^2$$

with $K > 0$. Note that this term raises the free energy of inhomogeneities ($\vec{\nabla} \psi \neq 0$). In summary, the complete Landau free energy is

$$F = \int d^d x \{f(\psi) + \frac{K}{2}|\nabla \psi|^2\} - H \int d^d x \psi \quad f(\psi) = \frac{r}{2}\psi^2 + \frac{u}{4}\psi^4 \quad (10.2)$$

In high energy physics, this form of the energy is called a ψ^4 field theory. The ∇^2 term corresponds to the kinetic energy of a particle; r is proportional to its mass; and ψ^4 describes interactions. van der Waals also introduced the same free energy (the case with $K \equiv 0$ is normally given in textbooks). He was thinking of a liquid-gas transitions, so that $\psi = n - n_c$, where n is density and n_c is the density at the critical point.

It is worth emphasizing that the whole machinery of quadratic, quartic, and square gradient terms in the free energy is necessary. For example, note that the critical point is defined in thermodynamics as

$$\left(\frac{\partial P}{\partial V}\right)_{T_c} = 0, \quad \left(\frac{\partial^2 P}{\partial V^2}\right)_{T_c} = 0.$$

for a liquid-gas system. In a ferromagnet, one similarly has,

$$\left(\frac{\partial H}{\partial m}\right)_{T_c} = 0, \quad \left(\frac{\partial^2 H}{\partial m^2}\right)_{T_c} = 0.$$

One can write a series expansion for the equation of state at $T = T_c$ in a liquid as follows,

$$p(n, T_c) = p_c + \left(\frac{\partial p}{\partial n}\right)_{T_c} (n - n_c) + \frac{1}{2} \left(\frac{\partial^2 p}{\partial n^2}\right)_{T_c} (n - n_c)^2 + \frac{1}{6} \left(\frac{\partial^3 p}{\partial n^3}\right)_{T_c} (n - n_c)^3 + \dots$$

and similarly for a ferromagnet

$$H(m, T_c) = H_c + \left(\frac{\partial H}{\partial m}\right)_{T_c} (m - m_c) + \frac{1}{2} \left(\frac{\partial^2 H}{\partial m^2}\right)_{T_c} m^2 + \frac{1}{6} \left(\frac{\partial^3 H}{\partial m^3}\right)_{T_c} m^3 + \dots$$

If the dependence of p of H on the order parameter is analytic at the critical point, then for both systems one must have

$$(p - p_c) \propto (n - n_c)^\delta \quad H \propto m^\delta$$

with $\delta = 3$. Experimentally, however, $\delta \approx 5.0$ in three dimensions, while $\delta = 15$ in $d = 2$. This simple argument indicates that there is quite a subtle problem to solve.

We return to classical Landau theory by considering first homogeneous states of the system (in which the gradients of the order parameter vanish). In this case, we need only consider

$$\frac{F}{V} = f(\psi) - H\psi = \frac{r_0}{2}(T - T_c)\psi^2 + \frac{u}{4}\psi^4 - H\psi. \quad (10.3)$$

The equilibrium state can be found by direct minimization of the free energy with respect to the order parameter. For $H = 0$ we find,

$$\frac{\partial f}{\partial \psi} = 0 = r_0(T - T_c)\psi + u\psi^3$$

This cubic equations has three solutions

$$\psi = 0, \quad \psi = \pm \sqrt{\frac{r_0(T_c - T)}{u}} \quad (10.4)$$

These three solutions are minima in some range of parameters, and maxima in others. This can be checked by computing the sign of $\partial^2 f / \partial \psi^2$. It is found that $\psi = 0$ is a minimum of the free energy for $T > T_c$, and $\psi = \pm \sqrt{\frac{r_0(T_c - T)}{u}}$ are minima for $T < T_c$. Therefore this phenomenological expansion of F does provide a phase transition at $T = T_c$ between a high temperature phase in which the order parameter vanishes, and a low temperature phase of nonzero order parameter. We also note that Eq. (10.4) implies $\beta = 1/2$, the same value predicted by mean field theory.

In the case $H \neq 0$, there is only one solution

$$\frac{\partial}{\partial \psi}(f(\psi) - H\psi) = 0 = \frac{\partial f}{\partial \psi} - H$$

or

$$H = r_0(T - T_c)\psi + u\psi^3$$

Note that this gives

$$1/\chi_T = \left(\frac{\partial H}{\partial \psi} \right)_T = r_0(T - T_c)$$

close to the critical point, or

$$\chi_T \propto |T - T_c|^{-1}$$

Hence, $\gamma = 1$ in Landau theory, also the same value given by the mean field approximation. At $T = T_c$ we have $H \propto \psi^3$, or $\delta = 3$.

Finally, since for $H = 0$

$$f = \frac{1}{2}r\psi^2 + \frac{1}{4}u\psi^4$$

and

$$\psi^2 = \begin{cases} 0, & T > T_c \\ -\frac{r}{u}, & T < T_c \end{cases}$$

we have,

$$f = \begin{cases} 0, & T > T_c \\ -\frac{1}{4}\frac{r^2}{u} \propto (T_c - T)^2, & T < T_c \end{cases}$$

Since the specific heat $c \propto \partial^2 f / \partial T^2$, we find

$$c \propto \begin{cases} 0, & T > T_c \\ \text{constant}, & T < T_c \end{cases}$$

The fact that there is no divergence implies $\alpha = 0$.

In summary, the values of the critical exponents given by Landau's theory are the same as those given by mean field theory:

$$\begin{array}{ll} \alpha & 0 \text{ (discontinuity)} \\ \beta & 1/2 \\ \gamma & 1 \\ \delta & 3 \end{array}$$

10.2 Ornstein-Zernicke theory of correlations

We next extend Landau's theory by allowing spatially dependent fluctuations. We follow the treatment by Ornstein and Zernicke. It works well except near the critical point. The essential point is that in a bulk phase for, say, $T \gg T_c$ and $H = 0$ ψ has an average value of zero $\langle \psi \rangle = 0$ so that we expect $\psi^4 \ll \psi^2$. Therefore above T_c we approximate the free energy by

$$F \approx \int d^d x \left[\frac{K}{2} (\nabla \psi)^2 + \frac{r}{2} \psi^2 \right]$$

Far below T_c , we expect ψ to be close to its average value

$$\langle \psi \rangle = \pm \sqrt{-\frac{r}{u}}$$

where $r = r_0(T - T_c) < 0$ below T_c . Let $\Delta\psi = \psi - \langle \psi \rangle$. 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

$$\begin{aligned} \psi^4 &= (\Delta\psi + \langle \psi \rangle)^4 \\ &= (\Delta\psi)^4 + 4\langle \psi \rangle (\Delta\psi)^3 + 6\langle \psi \rangle^2 (\Delta\psi)^2 + 4\langle \psi \rangle^3 (\Delta\psi) + \langle \psi \rangle^4. \end{aligned}$$

Grouping terms in powers of $\Delta\psi$, we find

$$\frac{r}{2} \psi^2 + \frac{u}{4} \psi^4 = \frac{r}{2} \langle \psi \rangle^2 + \frac{u}{4} \langle \psi \rangle^4 + (r\langle \psi \rangle + u\langle \psi \rangle^3) \Delta\psi + \left(\frac{r}{2} + \frac{3u}{2} \langle \psi \rangle^2 \right) (\Delta\psi)^2 + \mathcal{O}((\Delta\psi)^3)$$

But $\langle \psi \rangle = \pm \sqrt{-r/u}$, so

$$\frac{r}{2} \psi^2 + \frac{u}{4} \psi^4 = \frac{r}{2} \langle \psi \rangle^2 + \frac{u}{4} \langle \psi \rangle^4 - r(\Delta\psi)^2 + \mathcal{O}((\Delta\psi)^3).$$

The first two terms in the right hand side are a constant independent of the order parameter, and we omit them. Also $(\Delta\psi)^3 \ll (\Delta\psi)^2$, hence we approximate

$$F \approx \int d^d \vec{x} \left(\frac{K}{2} (\nabla \Delta\psi)^2 + |r| (\Delta\psi)^2 \right)$$

where we have used the fact that $r < 0$ below T_c .

Both approximate free energies (above and below T_c are formally the same as the energy for the fluctuating interface

$$\frac{\sigma}{2} \int d^{d-1} \vec{x} \left(\frac{\partial h}{\partial x} \right)^2$$

Averages and correlation functions can be calculated in exactly the same way as we did for that case. The definitions of Fourier transforms are the same (except that now $d^{d-1}x \rightarrow d^d x$ and $L^{d-1} \rightarrow L^d$, for example), and the arguments concerning translational invariance of space, and isotropy of space still hold. Let

$$C(x) = \langle \psi(x) \psi(0) \rangle$$

be the spin-spin correlation function, then

$$\hat{C}(\vec{k}) = \frac{k_B T}{K k^2 + r}, \quad T > T_c$$

and since

$$\langle \Delta\psi(x) \Delta\psi(0) \rangle = \langle \psi(x) \psi(0) \rangle - \frac{|r|}{u}$$

by using $\langle \psi \rangle^2 = |r|/u$, we have

$$\hat{C}(\vec{k}) = \frac{k_B T}{K k^2 + 2|r|} - \frac{|r|}{u} (2\pi)^d \delta(\vec{k}), \quad T < T_c$$

Inverse Fourier transformation will give the correlation function in real space. Recall that the Green's function of the Helmholtz equation is defined as

$$(\nabla^2 - \xi^{-2})G(x) = -\delta(x)$$

where $\xi > 0$ is a constant. The Fourier transform of this equation is

$$\hat{G}(k) = 1/(k^2 + \xi^{-2}),$$

exactly the same that we have obtained from the Ornstein-Zernicke theory. The Green's function of the Helmholtz equations is

$$G(x) = \begin{cases} \xi/2 e^{-x/\xi}, & d = 1 \\ \frac{1}{2\pi} K_0(x/\xi), & d = 2 \\ \frac{1}{4\pi x} e^{-x/\xi}, & d = 3 \end{cases}$$

$K_0(y)$ is the modified Bessel function of order *zero*th and satisfies $K_0(y \rightarrow 0) \sim -\ln y$, and $K_0(y \rightarrow \infty) \sim (\pi/2y)^{1/2} e^{-y}$. This completes our calculation of the correlation function as a function of the dimensionality of the system.

In discussions of critical phenomena, these three relations are consolidated into

$$G(x) \propto \frac{1}{x^{(d-1)/2}} \frac{1}{\xi^{(d-3)/2}} e^{-x/\xi}.$$

Hence

$$G(x) \sim e^{-x/\xi}$$

unless $\xi = \infty$, in which case

$$G(x) \sim \frac{1}{x^{d-2}}$$

for large x . In summary, up to a constant of proportionality, we have

$$C(x) = \begin{cases} e^{-x/\xi}, & T > T_c \\ \frac{1}{x^{d-2}}, & T = T_c \\ \langle \psi \rangle^2 + \mathcal{O}(e^{-x/\xi}), & T < T_c \end{cases}$$

where the correlation length ξ is

$$\xi = \begin{cases} \sqrt{\frac{K}{r}}, & T > T_c \\ \sqrt{\frac{K}{2|r|}}, & T < T_c \end{cases}$$

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

$$\langle \Delta\psi(x) \Delta\psi(0) \rangle = e^{-x/\xi}$$

As sketched in Fig. 10.2 the correlation length diverges as $\xi \sim 1/\sqrt{|r|}$ above and below T_c , so

$$\xi \sim |T - T_c|^{-1/2}$$

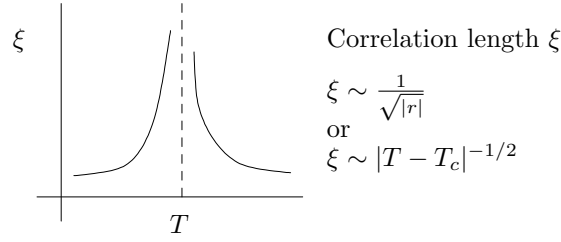
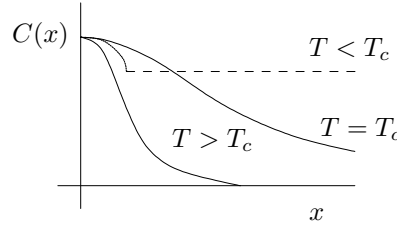


Figure 10.2:

Figure 10.3: Correlation function above at and below T_c

in Ornstein-Zernicke theory. The correlation function becomes a power law at T_c .

$$C(x) = \frac{1}{x^{d-2}}, \quad T = T_c, H = 0$$

Since in general we have defined the exponents

$$\xi \sim |T - T_c|^\nu$$

and

$$C(x) \sim \frac{1}{x^{d-2+\eta}}$$

we have that Ornstein-Zernicke theory predicts $\nu = 1/2$ and $\eta = 0$.

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

Therefore the continuum formulations of the Ising model given so far give exactly the same predictions as the mean field approximation, and hence disagree in fundamental ways with both experiments and numerical approximate solutions, as well as Onsager's exact solution in two dimensions.

10.3 Ginzburg criterion

In the Ornstein-Zernicke theory we have explicitly used the fact that fluctuations are small $(\Delta\psi)^3 \ll (\Delta\psi)^2$ so that the free energy considered is quadratic in the order parameter. We show, however, that this assumption is completely inconsistent with the values of the critical exponents given by the theory, and hence the results obtained can be considered suspect. In order to quantify the magnitude of fluctuations, we compute

$$\frac{\langle \Delta\psi^2 \rangle}{\langle \psi \rangle^2}, \quad (10.5)$$

in the vicinity of T_c ($r = 0$). We also define,

$$t = \left| \frac{T - T_c}{T_c} \right|$$

and examine the behavior of the ratio (10.5) as $t \rightarrow 0$.

By definition we have $\langle \psi \rangle \sim t^\beta$ as $t \rightarrow 0$. On the other hand, the magnetic susceptibility

$$\chi_T = \int d^d \vec{x} \langle \Delta \psi(\vec{x}) \Delta \psi(0) \rangle \sim t^{-\gamma},$$

by definition of γ . We estimate the order of magnitude of the integral as $\langle \Delta \psi^2 \rangle \xi^d$, and therefore by using the definition of the correlation length exponent $\xi \sim t^{-\nu}$, we find $\langle \Delta \psi^2 \rangle t^{-\nu d} \sim t^{-\gamma}$. Therefore,

$$\frac{\langle \Delta \psi^2 \rangle}{\langle \psi \rangle^2} \sim t^{-\gamma + \nu d - 2\beta}.$$

This relation is known as the Ginzburg criterion. Note that if we substitute the mean field (or Landau) values of the critical exponents, we find

$$\frac{\langle \Delta \psi^2 \rangle}{\langle \psi \rangle^2} \sim t^{(d-4)/2}.$$

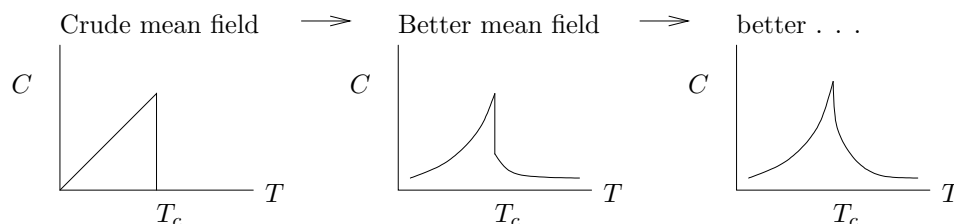
If $d > 4$ the approximation of neglecting higher order terms in the expansion of the fluctuations is justified, as the fluctuations become negligible relative to the average as the critical point is approached. For $d < 4$ on the other hand, the variance of the fluctuations diverges as the critical point is approached. The quadratic approximation for the free energy used in this chapter is therefore not valid.

Chapter 11

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 given earlier is quite crude. Refinements can be made by incorporating more local structure. Pathria's book on Statistical Mechanics reviews some of this work. Consider for example 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 , whereas our mean field calculations predicted a discontinuity at the critical temperature. Further refinements of the mean field approach have led to much better predictions of C away from but not close to T_c . At T_c , all mean field theories, no matter how refined, give C to be discontinuous.



All mean theories give $C \sim$ discontinuous

Figure 11.1:

Experiments also show a divergence as given in Fig. 11.2 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 critical points. As discussed in the previous chapter, Landau's theory gives

$$n - n_c \sim (T_c - T)^\beta$$

with $\beta = 1/2$. Experiments in three dimensions give $\beta \approx 0.3$ for many liquids (usually quoted as $1/3$ historically). Figure 11.3 shows this result. The figure is taken from the most comprehensive experimental survey given in Guggenheim, J. Chem. Phys. 13, 253 (1945) and reproduced in Stanley's book.

If one examines typical configurations of a ferromagnet near T_c (where $\xi \rightarrow \infty$) the result is striking. Figure 11.4 shows some quite old simulations of the Ising model in two dimensions, also taken from Stanley's book. Spins up are shown in white in the figure, and spins down in black. As T_c is approached (frames (e) and (f) in the figure) domains of either phase appear to adopt a fractal structure. The structure of the configurations is similar for a fluid near its critical point. Imagine we heat it through T_c while fixing the pressure so that it moves on the coexistence line

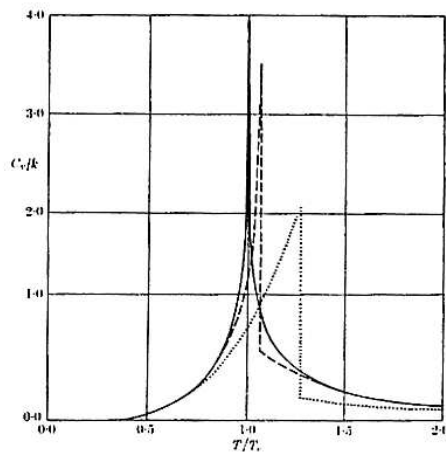


FIG. 1.13. The solid curve shows the specific heat of the two-dimensional Ising model as obtained from the exact solution of Onsager (solid curve), from the Bethe approximation (dotted curve), and the Kramers-Wannier and Kikuchi approximation (broken curve). After Domb (1960).

Figure 11.2:

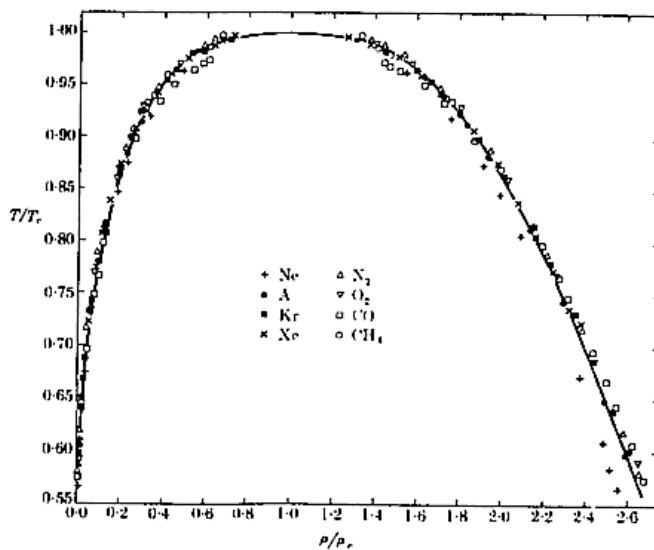


FIG. 1.8. Measurements on eight fluids of the coexistence curve (a reflection of the $P\rho T$ surface in the ρT plane analogous to Fig. 1.3). The solid curve corresponds to a fit to a cubic equation, i.e. to the choice $\beta = \frac{1}{3}$, where $\rho - \rho_c \sim (-\epsilon)^{\beta}$. From Guggenheim (1945).

Figure 11.3:

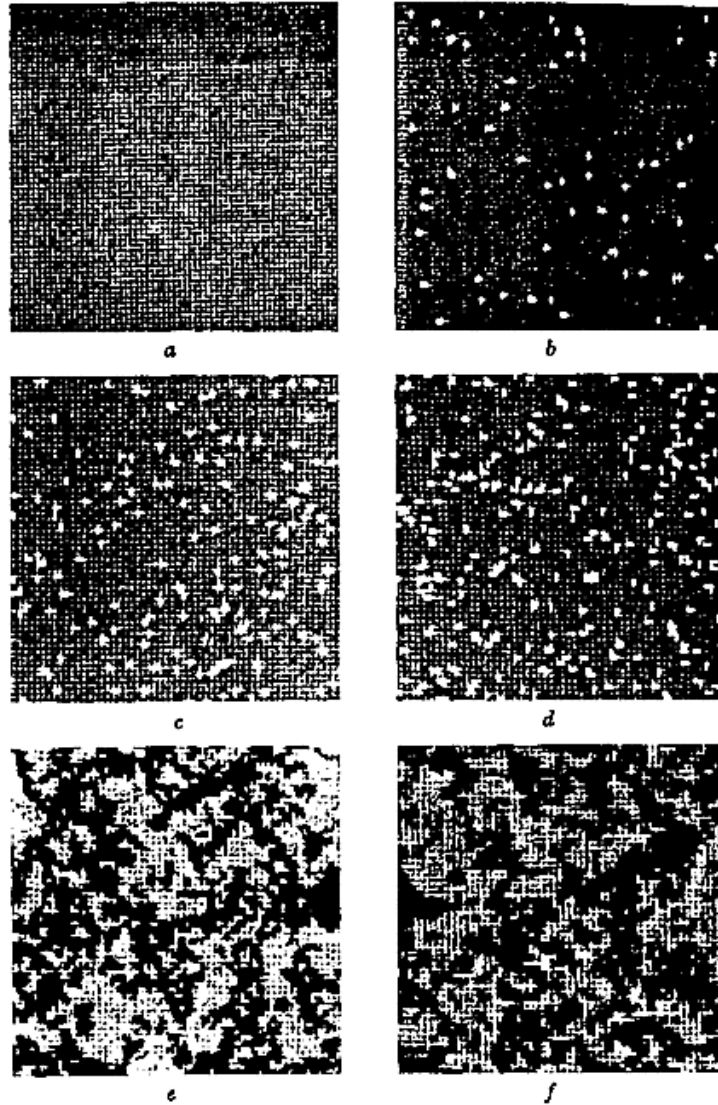


FIG. 1.5. Schematic indication of the lattice-gas model of a fluid system. Fig. 1.4 (which contains 8×20 cells) should be interpreted as representing a relatively small portion of this lattice (which contains 64×64 cells). (a) is the completely ordered state (which exists only at $T = 0$); (b), $T \simeq \frac{1}{2}T_c$; (c), $T \simeq \frac{1}{2}T_c$; (d), $T \simeq \frac{1}{2}T_c$; (e), $T \simeq T_c$; (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 11.4:

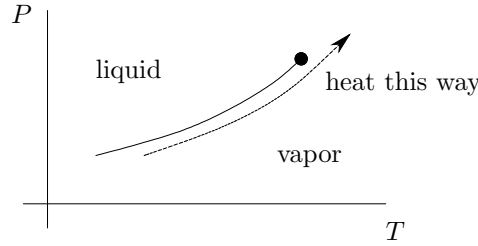


Figure 11.5:

of the phase diagram (Fig. 11.5). The appearance of the configurations will be quite different at different temperatures. The system is transparent, except for the interface and meniscus separating

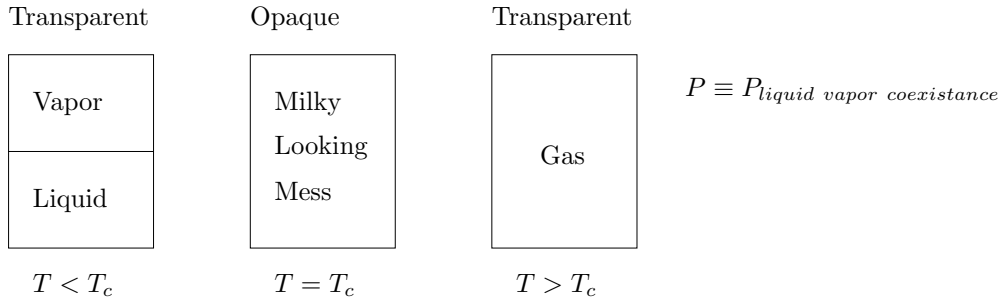


Figure 11.6:

liquid and vapor, for $T < T_c$ and $T > T_c$ (Fig. 11.6). However, as we approach T_c , it becomes turbid or milky looking. This is called “critical opalescence”. The correlation length ξ is so large that density fluctuations become of the order of the wavelength of light (or larger) and hence scatter visible light. This is not unlike the size of spin fluctuations in Fig. 11.4.

Given the difficulties in reconciling experimental observations near the critical point of many systems and theoretical treatments based on mean field approaches or Landau’s theory, a significant amount of effort was devoted to proving rigorous results that would follow from the laws of thermodynamics alone. Unfortunately, thermodynamic arguments rely on system stability at the critical point (entropy is maximized, free energy is minimized, etc.), and hence only inequalities can be obtained. These inequalities follow from the curvature of the thermodynamic potential at the critical point, and the fact that any fluctuation in the system has to obey such inequalities. Such studies led to rigorous inequalities among the six critical exponents. (One should mention the assumption that any divergence is the same for $T \rightarrow T_c^+$ or $T \rightarrow T_c^-$. This 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 results. Other inequalities were proved with some plausible assumptions. It turns out that all inequalities are satisfied for all the exponents quoted previously (be them theoretically derived or experimentally observed), but as equalities. In fact, all the thermodynamic inequalities derived turned out to be obeyed as equalities.

We summarize all the independent relations between exponents

$$\alpha + 2\beta + \gamma = 2$$

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

$$\gamma = \beta(\delta - 1)$$

and

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

Given the number of exponents involved, and the relationships given, it turns out that for the Ising model there are only two independent exponents. The last equation is called the “hyperscaling” relation as it is the only one that explicitly involves the spatial dimension d . The others are simply called “scaling” relations.

11.1 Scaling With ξ

A breakthrough in the theory of critical phenomena occurred when Widom introduced phenomenologically the idea of scaling, which lead to the derivation of the equality in the thermodynamic inequalities given above. While his ideas made possible also the interpretation of experimental evidence, his scaling assumption remains an assumption, and it did not predict the values of the exponents, only relationships among them. In essence, he assumed that all thermodynamic functions have a scale invariant contribution as $T \rightarrow T_c$, and that this scale invariant contribution becomes a generalized homogeneous function of its arguments. Consider, for example, the equilibrium correlation function $C = C(x)$ which in the case of the Ising model is a function of T and H as well (the two other intensive parameters on which the partition function depends). Instead of using T as the independent variable, let us introduce ξ instead. Near T_c they are both related through $\xi \sim |T - T_c|^{-\nu}$. The assumption that C is a generalized homogeneous function of its arguments is written as,

$$C(x, \xi, H) = \frac{1}{\xi^{d-2+\eta}} \tilde{C}(x/\xi, H\xi^\eta). \quad (11.1)$$

Note that physically this equation embodies the idea of scale invariance: correlations do not depend on x alone, but on the ratio x/ξ instead. Similarly, since blocks of spins are correlated over a distance ξ , only a little H is needed to significantly magnetize the system: the effect of H is magnified by ξ^η .

We can also introduce explicitly the temperature into the scaling relation (11.1). Since $\xi \sim |T - T_c|^{-\nu}$, which we write as $\xi \sim t^{-\nu}$ by introducing $t = |T - T_c|/T_c$, we have

$$C(x, T, H) = \frac{1}{x^{d-2+\eta}} \tilde{C}(xt^\nu, H/t^\Delta) \quad (11.2)$$

with

$$\Delta = y\nu$$

being called the “gap exponent”.

11.1.1 Derivation of exponent equalities

We now show how the scaling relation (11.2) allows one to derive equalities among critical exponents. First we start with the magnetic susceptibility sum rule

$$\begin{aligned} \chi_T &\propto \int d^d x C(x, T, H) \\ &= \int d^d x \frac{1}{x^{d-2+\eta}} \tilde{C}(xt^\nu, H/t^\Delta) \\ &= (t^\nu)^{-2+\eta} \int d^d (xt^\nu) \frac{\tilde{C}(xt^\nu, H/t^\Delta)}{(xt^\nu)^{d-2+\eta}} \end{aligned}$$

so that

$$\chi_T(T, H) = t^{(-2+\eta)\nu} \tilde{\chi}(H/t^\Delta)$$

where we have defined

$$\int_{-\infty}^{\infty} d^d(xt^\nu) \frac{1}{(xt^\nu)^{d-2+\eta}} \tilde{C}(xt^\nu, H/t^\Delta) = \tilde{X}\left(\frac{H}{t^\Delta}\right)$$

But by definition of the exponent γ ,

$$\chi(T, H=0) = t^{-\gamma},$$

hence,

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

This exponent equality is called Fisher's law.

Another equality can be derived by using the definition of magnetic susceptibility,

$$\frac{\partial \langle \psi(T, H) \rangle}{\partial H} = \chi_T(T, H)$$

relation that can be integrated to

$$\begin{aligned} \langle \psi(T, H) \rangle &= \int_0^H dH' \chi(T, H') \\ &= \int_0^H dH' t^{-\gamma} \tilde{\chi}\left(\frac{H'}{t^\Delta}\right) \\ &= t^{\Delta-\gamma} \int_0^H d\left(\frac{H'}{t^\Delta}\right) \tilde{\chi}\left(\frac{H'}{t^\Delta}\right) \\ &= t^{\Delta-\gamma} \tilde{\psi}(H/t^\Delta) \end{aligned}$$

But by definition of the exponent β we have

$$\langle \psi(T, H=0) \rangle = t^\beta$$

Therefore,

$$\beta = \Delta - \gamma$$

and the gap exponent

$$\Delta = \beta + \gamma$$

or equivalently, $y = (\beta + \gamma)/\nu$. So we have

$$\langle \psi(T, H) \rangle = t^\beta \tilde{\psi}\left(\frac{H}{t^{\beta+\gamma}}\right) \quad (11.3)$$

We mention next a useful transformation of the scaling relations that allows one to change the scaling variables, or scaling fields as they are sometimes called. We illustrate it with a manipulation of Eq. (11.3). We clearly do not know the functional form of $\tilde{\psi}$, all we know is that it is not a function of the two arguments separately, but rather a function of the combination $\frac{H}{t^{\beta+\gamma}}$. Therefore if we multiply or divide by any function of $H/t^{\beta+\gamma}$ into $\tilde{\psi}$, we obtain another unknown function, which still a function of the same argument $H/t^{\beta+\gamma}$. In particular, a crafty choice can cancel the t^β factor as follows:

$$\langle \psi(T, H) \rangle = t^\beta \left(\frac{H}{t^{\beta+\gamma}}\right)^{\frac{\beta}{\beta+\gamma}} \left(\frac{H}{t^{\beta+\gamma}}\right)^{\frac{-\beta}{\beta+\gamma}} \tilde{\psi}\left(\frac{H}{t^{\beta+\gamma}}\right)$$

We combine the last two terms on the right hand side into $\tilde{\psi}$, and write,

$$\langle \psi(T, H) \rangle = H^{\beta/\beta+\gamma} \tilde{\psi}\left(\frac{H}{t^{\beta+\gamma}}\right)$$

with $H^{-\beta/\beta+\gamma}\tilde{\psi}(H) = \tilde{\psi}(H)$.

This new form of the scaling relation allows us to determine yet another exponent equality. Recall that by definition of δ

$$\langle\psi(T_c, H)\rangle \sim H^{1/\delta}$$

gives the shape of the critical isotherm. Therefore $\delta = 1 + \gamma/\beta$, or

$$\gamma = \beta(\delta - 1).$$

This is called Widom's scaling law.

Yet one more relation can be derived as follows. Since $\langle\psi\rangle = \partial F/\partial H$,

$$\begin{aligned} F &= \int dH \langle\psi\rangle \\ &= \int dH t^\beta \tilde{\psi}\left(\frac{H}{t^{\beta+\gamma}}\right) \\ &= t^{2\beta+\gamma} \int d\left(\frac{H}{t^{\beta+\gamma}}\right) \tilde{\psi}\left(\frac{H}{t^{\beta+\gamma}}\right). \end{aligned}$$

Hence

$$F(T, H) = t^{2\beta+\gamma} \tilde{F}\left(\frac{H}{t^{\beta+\gamma}}\right)$$

But by definition of the exponent α ,

$$F(T, 0) = t^{2-\alpha}$$

therefore $2 - \alpha = 2\beta + \gamma$, or

$$\alpha + 2\beta + \gamma = 2$$

which is Rushbrooke's law.

Finally we analyze the behavior of successive derivatives of the free energy. This is important in the analysis we carried out earlier regarding a Taylor series expansion of thermodynamic variables, and how they failed to predict basic properties near the critical point. This calculation clarifies this issue. The form of the free energy is

$$F(T, H) = t^{2-\alpha} \tilde{F}\left(\frac{H}{t^{\beta+\gamma}}\right),$$

therefore

$$\lim_{H \rightarrow 0} \frac{\partial^n F}{\partial H^n} = F^{(n)} = t^{2-\alpha} t^{-n(\beta+\gamma)}$$

Therefore we obtain the result that successive derivatives of F are separated by a constant “gap” exponent $(\beta + \gamma)$, with the corresponding exponent being

$$y_n = 2 - \alpha - n(\beta + \gamma)$$

In short, the free energy is not analytic at the critical point, and hence cannot be expanded in Taylor series of its arguments.

We finally discuss two physically intuitive arguments to derive the hyperscaling relation, which are nevertheless not rigorous. The free energy is extensive and has units of energy, so

$$F \sim k_B T \frac{V}{\xi^d}$$

as one can think of V/ξ^d as the number of independent parts of the system. Since $\xi \sim |T - T_c|^{-\nu}$, then $F = |T - T_c|^{\nu d}$. But the singular behavior in F defines α via

$$F = |T - T_c|^{2-\alpha}$$

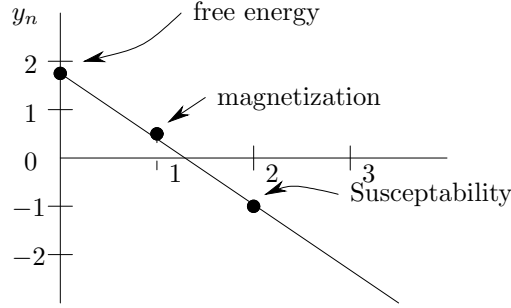


Figure 11.7: Gap Exponent

Hence

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

The hyperscaling relation works for all systems, except that mean field solutions must have $d = 4$. Another argument, not as compelling, but interesting, is as follows. One can imagine that near T_c fluctuations in the order parameter are of the same order as $\langle \psi \rangle$ itself. If so,

$$\frac{\langle (\Delta\psi)^2 \rangle}{\langle \psi \rangle^2} = \mathcal{O}(1)$$

This assumption is the weak part of the argument as we do not know the relative size of fluctuations to the average. As argued earlier,

$$\langle \psi \rangle^2 \sim t^{2\beta}$$

Also as given in the previous chapter,

$$\int d^d x \langle \Delta\psi(x) \Delta\psi(0) \rangle = \chi_T$$

The integral is only nonzero for $x \leq \xi$. In that range, $\langle \Delta\psi(x) \Delta\psi(0) \rangle \sim \mathcal{O}(\langle (\Delta\psi)^2 \rangle)$ so that

$$\chi_T = \int d^d x \langle \Delta\psi(x) \Delta\psi(0) \rangle \sim \xi^d \langle (\Delta\psi)^2 \rangle = \chi_T$$

Since $\chi_T \sim t^{-\gamma}$ and $\xi \sim t^{-\nu}$ then $\langle (\Delta\psi)^2 \rangle = t^{\nu d - \gamma}$. The relative fluctuation therefore obeys

$$\frac{\langle (\Delta\psi)^2 \rangle}{\langle \psi \rangle^2} \sim t^{\nu d - \gamma - 2\beta}$$

which we assumed is an order one quantity. Therefore, $\nu d = \gamma + 2\beta$. To put it into the same form as before, use Rushbrooke's scaling law, $\alpha + 2\beta + \gamma = 2$. This gives $\nu d = 2 - \alpha$.

Chapter 12

Renormalization Group

The renormalization group is a theory that builds on scaling ideas, and provides a procedure to calculate critical exponents. It exploits the observation that phase transitions can be described by a few long wavelength properties. These include the symmetry of the order parameter and the dimensionality of space (for systems without long range forces or quenched impurities). Short length scale properties are systematically eliminated by the renormalization group transformation.

The basic ideas behind the renormalization group were given by Kadanoff in 1966 in his “block spin” treatment of the Ising model. A more general method, also free of inconsistencies, was given by Wilson (who won the Nobel prize for his work). Furthermore, Wilson and Fisher devised a very useful implementation of the renormalization group in the so called $\epsilon (= 4 - d)$ expansion, where a small parameter ϵ is eventually extrapolated to unity to predict properties in $d = 3$.

At a critical point, Kadanoff reasoned, spins in the Ising model act together up to distances of the order of the correlation length. He argued that one could compute the partition function recursively, by progressively eliminating (or coarse graining) scales smaller than ξ . From the recursion relation describing scale changes, one can in fact solve the Ising model. This is only an asymptotic solution valid for long length scales, but this is sufficient to derive the correct thermodynamic behavior.

12.1 Decimation in one dimension

In order to illustrate these ideas with a specific example, we will consider a special coarse graining transformation in the one dimensional Ising model that can be carried out exactly. Of course, this is an academic exercise as there is no phase transition in one dimension, but the renormalization procedure can be followed in detail. We first introduce a coupling constant $K = J/k_B T$ and write the partition function as,

$$Z = \sum_{\{S_i\}} e^{\sum_i K S_i S_{i+1}}.$$

We first rewrite $e^{K S_i S_{i+1}} = \cosh K (1 + u S_i S_{i+1})$ with $u = \tanh K$. This relationship is easy to verify by explicitly substituting the values corresponding to all four combinations of spin values. We rewrite the partition function as

$$Z = \sum_{\{S_i\}} \prod_i \cosh K (1 + u S_i S_{i+1}).$$

We now introduce a specific renormalization transformation called “decimation”: in the sum leading to the partition function, all spins on odd sites $i = 2n + 1, n = 0, 1, \dots$ are singled out, and the sum carried out over them. This leads a sum that will depend only on the even numbered spins:

$$Z = \sum_{\{S_{2n}\}} \sum_{\{S_{2n+1}\}} \prod_i \cosh K (1 + u S_i S_{i+1}) = \sum_{\{S_{2n}\}} \prod_{i=2n} \sum_{\{S_{2n+1}\}} e^{K(S_{2n} S_{2n+1} + S_{2n+1} S_{2n+2})}.$$

The sum over, say S_3 can be done because both S_2 and S_4 are held fixed. That is, for any spin configuration of given S_2 and S_4 we can sum over all possible states of S_3 . The same can be done for all the odd spins when holding the even spins constant. Explicitly,

$$\begin{aligned} \sum_{\{S_{2n+1}\}} e^{K(S_{2n}S_{2n+1}+S_{2n+1}S_{2n+2})} &= \cosh^2 K \sum_{\{S_{2n+1}\}} (1 + uS_{2n}S_{2n+1})(1 + uS_{2n+1}S_{2n+2}) = \\ &= 2 \cosh^2 K (1 + u^2 S_{2n}S_{2n+2}). \end{aligned}$$

Substituting back into the partition function, one has now,

$$Z = \sum_{\{S_{2n}\}} \sum_{\{S_{2n+1}\}} \prod_{2n} 2 \cosh^2 K (1 + u^2 S_{2n}S_{2n+2}).$$

This is the partition function of the decimated system (odd spins summed over). Given the partial summation carried over, Z no longer includes any dependence on fluctuations at length scale of the order of the lattice spacing: the sums over them have already been computed. The remarkable observation is that this partition function is the same as the partition function of the original system except for a multiplicative constant $\cosh K \rightarrow 2 \cosh^2 K$ and replacing u by u^2 . In fact, in order to indicate this transformation of u after one iteration of our transformation, one writes,

$$u^{(1)} = u^2.$$

This is called a recursion relation as it relates how the interaction constant of the decimated spins is related to the original lattice constants. The remaining even spins interact with a coupling constant $u^{(1)}$ which already contains the effect of all the possible configurations of the intervening spins that have been eliminated. You can think of this as an effective coupling constant that includes the effects of the odd spins.

In the original variables, the recursion relation is $\tanh K^{(1)} = (\tanh K)^2$. Since the inverse of \tanh is known, we can write,

$$K^{(1)} = \frac{1}{2} \ln \cosh(2K).$$

This transformation can be iterated over and over again until carrying out the entire sum leading to the partition function. At each level of iteration, the effective length scale of the fluctuations in halved as only the remaining, effective, spins interact with an effective coupling constant $K^{(i)}$. We write this iteration as

$$K^{(i+1)} = R(K^{(i)}) \quad R(x) = \frac{1}{2} \ln \cosh(2x). \quad (12.1)$$

Note that under iteration,

$$\lim_{n \rightarrow \infty} K^{(n)} = 0,$$

unless $K^{(0)} = \infty$. Under iteration, the partition function is transformed into an Ising model of effective spins of ever decreasing coupling constant $K^{(n)}$. Since $K = J/k_B T$, this can be interpreted as that under iteration the remaining effective spins appear to be at higher and higher temperature. In short, the partition function is that of a disordered, infinite temperature Ising model. This is of course the only stable phase for the Ising model. The other solution of this recursion is the particular case in which the original system of spins has $K^{(0)} = \infty$, that is, the system is at $T = 0$. This is the perfectly ordered solution, which we know is only stable at $T = 0$.

With this simple transformation, we have already determined that the partition function of the Ising model only has two possibilities: one corresponding to the infinite temperature Ising model for any arbitrary starting Ising model (of fixed temperature), and the zero temperature Ising model. For later reference, the solutions of the recursion relation are called fixed points. A fixed point is one in which the system remains invariant under the transformation:

$$K^* = R(K^*).$$

In the case given, the high temperature fixed point is said to be stable trajectories with any starting K flow into it by the action of the transformation R , whereas the low temperature fixed point it said to be unstable. Only a trajectory under iteration that starts at the fixed point remains on it. Any small perturbation will cause the trajectory to flow toward the stable fixed point.

12.2 Block scaling renormalization group

Decimation can only be done exactly in one dimension. We turn here to a different transformation: block scaling. We write the partition function as

$$Z(N, K, h) = \sum_{\{S_i\}} e^{K \sum_{\langle ij \rangle} S_i S_j + h \sum_i S_i}$$

where $K = J/k_B T$ and $h = H/k_B T$. Now, let us reorganize all these sum as follows:

1. Divide the lattice into blocks of size b^b (we'll let $b = 2$ in the cartoon below, Fig. 12.3).
2. Each block will have an internal sum done on all the spins it contains leading to an effective block spin $S_i^{(1)} = \pm 1$. There are many different block transformations used, but for the purposes of the discussion we can think of a majority rule transformation defined as

$$S_i^{(1)} = \begin{cases} +1, & \sum_{i \in b^d} S_i > 0 \\ -1, & \sum_{i \in b^d} S_i < 0 \\ \pm 1, & \sum_{i \in b^d} S_i = 0 \\ \text{randomly} \end{cases}$$

Clearly, if N is the number of spins in the original lattice, only N/b^d spins will remain 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) recursion relation for ξ is

$$\xi^{(m+1)} = \xi^{(m)} / b$$

If the RG transformation has a fixed point, then

$$\xi^* = \xi^* / b$$

The equation defining the fixed point has only two possible solutions

$$\xi^* = \begin{cases} 0, & \text{Trivial fixed point : } T = 0, \infty \\ \infty, & \text{Nontrivial fixed point : } T = T_c \end{cases}$$

Note that the critical fixed point (the fixed point at T_c) is unstable, while the $T = 0$ and $T = \infty$

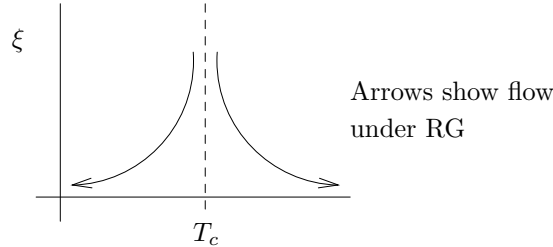


Figure 12.1:

fixed points are stable. If one does not start at a state in which the correlation length is exactly infinite, the flow defined by the blocking transformation is such that it will lead to the fixed point of $\xi = 0$. This is shown in Fig. 12.1.

As was the case with the decimation transformation, the coarse-graining eliminates information on small scales in a controlled way. Consider Fig. 12.2, and the blocking transformation shown in Fig. 12.3. The states in the new renormalized system have $S_i^{(m+1)} = \pm 1$, on sites $i = 1, 2, \dots, N^{(m+1)}$. Of course, $N^{(m+1)} = N^{(m)} / b^d$. In the figure, $N^{(m+1)} = 16$, $b^d = 2^2$, and $N^{(m)} = 64$.

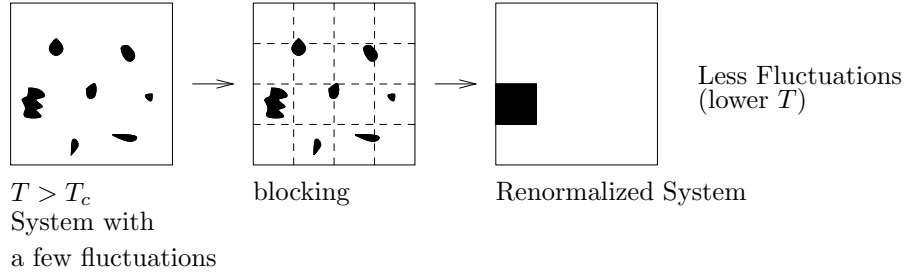


Figure 12.2:

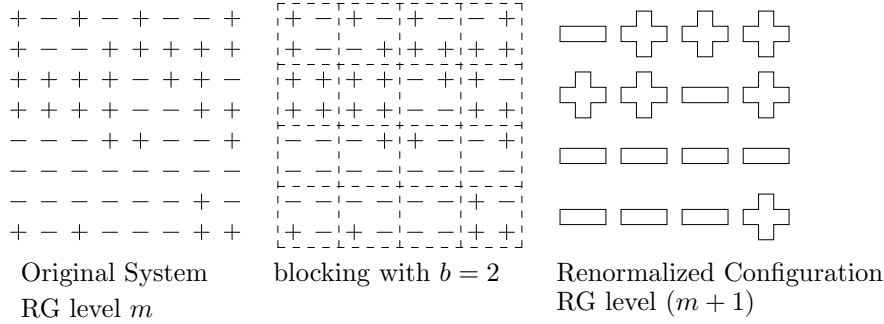


Figure 12.3:

It is simple to define and label all the states in the sum for Z as we renormalize, but how do the spins interact after renormalization? In other words, how do we determine the coupling constants in the renormalized lattice as a function of the original coupling constant. It is clear that, if the original system only has short ranged interactions, the renormalized system will as well.

At this point, historically, Kadanoff reasoned that if $E^{(m)}(\{S_i^{(m)}\})$ was that of an Ising model, then $E^{(m+1)}(\{S_i^{(m+1)}\})$ was as well. The coupling constants $K^{(m)}$ and $h^{(m)}$ would depend on the level of transformation (m), but not the functional form of the energy E . This is not correct as we will see shortly, but the idea and approach are dead on target. If this assumption were correct, the partition function after (m) iterations of the transformation would be simply,

$$Z^{(m)}(N^{(m)}, K^{(m)}, h^{(m)}) = \sum_{\{S_i^{(m)}\}} e^{K^{(m)} \sum_{\langle ij \rangle} S_i^{(m)} S_j^{(m)} + h^{(m)} \sum_i S_i^{(m)}}$$

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 since we are only doing partial sums to compute the partition function, it must be the case that $Z^{(m)} = Z^{(m+1)}$. Therefore,

$$N^{(m)} f^{(m)}(h^{(m)}, t^{(m)}) = N^{(m+1)} f^{(m+1)}(h^{(m+1)}, t^{(m+1)})$$

or,

$$f^{(m)}(h^{(m)}, t^{(m)}) = b^{-d} f^{(m+1)}(h^{(m+1)}, t^{(m+1)}) \quad (12.2)$$

This is the recursion relation for the free energy under the assumption that the original lattice and all the renormalized lattices have an interaction of the Ising type.

Since h and t change under RG and the transformation only depends on the parameter b , we assume

$$t^{(m+1)} = b^A t^{(m)} \quad h^{(m+1)} = b^B t^{(m)} \quad (12.3)$$

Hence,

$$f^{(m)}(h, t) = b^{-d} f^{(m+1)}(hb^B, tb^A)$$

Now Kadanoff assumed (incorrectly) that $f^{(m)}$ and $f^{(m+1)}$ are both simply the Ising model, and hence the same functionally. Hence

$$f(h, t) = b^{-d} f(hb^B, tb^A) \quad (12.4)$$

Hence, the free energy is a generalized homogeneous function of its arguments. This is the same result assumed by Widom. We do not have a derivation for the exponents A and B , and hence the possibility remains open for the existence of nontrivial exponents.

The recursion relation quickly leads to a scaling form for the free energy. Since b is arbitrary, choose $tb^A = 1$. Then

$$f(h, t) = t^{d/A} f(ht^{-B/A}, 1) = t^{d/A} \tilde{f}(ht^{-B/A})$$

which is of the same form as the scaling relation previously obtained for F .

It is now known that, in general, configuration blocking of the Ising model does not lead to Ising-type interactions among the block spins. Consequently, the arguments above do not apply, as at each iteration the functional form of the energy as a function of the block spins is different. In order to see why, let us consider the interactions among spins before and after application of the RG transformation. For simplicity, let $H = 0$. The original system has interactions spanning only

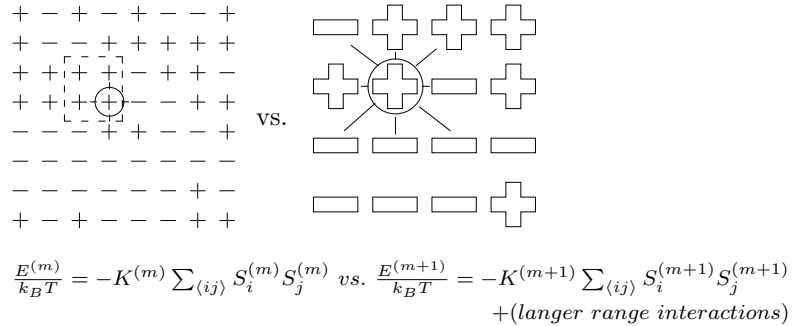


Figure 12.4:

its nearest neighbors as shown schematically in Fig. 12.4. The blocked spin treats four spins as interacting together, with identical strength, with other blocked spins. This interaction is shown in Fig. 12.5. Note that the direct part of the nearest-neighbor interaction to the right in the

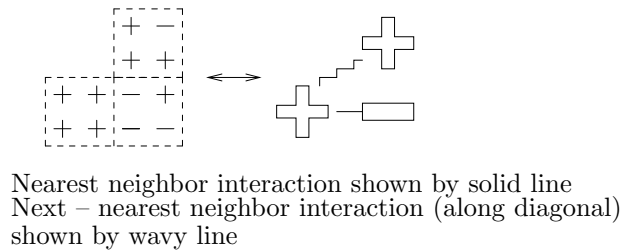


Figure 12.5:

renormalized system has 4 spins interacting equivalently with 4 other spins. There are 16 different configurations of the original system which are replaced by one global, effective, interaction of the block spin. Those sixteen cases have all $K^{(m+1)}$ contributions treated equally in block spin near

neighbor interaction despite the fact that,

$$\begin{cases} 2 \text{ interactions separated by 1 lattice constant} \\ 6 \text{ interactions separated by 2 lattice constants} \\ 8 \text{ interactions separated by 3 lattice constants} \\ 2 \text{ interactions separated by 4 lattice constants} \end{cases}$$

Let's compare this to next nearest neighbor interactions along the diagonals. There are again 16

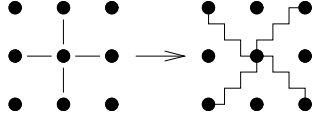


Figure 12.6: Nearest neighbor interaction vs. next nearest neighbor interaction

configurations of the original system that are treated equally in block spin representation:

$$\begin{cases} 1 \text{ interactions separated by 2 lattice constants} \\ 4 \text{ interactions separated by 3 lattice constants} \\ 6 \text{ interactions separated by 4 lattice constants} \\ 4 \text{ interactions separated by 5 lattice constants} \\ 1 \text{ interactions separated by 6 lattice constants} \end{cases}$$

Let us define the block spin next nearest neighbor coupling constant $K_{nn}^{(m+1)}$. It includes all the cases above with a single effective interaction. It is clear from looking at the interactions in the original system that underlie the block spin interactions that $K^{(m+1)}$ (nearest neighbor) and $K_{nn}^{(m+1)}$ (next nearest neighbor) are in fact quite similar, although it looks like $K_{nn}^{(m+1)} < K^{(m+1)}$. However, Kadanoff's assumption implies that $K_{nn}^{(m+1)} = 0$ (the renormalized spins have an Ising type interaction). This assumption turns out to be incorrect.

It is nowadays possible to numerically track what happens to the Ising model interactions under the block spin renormalization group. It turns out that longer range interactions between effective spins are generated by the block transformation, much as we have described above, as well as three and higher spin interactions (terms like $K_3 \sum S_i S_j S_k$ or higher). Although the interactions remain short ranged among the block spins, it turns out they become quite complicated.

The appearance of complex interaction terms under renormalization and their corresponding coupling constants can be most usefully illustrated as a flow in a space of coupling constants. For $H = 0$ and the original Ising model, the phase diagram is shown in Fig. 12.7. For a generalized Ising

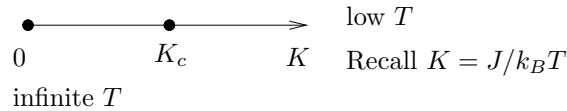


Figure 12.7: Usual Ising model $H = 0$

model with only next nearest neighbor interactions ($J = 0$) one gets the same phase diagram for $H = 0$ (Fig. 12.8). Imagine now what happens to the phase diagram for a generalized Ising 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 $H = 0$. The phase diagram is sketched in Fig. 12.9. These two phase diagrams illustrate the

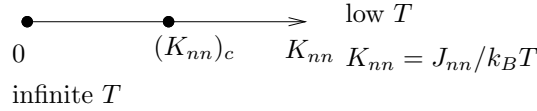


Figure 12.8: Ising model, no nearest neighbor interaction

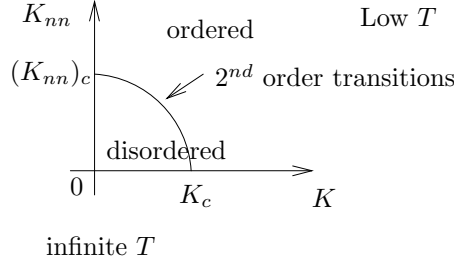


Figure 12.9:

complexity that can be expected under renormalization. If we start with the Ising model, the phase diagram is given in Fig. 12.7. Recall the existence of one unstable fixed point at T_c , and two trivial and stable fixed points. Under one iteration of the renormalization transformation, the energy of the block spins will contain next nearest neighbor interactions, but also next near neighbor, and possibly higher. In that case, the phase diagram of the renormalized system could be that of Fig. 12.9 (if only next nearest neighbor terms were generated), with a more complicated fixed point structure.

It is customary then to introduce a many dimensional space of coupling constants, and to examine the flow under renormalization in this space. Figure 12.10 illustrates possible flows on a two dimensional projection (for simplicity). Three fixed points are schematically indicated: $(K^*, K_{nn}^*) = (0, 0)$ (the infinite temperature, disordered fixed point), $(K^*, K_{nn}^*) = (\infty, \infty)$ (the zero temperature fixed point), and (K^*, K_{nn}^*) , a fixed point which is stable along the direction tangent to the line, but unstable in the normal direction. This is the fixed point that corresponds to the critical point of the renormalized model, and hence of the system. It is from the fixed point structure, and from the flows in coupling constant space that one can compute scaling functions and associated exponents.

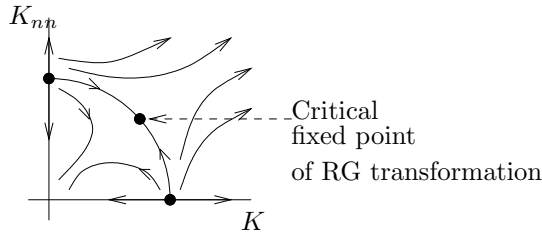


Figure 12.10:

One should point out that the effective interactions developed by the RG depend on the specific RG transformation. For example, blocking with $b = 3$ would give a different sequence of interactions, and a different fixed point result than $b = 2$. But, it turns out, that since we are computing the same partition function by doing the partial sums in different ways, the exponents obtained must be all the same.

The way to formalize our discussion so far was developed by Wilson. The RG transformation takes us from one set of coupling constants $\vec{K}^{(m)}$ to another $\vec{K}^{(m+1)}$ via a length scale change of b entering a recursion relation: $\vec{K}^{(m+1)} = R(\vec{K}^{(m)})$. We have introduced a vector notation to accommodate the many axes used in coupling constant space, $\vec{K} = (K, K_{nn}, K_3, \dots)$. By iteration, one may reach a fixed point $\vec{K}^* = R(\vec{K}^*)$. Let us now imagine that after a sufficient number of iterations, we are close enough to the fixed point \vec{K}^* that we can linearize the transformation R

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 quite complex. Usually it is done by numerical methods, or by the $\epsilon = 4 - d$ expansion described later.

Linearization of the recursion relation near the fixed point can be obtained from (we use Greek subindices to denote components of \vec{K})

$$K_\alpha^{(m+1)} - K_\alpha^* = R_\alpha(\vec{K}^{(m)}) - R_\alpha(\vec{K}^*)$$

so that by linearization,

$$\delta K_\alpha^{(m+1)} = \sum_\beta M_{\alpha\beta} \delta K_\beta^{(m)}$$

with

$$M_{\alpha\beta} = (\partial R_\alpha / \partial K_\beta)^*$$

and $\delta K_\alpha = K_\alpha - K_\alpha^*$

12.2.1 Monte Carlo Renormalization Group

Before we continue our analysis of the transformation, it is useful to consider a purely numerical method to obtain the matrix $M_{\alpha\beta}$. This method was developed by Swendsen (Phys. Rev. Lett. **14**, 859 (1979)). First, write the energy formally as

$$E = \sum_\alpha K_\alpha S_\alpha$$

where K_α are the components of \vec{K} , and hence $S_1 = \sum_{\langle i,j \rangle} S_i S_j$, and higher order interactions. By using the chain rule,

$$\frac{\partial \langle S_\gamma^{(m+1)} \rangle}{\partial K_\beta^{(m)}} = \sum_\alpha \frac{\partial K_\alpha^{(m+1)}}{\partial K_\beta^{(m)}} \frac{\partial \langle S_\gamma^{(m+1)} \rangle}{\partial K_\alpha^{(m+1)}}.$$

This relation involves the matrix elements that we need to compute. The other derivatives are in fact correlation functions:

$$\frac{\partial \langle S_\gamma^{(m+1)} \rangle}{\partial K_\alpha^{(m+1)}} = \frac{\partial}{\partial K_\alpha^{(m+1)}} \left[\frac{1}{Z} \sum S_\gamma^{(m+1)} e^{\sum_\alpha K_\alpha S_\alpha} \right].$$

Explicitly taking the derivatives leads to,

$$\begin{aligned} \frac{\partial \langle S_\gamma^{(m+1)} \rangle}{\partial K_\alpha^{(m+1)}} &= -\frac{1}{Z} \frac{\partial Z}{\partial K_\alpha^{(m+1)}} \langle S_\gamma^{(m+1)} \rangle + \frac{1}{Z} \sum S_\gamma^{(m+1)} S_\alpha^{(m+1)} e^{\sum_\alpha K_\alpha S_\alpha} = \\ &= -\langle S_\gamma^{(m+1)} \rangle \langle S_\alpha^{(m+1)} \rangle + \langle S_\gamma^{(m+1)} S_\alpha^{(m+1)} \rangle. \end{aligned}$$

The right hand side is an equilibrium correlation function that can be calculated by simulation (the details of the simulation method are described in the chapter on Monte Carlo methods). In essence, the Ising model is simulated, the the block transformation is carried out numerically so that the correlation functions for the block configurations can also be computed.

Analogously, one can also find that,

$$\frac{\partial \langle S_\gamma^{(m+1)} \rangle}{\partial K_\beta^{(m)}} = -\langle S_\gamma^{(m+1)} \rangle \langle S_\beta^{(m)} \rangle + \langle S_\gamma^{(m+1)} S_\beta^{(m)} \rangle.$$

Therefore, the simulation plus a judicious truncation in the number of components of \vec{K} to be considered (neglecting very long range interactions that will remain small) leads to a numerical estimate of the matrix elements $M_{\alpha\beta}$ that are required.

12.2.2 Scaling form, and critical exponents

Assume now that we have obtained a fixed point of a given RG transformation, and computed the matrix $M_{\alpha\beta}$. We now diagonalize M , and call λ_α the eigenvalues. Assume for simplicity of exposition that $\delta\vec{K}$ above are directly the eigenvectors (that is, that the original matrix is already diagonal. Otherwise, the discussion that follows applies to the actual eigenvectors). Then,

$$M_{\alpha\beta} = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \dots & \\ 0 & & & \dots \end{bmatrix}$$

so that

$$\delta K_\alpha^{(m+1)} = \lambda_\alpha \delta K_\alpha^{(m)}$$

The functional dependence of λ_α is easy to find. 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:

$$M(b)M(b') = M(bb')$$

This is precisely one of the requirements for the transformation to be a “group” in the mathematical sense. Or,

$$\lambda(b)\lambda(b') = \lambda(bb')$$

The solution to this equations is $\lambda = b^y$ since $(bb')^y = b^y b'^y$. Therefore the renormalization group transformation leads to

$$\delta K_\alpha^{(m+1)} = b^{y_\alpha} \delta K_\alpha^{(m)} \quad (12.5)$$

near the fixed point.

If $y_\alpha > 0$, δK_α is called relevant, i.e., its magnitude increases under renormalization. If $y_\alpha = 0$, δK_α is called marginal as it neither increases nor decreases under renormalization. If $y_\alpha < 0$, then δK_α is called irrelevant, as it becomes smaller under renormalization. Under infinite interactions, this coupling constant would disappear from the energy for the block spins.

The critical part of the argument is that we require that the transformation preserves the partition function. This partition function is after all what we wish to compute by the partial summation involved in the renormalization transformation. This requirement and the recursion relation (12.5) imply for the free energy

$$f^{(m)}(\delta\vec{K}^{(m)}) = b^{-d} f^{(m+1)}(\delta\vec{K}^{(m+1)}),$$

or at the fixed point,

$$f^*(\delta K_1, \delta K_2, \dots) = b^{-d} f^*(b^{y_1} \delta K_1, b^{y_2} \delta K_2, \dots)$$

which proves that the free energy is a generalized homogeneous function of its arguments, and provides explicit values of the scaling exponents from the eigenvalues of the matrix M .

In practice, the problem is to find the fixed point. A complete solution was provided with the so called $\epsilon = 4 - d$ expansion. This is discussed later in this chapter. From our discussion above about the Ginzburg criterion, mean field theory works for $d > 4$ (it is marginal with logarithmic corrections in $d = 4$). Therefore it is expected there has to be a very easy to find and easy to analyze fixed point in $d \geq 4$. The assumption of this theory is that the fixed point analysis in $d = 4$ can be analytically continued to $d < 4$ in a controlled $\epsilon = 4 - d$ expansion.

We finally list some results that extend our calculation to the general case in which the matrix M is not diagonal, which is normally the case. We assume that M 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

$$\delta \hat{K}_\gamma^{(m+1)} = \lambda^{(\alpha)} \sum_\beta e_\beta^{(\gamma)} \delta K_\beta^{(m)}$$

so,

$$\delta \hat{K}_\gamma^{(m+1)} = \lambda^{(\alpha)} \delta \hat{K}_\gamma^{(m)}$$

This is what we had (for δK 's rather than $\delta \hat{K}$'s) a few pages back. Everything follows the same from there.

12.3 ϕ^4 model: dimensional analysis

We start from

$$F[\psi] = \int d^d \vec{x} \left\{ \frac{K}{2} (\nabla \psi)^2 + \frac{r}{2} \psi^2 + \frac{u}{4} \psi^4 \right\}$$

near T_c , where K and u are constants and $r \propto T - T_c$. We introduce the following change of variables $F' = \beta F$, $\psi' = \sqrt{\beta K} \psi$, $r' = r/K$, $u' = u/\beta K^2$. After some algebra, removing the primes, and renaming $r' = r_0$ and $u' = u_0$ we have,

$$F = \int d^d \vec{x} \left[\frac{1}{2} |\nabla \psi|^2 + \frac{r_0}{2} \psi^2 + \frac{u_0}{4} \psi^4 \right].$$

Let us examine the dimensions of the various terms (denoted by $[.]$. By definition, F is dimensionless $[F] = 1$. Therefore

$$\left[\int d^d \vec{x} |\nabla \psi|^2 \right] = L^d L^{-2} [\psi]^2 = 1,$$

or $[\psi] = L^{1-d/2}$. Furthermore, $[r_0] = L^{-2}$, and $[u_0] = L^{d-4}$. These are called the engineering dimensions.

Critical phenomena involve different scalings with length as we have seen, and therefore the actual dimensions of variables near the critical point do not agree with the engineering dimensions. This is one of the manifestations of the complexity of critical phenomena. For example,

$$[\psi^2] = \frac{1}{L^{d-2}}$$

but we know that T_c

$$\langle \psi(0) \psi(x) \rangle = \frac{1}{x^{d-2+\eta}}$$

where $\eta \neq 0$ as the engineering dimension would lead one to believe. 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

$$\xi = (T - T_c)^{-\nu}$$

with $\nu \neq 1/2$. The results of our scaling theory give as scaling dimensions

$$[\psi]_{\text{scaling}} = \frac{1}{L^{(d-2+\eta)/2}} = \frac{1}{L^{\frac{d}{2}-1}} \frac{1}{L^{\eta/2}}$$

and

$$[r]_{\text{scaling}} = \frac{1}{L^{1/\nu}} = \frac{1}{L^2} \frac{1}{L^{\frac{1}{\nu}-2}}$$

One can interpret these anomalous dimensions as following from a dependence on another length in the problem: the lattice constant $a \approx 5\text{\AA}$, and rewrite

$$\langle \psi^2 \rangle = \frac{1}{x^{d-2+\eta}} a^\eta$$

so that the dimensions of ψ are those given by dimensional analysis (engineering dimensions). The exponent $\eta/2$ is called ψ 's anomalous dimension, while $1/\eta - 2$ is the anomalous dimension of r .

It is useful at this point to rescale variables again by using $\mathcal{L} = r_0^{-1/2}$ as the unit of length. We define $\psi' = \psi/\mathcal{L}^{1-d/2}$, $x' = x/\mathcal{L}$ and $u' = u_0/\mathcal{L}^{d-4}$. Simple algebraic substitutions lead to (removing primes, except in u),

$$F = \int d^d \vec{x} \left[\frac{1}{2} |\nabla \psi|^2 + \frac{1}{2} \psi^2 + \frac{u'}{4} \psi^4 \right],$$

Therefore the partition function of the ϕ^4 model will be a function of the single parameter u' . Since it is the coefficient of the quartic term in the energy, the first approximation is to treat it as a perturbation. This leads to the Ornstein-Zernicke theory, and hence to mean field exponents. However, since $r_0 \propto (T - T_c)$ near the critical temperature, one has $u' \sim t^{(d-4)/2}$. This simple analysis shows that for $d > 4$, $u' \rightarrow 0$ as T_c is approached, and hence the quartic term in the energy can be truly neglected. Therefore the Ornstein-Zernicke theory can be expected to hold. On the other hand, for $d < 4$, the nonlinear term diverges and this simplification is incorrect. This is in fact a restatement of the Ginzburg criterion.

This analysis if the starting point of the renormalization group treatment of the ϕ^4 model. We expect that a ‘‘Gaussian fixed point’’ will exist given by $r^* = 0, u^* = 0$, and $h^* = 0$, which will be correct for $d > 4$. The success of the ϵ expansion is to show that there exists a new fixed point for $d < 4$ so that the Gaussian fixed point becomes unstable in this range. This new fixed point does provide with the exponents observed experimentally.

12.4 Momentum shell RG. Gaussian fixed point

Before proceeding to a more complete calculation, we illustrate the technique used (momentum shell RG) for the case of the Gaussian fixed point. Consider,

$$F = \int d^d \vec{x} \left[\frac{1}{2} |\nabla \psi|^2 + \frac{r_0}{2} \psi^2 \right] = \int_0^\Lambda \frac{d^d \vec{k}}{(2\pi)^d} \frac{1}{2} (r_0 + k^2) |\hat{\psi}_{\vec{k}}|^2.$$

We now introduce the following transformation: Replace the integral \int_0^Λ into a long wavelength component $\int_0^{\Lambda/l}$ and a short wavelength component $\int_{\Lambda/l}^\Lambda$. With the respective integrals we define a long wavelength field and a perturbation,

$$\psi(\vec{x}) = \bar{\psi}(x) + \delta\psi(x) = \int_0^{\Lambda/l} \frac{d^d \vec{k}}{(2\pi)^d} \hat{\psi}_{\vec{k}} e^{i\vec{k} \cdot \vec{x}} + \int_{\Lambda/l}^\Lambda \frac{d^d \vec{k}}{(2\pi)^d} \hat{\psi}_{\vec{k}} e^{i\vec{k} \cdot \vec{x}}.$$

In the example of block spins, the field $\bar{\psi}$ would correspond to block spins, whereas $\delta\psi$ would describe fluctuations within blocks. Since at the Gaussian level all modes decouple, we can write the partition function as

$$Z = \int \prod_{0 < k < \Lambda/l} d\bar{\psi}_{\vec{k}} \prod_{\Lambda/l < k < \Lambda} d\delta\psi_{\vec{k}} e^{\int_0^{\Lambda/l} \frac{d^d \vec{k}}{(2\pi)^d} \frac{1}{2} (r_0 + k^2) |\bar{\psi}_{\vec{k}}|^2} e^{\int_{\Lambda/l}^{\Lambda} \frac{d^d \vec{k}}{(2\pi)^d} \frac{1}{2} (r_0 + k^2) |\delta\psi_{\vec{k}}|^2}.$$

The integral over $\delta\psi$ is Gaussian and can be done. This is the “renormalization” step of the transformation that removes a subset of the degrees of freedom.

The second step is rescaling the variables so that the renormalized system is of the same size as the original system. We introduce,

$$k_l = lk \quad \bar{\psi}_l(k_l) = \frac{1}{z} \bar{\psi}_k.$$

Calling $Z_{\delta\psi}$ the partial partition function arising from the integration over $\delta\psi$, we find,

$$Z = Z_{\delta\psi} \int \prod_k d\psi_l(k) e^{\int_0^{\Lambda} l^{-d} \frac{d^d k_l}{(2\pi)^d} \left(r_0 + \frac{k_l^2}{l^2} \right) |\psi_l(k)|^2 z^2}.$$

If we now require that this partition function reduce to the same form as the original partition function, we require $l^{-d} z^2 / l^2 = 1$, or a factor $z = l^{1+d/2}$. Then

$$Z = Z_{\delta\psi} \int \prod_k d\psi_l(k) e^{\int_0^{\Lambda} \frac{d^d k_l}{(2\pi)^d} (l^{-d} r_0 z^2 + k_l^2) |\psi_l(k)|^2 z^2},$$

which is the original partition function if we define

$$r_l = l^{-d} z^2 r_0 \quad r_l = l^2 r_0.$$

The partial summation and rescaling leaves the partition function invariant if we allow a scale dependent coefficient r . In effect, this is the recursion relation for this transformation.

We immediately find two Gaussian fixed points: $r^* = 0$ (the critical points) and $r^* \infty$. Since the equation is already linearized around the fixed point $r^* = 0$, we find that the critical exponent is $y_r = 2$.

Therefore we would write for the free energy,

$$f^*(\delta r) = l^{-d} f^*(l^2 \delta r).$$

By definition or r , $\delta r \sim t$, the reduced temperature away from the critical point, and since l is a length we can write,

$$f^*(\delta r) = \xi^{-d} f^*(\xi^2 t).$$

From this scaling form we conclude that

$$\xi^2 \sim t^{-1} \quad \text{or} \quad \xi \sim t^{-1/2},$$

the known mean field result.

12.5 (*) ϵ expansion

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”:

$$\begin{aligned} -L \frac{du'}{dL} &= -(4-d) L^{4-d} u \\ &= -(4-d) u' \end{aligned}$$

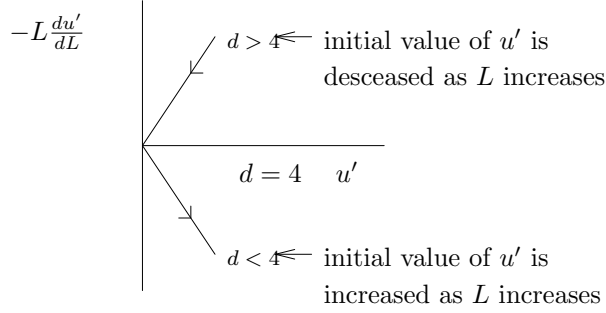


Figure 12.11:

The direction of the arrows is from

$$\begin{aligned}\frac{du'}{d \ln L} &= -(d-4)u' \quad , \quad d > 4 \\ \frac{du'}{d \ln L} &= (4-d)u' \quad , \quad d < 4\end{aligned}$$

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. Clearly

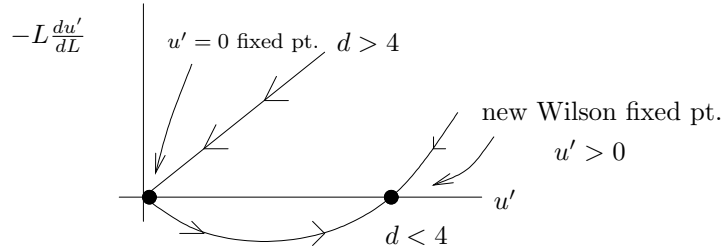


Figure 12.12:

$$-L \frac{du'}{dL} = (\text{Dim. of } u) u'$$

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 small length scale behavior, which, in Fourier space, means integrating out large wavenumber behavior. Schematically shown in fig. 12.13. Let

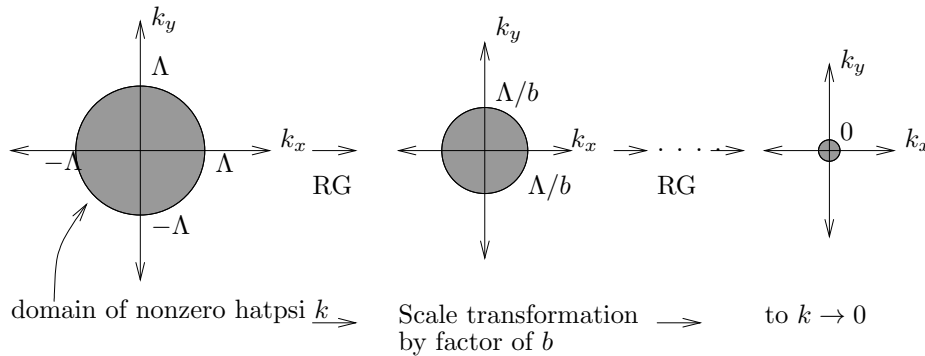


Figure 12.13:

$$\begin{array}{lll}
\psi(x) & = & \bar{\psi}(x) + \delta\psi(x) \\
\text{all degrees of freedom} & \text{degrees of freedom} & \text{degrees of freedom} \\
\text{after several} & \text{for “slow” modes on} & \text{for “fast” modes on} \\
\text{transformations} & |k| < \Lambda/b & \Lambda/b < |k| < \Lambda \\
& |k| < \Lambda & \text{to be integrated out}
\end{array}$$

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. 12.14. This is called the “momentum-shell” RG. We do this in momentum

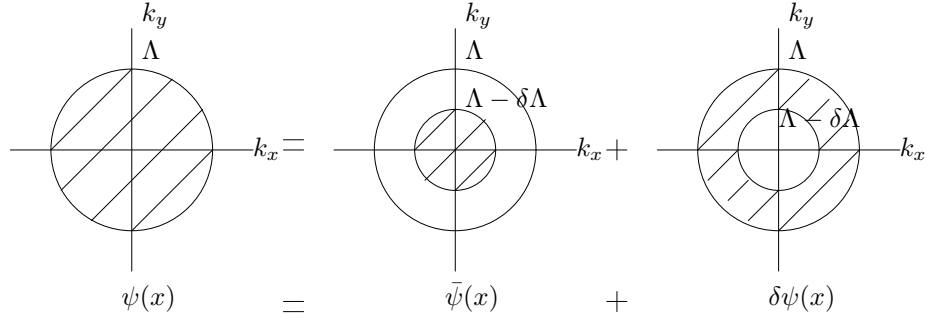


Figure 12.14:

space rather than real space because it is convenient for factors of

$$\int \underbrace{d^d x (\nabla \psi)^2}_{\text{coupling different } \psi(x)\text{'s}} = \int \underbrace{\frac{d^d k}{(2\pi)^d} k^2 |\hat{\psi}_k|^2}_{\text{Fourier modes uncoupled}}$$

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

$$\begin{aligned} \frac{1}{2} [\nabla(\bar{\psi} + \delta\psi)]^2 &= \frac{1}{2} (\nabla\bar{\psi})^2 + \frac{1}{2} (\nabla\delta\psi)^2 + \text{odd term} \\ \frac{r}{2} [\bar{\psi} + \delta\psi]^2 &= \frac{r}{2} \bar{\psi}^2 + \frac{r}{2} (\delta\psi)^2 + \text{odd term} \\ \frac{u}{4} [\bar{\psi} + \delta\psi]^4 &= \frac{u}{4} \bar{\psi}^4 \left[1 + \frac{\delta\psi}{\bar{\psi}} \right]^4 \\ &= \frac{u}{4} \bar{\psi}^4 \left[1 + 4 \frac{\delta\psi}{\bar{\psi}} + 6 \frac{(\delta\psi)^2}{\bar{\psi}^2} + 4 \frac{(\delta\psi)^3}{\bar{\psi}^3} + \frac{(\delta\psi)^4}{\bar{\psi}^4} \right] \\ &= \frac{u}{4} \bar{\psi}^4 + \frac{3u}{2} \bar{\psi}^2 (\delta\psi)^2 + \mathcal{O}(\delta\psi)^4 + \text{odd term} \end{aligned}$$

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

$$F[\bar{\psi} + \delta\psi] = \underbrace{F\{\bar{\psi}\}}_{|k| < \Lambda - \delta\Lambda} + \int d^d x \left[\frac{1}{2} (\nabla\delta\psi)^2 + \frac{1}{2} (r + 3u\bar{\psi}^2) (\delta\psi)^2 \right]$$

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 $\ln(1 + \epsilon) = \epsilon - \frac{\epsilon^2}{2} + \frac{\epsilon^3}{3}$

$$\begin{aligned} &= \frac{1}{2} \left\{ \frac{r}{\Lambda^2} + \frac{3u}{\Lambda^2} \bar{\psi}^2 - \frac{1}{2} \left(\frac{r}{\Lambda^2} + \frac{3u}{\Lambda^2} \bar{\psi}^2 \right)^2 + \dots \right\} \\ &= \frac{1}{2} \left\{ \cancel{\frac{r}{\Lambda^2}}^{\text{const.}} + \frac{3u}{\Lambda^2} \bar{\psi}^2 - \frac{1}{2} \left(\cancel{\frac{r}{\Lambda^2}}^{\text{const.}} \right)^2 - \frac{3ur}{\Lambda^2} \bar{\psi}^2 - \frac{9u^2}{2\Lambda^4} \bar{\psi}^4 + \dots \right\} \\ &= \frac{1}{2} \left(\frac{3u}{\Lambda^2} - \frac{3ur}{\Lambda^4} \right) \bar{\psi}^2 + \frac{1}{4} \left(\frac{-9u^2}{\Lambda^4} \right) \bar{\psi}^4 \end{aligned}$$

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

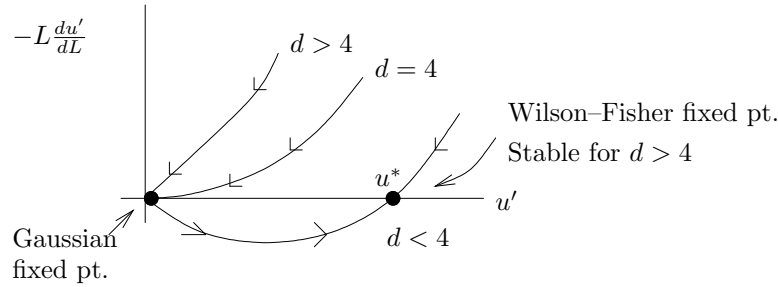


Figure 12.15:

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

$$r^* = -\frac{\epsilon}{6}$$

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,

$$-L \frac{d\delta r}{dL} = -2(1 - \frac{\epsilon}{6})\delta r \quad \text{Wilson-Fisher fixed pt. (stable } d < 4)$$

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...

Chapter 13

Master Equation and the Monte Carlo method

The Monte Carlo method is a class of random or stochastic methods that are used to simulate a large variety of physical phenomena. It has wide applicability in Statistical Mechanics, as it affords an almost unique route to approximately compute statistical averages in complex nonlinear systems. We focus below on its most common implementation in Statistical Mechanics, and leave out applications to other fields.

We have defined a canonical average as

$$\langle A \rangle = \frac{1}{Z} \sum_{\nu} A_{\nu} e^{-\beta E_{\nu}}. \quad (13.1)$$

We mentioned earlier that it is not necessary in order to compute macroscopic averages to solve for the time dependent trajectory of the system in its phase space, and then compute temporal averages of the quantities of interest. Rather, we compute averages over known probability measures, such as that given in Eq. (13.1).

We can easily devise a scheme to approximately compute the average in Eq. (13.1) by numerical means. The scheme could be as follows: given a specific physical system and the energy as a function of its microscopic variables, choose a configuration ν completely at random, and evaluate A_{ν} and $p_{\nu} \sim e^{-\beta E_{\nu}}$. Repeat for a large number of times until a good estimate of $\langle A \rangle$ has been obtained. As you may imagine, the problem with this approach is that the sample space is huge ($\sim e^{10^{23}}$) for a macroscopic system. Furthermore, since E is extensive ($E \sim N \sim 10^{23}$), the probability of each configuration p_{ν} extremely small. Therefore the procedure just outlined would require adding an extremely large number of extremely small quantities.

Both difficulties could be alleviated if we had a way of sampling only the most likely states, those that contribute the most to the average. The Monte Carlo method provides a way to precisely do that: sample states not with uniform probability as outlined above, but rather with probability precisely equal to p_{ν} . If a sequence of N states $\{\nu'\}$ is created, which are distributed according to the canonical distribution function, then

$$\langle A \rangle = \frac{1}{N} \sum_{\nu'} A_{\nu'}$$

The Monte Carlo method which we describe next is a particular way of generating configurations that are distributed according to a predefined probability distribution function. Imagine that the system of interest is initially in a state S (we use S instead of ν in the following). This state might, for example, be all the spins $+1$ if the system in question is the Ising model. We imagine next that we generate a process in configuration space that prescribes transitions from one state to another. We define a transition probability of going from $S \rightarrow S'$, in a unit time as

$$W(S, S') \underset{\text{(Read As)}}{\equiv} W(S \leftarrow S') \quad (13.2)$$

This transition is schematically shown in Fig. 13.1. Of course, one needs to keep in mind transitions

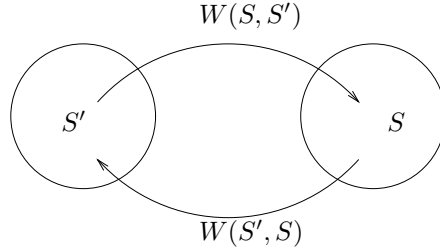


Figure 13.1:

out of S into any other state S' with probability

$$W(S', S) \underset{\text{(Read As)}}{\equiv} W(S' \leftarrow S). \quad (13.3)$$

The master equation is the equation that governs the evolution of the probability distribution $P(S)$ under the transitions just discussed. It is customary to assume that the W are transition rates per unit time, and hence that $P(S)$ changes as a function of time. However, time can be a fictitious variable simply denoting a sequence of transitions in phase space. The equation reads,

$$\frac{\partial P(S, t)}{\partial t} = \sum_{S'} [W(S, S') P(S') - W(S', S) P(S)]. \quad (13.4)$$

The transition rates in the master equation are assumed to depend only on the current state, but not on the prior trajectory. This defines a so called Markov process. Note also that the master equation is not invariant under time reversal.

We now examine restrictions on the master equation to generate states distributed according to the canonical probability distribution. First, we want $p(S)$ to be a time independent distribution, and so we require that

$$\frac{\partial P_{eq}(S)}{\partial t} = 0. \quad (13.5)$$

A sufficient condition for the right hand side of Eq. (13.4) to be zero is,

$$W(S, S') P_0(S') = W(S', S) P_0(S) \quad (13.6)$$

where P_0 is the stationary distribution function. This equation is known as the detailed balance condition. We now require that $P_0(S) = (1/Z)e^{-\beta E(S)}$. Therefore,

$$\frac{W(S, S')}{W(S', S)} = e^{-(E(S) - E(S'))/k_B T} \quad (13.7)$$

This is a surprisingly simple result for the transition rates W : it is sufficient (but not necessary) that they satisfy Eq. (13.7) for the process to asymptotically yield a stationary distribution of states $P(S)$ which equals the canonical distribution. In fact, if we are only concerned with equilibrium properties, any set of W 's which satisfy detailed balance will produce a sequence of states distributed according to the canonical distribution. All one needs to do is pick a convenient W .

The two most common choices of W 's used in numerical work are the Metropolis rule

$$W(S, S') = \begin{cases} e^{-\Delta E/k_B T}, & \Delta E > 0 \\ 1, & \Delta E \leq 0 \end{cases} \quad (13.8)$$

where $\Delta E = E(S) - E(S')$, and the Glauber rule

$$W(S, S') = \frac{1}{2} (1 - \tanh \frac{\Delta E}{2k_B T}) \quad (13.9)$$

Let us quickly verify that W for the Metropolis rule satisfies the detailed balance condition. 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 the detailed balance condition is satisfied by this transition rule. You can check yourself that the Glauber rule also satisfies detailed balance. The form of the transition rate as a function of the energy change is shown in Fig. 13.2 and 13.3.

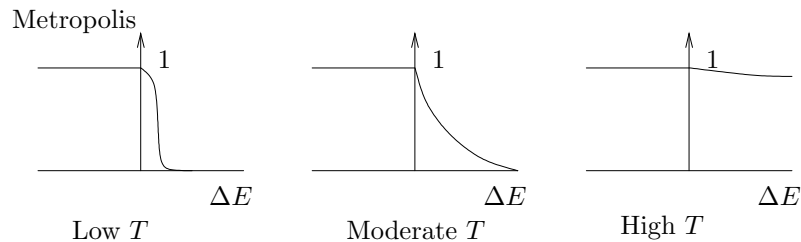


Figure 13.2:

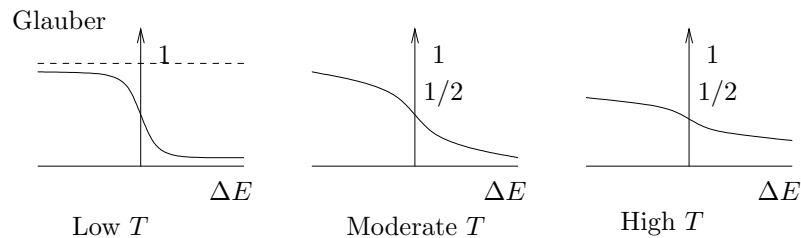


Figure 13.3:

In practice, one introduces transitions from state $\{S'\} \rightarrow \{S\}$ that take place at the rate $W(\{S\}, \{S'\})$. One usually wants the states $\{S'\}$ and $\{S\}$ to be “close” in configuration space, so that W is not too small (close in energy for the Metropolis and Glauber rules). For example, for 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. 13.4.

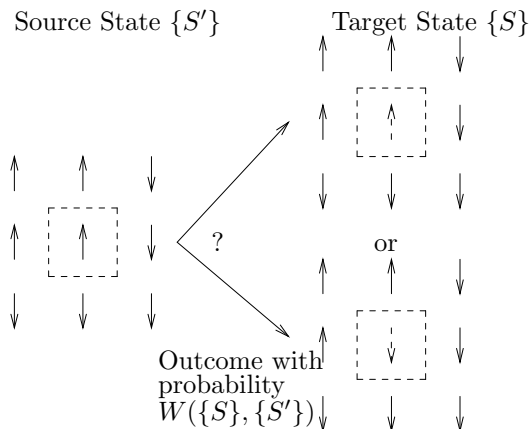


Figure 13.4:

Let us consider the explicit steps for the 2 dimensional Ising model in zero field. The energy of any configuration is

$$E_\nu = -J \sum_{\langle ij \rangle} S_i S_j \quad (13.10)$$

where $S_i = \pm 1$ are the spins on sites $i = 1, 2, \dots, N$ spins, and $\langle ij \rangle$ means only nearest neighbors are summed over. The positive constant J is the coupling interaction. If the transitions to be explored involve only one spin, there are only five cases to consider as shown in Fig. (13.5).

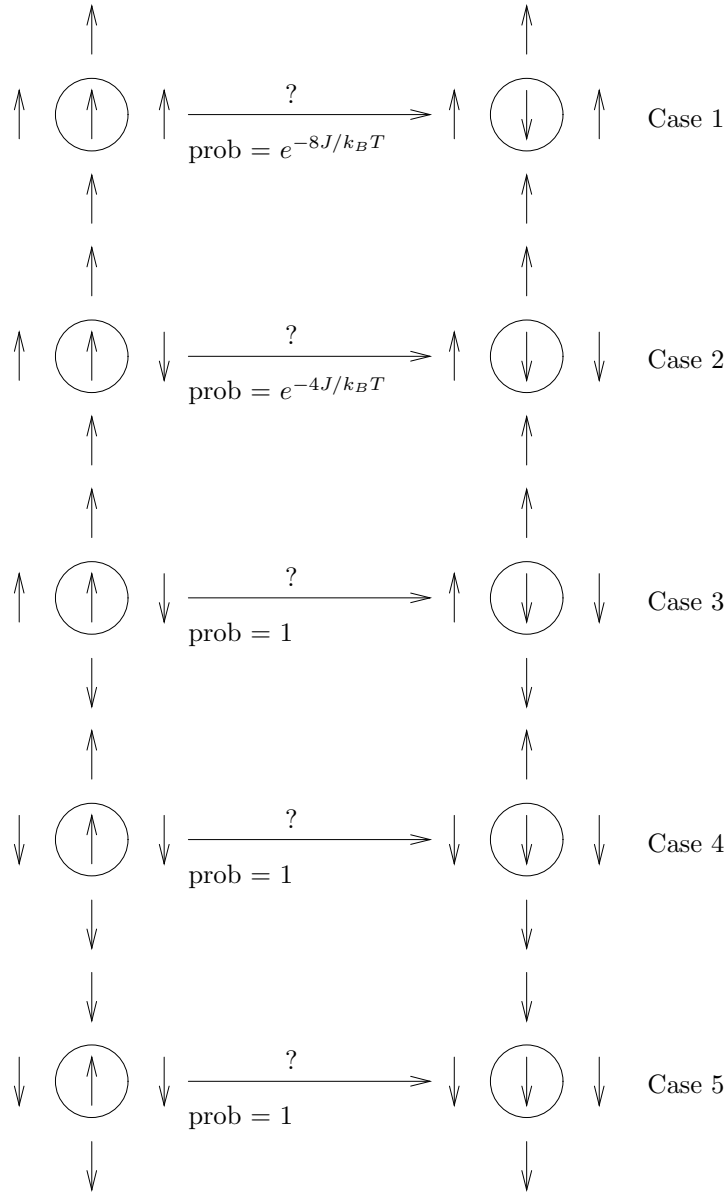


Figure 13.5:

The resulting transition probabilities are as follows (assuming we use the Metropolis rule):

For Case 1:

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

For Case 2:

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

For Case 3:

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

For Case 4:

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

For Case 5:

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

What to do next is simple if $W = 1$, the new state S is accepted with probability one. If on the other hand $W = e^{-4J/k_B T}$, say for example $W = 0.1$, the spin flips 10% of the time on average. In order to implement this the standard procedure is to compare $W = 0.1$ to a random number uniformly distributed between 0 and 1. If the random number is less than or equal to W , then you flip the spin. With this rule, the spin has exactly a probability of 0.1 of being flipped, exactly as required. Of course, this is repeated over and over again until a sufficiently large number of configurations have been generated. Because detailed balance is satisfied, the configurations given by this algorithm will be distributed according to the canonical distribution as desired.

Here is an outline of a computer code. If you set the temperature of the system to be high and initially set $S_i = 1$ for all spins, you will see that the magnetization quickly relaxes to zero (Fig. 13.7). If the system temperature is below T_c , the magnetization quickly relaxes to its equilibrium value (see Fig. 13.8). Doing this for many temperatures, for a large enough lattice, gives the phase diagram shown in Fig. 13.9. One can also compute the running average of the energy of the system, to find the curve shown in Fig. 13.10. By taking the numerical derivative, one can estimate the heat capacity per spin

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

Try it. Your results will be very close to Onsager's results even with fairly small systems $N = 10 \times 10$, or 100×100 over a thousand or so time steps, and you'll do much better than the mean field theory.

As an exercise, think about the code that you would write to simulate the discrete model of interface fluctuations:

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

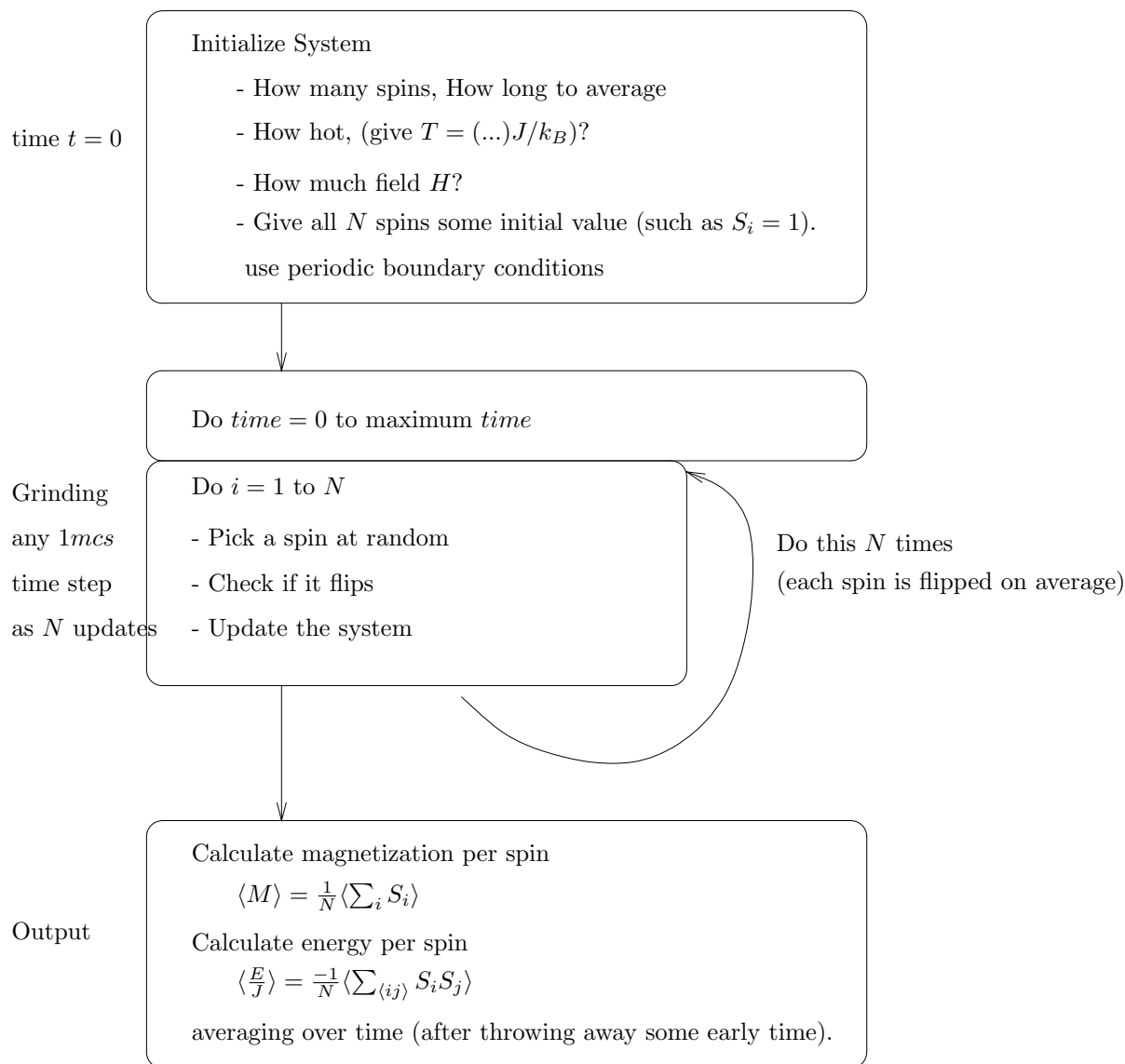
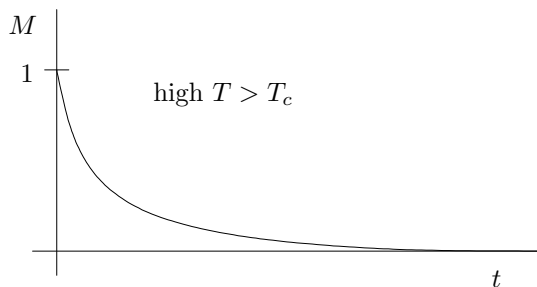


Figure 13.6:



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

Figure 13.7:

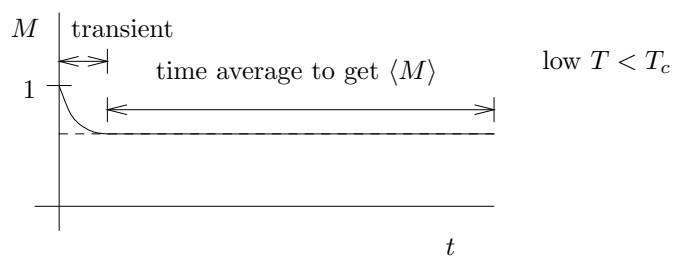


Figure 13.8:

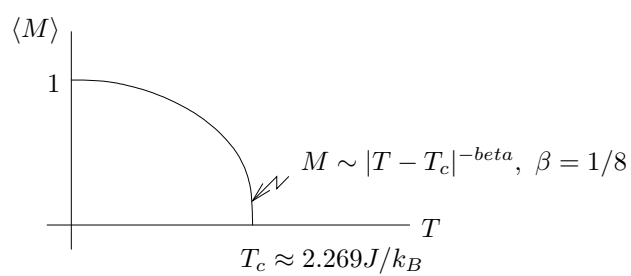


Figure 13.9:

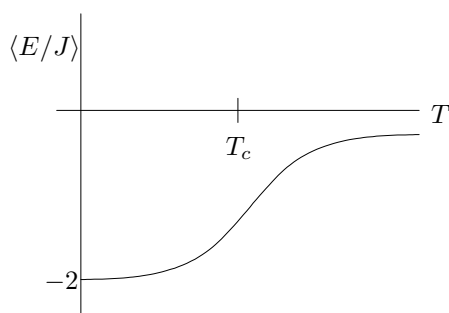


Figure 13.10:

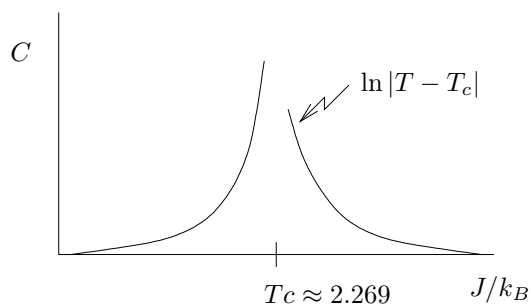


Figure 13.11:

where

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

where h_i is the interface displacement at site i .

13.1 Other ensembles

The Monte Carlo method outlines above can be generalized to other ensembles beyond the canonical distribution considered above. For example, if we have an ensemble at fixed (N, p, T) , in which both the volume and the energy of the system can fluctuate, the corresponding probability distribution is,

$$p_\nu = \frac{1}{Z} e^{-\beta(E_\nu + pV_\nu)}.$$

The condition of detailed balance for this case would be,

$$\frac{W(\nu', \nu)}{W(\nu, \nu')} = e^{-\beta(E_{\nu'} - E_\nu) - \beta p(V_{\nu'} - V_\nu)}.$$

The sequence of configurations to be generated by the algorithm would change not only the configuration of the system but also its overall volume. This is, of course, not quite physical in many cases, but nevertheless it will produce a sequence of states distributed according to the correct equilibrium probability distribution.