

Lecture Notes for the C6 Theory Option

F.H.L. Essler

The Rudolf Peierls Centre for Theoretical Physics
Oxford University, Oxford OX1 3NP, UK

March 9, 2015

PRELIMINARY VERSION (NEEDS PROOF READING).

Please report errors and typos to fab@thphys.ox.ac.uk

©2013 F.H.L. Essler

Part I

STATISTICAL MECHANICS AND PHASE TRANSITIONS

1 BRIEF REVIEW OF SOME RELEVANT QUANTITIES

Consider a classical many-particle system coupled to a heat bath at temperature T . The *partition function* is defined as

$$Z = \sum_{\text{configurations } C} e^{-\beta E(C)}, \quad \beta = \frac{1}{k_B T}. \quad (1)$$

Here the sum is over all possible configurations C , and $E(C)$ is the corresponding energy. Thermal averages of observables are given by

$$\langle \mathcal{O} \rangle_\beta = \frac{1}{Z} \sum_{\text{configurations } C} \mathcal{O}(C) e^{-\beta E(C)}, \quad (2)$$

where $\mathcal{O}(C)$ is the value of the observable \mathcal{O} in configuration C . The average energy is

$$E = \frac{1}{Z} \sum_{\text{configurations } C} E(C) e^{-\beta E(C)} = -\frac{\partial}{\partial \beta} \ln(Z). \quad (3)$$

The *free energy* is

$$F = -k_B T \ln(Z). \quad (4)$$

The *entropy* is

$$S = \frac{E - F}{T} = k_B \ln(Z) - k_B \beta \frac{\partial}{\partial \beta} \ln(Z) = k_B \frac{\partial}{\partial T} [T \ln(Z)]. \quad (5)$$

2 THE ISING MODEL

Ferromagnetism is an interesting phenomenon in solids. Some metals (like Fe or Ni) are observed to acquire a finite magnetization below a certain temperature. Ferromagnetism is a fundamentally quantum mechanical effect, and arises when electron spins spontaneously align along a certain direction. The Ising model is a very crude attempt to model this phenomenon. It is defined as follows. We have a lattice in D dimensions with N sites. On each site j of this lattice sits a “spin” variable σ_j , which can take the two values ± 1 .

These are referred to as “spin-up” and “spin-down” respectively. A given set $\{\sigma_1, \sigma_2, \dots, \sigma_N\}$ specifies a *configuration*. The corresponding energy is taken to be of the form

$$E(\{\sigma_j\}) = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j - h \sum_{j=1}^N \sigma_j, \quad (6)$$

where $\langle ij \rangle$ denote nearest-neighbour bonds on our lattice and $J > 0$. The first term favours alignment on neighbouring spins, while h is like an applied magnetic field. Clearly, when $h = 0$ the lowest energy states are obtained by choosing all spins to be either up or down. The question of interest is whether the Ising model displays a finite temperature phase transition between a ferromagnetically ordered phase at low temperatures, and a paramagnetic phase at high temperatures. The partition function of the model is

$$Z = \sum_{\sigma_1=\pm 1} \sum_{\sigma_2=\pm 1} \dots \sum_{\sigma_N=\pm 1} e^{-\beta E(\{\sigma_j\})}. \quad (7)$$

The *magnetization per site* is given by

$$m(h) = \frac{1}{N} \langle \sum_{j=1}^N \sigma_j \rangle_\beta = \frac{1}{N\beta} \frac{\partial}{\partial h} \ln(Z). \quad (8)$$

The *magnetic susceptibility* is defined as

$$\chi(h) = \frac{\partial m(h)}{\partial h} = \frac{1}{N\beta} \frac{\partial^2}{\partial h^2} \ln(Z). \quad (9)$$

Substituting the expression (7) for the partition function and then carrying out the derivatives it can be expressed in the form

$$\chi(h) = \frac{\beta}{N} \sum_{l,m=1}^N \langle \sigma_l \sigma_m \rangle_\beta - \langle \sigma_l \rangle_\beta \langle \sigma_m \rangle_\beta. \quad (10)$$

2.1 THE ONE-DIMENSIONAL ISING MODEL

The simplest case is when our lattice is one-dimensional, and we impose periodic boundary conditions. The energy then reads

$$E = \sum_{j=1}^N \left[-J \sigma_j \sigma_{j+1} - \frac{h}{2} (\sigma_j + \sigma_{j+1}) \right] \equiv \sum_{j=1}^N E(\sigma_j, \sigma_{j+1}), \quad (11)$$

where we have defined

$$\sigma_{N+1} = \sigma_1. \quad (12)$$

The partition function can be expressed in the form

$$Z = \sum_{\sigma_1, \dots, \sigma_N} \prod_{j=1}^N e^{-\beta E(\sigma_j, \sigma_{j+1})}. \quad (13)$$

It can be evaluated exactly by means of the *transfer matrix method*.

2.1.1 TRANSFER MATRIX APPROACH

The general idea is to rewrite Z as a product of matrices. The transfer matrix T is taken to be a 2×2 matrix with elements

$$T_{\sigma\sigma'} = e^{-\beta E(\sigma, \sigma')}. \quad (14)$$

Its explicit form is

$$T = \begin{pmatrix} T_{++} & T_{+-} \\ T_{-+} & T_{--} \end{pmatrix} = \begin{pmatrix} e^{\beta(J+h)} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta(J-h)} \end{pmatrix}. \quad (15)$$

The partition function can be expressed in terms of the transfer matrix as follows

$$Z = \sum_{\sigma_1, \dots, \sigma_N} T_{\sigma_1 \sigma_2} T_{\sigma_2 \sigma_3} \dots T_{\sigma_{N-1} \sigma_N} T_{\sigma_N \sigma_1} \quad (16)$$

As desired, this has the structure of a matrix multiplication

$$\boxed{Z = \text{Tr}(T^N)}. \quad (17)$$

The trace arises because we have imposed periodic boundary conditions. As T is a real symmetric matrix, it can be diagonalized, i.e.

$$U^\dagger T U = \begin{pmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{pmatrix}, \quad (18)$$

where U is a unitary matrix and

$$\lambda_{\pm} = e^{\beta J} \cosh(\beta h) \pm \sqrt{e^{2\beta J} \sinh^2(\beta h) + e^{-2\beta J}}. \quad (19)$$

Using the cyclicity of the trace and $U U^\dagger = I$, we have

$$Z = \text{Tr}(U U^\dagger T^N) = \text{Tr}(U^\dagger T^N U) = \text{Tr}([U^\dagger T U]^N) = \text{Tr} \begin{pmatrix} \lambda_+^N & 0 \\ 0 & \lambda_-^N \end{pmatrix} = \lambda_+^N + \lambda_-^N. \quad (20)$$

But as $\lambda_- < \lambda_+$ we have

$$Z = \lambda_+^N \left(1 + \left[\frac{\lambda_-}{\lambda_+} \right]^N \right) = \lambda_+^N \left(1 + e^{-N \ln(\lambda_+/\lambda_-)} \right). \quad (21)$$

So for large N , which is the case we are interested in, we have with exponential accuracy

$$\boxed{Z \simeq \lambda_+^N}. \quad (22)$$

Given the partition function, we can now easily calculate the magnetization per site

$$m(h) = \frac{1}{N\beta} \frac{\partial}{\partial h} \ln(Z). \quad (23)$$

In Fig. 1 we plot $m(h)$ as a function of inverse temperature $\beta = 1/k_B T$ for two values of magnetic field h . We see that for non-zero h , the magnetization per site takes its maximum value $m = 1$ at low temperatures. At high temperatures it goes to zero. This is as expected, as at low T the spins align along the direction of the applied field. However, as we decrease the field, the temperature below which $m(h)$ approaches unity decreases. In the limit $h \rightarrow 0$, the magnetization per site vanishes at all finite temperatures. Hence there is no phase transition to a ferromagnetically ordered state in the one dimensional Ising model.

2.1.2 AVERAGES OF OBSERVABLES IN THE TRANSFER MATRIX FORMALISM

The average magnetization at site j is

$$\langle \sigma_j \rangle_\beta = \frac{1}{Z} \sum_{\sigma_1, \dots, \sigma_N} \sigma_j e^{-\beta E(\{\sigma_j\})}. \quad (24)$$

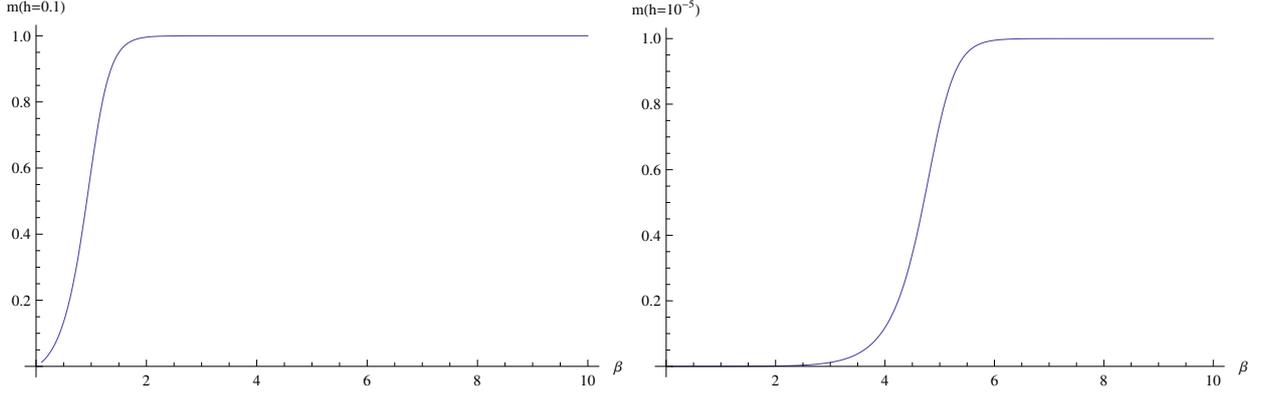


Figure 1: Magnetization per site as a function of inverse temperature for two values of applied magnetic field. We see that when we reduce the magnetic field, the temperature region in which the magnetization is essentially zero grows.

We can express this in terms of the transfer matrix as

$$\langle \sigma_j \rangle_\beta = \frac{1}{Z} \sum_{\sigma_1, \dots, \sigma_N} T_{\sigma_1 \sigma_2} T_{\sigma_2 \sigma_3} \dots T_{\sigma_{j-1} \sigma_j} \sigma_j T_{\sigma_j \sigma_{j+1}} \dots T_{\sigma_N \sigma_1}. \quad (25)$$

Using that

$$(T \sigma^z)_{\sigma_{j-1} \sigma_j} = T_{\sigma_{j-1} \sigma_j} \sigma_j, \quad (26)$$

where $\sigma^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ is the Pauli matrix, we obtain

$$\langle \sigma_j \rangle_\beta = \frac{1}{Z} \text{Tr} [T^{j-1} \sigma^z T^{N-j+1}] = \frac{1}{Z} \text{Tr} [T^N \sigma^z]. \quad (27)$$

Diagonalizing T by means of a unitary transformation as before, this becomes

$$\langle \sigma_j \rangle_\beta = \frac{1}{Z} \text{Tr} [U^\dagger T^N U U^\dagger \sigma^z U] = \frac{1}{Z} \text{Tr} \left[\begin{pmatrix} \lambda_+^N & 0 \\ 0 & \lambda_-^N \end{pmatrix} U^\dagger \sigma^z U \right]. \quad (28)$$

The matrix U is given in terms of the normalized eigenvectors of T

$$T|\pm\rangle = \lambda_\pm |\pm\rangle \quad (29)$$

as

$$U = (|+\rangle, |-\rangle). \quad (30)$$

For $h = 0$ we have

$$U|_{h=0} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \quad (31)$$

This gives

$$\langle \sigma_j \rangle_\beta \Big|_{h=0} = 0. \quad (32)$$

For general h the expression is more complicated

$$U = \begin{pmatrix} \frac{\alpha_+}{\sqrt{1+\alpha_+^2}} & \frac{\alpha_-}{\sqrt{1+\alpha_-^2}} \\ \frac{1}{\sqrt{1+\alpha_+^2}} & \frac{1}{\sqrt{1+\alpha_-^2}} \end{pmatrix}, \quad \alpha_\pm = \sqrt{1 + e^{4\beta J} \sinh^2(\beta h)} \pm e^{2\beta J} \sinh(\beta h). \quad (33)$$

The magnetization per site in the thermodynamic limit is then

$$\lim_{N \rightarrow \infty} \langle \sigma_j \rangle_\beta = \lim_{N \rightarrow \infty} \frac{\left(\frac{\alpha_+^2 - 1}{\alpha_+^2 + 1}\right) \lambda_+^N + \left(\frac{\alpha_-^2 - 1}{\alpha_-^2 + 1}\right) \lambda_-^N}{\lambda_+^N + \lambda_-^N} = \left(\frac{\alpha_+^2 - 1}{\alpha_+^2 + 1}\right). \quad (34)$$

This now allows us to prove, that in the one dimensional Ising model there is no phase transition at any finite temperature:

$$\boxed{\lim_{h \rightarrow 0} \lim_{N \rightarrow \infty} \langle \sigma_j \rangle_\beta = 0, \quad \beta < \infty.} \quad (35)$$

Note the order of the limits here: we first take the infinite volume limit at finite h , and only afterwards take h to zero. This procedure allows for spontaneous symmetry breaking to occur, but the outcome of our calculation is that the spin reversal symmetry remains unbroken at any finite temperature.

Similarly, we find

$$\langle \sigma_j \sigma_{j+r} \rangle_\beta = \frac{1}{Z} \text{Tr} [T^{j-1} \sigma^z T^r \sigma^z T^{N+1-j-r}] = \frac{1}{Z} \text{Tr} \left[U^\dagger \sigma^z U \begin{pmatrix} \lambda_+^r & 0 \\ 0 & \lambda_-^r \end{pmatrix} U^\dagger \sigma^z U \begin{pmatrix} \lambda_+^{N-r} & 0 \\ 0 & \lambda_-^{N-r} \end{pmatrix} \right]. \quad (36)$$

We can evaluate this for zero field $h = 0$

$$\langle \sigma_j \sigma_{j+r} \rangle_\beta \Big|_{h=0} = \frac{\lambda_+^{N-r} \lambda_-^r + \lambda_-^{N-r} \lambda_+^r}{\lambda_+^N + \lambda_-^N} \approx \left[\frac{\lambda_-}{\lambda_+} \right]^r = e^{-r/\xi}. \quad (37)$$

So in zero field the two-point function decays exponentially with *correlation length*

$$\xi = \frac{1}{\ln \coth(\beta J)}. \quad (38)$$

2.2 THE TWO-DIMENSIONAL ISING MODEL

We now turn to the 2D Ising model on a square lattice with periodic boundary conditions. The spin variables have now two indices corresponding to rows and columns of the square lattice respectively

$$\sigma_{j,k} = \pm 1, \quad j, k = 1, \dots, N. \quad (39)$$

The boundary conditions are $\sigma_{k,N+1} = \sigma_{k,1}$ and $\sigma_{N+1,j} = \sigma_{1,j}$, which correspond to the lattice “living” on

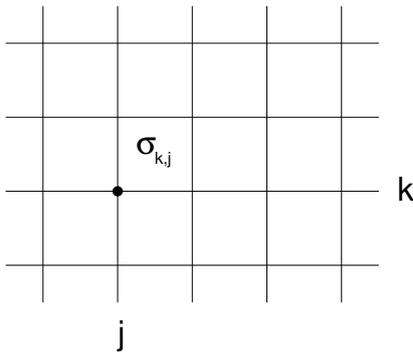


Figure 2: Ising model on the square lattice.

the surface of a torus. The energy in zero field is

$$E(\{\sigma_{k,j}\}) = -J \sum_{j,k} \sigma_{k,j} \sigma_{k,j+1} + \sigma_{k,j} \sigma_{k+1,j}. \quad (40)$$

2.2.1 TRANSFER MATRIX METHOD

The partition function is given by

$$Z = \sum_{\{\sigma_{j,k}\}} e^{-\beta E(\{\sigma_{k,j}\})}. \quad (41)$$

The idea of the transfer matrix method is again to write this in terms of matrix multiplications. The difference to the one dimensional case is that the transfer matrix will now be much larger. We start by expressing the partition function in the form

$$Z = \sum_{\{\sigma_{j,k}\}} e^{-\beta \sum_{k=1}^N E(k;k+1)}, \quad (42)$$

where

$$E(k;k+1) = -J \sum_{j=1}^N \sigma_{k,j} \sigma_{k+1,j} + \frac{1}{2} [\sigma_{k,j} \sigma_{k,j+1} + \sigma_{k+1,j} \sigma_{k+1,j+1}]. \quad (43)$$

This energy depends only on the configurations of spins on rows k and $k+1$, i.e. on spins $\sigma_{k,1}, \dots, \sigma_{k,N}$ and $\sigma_{k+1,1}, \dots, \sigma_{k+1,N}$. Each configuration of spins on a given row specifies a sequence s_1, s_2, \dots, s_N with $s_j = \pm 1$. Let us associate a vector

$$|\mathbf{s}\rangle \quad (44)$$

with each such sequence. By construction there 2^N such vectors. We then define a scalar product on the space spanned by these vectors by

$$\langle \mathbf{t} | \mathbf{s} \rangle = \prod_{j=1}^N \delta_{t_j, s_j}. \quad (45)$$

With this definition, the vectors $\{|\mathbf{s}\rangle\}$ form an orthonormal basis of a 2^N dimensional linear vector space. In particular we have

$$I = \sum_{\mathbf{s}} |\mathbf{s}\rangle \langle \mathbf{s}|. \quad (46)$$

Finally, we define a $2^N \times 2^N$ transfer matrix T by

$$\langle \sigma_k | T | \sigma_{k+1} \rangle = e^{-\beta E(k;k+1)}. \quad (47)$$

The point of this construction is that the partition function can now be written in the form

$$Z = \sum_{\sigma_1} \sum_{\sigma_2} \cdots \sum_{\sigma_N} \langle \sigma_1 | T | \sigma_2 \rangle \langle \sigma_2 | T | \sigma_3 \rangle \cdots \langle \sigma_{N-1} | T | \sigma_N \rangle \langle \sigma_N | T | \sigma_1 \rangle \quad (48)$$

We now may use (46) to carry out the sums over spins, which gives

$$\boxed{Z = \text{Tr} [T^N]} \quad (49)$$

where the trace is over our basis $\{|\mathbf{s}\rangle | s_j = \pm 1\}$ of our 2^N dimensional vector space. Like in the 1D case, thermodynamic properties involve only the largest eigenvalues of T . Indeed, we have

$$Z = \sum_{j=1}^{2^N} \lambda_j^N, \quad (50)$$

where λ_j are the eigenvalues of T . The free energy is then

$$F = -k_B T \ln(Z) = -k_B T \ln \left[\lambda_{\max}^N \sum_{j=1}^{2^N} \left(\frac{\lambda_j}{\lambda_{\max}} \right)^N \right] = -k_B T N \ln(\lambda_{\max}) - k_B T \ln \left[\sum_j \left(\frac{\lambda_j}{\lambda_{\max}} \right)^N \right], \quad (51)$$

where λ_{\max} is the largest eigenvalue of T , which we assume to be unique. As $|\lambda_j/\lambda_{\max}| < 1$, the second contribution in (51) is bounded by $-k_B T N \ln(2)$, and we see that in the thermodynamic limit the free energy per site is

$$f = \lim_{N \rightarrow \infty} \frac{F}{N^2} = \lim_{N \rightarrow \infty} -\frac{k_B T}{N} \ln(\lambda_{\max}). \quad (52)$$

Thermodynamic quantities are obtained by taking derivatives of f and hence only involve the largest eigenvalue of T . The main complication we have to deal with is that T is still a very large matrix. This poses the question, why we should bother to use a transfer matrix description anyway? Calculating Z from its basic definition (41) involves a sum with 2^{N^2} terms, i.e. at least 2^{N^2} operations on a computer. Finding the largest eigenvalue of a $M \times M$ matrix involves $\mathcal{O}(M^2)$ operations, which in our case amounts to $\mathcal{O}(2^{2N})$. For large values of N this amounts to an enormous simplification.

2.2.2 SPONTANEOUS SYMMETRY BREAKING

Surprisingly, the transfer matrix of the 2D Ising model can be diagonalized exactly. Unfortunately we don't have the time to go through the somewhat complicated procedure here, but the upshot is that the 2D Ising model can be solved *exactly*. Perhaps the most important result is that in the thermodynamic limit the square lattice Ising model has a *finite temperature phase transition* between a paramagnetic and a ferromagnetic phase. The magnetization per site behaves as shown in Fig.3. At low temperatures $T < T_c$

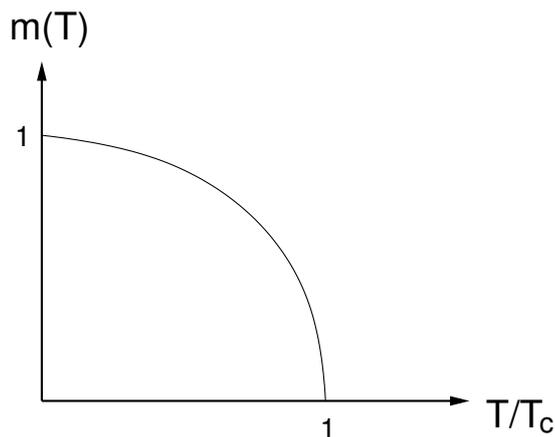


Figure 3: Phase Transition in the square lattice Ising model.

there is a non-zero magnetization per site, even though we did not apply a magnetic field. This is surprising, because our energy (40) is unchanged if we flip all spins

$$\sigma_{j,k} \rightarrow -\sigma_{j,k}. \quad (53)$$

The operation (53) is a discrete (two-fold) symmetry of the Ising model. Because we have translational invariance, the magnetization per site is

$$m = \langle \sigma_{j,k} \rangle_{\beta}. \quad (54)$$

Hence a non-zero value of m signifies the *spontaneous breaking* of the discrete symmetry (53). In order to describe this effect mathematically, we have to invoke a bit of trickery. Let us consider zero temperature. Then there are exactly two degenerate lowest energy states: the one with all spins $\sigma_{j,k} = +1$ and the one with all spins $\sigma_{j,k} = -1$. We now apply a very small magnetic field to the system, i.e. add a term

$$\delta E = -\epsilon \sum_{j,k} \sigma_{j,k} \quad (55)$$

to the energy. This splits the two states, which now have energies

$$E_{\pm} = -JN_B \mp \epsilon N, \quad (56)$$

where N_B is the number of bonds. The next step is key: we now *define* the thermodynamic limit of the free energy per site as

$$f(T) \equiv \lim_{\epsilon \rightarrow 0} \lim_{N \rightarrow \infty} \frac{-k_B T \ln(Z)}{N^2}. \quad (57)$$

The point is that the contributions $Z_{\pm} = e^{-\beta E_{\pm}}$ of the two states to Z are such that

$$\frac{Z_-}{Z_+} = e^{-2\epsilon N/k_B T}. \quad (58)$$

This goes to zero when we take N to infinity! So in the above sequence of limits, only the state with all spins up contributes to the partition function, and this provides a way of describing spontaneous symmetry breaking! The key to this procedure is that

$$\boxed{\lim_{\epsilon \rightarrow 0} \lim_{N \rightarrow \infty} Z \neq \lim_{N \rightarrow \infty} \lim_{\epsilon \rightarrow 0} Z.} \quad (59)$$

The procedure we have outlined above, i.e. introducing a symmetry breaking field, then taking the infinite volume limit, and finally removing the field, is very general and applies to all instances where spontaneous symmetry breaking occurs.

Exercises:

Question 1. A lattice model for non-ideal gas is defined as follows. The sites i of a lattice may be empty or occupied by at most one atom, and the variable n_i takes the values $n_i = 0$ and $n_i = 1$ in the two cases. There is an attractive interaction energy J between atoms that occupy neighbouring sites, and a chemical potential μ . The model Hamiltonian is

$$\mathcal{H} = -J \sum_{\langle ij \rangle} n_i n_j - \mu \sum_i n_i, \quad (60)$$

where $\sum_{\langle ij \rangle}$ is a sum over neighbouring pairs of sites.

(a) Describe briefly how the *transfer matrix method* may be used to calculate the statistical-mechanical properties of one-dimensional lattice models with short range interactions. Illustrate your answer by explaining how the partition function for a one-dimensional version of the lattice gas, Eq. (1), defined on a lattice of N sites with periodic boundary conditions, may be evaluated using the matrix

$$\mathbf{T} = \begin{pmatrix} 1 & e^{\beta\mu/2} \\ e^{\beta\mu/2} & e^{\beta(J+\mu)} \end{pmatrix}.$$

(b) Derive an expression for $\langle n_i \rangle$ in the limit $N \rightarrow \infty$, in terms of elements of the eigenvectors of this matrix.

(c) Show that

$$\langle n_i \rangle = \frac{1}{1 + e^{-2\theta}},$$

where

$$\sinh(\theta) = \exp(\beta J/2) \sinh(\beta[J + \mu]/2).$$

Sketch $\langle n_i \rangle$ as a function of μ for $\beta J \gg 1$, and comment on the physical significance of your result.

Question 2. The one-dimensional 3-state Potts model is defined as follows. At lattice sites $i = 0, 1, \dots, L$ “spin” variables σ_i take integer values $\sigma_i = 1, 2, 3$. The Hamiltonian is then given by

$$H = -J \sum_{i=0}^{L-1} \delta_{\sigma_i, \sigma_{i+1}}, \quad (61)$$

where $\delta_{a,b}$ is the Kronecker delta, $J > 0$.

(a) What are the ground states and first excited states for this model?

(b) Write down the transfer matrix for (61). Derive an expression for the free energy per site f in the limit of large L in terms of the transfer matrix eigenvalues. Show that vectors of the form $(1, z, z^2)$ with $z^3 = 1$ are eigenvectors, and hence find the corresponding eigenvalues. Show that at temperature T (with $\beta = 1/k_B T$) and in the limit $L \rightarrow \infty$

$$f = -k_B T \ln \left(3 + e^{\beta J} - 1 \right). \quad (62)$$

(c) The boundary variable σ_0 is fixed in the state $\sigma_0 = 1$. Derive an expression (for large L), that the variable at site $\ell \gg 1$ is in the same state, in terms of the transfer matrix eigenvalues and eigenvectors. Show that your result has the form

$$\langle \delta_{\sigma_\ell, 1} \rangle = \frac{1}{3} + \frac{2}{3} e^{-\ell/\xi}. \quad (63)$$

How does ξ behave in the low and high temperature limits?

Question 3. Consider a one dimensional Ising model on an open chain with N sites, where N is odd. On all even sites a magnetic field $2h$ is applied, see Fig. 4. The energy is

$$E = -J \sum_{j=1}^{N-1} \sigma_j \sigma_{j+1} + 2h \sum_{j=1}^{(N-1)/2} \sigma_{2j}. \quad (64)$$

(a) Show that the partition function can be written in the form

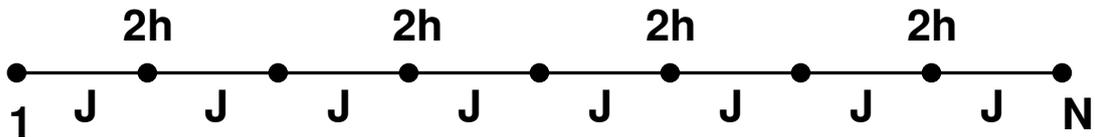


Figure 4: Open Ising chain with magnetic field applied to all even sites.

$$Z = \langle u | T^{(N-1)/2} | v \rangle, \quad (65)$$

where T is an appropriately constructed transfer matrix, and $|u\rangle$ and $|v\rangle$ two dimensional vectors. Give explicit expressions for T , $|u\rangle$ and $|v\rangle$.

(b) Calculate Z for the case $h = 0$.

2.2.3 PEIERLS ARGUMENT

The Peierls argument is a nice way of establishing that the 2D square lattice Ising model has magnetic long-range order at sufficiently low temperatures *without actually solving the model*. Given that at very high temperature there is no magnetic order, this shows that there must be at least one phase transition at a finite temperature.

Here the thermal average of $X_{\sigma}(b, i)$ can be written as

$$\langle X_{\sigma}(b, i) \rangle_{\beta} = \frac{1}{Z} \sum'_{\sigma} e^{-\beta E(\sigma)}, \quad (70)$$

where the sum is only over configurations, in which (b, i) occurs. Now consider the configuration σ' obtained from σ by reversing the spins inside the domain wall (b, i) . Clearly the energies of the two configurations

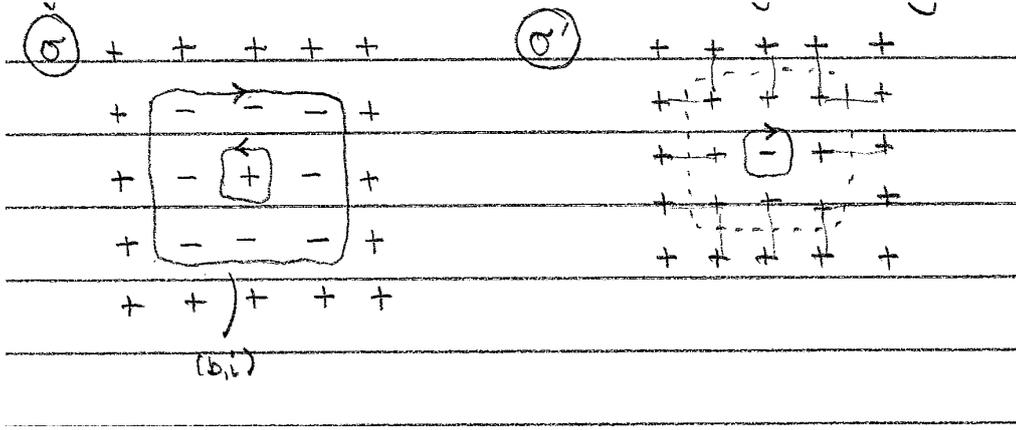


Figure 6: Configurations σ and σ' related by reversing all spins inside the domain wall (b, i) . Shown are all the bonds whose energies have been changed from $-J$ to J .

are related by

$$E(\sigma) = E(\sigma') + 2bJ. \quad (71)$$

This can be used to obtain a bound on Z

$$Z \geq \sum'_{\sigma'} e^{-\beta E(\sigma')} \geq \sum'_{\sigma} e^{-\beta E(\sigma)} e^{2bJ\beta}, \quad (72)$$

where the first sum is only over configurations in which (b, i) occurs, and where we then have flipped all spins inside the domain wall. This gives us a bound on

$$\langle X_{\sigma}(b, i) \rangle_{\beta} = \frac{1}{Z} \sum'_{\sigma} e^{-\beta E(\sigma)} \leq e^{-2bJ\beta}. \quad (73)$$

Now we put everything together

$$\langle N_- \rangle_{\beta} \leq \sum_b \frac{b^2}{16} \sum_{i=1}^{m(b)} e^{-2\beta Jb} \leq \sum_b \frac{b^2}{16} 4N_t 3^{b-1} e^{-2\beta Jb} = \frac{N_t}{12} \sum_{b=4,6,8,\dots} b^2 \left[3e^{-2\beta J} \right]^b. \quad (74)$$

The sum over b can now be easily carried out, and the results at small T (large β) is

$$\langle N_- \rangle_\beta \leq 108 N_t e^{-8\beta J}. \quad (75)$$

So, at low temperatures we have

$$\frac{\langle N_- \rangle_\beta}{N_t} \ll \frac{1}{2}. \quad (76)$$

This proves the existence of a spontaneous magnetization at low temperatures.

3 MEAN FIELD THEORY

Consider the Ising model on a D -dimensional lattice with coordination number (number of nearest neighbour sites) z

$$E = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j - h \sum_j \sigma_j. \quad (77)$$

Here $\langle ij \rangle$ denotes nearest neighbour bonds, and each bond is counted once. The magnetization per site is

$$m = \frac{1}{N} \sum_{j=1}^N \langle \sigma_j \rangle_\beta. \quad (78)$$

We now rewrite the energy using

$$\sigma_i = m + (\sigma_i - m). \quad (79)$$

In particular we have

$$\sigma_i \sigma_j = m^2 + m(\sigma_j - m) + m(\sigma_i - m) + (\sigma_i - m)(\sigma_j - m). \quad (80)$$

The idea of the *mean-field approximation* is to assume that the deviations $\sigma_j - m$ of the spins from their average values are small, and to neglect the terms quadratic in these fluctuations. This gives

$$E_{\text{MF}} = -J \sum_{\langle ij \rangle} -m^2 + m(\sigma_i + \sigma_j) - h \sum_j \sigma_j. \quad (81)$$

Physically, what we have done is to replace the interaction of a given spin with its neighbours by *an average magnetic field*. We can simplify (81) further by noting that

$$\begin{aligned} -J \sum_{\langle ij \rangle} -m^2 &= Jm^2 \frac{Nz}{2}, \\ \sum_{\langle ij \rangle} \sigma_i + \sigma_j &= z \sum_j \sigma_j. \end{aligned} \quad (82)$$

The mean-field energy then becomes

$$E_{\text{MF}} = \frac{JNz}{2} m^2 - (Jzm + h) \sum_{j=1}^N \sigma_j. \quad (83)$$

The partition function in the mean-field approximation is

$$\begin{aligned}
Z_{\text{MF}} = \sum_{\{\sigma_j\}} e^{-\beta E_{\text{MF}}} &= e^{-\frac{NJz\beta m^2}{2}} \sum_{\sigma_1} \dots \sum_{\sigma_N} \prod_{j=1}^N e^{\beta(Jzm+h)\sigma_j} \\
&= e^{-\frac{NJz\beta m^2}{2}} \left[\sum_{\sigma_1} e^{\beta(Jzm+h)\sigma_1} \right] \dots \left[\sum_{\sigma_N} e^{\beta(Jzm+h)\sigma_N} \right] \\
&= e^{-\frac{NJz\beta m^2}{2}} [2 \cosh(Jzm\beta + h\beta)]^N.
\end{aligned} \tag{84}$$

The magnetization per site is

$$m = \frac{1}{N} \sum_{j=1}^N \langle \sigma_j \rangle_{\beta} = \langle \sigma_N \rangle_{\beta}, \tag{85}$$

where we have used translational invariance in the last step. In mean field theory we have

$$\boxed{m = \langle \sigma_j \rangle_{\beta} = \frac{1}{Z_{\text{MF}}} e^{-\frac{NJz\beta m^2}{2}} \sum_{\sigma_1} \dots \sum_{\sigma_N} \sigma_N \prod_{j=1}^N e^{\beta(Jzm+h)\sigma_j} = \tanh(Jzm\beta + \beta h).} \tag{86}$$

This is a *self-consistency equation* for m .

3.1 SOLUTION OF THE SELF-CONSISTENCY EQUATION FOR $h = 0$

For zero field the self-consistency equation reads

$$m = \tanh(Jzm\beta). \tag{87}$$

This can be solved graphically by looking for intersections of the functions $g_1(m) = \tanh(Jzm\beta)$ and $g_2(m) = m$. There are either one or three solutions

$$m = \begin{cases} 0 & \text{if } Jz\beta < 1 \\ \pm m_0, 0 & \text{if } Jz\beta > 1 \end{cases}. \tag{88}$$

We still have to check whether these solutions correspond to minima of the free energy per site. The latter is

$$f_{\text{MF}} = -\frac{1}{\beta N} \ln(Z_{\text{MF}}) = \frac{Jzm^2}{2} - \frac{1}{\beta} \ln[2 \cosh(Jzm\beta)]. \tag{89}$$

We have

$$\frac{\partial^2 f_{\text{MF}}}{\partial m^2} = Jz \left[1 - \frac{Jz\beta}{\cosh(Jzm\beta)} \right]. \tag{90}$$

This is negative for $m = 0$ and $Jz\beta > 1$, and hence this solution corresponds to a *maximum* of the free energy and hence must be discarded. This leaves us with

$$\boxed{m = \begin{cases} 0 & \text{if } T > T_c \\ \pm m_0 & \text{if } T < T_c \end{cases}}, \tag{91}$$

where the transition temperature is

$$T_c = \frac{Jz}{k_B}. \tag{92}$$

3.2 VICINITY OF THE PHASE TRANSITION

Let us define a dimensionless variable, that measures the distance in temperature to the phase transition

$$t = \frac{T - T_c}{T_c}. \quad (93)$$

For $|t| \ll 1$ we obtain the following results

1. Magnetization per site

$$m|_{h=0} \simeq \begin{cases} 0 & \text{if } T > T_c, \\ \sqrt{-3t} & \text{if } T < T_c. \end{cases} \quad (94)$$

2. Magnetic susceptibility in zero field

$$\chi = \frac{\partial m}{\partial h}|_{h=0} \simeq \begin{cases} \frac{1}{k_B T_c} t^{-1} & \text{if } T > T_c, \\ \frac{1}{2k_B T_c} (-t)^{-1} & \text{if } T < T_c, \end{cases} \quad (95)$$

3. Free energy per site and heat capacity per volume

$$f_{\text{MF}}|_{h=0} \simeq -k_B T \ln 2 + \begin{cases} 0 & \text{if } T > T_c \\ -\frac{3k_B T_c}{4} t^2 & \text{if } T < T_c. \end{cases} \quad (96)$$

$$\frac{C}{V} = -T \frac{\partial^2 f_{\text{MF}}}{\partial T^2}|_{h=0} \simeq \begin{cases} 0 & \text{if } T > T_c \\ \frac{3k_B}{2} & \text{if } T < T_c. \end{cases} \quad (97)$$

4 LANDAU THEORY OF PHASE TRANSITIONS

4.1 PHASE TRANSITIONS

Physically a phase transition is a point in parameter space, where the physical properties of a many-particle system undergo a sudden change. An example is the paramagnet to ferromagnet transition in Fe or Ni shown in Fig. 7.

Mathematically a phase transition is a point in parameter space, where the free energy $F = -k_B T \ln(Z)$ becomes a nonanalytic function of one of its parameters (i.e. F or some of its derivatives becomes singular or discontinuous) in the *thermodynamic limit*.

For a finite system this can never happen, because

$$Z = \sum_{\text{configurations } C} e^{-E(C)/k_B T} \quad (98)$$

is a finite sum over finite, positive terms. Hence all derivatives are finite and well defined as well.

Phase transitions are usually divided into two categories:

1. *First Order Phase Transitions.*

Here the free energy is continuous, but a first derivative is discontinuous. At the transition there is *phase coexistence*. The magnetization per site is a first order derivative of the free energy with respect to the magnetic field h . Therefore the phase transition at $h = 0$ and $T < T_c$ in Fig. 7 is first order.

2. *Second Order Phase Transitions.*

These are characterized by a divergence in one of the higher order derivatives (“susceptibilities”) of the free energy. The phase transition as a function of T for $h = 0$ in Fig. 7 is second order.

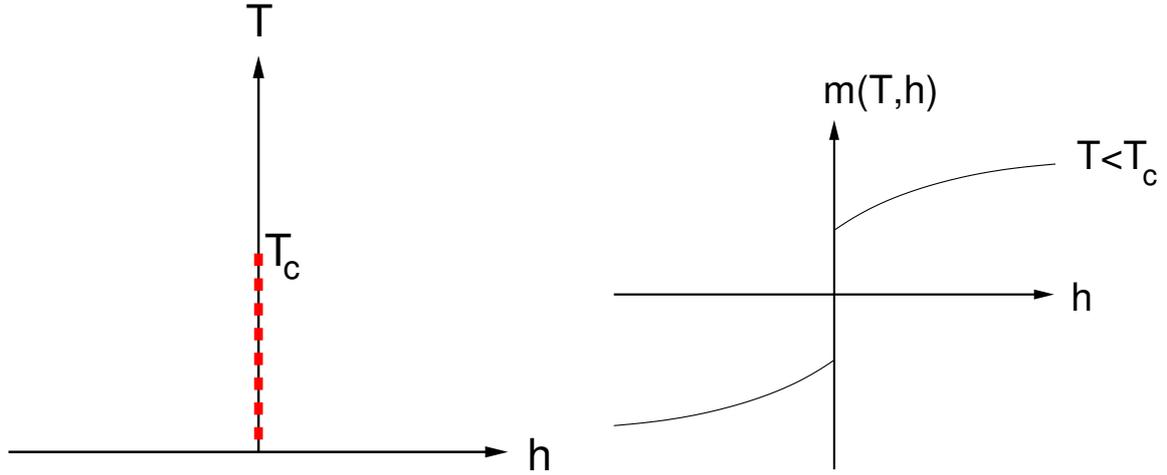


Figure 7: Phase diagram for the paramagnetic to ferromagnetic transition. The *magnetization per site* $m(T, h)$ jumps when crossing zero for $T < T_c$.

4.2 CRITICAL BEHAVIOUR AND UNIVERSALITY

Close to a critical point thermodynamic functions display power-law behaviours characterized by *critical exponents*. We will now discuss various such exponents, using as a specific example the paramagnet to ferromagnet transition.

1. ORDER PARAMETER

This is a quantity that is different in the various phases and can be used to characterize the phase transition. For the paramagnet to ferromagnet transition in zero magnetic field the appropriate order parameter is the magnetization per site

$$m(T) = \lim_{h \rightarrow +0} \lim_{V \rightarrow \infty} \frac{1}{V} M(h, T). \quad (99)$$

Here $M(h, T)$ is the magnetization. Where $T \approx T_c$, one has

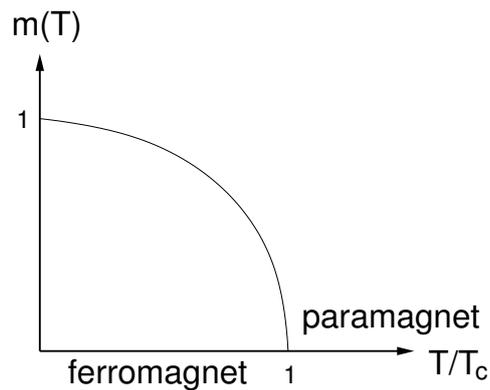


Figure 8: Order parameter for the paramagnet to ferromagnet transition in zero field.

$$m(T) \sim \begin{cases} 0 & \text{if } T > T_c \\ |t|^\beta & \text{if } T < T_c \end{cases} \quad t = \frac{T - T_c}{T_c}. \quad (100)$$

β is a *critical exponent*.

2. SUSCEPTIBILITIES

At the critical point the system is very sensitive to external perturbations. The singularity in the response of the order parameter to a field “conjugate” to it is characterized by critical exponents γ_{\pm} . For our magnet

$$\chi_{\pm}(T) = \left. \frac{\partial}{\partial h} \right|_{h=0} \lim_{V \rightarrow \infty} \frac{1}{V} M(h, T) \sim |t|^{-\gamma_{\pm}}. \quad (101)$$

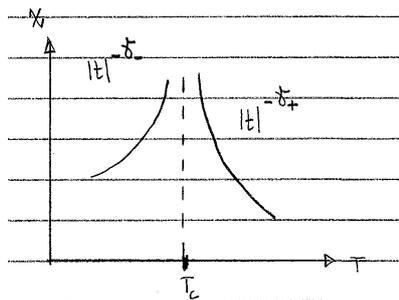


Figure 9: Critical behaviour of the magnetic susceptibility.

3. HEAT CAPACITY

A third critical exponent is associated with the *heat capacity*

$$C(T) = -T \frac{\partial^2 F}{\partial T^2} \sim \begin{cases} A_+ |t|^{-\alpha_+} & \text{if } T > T_c \\ A_- |t|^{-\alpha_-} & \text{if } T < T_c. \end{cases} \quad (102)$$

Depending on the signs of α_{\pm} this may or may not be singular, see Fig. 10.

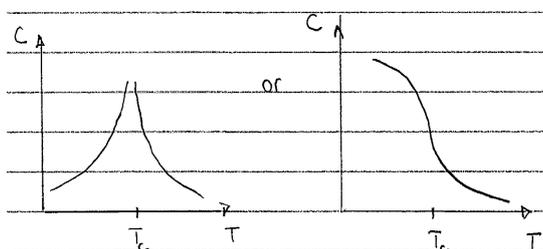


Figure 10: Critical behaviour of the heat capacity.

4.2.1 UNIVERSALITY

The critical exponents are *insensitive* to microscopic details of the system under consideration and are characteristic of the critical point. A consequence of this is that completely different systems can exhibit the same critical behaviour!

4.3 LANDAU THEORY

Landau Theory is a general approach to phase transitions that

- is phenomenological in nature and deals only with macroscopic quantities;
- applies only to the neighbourhood of a critical point, where the order parameter is small.

Landau theory is constructed as follows.

1. Identify the order parameter(s) $M(\mathbf{r})$ characterizing the phase transition. Depending on which phase transition we are dealing with M can be a real scalar, a complex scalar, a real or complex vector or something more complicated. For the paramagnet to ferromagnet transition the order parameter is the magnetization per site, i.e. a real number.
2. Form a “coarse-grained” order parameter density $\Phi(\mathbf{r})$. Think of this as the microscopic order parameter averaged over atomic distances. This is a *continuum field*. For the example of the paramagnet to ferromagnet transition we are dealing with a real scalar field.
3. Consider the free energy density to be a *functional* of the order parameter field $\Phi(\mathbf{r})$. The free energy is then

$$\boxed{\beta F = \int d^D \mathbf{r} f[\Phi(\mathbf{r})].} \quad (103)$$

4. By construction of the order parameter(s), the latter is (are) *small* close to our critical point. This allows us to expand $f[\Phi(\mathbf{r})]$ as a power series around $\Phi = 0$. From now on we will focus on the simplest case of a real scalar order parameter $\phi(\mathbf{r})$. Then the series expansion is

$$\boxed{f[\phi] \simeq \text{const} - h\phi + \alpha_2\phi^2 + \frac{1}{2}|\nabla\phi|^2 + \alpha_3\phi^3 + \alpha_4\phi^4 + \dots} \quad (104)$$

where the coefficient of the gradient term is fixed by convention to be 1/2. This makes ϕ in general dimensionful

$$\dim[\phi(\mathbf{r})] = (\text{length})^{1-D/2}. \quad (105)$$

The only linear term that is not a total derivative is $-h\phi$, where h is an external field (a “source”) coupling to the order parameter. Total derivative terms can be dropped, because they only give boundary contributions to F . *The coefficients α_n are a priori all functions of temperature.*

5. In translationally invariant systems the free energy is minimized by \mathbf{r} -independent order parameters (i.e. $\nabla\phi(\mathbf{r}) = 0$). The reason is that $\frac{1}{2}|\nabla\phi|^2 \geq 0$, and hence this contribution to F is minimized by constant solutions. For constant field h the potential $V(\phi(\mathbf{r})) = -h\phi + \alpha_2\phi^2 + \alpha_4\phi^4$ is also minimized by constant solutions. In order to understand the nature of the phase transition, we therefore can simply look at the minima of the potential $V(\phi)$.
6. Finally, we use symmetries and the fact that we are interested in the vicinity of a critical point to constrain the α_j .

- If we truncate our expansion at order ϕ^4 , then thermodynamic stability requires

$$\boxed{\alpha_4 > 0.} \quad (106)$$

If $\alpha_4 < 0$ the free energy density would be unbounded from below and could become infinitely negative, which is forbidden.

- If we know that the system is invariant under certain symmetry operations, e.g.

$$\phi \rightarrow -\phi, \quad (107)$$

then the free energy must respect this symmetry. A ferromagnet has the symmetry (107) in absence of a magnetic field because of *time-reversal invariance*. Hence we must have $\alpha_3 = 0$ in this case.

- In the case $h = \alpha_3 = 0$, for a translationally invariant system, we can obtain the temperature dependence of α_2 as follows. As discussed above, the nature of the phase transition can be inferred from the minima of the potential $V(\phi) = \alpha_2\phi^2 + \alpha_4\phi^4$. This is done in Fig. 11. We see that the

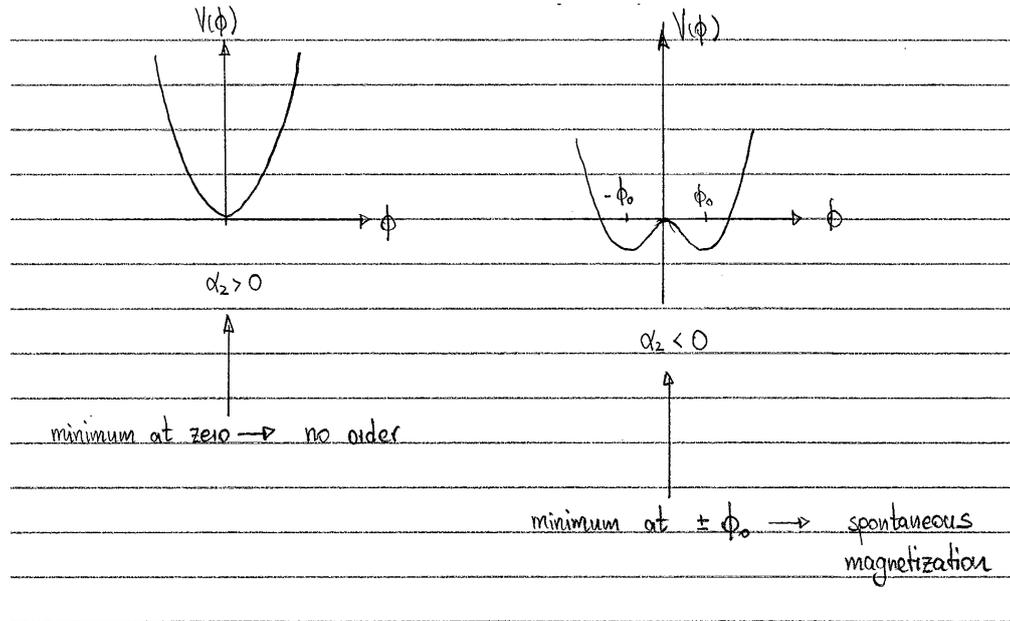


Figure 11: Minima of the potential $V(\phi)$. For $\alpha_2 > 0$ the minima occurs at $\phi = 0$, so there is no ferromagnetic order. For $\alpha_2 < 0$ there are two minima at $\phi = \pm\phi_0$, corresponding to the emergence of a spontaneous magnetization.

phase transition corresponds to α_2 changing sign at $T = T_c$. So in the vicinity of the transition we have (by Taylor expanding α_2 in $T - T_c$)

$$\boxed{\alpha_2(t) = At + \mathcal{O}(t^2), \quad t = \frac{T - T_c}{T_c}, \quad A > 0.} \quad (108)$$

The parameter α_4 is also temperature dependent, but this dependence is subleading

$$\alpha_4(t) = \alpha_4(0) + \mathcal{O}(t). \quad (109)$$

7. If we have $\alpha_3 < 0$ the transition is generically first order. To see this we again use that in a translationally invariant system the minima of the free energy density will be \mathbf{r} -independent, so that we merely need to scrutinize the potential $V(\phi)$ to understand the nature of the phase transition. In Fig. 12 we plot $V(\phi)$ when α_2 is decreased at fixed α_3, α_4 . We see that initially the minimum occurs at $\phi = 0$ (no order), and at some critical value of α_2 then jumps from zero to a finite value ϕ_0 . This is characteristic of a first order transition.

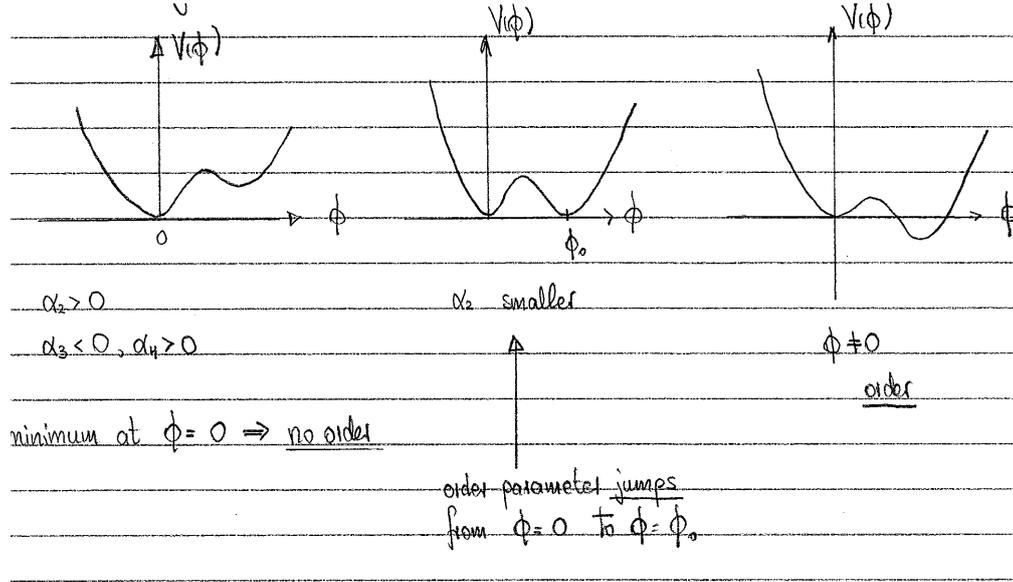


Figure 12: Minima of the potential $V(\phi)$ for $\alpha_3 < 0$. Decreasing the value of α_2 leads to a discontinuous jump in the order parameter at some critical value. The transition is therefore first order.

4.3.1 THERMODYNAMIC EQUILIBRIUM

The state of thermodynamic equilibrium is obtained by minimizing the free energy (we assume that $\alpha_3 = 0$)

$$\beta F = \int d^D \mathbf{r} \left[-h\phi + \alpha_2 \phi^2 + \frac{1}{2} |\nabla \phi|^2 + \alpha_4 \phi^4 \right]. \quad (110)$$

In our case we are searching for the order parameter configuration $\phi(\mathbf{r})$ that gives the smallest contribution to βF . This is found by functional extremization

$$\frac{\delta F}{\delta \phi(\mathbf{r})} = 0. \quad (111)$$

The resulting nonlinear differential equation is

$$\boxed{-\nabla^2 \phi(\mathbf{r}) + 2\alpha_2 \phi(\mathbf{r}) + 4\alpha_4 \phi^3(\mathbf{r}) - h = 0.} \quad (112)$$

It is easy to see (see the discussion above) that the solutions to (112) that minimize the free energy are in fact \mathbf{r} -independent (if we ignore boundary conditions). Hence for zero field $h = 0$ the order parameter configuration that minimizes the free energy is

$$\phi(\mathbf{r}) = \begin{cases} 0 & \text{if } \alpha_2 > 0 \leftrightarrow T > T_c \\ \pm\phi_0 = \pm\sqrt{-\frac{\alpha_2}{2\alpha_4}} & \text{if } \alpha_2 < 0 \leftrightarrow T < T_c \end{cases} . \quad (113)$$

We observe that we are dealing with a second order phase transition (because ϕ_0 vanishes at the critical point) from a paramagnetic to a ferromagnetic phase.

4.3.2 BEYOND THE LANDAU FREE ENERGY

So far we have focussed entirely on the state of thermodynamic equilibrium. We now want to extend Landau theory to take into account fluctuations as well. Let us go back to the microscopic model underlying our Landau free energy. The partition function for this microscopic theory is

$$Z_{\text{micro}} = \sum_{\text{configurations } C} e^{-\beta E(C)} . \quad (114)$$

In order to define our order-parameter field, we used a coarse-graining procedure. Hence after coarse-graining Z should become

$$Z \longrightarrow \sum_{\substack{\text{order parameter} \\ \text{configurations}}} e^{-\beta\mathcal{H}} . \quad (115)$$

Because under the coarse-graining procedure many microscopic configurations C map onto the same order parameter configuration $\phi(\mathbf{r})$, the ‘‘Landau-Ginzburg Hamiltonian’’ $\beta\mathcal{H}$ incorporates certain entropic effects. By employing the same logic as before, we can argue that

$$\beta\mathcal{H} = \int d^D\mathbf{r} f[\phi(\mathbf{r})] , \quad (116)$$

where $f[\phi(\mathbf{r})]$ is the same functional we constructed when considering the Landau free energy. As the order parameter is really a continuous field, what we mean by the sum in (115) is really the *functional integral*

$$Z = \int \mathcal{D}\phi(\mathbf{r}) e^{-\beta\mathcal{H}} . \quad (117)$$

The latter is defined as follows:

- We first discretize our D -dimensional spatial variable

$$\mathbf{r} \longrightarrow a_0\mathbf{n} = a_0(n_1, \dots, n_D), \quad (118)$$

where a_0 is a lattice spacing and the total number of points on our discrete grid is N^D .

- We then discretize the order parameter field and its derivatives

$$\begin{aligned} \phi(\mathbf{r}) &\longrightarrow \phi_{\mathbf{n}} , \\ \partial_{r_j}\phi(\mathbf{r}) &\longrightarrow \frac{\phi_{\mathbf{n}+\mathbf{e}_j} - \phi_{\mathbf{n}}}{a_0} , \end{aligned} \quad (119)$$

where \mathbf{e}_j are unit vectors in the j -direction.

- The Landau-Ginzburg Hamiltonian is discretized as

$$\beta\mathcal{H} \longrightarrow \sum_{\mathbf{n}} f[\phi_{\mathbf{n}}]a_0^D . \quad (120)$$

- The functional integral is then defined as follows

$$Z = \int \mathcal{D}\phi(\mathbf{r}) e^{-\beta\mathcal{H}} \equiv \lim_{N \rightarrow \infty} \int \prod_{\mathbf{n}} d\phi_{\mathbf{n}} e^{-\sum_{\mathbf{m}} f[\phi_{\mathbf{m}}] a_0^D}. \quad (121)$$

Crucially, the functional integral defined in this way can be manipulated according to the same rules we derived for path-integrals in Quantum Mechanics. In this new way of looking at things we now can analyze properties that are not directly related to the free energy. For example, we may ask about properties of correlation functions like

$$\langle \phi(\mathbf{r})\phi(\mathbf{0}) \rangle_{\beta} \equiv \frac{1}{Z} \int \mathcal{D}\phi \phi(\mathbf{r})\phi(\mathbf{0}) e^{-\beta\mathcal{H}}. \quad (122)$$

4.3.3 SADDLE POINT APPROXIMATION

The Landau-Ginzburg field theory

$$Z = \int \mathcal{D}\phi(\mathbf{r}) e^{-\int d^D\mathbf{r} f[\phi(\mathbf{r})]}, \quad (123)$$

is still difficult to analyze. For the example we have discussed, it reduces to the Euclidean space version of the $\lambda\phi^4$ theory you have encountered in the field theory part of the course. In order to proceed we therefore resort to further approximations. The *saddle-point approximation* takes into account the thermodynamically most likely configuration $\phi(\mathbf{r})$, i.e. the configuration that minimizes

$$\beta\mathcal{H} = \int d^D\mathbf{r} f[\phi(\mathbf{r})]. \quad (124)$$

We see that the saddle-point approximation precisely recovers the results of the Landau free energy approach! However, using our new formulation we now go beyond this approximation and take into account fluctuations. We will see below how to do this.

4.3.4 MEAN FIELD EXPONENTS

Using the saddle point solution we can determine the corresponding approximation for the critical exponents.

- Order parameter.

Using that $\alpha_2 = At$ for $t = (T - T_c)/T_c < 0$, we have

$$\phi_0 = \pm \sqrt{\frac{A}{2\alpha_4}} |t|^{\frac{1}{2}}. \quad (125)$$

This gives the critical exponent

$$\beta = \frac{1}{2}. \quad (126)$$

- Magnetic susceptibility.

Differentiating (112) with respect to h gives for \mathbf{r} -independent solutions

$$2\alpha_2 \frac{\partial \phi}{\partial h} + 12\alpha_4 \frac{\partial \phi}{\partial h} \phi^2 = 1. \quad (127)$$

The zero-field susceptibility is thus

$$\chi = \left. \frac{\partial \phi}{\partial h} \right|_{h=0} = \frac{1}{2\alpha_2 + 12\alpha_4 \phi^2}. \quad (128)$$

Using that $\alpha_2 = At$ and $\phi_0^2 = A|t|/2\alpha_4$ this becomes

$$\chi = \begin{cases} \frac{1}{2At} & \text{if } t > 0, \\ \frac{1}{4A|t|} & \text{if } t < 0. \end{cases} \quad (129)$$

This gives the critical exponents

$$\gamma_{\pm} = 1. \quad (130)$$

- Heat capacity.

The heat capacity is defined by

$$C(T, h = 0) = -T \frac{\partial^2 F}{\partial T^2}. \quad (131)$$

The saddle point contribution to the free energy is

$$\frac{F}{V} \sim \begin{cases} 0 & \text{if } t > 0, \\ -\frac{A^2 k_B T_c t^2}{4\alpha_4} & \text{if } t < 0, \end{cases} \quad (132)$$

giving

$$C(T, h = 0) \sim \begin{cases} 0 & \text{if } t > 0, \\ \frac{A^2 k_B T_c^2 V}{2\alpha_4} & \text{if } t < 0, \end{cases} \quad (133)$$

This has a finite jump at T_c , which corresponds to the critical exponent

$$\alpha = 0. \quad (134)$$

- Correlation length exponent.

The exponents described above can all be obtained from the saddle point solution, or equivalently the Landau free energy. This is not the case for the *correlation length exponent* ν , which is related to fluctuations around the saddle point.

Away from the critical point the (connected) order-parameter two-point function decays exponentially with distance

$$\langle \phi(\mathbf{r})\phi(\mathbf{0}) \rangle_{\beta} - \langle \phi(\mathbf{r}) \rangle_{\beta} \langle \phi(\mathbf{0}) \rangle_{\beta} = e^{-|\mathbf{r}|/\xi}, \quad |\mathbf{r}| \rightarrow \infty. \quad (135)$$

The correlation length ξ diverges when the critical point is approached

$$\xi \sim |t|^{-\nu}. \quad (136)$$

The relation (136) defines the exponent ν . We now determine ν in what is known as the *Gaussian approximation*. In the disordered phase this amounts to simply dropping the ϕ^4 term in the free energy density, i.e. by setting

$$\langle \phi(\mathbf{r})\phi(\mathbf{0}) \rangle_{\beta} \approx \frac{1}{Z} \int \mathcal{D}\phi \phi(\mathbf{r})\phi(\mathbf{0}) \exp \left(- \int d^D \mathbf{r}' \left[\frac{1}{2} |\nabla \phi(\mathbf{r}')|^2 + \alpha_2 \phi^2(\mathbf{r}') \right] \right). \quad (137)$$

It is not a priori clear that the Gaussian approximation will give a good account of the two point function. It turns out to be good if the spatial dimensionality D is sufficiently high.

The two-point function (137) can be calculated using a *generating functional*

$$Z[h] = \int \mathcal{D}\phi \exp \left(- \int d^D \mathbf{r}' \left[\frac{1}{2} |\nabla \phi(\mathbf{r}')|^2 + \alpha_2 \phi^2(\mathbf{r}') - h(\mathbf{r}')\phi(\mathbf{r}') \right] \right). \quad (138)$$

Indeed, we have

$$\langle \phi(\mathbf{r})\phi(\mathbf{0}) \rangle_\beta = \frac{\delta}{\delta h(\mathbf{r})} \frac{\delta}{\delta h(\mathbf{0})} \Big|_{h=0} \ln(Z[h]). \quad (139)$$

We calculate the generating functional by going to Fourier space

$$\phi(\mathbf{r}) = \int \frac{d^D \mathbf{p}}{(2\pi)^D} e^{-i\mathbf{p}\cdot\mathbf{r}} \tilde{\phi}(\mathbf{p}), \quad h(\mathbf{r}) = \int \frac{d^D \mathbf{p}}{(2\pi)^D} e^{-i\mathbf{p}\cdot\mathbf{r}} \tilde{h}(\mathbf{p}). \quad (140)$$

This gives

$$\beta\mathcal{H} = \int \frac{d^D \mathbf{p}}{(2\pi)^D} \left[\left(\frac{\mathbf{p}^2}{2} + \alpha_2 \right) \tilde{\phi}(\mathbf{p})\tilde{\phi}(-\mathbf{p}) - h(\mathbf{p})\tilde{\phi}(-\mathbf{p}) \right]. \quad (141)$$

Next we “complete the square” by changing variables to

$$\tilde{\varphi}(\mathbf{p}) = \tilde{\phi}(\mathbf{p}) - \frac{\tilde{h}(\mathbf{p})}{\mathbf{p}^2 + 2\alpha_2}. \quad (142)$$

As the Jacobian of the change of variables is 1, this gives

$$Z[h] = \int \mathcal{D}\tilde{\varphi} \exp \left(- \int \frac{d^D \mathbf{p}}{(2\pi)^D} \left[\frac{\mathbf{p}^2}{2} + \alpha_2 \right] \tilde{\varphi}(\mathbf{p})\tilde{\varphi}(-\mathbf{p}) \right) \exp \left(\frac{1}{2} \int \frac{d^D \mathbf{p}}{(2\pi)^D} \frac{\tilde{h}(\mathbf{p})\tilde{h}(-\mathbf{p})}{\mathbf{p}^2 + 2\alpha_2} \right). \quad (143)$$

The first factor is merely a constant, which we will denote by \mathcal{N} , while the second factor is rewritten as

$$\begin{aligned} Z[h] &= \mathcal{N} \exp \left(\frac{1}{2} \int d^D \mathbf{r} d^D \mathbf{r}' h(\mathbf{r}) G(\mathbf{r} - \mathbf{r}') h(\mathbf{r}') \right), \\ G(\mathbf{r}) &= \int \frac{d^D \mathbf{p}}{(2\pi)^D} \frac{e^{-i\mathbf{p}\cdot\mathbf{r}}}{\mathbf{p}^2 + 2\alpha_2}. \end{aligned} \quad (144)$$

Taking functional derivatives we have

$$\langle \phi(\mathbf{r})\phi(\mathbf{0}) \rangle_\beta = G(\mathbf{r}) \sim \frac{e^{-|\mathbf{r}|\sqrt{2\alpha_2}}}{|\mathbf{r}|^{\frac{D-1}{2}}}, \quad |\mathbf{r}| \rightarrow \infty. \quad (145)$$

This gives the correlation length

$$\xi = \frac{1}{\sqrt{2\alpha_2}} \sim \frac{1}{t^{\frac{1}{2}}}, \quad (146)$$

and thus the critical exponent

$$\boxed{\nu = \frac{1}{2}}. \quad (147)$$

Given the explicit calculation we have just done, we are now in the position to introduce a shortcut for obtaining the two point function in the Gaussian approximation in similar situations. In absence of a source $h(\mathbf{r})$ the Landau-Ginzburg Hamiltonian (and the Landau free energy βF for that matter) is written as

$$\beta\mathcal{H} = \frac{1}{2} \int \frac{d^D \mathbf{p}}{(2\pi)^D} \left[(\mathbf{p}^2 + 2\alpha_2) \tilde{\phi}(\mathbf{p})\tilde{\phi}(-\mathbf{p}) + \text{quartic} \right]. \quad (148)$$

From this expression we can simply *read off* the result for the two point function in the Gaussian approximation

$$\langle \tilde{\phi}(\mathbf{p})\tilde{\phi}(\mathbf{q}) \rangle_\beta = \frac{(2\pi)^D \delta^D(\mathbf{p} - \mathbf{q})}{\mathbf{p}^2 + 2\alpha_2}. \quad (149)$$

Here the denominator is given by whatever factor multiplies the quadratic term in $\beta\mathcal{H}$. The delta-function expresses momentum conservation.

In the ordered phase $t < 0$ we expand $f[\phi(\mathbf{r})]$ around one of the minima at $\pm\phi_0$. The choice of minimum implements *spontaneous symmetry breaking*. We have

$$V(\phi) = \alpha_2\phi^2 + \alpha_4\phi^4 \simeq \alpha_2\phi_0^2 + \alpha_4\phi_0^4 + (\alpha_2 + 6\alpha_4\phi_0^2)(\phi - \phi_0)^2 + \dots \quad (150)$$

We may drop the constant and retain only the contribution quadratic in $\delta\phi = \phi - \phi_0$ (this is the Gaussian approximation in the ordered phase), which gives

$$f[\delta\phi(\mathbf{r})] \simeq \frac{1}{2}|\nabla\delta\phi|^2 + \tilde{\alpha}_2\delta\phi^2. \quad (151)$$

Here $\tilde{\alpha}_2 = -2\alpha_2 > 0$. We may now copy the calculation in the disordered phase and obtain for the *connected correlation function*

$$\langle\delta\phi(\mathbf{r})\delta\phi(\mathbf{0})\rangle_\beta = \langle\phi(\mathbf{r})\phi(\mathbf{0})\rangle_\beta - \langle\phi(\mathbf{r})\rangle_\beta\langle\phi(\mathbf{0})\rangle_\beta \sim \frac{e^{-|\mathbf{r}|\sqrt{-4\alpha_2}}}{|\mathbf{r}|^{\frac{D-1}{2}}}, \quad |\mathbf{r}| \rightarrow \infty. \quad (152)$$

We see that the correlation length scales as $\xi \propto |t|^{-1/2}$, giving again $\nu = 1/2$.

4.4 OTHER EXAMPLES OF PHASE TRANSITIONS

4.4.1 ISOTROPIC-NEMATIC TRANSITION IN LIQUID CRYSTALS

Liquid crystals are fluids of rod-like molecules. At high temperatures their centres of mass are randomly distributed and the rods are randomly oriented. At low temperatures, in the *nematic phase* the rods spontaneously align along a common axis, see Fig. 13. What is the order parameter characterizing this

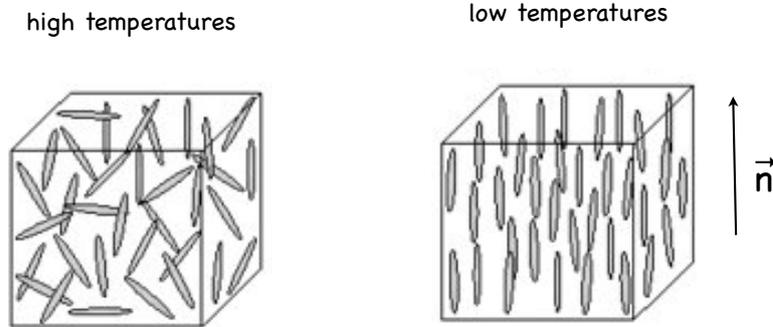


Figure 13: At low temperatures the rod-like molecules spontaneously align along an axis \vec{n} .

transition?

Let us associate a unit vector $\vec{n}(\mathbf{r})$ with a molecule at position \mathbf{r} . The first guess one may have is to take $\langle\vec{n}(\mathbf{r})\rangle_\beta$ as the order parameter. This will not work, because the two vectors $\vec{n}(\mathbf{r})$ and $-\vec{n}(\mathbf{r})$ describe the same orientation of the molecule. Hence the order parameter must be invariant under

$$\vec{n}(\mathbf{r}) \longrightarrow -\vec{n}(\mathbf{r}). \quad (153)$$

So how about something quadratic like

$$\langle n_i(\mathbf{r})n_j(\mathbf{r}) \rangle_\beta. \quad (154)$$

The problem with this expression is that it is different from zero even for randomly oriented molecules (which is what one has at very high temperatures). Indeed, using a parametrization of the unit vector for a single molecule in terms of polar coordinates we have

$$\vec{n} = \begin{pmatrix} \sin \theta \cos \varphi \\ \sin \theta \sin \varphi \\ \cos \theta \end{pmatrix}. \quad (155)$$

Then averaging over all possible orientations gives

$$\langle \langle n_i n_j \rangle \rangle = \frac{1}{4\pi} \int_0^\pi d\theta \int_0^{2\pi} d\varphi \sin \theta n_i n_j = \frac{1}{3} \delta_{i,j} \neq 0. \quad (156)$$

This consideration suggests to try

$$Q_{ij} = \langle n_i n_j \rangle_\beta - \frac{1}{3} \delta_{i,j} \quad (157)$$

as our order parameter. At very high temperatures, when molecules have random orientations, this is zero. On the other hand, if the molecules are aligned in the z -direction, i.e. $\vec{n} = \vec{e}_z$, we have

$$Q = \begin{pmatrix} -\frac{1}{3} & 0 & 0 \\ 0 & -\frac{1}{3} & 0 \\ 0 & 0 & \frac{2}{3} \end{pmatrix}. \quad (158)$$

So this seems to work. Given the order parameter, how do we then determine the free energy? In the high temperature phase the free energy must be invariant under rotations of the molecules, i.e. under

$$Q(\mathbf{r}) \longrightarrow RQ(\mathbf{r})R^T. \quad (159)$$

This suggests the following expansion for the free energy density

$$f[Q(\mathbf{r})] = \frac{1}{2} |\nabla Q|^2 + \alpha_2 \text{Tr}[Q^2] + \alpha_3 \text{Tr}[Q^3] + \alpha_4 (\text{Tr}[Q^2])^2 + \dots \quad (160)$$

Here $|\nabla Q|^2 = \sum_{i,j,k=1}^3 (\partial_k Q_{ij})^2$. The presence of a cubic term suggests that the transition is first order, which is indeed correct.

4.4.2 SUPERFLUID TRANSITION IN WEAKLY INTERACTING BOSONS

Let us recall the second quantized Hamiltonian for weakly repulsive bosons

$$H = \int d^4\mathbf{r} \left[c^\dagger(\mathbf{r}) \left(-\frac{\nabla^2}{2m} \right) c(\mathbf{r}) + \frac{U}{2} c^\dagger(\mathbf{r}) c^\dagger(\mathbf{r}) c(\mathbf{r}) c(\mathbf{r}) \right]. \quad (161)$$

In the superfluid phase we have macroscopic occupation of the zero momentum single-particle state

$$\langle c(\mathbf{p} = 0) \rangle_\beta = \sqrt{N_0}. \quad (162)$$

This implies that

$$\psi(\mathbf{r}) = \langle c(\mathbf{r}) \rangle_\beta \neq 0, \quad T < T_c. \quad (163)$$

At high temperatures the expectation value vanishes, so that we may take $\psi(\mathbf{r})$ as a *complex valued order parameter*. The Landau free energy describing the transition can then be constructed by noting that it should be invariant under $U(1)$ transformations

$$\psi(\mathbf{r}) \longrightarrow e^{i\theta} \psi(\mathbf{r}). \quad (164)$$

This gives

$$F[\psi] = \int d^D \mathbf{r} \left[\frac{1}{2m} |\nabla \psi|^2 + \alpha_2 |\psi(\mathbf{r})|^2 + \alpha_4 |\psi|^4 \right]. \quad (165)$$

4.5 ISING MODEL AND Φ^4 THEORY

The 3-dimensional Ising model has a paramagnet to ferromagnet phase transition that can be described by Landau theory. An obvious question is then whether it is possible to *derive* the appropriate Landau theory starting from the lattice model. Let us therefore consider a general Ising model with Hamiltonian

$$H = - \sum_{i,j} \sigma_i \tilde{C}_{ij} \sigma_j - \tilde{h} \sum_i \sigma_i. \quad (166)$$

Let's work on a hypercubic lattice in D dimensions, and take

$$\tilde{C}_{ij} = \tilde{C}(|\mathbf{r}_i - \mathbf{r}_j|) \quad (167)$$

to be translationally invariant. The partition function is

$$Z = \sum_{\{\sigma_j\}} e^{\sum_{j,\ell} \sigma_j C_{j\ell} \sigma_\ell + h \sum_\ell \sigma_\ell}, \quad (168)$$

where we have defined $h = \beta \tilde{h}$ and $C_{ij} = \beta \tilde{C}_{ij}$. Recalling the Gaussian integration identity

$$1 = \underbrace{\sqrt{\det(4\pi C^{-1})}}_{\mathcal{N}} \int \underbrace{\prod_{\ell} d\psi'_\ell}_{\mathcal{D}\psi'} e^{-\frac{1}{4} \sum_{j,k} \psi'_j C_{jk}^{-1} \psi'_k}, \quad (169)$$

and then shifting the integration variables

$$\psi'_j = \psi_j - 2 \sum_k C_{jk} \sigma_k, \quad (170)$$

we arrive at the following identity

$$1 = \mathcal{N} \int \mathcal{D}\psi e^{-\frac{1}{4} \sum_{j,k} \psi_j C_{jk}^{-1} \psi_k + \sum_j \sigma_j \psi_j - \sum_{j,k} \sigma_j C_{jk} \sigma_k}. \quad (171)$$

Substituting this into our expression for the partition function gives

$$Z = \mathcal{N} \int \mathcal{D}\psi \sum_{\{\sigma_j\}} e^{-\frac{1}{4} \sum_{j,k} \psi_j C_{jk}^{-1} \psi_k + \sum_j \sigma_j (\psi_j + h)}. \quad (172)$$

What we have done is to remove the interaction between the spins by introducing auxiliary variables ψ_j at every site. This is an example of a *Hubbard-Stratonovich transformation*. Now we can perform the summations over spin variables

$$\sum_{\{\sigma_j\}} e^{\sum_k \sigma_k (\psi_k + h)} = \prod_k [2 \cosh(\psi_k + h)]. \quad (173)$$

Thus

$$Z = \mathcal{N} \int \mathcal{D}\psi e^{-\frac{1}{4} \sum_{j,k} \psi_j C_{jk}^{-1} \psi_k + \sum_j \ln(2 \cosh(h + \psi_j))} \quad (174)$$

Finally we change variables to

$$\phi_j = \frac{1}{2} \sum_k C_{jk}^{-1} (\psi_k + h), \quad (175)$$

which leads to the following expression for the partition function

$$Z = \mathcal{N}' \int \mathcal{D}\phi e^{-\sum_{j,k} \phi_j C_{jk} \phi_k + h \sum_j \phi_j + \sum_j \ln(\cosh(2 \sum_k C_{jk} \phi_k))}. \quad (176)$$

This is still at the level of the lattice model. We now want to focus on large distances/small energies, in order to obtain a field theory description. We furthermore will now assume that the matrix C , which depends on temperature, is such that large fluctuations of ϕ_j are suppressed. We then can take ϕ_j to be small and expand

$$\ln[\cosh(2 \underbrace{\sum_k C_{jk} \phi_k}_X)] \simeq \frac{X^2}{2} - \frac{X^4}{12} + \dots \quad (177)$$

Finally, we go to Fourier space using

$$\phi_j = \frac{1}{N} \sum_{\mathbf{k}} e^{-i\mathbf{k} \cdot \mathbf{r}_j} \phi(\mathbf{k}), \quad C_{jk} = \frac{1}{N} \sum_{\mathbf{k}} e^{-i\mathbf{k} \cdot (\mathbf{r}_j - \mathbf{r}_k)} C(\mathbf{k}), \quad (178)$$

and focus on the small- $|\mathbf{k}|$ regime. This gives

$$Z = \mathcal{N}'' \int \mathcal{D}\phi e^{-\frac{1}{N} \sum_{\mathbf{k}} \phi(\mathbf{k}) [c_1 + c_2 \mathbf{k}^2 + \dots] \phi(-\mathbf{k}) + c_3 h \phi(\mathbf{0}) + \frac{c_4}{N^3} \sum_{\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3} \phi(\mathbf{k}_1) \phi(\mathbf{k}_2) \phi(\mathbf{k}_3) \phi(-\mathbf{k}_1 - \mathbf{k}_2 - \mathbf{k}_3) + \dots}, \quad (179)$$

where the coefficients c_l depend on the matrix \tilde{C}_{jk} specifying the interactions of Ising spins. Going back to real space we recover Euclidean Φ^4 theory

$$Z = \mathcal{N}''' \int \mathcal{D}\phi e^{-\int d^D \mathbf{x} [c_1 (\nabla \phi(\mathbf{x}))^2 + c_2 \phi^2(\mathbf{x}) + c_3 h \phi(\mathbf{x}) + c_4 \phi^4(\mathbf{x})]}. \quad (180)$$

Voilà. Note however, that the relation of the field $\phi(\mathbf{x})$ to the original Ising spin is rather indirect. In order to get a better feeling of “what happens to the spin” under all the above transformations, we could repeat the above steps for the generating functional rather than the partition function. This will show us that Φ is indeed a good order parameter for the Ising transition.

Exercises:

Question 4. Consider a Landau expansion of the free energy of the form

$$F = \frac{a}{2} m^2 + \frac{b}{4} m^4 + \frac{c}{6} m^6$$

with $c > 0$. Examine the phase diagram in the $a - b$ plane, and show that there is a line of critical transitions $a = 0, b > 0$ which joins a line of first order transitions $b = -4(ca/3)^{1/2}$ at a point $a = b = 0$ known as a tricritical point.

Supposing that a varies linearly with temperature and that b is independent of temperature, compare the value of the exponent β at the tricritical point with its value on the critical line.

From Yeomans, *Statistical Mechanics of Phase Transitions*

Question 5.

- (a) Discuss how an order parameter may be used to characterise symmetry breaking at a phase transition.
 (b) Argue that the uniaxial ferromagnet-paramagnet transition can be described by a Landau free energy of the form

$$F = \int d^3\mathbf{r} \left[\frac{1}{2} |\nabla\phi(\mathbf{r})|^2 - h\phi(\mathbf{r}) + \alpha_2\phi^2(\mathbf{r}) + \alpha_3\phi^3(\mathbf{r}) + \alpha_4\phi^4(\mathbf{r}) \right]. \quad (181)$$

What can you say about α_4 ?

- (c) What is the nature of the transition for $h = 0$ if $\alpha_3 \neq 0$? Explain your answer.
 (d) Now assume that $\alpha_3 = h = 0$. Argue that close to the critical point

$$\alpha_2 = At, \quad t = \frac{T - T_c}{T_c} \text{ and } A > 0. \quad (182)$$

- (e) Derive the equation characterizing the saddle point solution for $\alpha_3 = h = 0$. What are the configurations ϕ with the lowest free energy for $h = 0$, at $T > T_c$ and at $T < T_c$? Why are these \mathbf{r} independent?
 (f) Now consider more general solutions to the saddle point equation in the low-temperature phase. With suitable boundary conditions the saddle point solutions for the order parameter are functions of x only, i.e. $\phi = \phi(x)$. Show that in this case

$$E = \frac{1}{2} \left[\frac{d\phi(x)}{dx} \right]^2 - \alpha_2\phi^2 - \alpha_4\phi^4 \quad (183)$$

is independent of x . Construct a solution $\phi(x)$ such that

$$\lim_{x \rightarrow \infty} \phi(x) = \phi_1, \quad \lim_{x \rightarrow -\infty} \phi(x) = \phi_2, \quad (184)$$

where $\phi_{1,2}$ are the solutions found in (d). Hint: determine E for such solutions first.

Question 6. A system with a real, two-component order parameter $(\phi_1(\mathbf{r}), \phi_2(\mathbf{r}))$ has a free energy

$$F = \int d^d\mathbf{r} \left[\frac{1}{2} |\nabla\phi_1(\mathbf{r})|^2 + \frac{1}{2} |\nabla\phi_2(\mathbf{r})|^2 - \frac{1}{2} (\phi_1^2(\mathbf{r}) + \phi_2^2(\mathbf{r})) + \frac{1}{4} (\phi_1^2(\mathbf{r}) + \phi_2^2(\mathbf{r}))^2 \right].$$

Find the order-parameter values Φ_1, Φ_2 that minimise this free energy. Consider small fluctuations around such state, with $(\phi_1(\mathbf{r}), \phi_2(\mathbf{r})) = (\Phi_1 + \varphi_1(\mathbf{r}), \Phi_2 + \varphi_2(\mathbf{r}))$ and expand F to second order in φ .

Assuming that the statistical weight of thermal fluctuations is proportional to $\exp(-F)$, calculate approximately the correlation function

$$\langle \varphi_1(\mathbf{r})\varphi_1(\mathbf{0}) + \varphi_2(\mathbf{r})\varphi_2(\mathbf{0}) \rangle$$

by evaluating a Gaussian functional integral. How does your result depend on the dimensionality d of the system?

Part II**RANDOM SYSTEMS AND STOCHASTIC PROCESSES**

In nature there are many phenomena, in which some quantity varies in a random way. An example is *Brownian motion*, which refers to the motion of a small particle suspended in a fluid. The motion, observed under a microscope, looks *random*. It is hopeless to try to compute the position in detail, but certain *average features* obey simple laws. Averaging over a suitable time interval is difficult and one therefore replaces time averaging of a single irregularly varying function of time by averaging over an *ensemble of functions*. The latter must of course be chosen in such a way that the two results agree

$$\text{time average} \longrightarrow \text{ensemble average}. \quad (185)$$

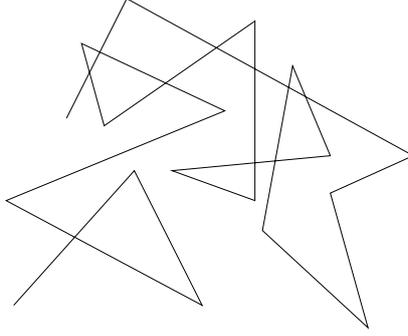


Figure 14: Cartoon of a path of a particle undergoing Brownian motion.

5 RANDOM VARIABLES

A random variable is an object X defined by

1. a set $\{x_j\}$ of possible values (either discrete or continuous);
2. a probability distribution $P_X(x_j)$ over this set

$$P_X(x_j) \geq 0, \quad \sum_j P_X(x_j) = 1 \text{ (this becomes an integral in the continuous case.)} \quad (186)$$

Example: Let X be the number of points obtained by casting a die

$$\{x_j\} = \{1, 2, 3, 4, 5, 6\}, \quad P_X(x_j) = \frac{1}{6}. \quad (187)$$

5.1 SOME DEFINITIONS

The probability distribution $P_X(x)$ can be characterized by the moments of X

$$\langle X^n \rangle \equiv \int dx x^n P_X(x). \quad (188)$$

The average value of X is the first moment $\langle X \rangle$, while the variance is $\sigma^2 = \langle X^2 \rangle - \langle X \rangle^2$.

The Fourier transform of $P_X(x)$ is called the characteristic function

$$\phi_X(k) = \int dx P_X(x) e^{ikx} = \sum_{n=0}^{\infty} \frac{(ik)^n}{n!} \langle X^n \rangle \equiv \langle e^{ikX} \rangle. \quad (189)$$

The last equality shows that the characteristic function is the generating function for the moments

$$\langle X^n \rangle = (-i)^n \left. \frac{d^n}{dk^n} \right|_{k=0} \phi_X(k). \quad (190)$$

The cumulants of $P_X(x)$ are defined as

$$C_n(X) = (-i)^n \left. \frac{d^n}{dk^n} \right|_{k=0} \ln(\phi_X(k)). \quad (191)$$

The first few cumulants are

$$C_1(X) = \langle X \rangle, \quad C_2(X) = \langle X^2 \rangle - \langle X \rangle^2, \quad C_3(X) = \langle X^3 \rangle - 3\langle X^2 \rangle \langle X \rangle + 2\langle X \rangle^3. \quad (192)$$

The Gaussian distribution is defined by

$$P_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-x_0)^2}{2\sigma^2}}. \quad (193)$$

Its characteristic function is

$$\phi_X(k) = e^{ikx_0 - \frac{1}{2}k^2\sigma^2}. \quad (194)$$

The cumulants are $C_1(X) = x_0$, $C_2(X) = \sigma^2$, $C_{n>2}(X) = 0$. Hence the Gaussian distribution is completely determined by its first two cumulants.

5.2 DISCRETE-TIME RANDOM WALK

To see the above definitions in action we consider a tutor on (without loss of generality) his way back to Summertown after a long night out (needless to say it must be a humanities tutor). He moves along Banbury road, by making each second a step forward/backward with equal probability. Modelling Banbury road by a line, his possible positions are all integers $-\infty < n < \infty$ (assuming for simplicity that Banbury road is infinitely long, which is probably how it feels to our tutor), and we want to know the probability $P_N(n)$ for her to be at position n after N steps, starting from $n = 0$.

- Each step is a random variable X_j ($j = 1, \dots, N$) taking the values ± 1 with probabilities $1/2$.
- The position after r steps is $Y = \sum_{j=1}^N X_j$.

Clearly we have

$$\langle Y \rangle = 0, \quad \langle Y^2 \rangle = N\langle X_j^2 \rangle = N, \quad (195)$$

where we have used that the steps are mutually independent. To obtain $p_N(n)$ we employ the characteristic function

$$\phi_{X_j}(k) = P_{X_j}(1)e^{ik1} + P_{X_j}(-1)e^{ik(-1)} = \cos(k). \quad (196)$$

The characteristic function of the random variable Y is

$$\begin{aligned} \phi_Y(k) &= \langle e^{ikY} \rangle = \langle e^{ik \sum_{j=1}^N X_j} \rangle = \prod_{j=1}^N \langle e^{ikX_j} \rangle = (\phi_{X_j}(k))^N \\ &= \frac{1}{2^N} (e^{ik} + e^{-ik})^N = \frac{1}{2^N} \sum_{r=0}^N \binom{N}{r} e^{ik(N-2r)}. \end{aligned} \quad (197)$$

On the other hand we have by definition of the characteristic function

$$\phi_Y(k) = \sum_n P_Y(n) e^{ikn} = \sum_{n=-N}^N p_N(n) e^{ikn}. \quad (198)$$

Equating (197) with (198) gives

$$p_N(n) = \frac{1}{2^N} \binom{N}{\frac{N-n}{2}}, \quad (199)$$

where the binomial coefficient is taken to be zero if $(N-n)/2$ is not an integer between 0 and N .

5.3 THE CENTRAL LIMIT THEOREM

Let X_j $j = 1, \dots, N$ be independent random variables with identical distributions $P_X(x)$ and consider their arithmetic mean

$$S_N = \frac{1}{N} \sum_{j=1}^N X_j. \quad (200)$$

Then for large N , S_N tends towards a *Gaussian*, irrespective of the form of $P_X(x)$, provided that $\langle X \rangle$ and $\langle X^2 \rangle$ are finite.

To see that this is the case, let us consider the characteristic function

$$\phi_{S_N}(k) = \langle e^{ikS_N} \rangle = \langle e^{i\frac{k}{N} \sum_{j=1}^N X_j} \rangle = \langle e^{ik\frac{X}{N}} \rangle^N = \left(\phi_X\left(\frac{k}{N}\right) \right)^N, \quad (201)$$

where we have used that the random variables are independent and have the same distribution. The cumulants of S_N are

$$C_n(S_N) = (-i)^n \left. \frac{d^n}{dk^n} \right|_{k=0} N \ln \left(\phi_X\left(\frac{k}{N}\right) \right). \quad (202)$$

We see that these cumulants are related to the cumulants $C_n(X)$ of $P_X(x)$ by

$$\boxed{C_n(S_N) = N^{1-n} C_n(X).} \quad (203)$$

Hence, for very large N $P_{S_N}(s)$ tends to a Gaussian distribution with average $\langle X \rangle$ and variance $C_2(X)/N = \sigma^2/N$, i.e.

$$\boxed{P_{S_N}(s) \longrightarrow \left(\frac{N}{2\pi\sigma^2} \right)^{\frac{1}{2}} \exp \left[-\frac{N}{2\sigma^2} (s - \langle X \rangle)^2 \right].} \quad (204)$$

6 STOCHASTIC PROCESSES: DEFINITIONS

A function $Y_X(t)$ of time t and a random variable X is called a *stochastic process* (SP). Examples are the position $x(t)$ or the velocity $v(t)$ of a particle undergoing Brownian motion. A SP is characterized by probability densities

$$\begin{aligned} P_1(y_1, t_1) &= \text{probability that } Y(t_1) = y_1; \\ &\vdots \\ P_n(y_1, t_1; \dots; y_n, t_n) &= \text{joint probability that } Y(t_1) = y_1, \dots, Y(t_n) = y_n. \end{aligned} \quad (205)$$

These are

- normalized

$$\int dy_1 \dots dy_n P_n(y_1, t_1; \dots; y_n, t_n) = 1; \quad (206)$$

- reducible

$$\int dy_n P_n(y_1, t_1; \dots; y_n, t_n) = P_{n-1}(y_1, t_1; \dots; y_{n-1}, t_{n-1}). \quad (207)$$

Another way of characterizing a SP is through *correlation functions*

$$\begin{aligned} \langle Y(t_1) \rangle &= \int dy_1 y_1 P(y_1, t_1), \\ \langle Y(t_1)Y(t_2) \rangle &= \int dy_1 dy_2 y_1 y_2 P_2(y_1, t_1; y_2, t_2), \\ &\dots \end{aligned} \quad (208)$$

A stochastic process is called *stationary* if for all n and any t

$$P_n(y_1, t_1; \dots; y_n, t_n) = P_n(y_1, t_1 + \tau; \dots; y_n, t_n + \tau). \quad (209)$$

The *conditional probability* that $Y = y_2$ at time t_2 , given that $Y = y_1$ at time t_1 is denoted by

$$P_{1|1}(y_2, t_2|y_1, t_1). \quad (210)$$

This is normalized

$$\int dy_2 P_{1|1}(y_2, t_2|y_1, t_1) = 1, \quad (211)$$

and related to unconditional probabilities by

$$P_2(y_1, t_1; y_2, t_2) = P_{1|1}(y_2, t_2|y_1, t_1) P_1(y_1, t_1). \quad (212)$$

The *conditional probability* that $Y = y_n$ at time t_n , given that $Y = y_j$ at time t_j for $j = 1, \dots, n-1$ is denoted by

$$P_{1|n-1}(y_n, t_n|y_1, t_1; y_2, t_2; \dots; y_{n-1}, t_{n-1}). \quad (213)$$

7 MARKOV PROCESSES

Perhaps the most important stochastic processes are so-called *Markov processes* (MP). Their defining property is that $t_1 < t_2 < \dots < t_n$

$$\boxed{P_{1|n-1}(y_n, t_n|y_1, t_1; y_2, t_2; \dots; y_{n-1}, t_{n-1}) = P_{1|1}(y_n, t_n|y_{n-1}, t_{n-1})}. \quad (214)$$

This means that at time t_{n-1} one can predict the state of the system at time t_n on the basis of *present* information, i.e. y_{n-1} , only! The history of how the system arrived at y_{n-1} at time t_{n-1} is irrelevant.

A Markov process is completely determined by the two functions $P_1(y_1, t_1)$ and $P_{1|1}(y_2, t_2|y_1, t_1)$, and this makes Markov processes tractable. Any $P_1(y_1, t_1)$ and $P_{1|1}(y_2, t_2|y_1, t_1)$ define a MP, provided that they fulfil the following two consistency conditions

- CHAPMAN-KOLMOGOROV EQUATION

$$\boxed{P_{1|1}(y_3, t_3|y_1, t_1) = \int dy_2 P_{1|1}(y_3, t_3|y_2, t_2) P_{1|1}(y_2, t_2|y_1, t_1), \quad t_3 > t_2 > t_1}. \quad (215)$$

- EVOLUTION EQUATION

$$\boxed{P_1(y_2, t_2) = \int dy_1 P_{1|1}(y_2, t_2|y_1, t_1) P_1(y_1, t_1), \quad t_2 > t_1}. \quad (216)$$

A MP is called *stationary* if $P_1(y, t)$ is time-independent and $P_{1|1}(y_2, t_2|y_1, t_1)$ depends only on the time difference $t_2 - t_1$ (and $y_{1,2}$).

7.1 EXAMPLES OF MARKOV PROCESSES

- The discrete-time random walk is a MP.
- The *Wiener process*, defined by

$$\begin{aligned} P_{1|1}(y_2, t_2|y_1, t_1) &= \frac{1}{\sqrt{2\pi(t_2 - t_1)}} e^{-\frac{(y_2 - y_1)^2}{2(t_2 - t_1)}}, \\ P_1(y, t) &= \frac{1}{\sqrt{2\pi t}} e^{-\frac{y^2}{2t}}, \end{aligned} \quad (217)$$

is a non-stationary Markov process. It was originally invented for describing the stochastic behaviour of the position of a Brownian particle.

7.2 MARKOV CHAINS

A Markov chain is a MP, in which the random variable only takes a finite number of values and involves discrete time steps. The probability $P_1(y, t)$ can then be represented as a N -component vector, and $P_{1|1}(y_2, t_2|y_1, t_1) \equiv T$ as an $N \times N$ matrix. The normalization condition (211) becomes

$$\boxed{\sum_{j=1}^N T_{jk} = 1,} \quad (218)$$

which is often referred to as probability conservation.

Example: Two state process with $Y = 1$ or $Y = 2$, with conditional probabilities

$$\begin{aligned} P_{1|1}(1, t+1; 1, t) &= 1 - q, \\ P_{1|1}(2, t+1; 1, t) &= q, \\ P_{1|1}(1, t+1; 2, t) &= r, \\ P_{1|1}(2, t+1; 2, t) &= 1 - r. \end{aligned} \quad (219)$$

Introducing a two-component vector

$$\vec{p}(t) = \begin{pmatrix} P_1(1, t) \\ P_1(2, t) \end{pmatrix}, \quad (220)$$

the evolution equation for the process can be expressed as a vector equation

$$\boxed{\vec{p}(t+1) = \underbrace{\begin{pmatrix} 1-q & r \\ q & 1-r \end{pmatrix}}_T \vec{p}(t).} \quad (221)$$

T is a square matrix with non-negative entries, that is in general not symmetric. The matrix elements T_{ij} are the rates for transitions from state j to state i . It is useful to rewrite the equation in components

$$p_n(t+1) = \sum_j T_{nj} p_j(t). \quad (222)$$

Then

$$p_n(t+1) - p_n(t) = \sum_j T_{nj} p_j(t) - \underbrace{\sum_j T_{jn}}_1 p_n(t) = \sum_j T_{nj} p_j(t) - T_{jn} p_n(t). \quad (223)$$

Now consider our time interval to be small instead of 1. Then (223) turns into a differential equation, called a *Master equation*

$$\boxed{\frac{dp_n(t)}{dt} = \sum_j T_{nj} p_j(t) - T_{jn} p_n(t).} \quad (224)$$

This has a nice physical interpretation as a “loss/gain” equation for probabilities: the first term on the right-hand side is the rate of transitions from state j to state n times the probability of j being realized, i.e. the total gain of probability for state n . The second term on the right-hand side is the rate of transitions from state n to state j times the probability of n being realized, i.e. the total loss of probability for state n . Let us now return to the discrete form (221), which can be iterated to give

$$\vec{p}(t+1) = T^{t+1} \vec{p}(0). \quad (225)$$

While T is generally not symmetric, it is nevertheless often diagonalizable. Then there exist left and right eigenvectors such that

$$T|R_\alpha\rangle = \lambda_\alpha|R_\alpha\rangle, \quad \langle L_\alpha|T = \lambda_\alpha\langle L_\alpha|, \quad \langle L_\alpha|R_\beta\rangle = \delta_{\alpha,\beta}, \quad (226)$$

and T can be represented in the form

$$T = \sum_{\alpha} \lambda_{\alpha} |R_{\alpha}\rangle\langle L_{\alpha}|. \quad (227)$$

Using the orthonormality of left and right eigenvectors we have

$$\boxed{T^{t+1} = \sum_{\alpha} \lambda_{\alpha}^{t+1} |R_{\alpha}\rangle\langle L_{\alpha}|.} \quad (228)$$

In our example

$$\lambda_1 = 1, \quad \langle L_1| = (1, 1), \quad |R_1\rangle = \frac{1}{1+r/q} \begin{pmatrix} r/q \\ 1 \end{pmatrix}. \quad (229)$$

$$\lambda_2 = 1 - q - r, \quad \langle L_2| = \left(-\frac{q}{r}, 1\right), \quad |R_2\rangle = \frac{1}{1+r/q} \begin{pmatrix} -1 \\ 1 \end{pmatrix}. \quad (230)$$

So for large t we have

$$T^{t+1} \approx |R_1\rangle\langle L_1| = \frac{1}{r+q} \begin{pmatrix} r & r \\ q & q \end{pmatrix}, \quad (231)$$

and hence

$$\vec{p}(\infty) = \frac{1}{r+q} \begin{pmatrix} r & r \\ q & q \end{pmatrix} \vec{p}(0). \quad (232)$$

8 BROWNIAN MOTION

We now want to think of Brownian motion as a Markov process. Let v_1, v_2, \dots be the velocities of the particle at different time steps. Then v_{k+1} depends only on v_k , but not on v_1, \dots, v_{k-1} .

8.1 LANGEVIN EQUATION

One approach to Brownian motion is via a *stochastic differential equation*, the *Langevin equation* for the velocity $v(t)$ (more precisely the velocity in $D = 1$ or a component of the velocity in $D > 1$)

$$\boxed{\frac{dv(t)}{dt} = -\gamma v(t) + \eta(t).} \quad (233)$$

Here the first term on the right hand side is damping term linear in v , while the second term represents the remaining *random force* with zero average $\langle \eta(t) \rangle = 0$. This is often referred to as “noise”. For simplicity we will assume collisions to be instantaneous, so that forces at different times are uncorrelated

$$\langle \eta(t)\eta(t') \rangle = \Gamma \delta(t - t'). \quad (234)$$

Given our assumptions about the noise, we can calculate noise-averaged quantities quite easily. We have

$$\frac{d}{dt'} \left(v(t') e^{\gamma t'} \right) = \left(\frac{dv}{dt'} + \gamma v \right) e^{\gamma t'} = \eta(t') e^{\gamma t'}, \quad (235)$$

where in the last step we used the Langevin equation (233). Integrating both sides of this equation between $t = 0$ and $t' = t$, we obtain

$$\boxed{v(t) = v(0)e^{-\gamma t} + \int_0^t dt' \eta(t')e^{-\gamma(t-t')}}. \quad (236)$$

Averaging this over the noise, we find the average velocity

$$\boxed{\langle v(t) \rangle = v(0)e^{-\gamma t} + \int_0^t dt' \underbrace{\langle \eta(t') \rangle}_{=0} e^{-\gamma(t-t')} = v(0)e^{-\gamma t}}. \quad (237)$$

Similarly we have

$$\begin{aligned} \langle v^2(t) \rangle &= \left\langle \left[v(0)e^{-\gamma t} + \int_0^t dt' \eta(t')e^{-\gamma(t-t')} \right] \left[v(0)e^{-\gamma t} + \int_0^t dt'' \eta(t'')e^{-\gamma(t-t'')} \right] \right\rangle \\ &= v^2(0)e^{-2\gamma t} + e^{-2\gamma t} \int_0^t dt' dt'' \underbrace{\langle \eta(t')\eta(t'') \rangle}_{\Gamma\delta(t'-t'')} e^{\gamma(t'+t'')} \end{aligned} \quad (238)$$

Carrying out the time integrals this becomes

$$\boxed{\langle v^2(t) \rangle = v^2(0)e^{-2\gamma t} + \frac{\Gamma}{2\gamma}(1 - e^{-2\gamma t})}. \quad (239)$$

The displacement of the particle is

$$x(t) - x(0) = \int_0^t dt' v(t') = \frac{v(0)}{\gamma}(1 - e^{-\gamma t}) + \int_0^t dt' \int_0^{t'} dt'' \eta(t'')e^{-\gamma(t'-t'')}. \quad (240)$$

Assuming $x(0) = 0$, the average position of the particle is

$$\boxed{\langle x(t) \rangle = \frac{v(0)}{\gamma}(1 - e^{-\gamma t})}. \quad (241)$$

Finally, we want to determine the particle's mean square deviation

$$\langle [x(t) - \langle x(t) \rangle]^2 \rangle = \langle x^2(t) \rangle - \langle x(t) \rangle^2. \quad (242)$$

Substituting (240) and using again that $x(0) = 0$ this becomes

$$\int_0^t dt_1 \int_0^{t_1} dt_2 e^{-\gamma(t_1-t_2)} \underbrace{\int_0^t dt'_1 \int_0^{t'_1} dt'_2 e^{-\gamma(t'_1-t'_2)} \langle \eta(t'_2)\eta(t'_1) \rangle}_{\Gamma\Theta(t'_1-t'_2)e^{-\gamma(t'_1-t'_2)}} = \int_0^t dt_1 \int_0^{t_1} dt_2 e^{-\gamma(t_1-2t_2)} \underbrace{\int_{t_2}^t dt'_1 e^{-\gamma t'_1}}_{\gamma^{-1}(e^{-\gamma t_2} - e^{-\gamma t})}. \quad (243)$$

Carrying out the remaining two integrals we find

$$\boxed{\langle [x(t) - \langle x(t) \rangle]^2 \rangle = \frac{\Gamma}{\gamma^2}t - \frac{\Gamma}{\gamma^3}(1 - e^{-\gamma t}) - \frac{\Gamma}{2\gamma^3}(1 - e^{-\gamma t})^2}. \quad (244)$$

So at very late times we have

$$\boxed{\langle x^2(t) \rangle = \frac{\Gamma}{\gamma^2}t + \dots} \quad (245)$$

The displacement grows like \sqrt{t} , which is characteristic of *diffusion*. Finally, we may relate Γ/γ to the temperature of the fluid by noting that

$$\langle v^2(t \rightarrow \infty) \rangle = \frac{\Gamma}{2\gamma}. \quad (246)$$

On the other hand, by equipartition we have

$$\frac{m}{2} \langle v^2 \rangle \sim \frac{k_B T}{2} \quad (247)$$

Combining these two equations, we arrive at

$$\boxed{\frac{\Gamma}{2\gamma} = \frac{k_B T}{m}}. \quad (248)$$

8.2 FOKKER-PLANCK EQUATION

We now want to derive a differential equation for the probability $P_1(v, t)$ of our particle having velocity v at time t from the Langevin equation

$$v(t) = v(0)e^{-\gamma t} + \int_0^t dt' e^{-\gamma(t-t')} \eta(t'). \quad (249)$$

Our starting point is the general evolution equation

$$P_1(v, t + \Delta t) = \int du P_{1|1}(v, t + \Delta t | u, t) P_1(u, t). \quad (250)$$

It is convenient to consider the integral

$$\Omega = \int dv [P_1(v, t + \Delta t) - P_1(v, t)] h(v), \quad (251)$$

where $h(v)$ is is test function (infinitely many time differentiable, $h(v)$ and all of its derivatives going to zero at infinity etc). On the one hand, we have to linear order in Δt

$$\int dv [P_1(v, t + \Delta t) - P_1(v, t)] h(v) = \int dv \frac{\partial P_1(v, t)}{\partial t} \Delta t h(v). \quad (252)$$

On the other hand, using (250) we have

$$\Omega = \int du dv h(v) P_{1|1}(v, t + \Delta t | u, t) P_1(u, t) - \int dv h(v) P_1(v, t). \quad (253)$$

Relabelling the integration variable from v to u in the second term, and using that normalization condition $\int dv P_{1|1}(v, t + \Delta t | u, t) = 1$, we obtain

$$\Omega = \int du P_1(u, t) \int dv P_{1|1}(v, t + \Delta t | u, t) [h(v) - h(u)]. \quad (254)$$

Expanding $h(v)$ around u in a Taylor series gives

$$\Omega = \int du P_1(u, t) \sum_{n=1}^{\infty} h^{(n)}(u) \underbrace{\int dv P_{1|1}(v, t + \Delta t | u, t) \frac{(v-u)^n}{n!}}_{D^{(n)}(u)}. \quad (255)$$

Integrating the n'th term in the sum n times by parts, and using the nice properties of the function $h(u)$, then leads to the following expression

$$\Omega = \int du h(u) \sum_{n=1}^{\infty} \left(-\frac{\partial}{\partial u}\right)^n P_1(u, t) D^{(n)}(u). \quad (256)$$

Using that (252) and (256) have to be equal for *any* test function $h(u)$, we conclude that

$$\boxed{\frac{\partial P_1(v, t)}{\partial t} \Delta t = \sum_{n=1}^{\infty} \left(-\frac{\partial}{\partial v}\right)^n P_1(v, t) D^{(n)}(v).} \quad (257)$$

This starts looking like our desired differential equation. What remains is to determine the quantities $D^{(n)}(v)$

$$\begin{aligned} D^{(n)}(v) &= \int dw P_{1|1}(w, t + \Delta t | v, t) \frac{(w - v)^n}{n!} = \int dz P_{1|1}(v + z, t + \Delta t | v, t) \frac{z^n}{n!} \\ &= \frac{1}{n!} \langle [v(t + \Delta t) - v(t)]^n \rangle. \end{aligned} \quad (258)$$

We see that $D^{(n)}(v)$ are related to the moments of the velocity difference distribution! We can use the Langevin equation to determine them, and the result is

$$\begin{aligned} \langle v(t + \Delta t) - v(t) \rangle &= -\gamma v(t) e^{-\gamma \Delta t}, \\ \langle (v(t + \Delta t) - v(t))^2 \rangle &= \Gamma \Delta t + \mathcal{O}((\Delta t)^2), \\ \langle (v(t + \Delta t) - v(t))^n \rangle &= \mathcal{O}((\Delta t)^2), \quad n \geq 3. \end{aligned} \quad (259)$$

Substituting these into (257) and then taking the limit $\Delta t \rightarrow 0$, we arrive at the *Fokker-Planck equation*

$$\boxed{\frac{\partial}{\partial t} P_1(v, t) = \gamma \frac{\partial}{\partial v} v P_1(v, t) + \frac{\Gamma}{2} \frac{\partial^2}{\partial v^2} P_1(v, t).} \quad (260)$$

This is a second order linear PDE for $P_1(v, t)$ and can be solved by standard methods. For initial conditions $P_1(v, 0) = \delta(v - v_0)$ we find

$$\boxed{P_1(v, t) = \frac{1}{\sqrt{2\pi\sigma^2(t)}} \exp\left(-\frac{(v - \bar{v}(t))^2}{2\sigma^2(t)}\right),} \quad (261)$$

where

$$\sigma^2(t) = \frac{\Gamma}{2\gamma} (1 - e^{-2\gamma t}), \quad \bar{v}(t) = v_0 e^{-\gamma t}. \quad (262)$$

In the limit $t \rightarrow \infty$ this turns into the *Maxwell distribution*

$$P_1(v, t) = \sqrt{\frac{\gamma}{\pi\Gamma}} e^{-\frac{\gamma v^2}{\Gamma}} = \sqrt{\frac{m}{2\pi k_B T}} e^{-\frac{mv^2}{2k_B T}}. \quad (263)$$

8.2.1 MOMENTS OF THE VELOCITY DIFFERENCE PROBABILITY DISTRIBUTION

Let us see how to derive (259), starting from the Langevin equation

$$v(t) = v(0) e^{-\gamma t} + \int_0^t dt' e^{-\gamma(t-t')} \eta(t'). \quad (264)$$

For a very small time interval Δt we have

$$\begin{aligned}
v(t + \Delta t) &= v(0)e^{-\gamma(t+\Delta t)} + \int_0^{t+\Delta t} dt' e^{-\gamma(t+\Delta t-t')} \eta(t') \\
&= v(0)e^{-\gamma t} (1 - \gamma \Delta t) + \int_0^t dt' e^{-\gamma(t-t')} \eta(t') (1 - \gamma \Delta t) + \int_t^{t+\Delta t} dt' e^{-\gamma(t-t')} \eta(t') + \mathcal{O}((\Delta t)^2) \\
&= (1 - \gamma \Delta t) v(t) + \int_t^{t+\Delta t} dt' e^{-\gamma(t-t')} \eta(t') + \mathcal{O}((\Delta t)^2).
\end{aligned} \tag{265}$$

Hence $\Delta v(t) = v(t + \Delta t) - v(t)$ is given by

$$\Delta v = -\gamma v \Delta t + \int_t^{t+\Delta t} dt' e^{-\gamma(t-t')} \eta(t') + \mathcal{O}((\Delta t)^2). \tag{266}$$

To derive (259) we rewrite (266) in the form

$$\Delta v(t) = v_0(t) \Delta t + \int dt' K(t, t') \eta(t'), \tag{267}$$

where

$$\begin{aligned}
v_0(t) &= -\gamma v(0) e^{-\gamma t}, \\
K(t, t') &= -\gamma \Theta(t - t') \Theta(t') e^{-\gamma(t-t')} \Delta t + \Theta(t' - t) \Theta(t + \Delta t - t') e^{-\gamma(t-t')}.
\end{aligned} \tag{268}$$

Let us now assume that $\eta(t)$ is Gaussian distributed. Then the probability distribution for the noise is the functional

$$P[\eta(t)] = e^{-\frac{1}{2\Gamma} \int dt \eta^2(t)}. \tag{269}$$

Averages are then given by the path integral

$$\langle \eta(t_1) \dots \eta(t_n) \rangle = \int \mathcal{D}\eta(t) \eta(t_1) \dots \eta(t_n) e^{-\frac{1}{2\Gamma} \int dt' \eta^2(t')}. \tag{270}$$

As this is Gaussian, we may use Wick's theorem to calculate averages

$$\begin{aligned}
\langle \eta(t) \rangle &= 0, \\
\langle \eta(t) \eta(t') \rangle &= \Gamma \delta(t - t'), \\
\langle \eta(t_1) \eta(t_2) \eta(t_3) \rangle &= \langle \eta(t_1) \eta(t_2) \rangle \langle \eta(t_3) \rangle + \langle \eta(t_3) \eta(t_1) \rangle \langle \eta(t_2) \rangle + \langle \eta(t_2) \eta(t_3) \rangle \langle \eta(t_1) \rangle = 0, \\
\langle \eta(t_1) \eta(t_2) \eta(t_3) \eta(t_4) \rangle &= \langle \eta(t_1) \eta(t_2) \rangle \langle \eta(t_3) \eta(t_4) \rangle + \dots
\end{aligned} \tag{271}$$

The probability distribution for $\Delta v(t)$ can be obtained from the generating function

$$Z(\lambda) = \langle e^{\lambda \Delta v(t)} \rangle = e^{\lambda v_0(t) \Delta t} \int \mathcal{D}\eta(t) e^{-\frac{1}{2\Gamma} \int dt' [\eta^2(t') - 2\lambda \Gamma K(t, t') \eta(t')]}. \tag{272}$$

Changing variables to

$$\tilde{\eta}(t') = \eta(t') - \lambda \Gamma K(t, t'), \tag{273}$$

the generating function becomes

$$Z(\lambda) = e^{\lambda v_0(t) \Delta t + \frac{\lambda^2 \Gamma}{2} \int dt' K^2(t, t')}. \tag{274}$$

It is then a straightforward matter to calculate the moments

$$\begin{aligned}
\langle \Delta v \rangle &= \left. \frac{dZ}{d\lambda} \right|_{\lambda=0} = v_0(t) \Delta t, \\
\langle (\Delta v)^2 \rangle &= \left. \frac{d^2 Z}{d\lambda^2} \right|_{\lambda=0} = \Gamma \int dt' K^2(t, t') = \Gamma \int_t^{t+\Delta t} dt' e^{-\gamma(t-t')} + \mathcal{O}((\Delta t)^2) \\
&= \Gamma \Delta t + \mathcal{O}((\Delta t)^2), \\
\langle (\Delta v)^n \rangle &= \left. \frac{d^n Z}{d\lambda^n} \right|_{\lambda=0} = \mathcal{O}((\Delta t)^2), \quad n \geq 3.
\end{aligned} \tag{275}$$

8.3 DIFFUSION EQUATION

Finally, we would like to obtain a differential equation for $P_1(x, t)$. The difficulty is that $x(t)$ is *not* a Markov process, because $x(t + \Delta t)$ depends on both $x(t)$ and $v(t)$. One can treat the combined evolution of $x(t)$ and $v(t)$, but we will follow a simpler route. Recalling that

$$v(t) = v(0)e^{-\gamma t} + \int_0^t dt' e^{-\gamma(t-t')} \eta(t'), \quad (276)$$

we see that $v(t)$ becomes a random variable for $t \gg \gamma^{-1}$ with

$$\langle v(t) \rangle = 0, \quad \langle v(t)v(t') \rangle = \frac{\Gamma}{2\gamma} e^{-\gamma|t-t'|}. \quad (277)$$

Now let us imagine that we observe the Brownian particle only at sufficiently long time intervals $t, t' \gg \gamma^{-1}$, and describe only these coarse grained positions. Then we may replace

$$\langle v(t)v(t') \rangle = \frac{\Gamma}{2\gamma} e^{-\gamma|t-t'|} \rightarrow \frac{\Gamma}{\gamma^2} \delta(t-t'). \quad (278)$$

This is because $\frac{\Gamma}{2\gamma} e^{-\gamma|t-t'|}$ is substantially different from zero only if $|t-t'| < \gamma^{-1}$, and

$$\int_{-\infty}^{\infty} dt \frac{\Gamma}{2\gamma} e^{-\gamma|t-t'|} = \frac{\Gamma}{\gamma^2}. \quad (279)$$

The differential equation for the position

$$\frac{dx(t)}{dt} = v(t), \quad (280)$$

then turns into a special case (of no damping) of the Langevin equation we solved for the velocity. We therefore can use our previous results to conclude that

$$\boxed{\frac{\partial}{\partial t} P_1(x, t) = \frac{\Gamma}{2\gamma^2} \frac{\partial^2}{\partial x^2} P_1(x, t).} \quad (281)$$

This is the *diffusion equation* with diffusion coefficient

$$D = \frac{\Gamma}{2\gamma^2}. \quad (282)$$

Its solution for initial conditions $P_1(x, 0) = \delta(x - x_0)$ is

$$\boxed{P_1(x, t) = \frac{1}{\sqrt{2\pi D|t-t_0|}} \exp\left(-\frac{(x-x_0)^2}{4D|t-t_0|}\right).} \quad (283)$$

Exercises:

Question 7. This question is concerned with the central limit theorem.

(i) Show explicitly that for $N \gg 1, pN \gg 1$ the binomial distribution

$$P_N(n) = \frac{N!}{n!(N-n)!} p^n q^{N-n}, \quad p + q = 1$$

becomes

$$P_N(n) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(n - \langle n \rangle)^2}{2\sigma^2}\right)$$

where $\sigma^2 = Npq$. Check that the same result follows from the central limit theorem.

(ii) Consider a random walk in one dimension, for which the probability of taking a step of length $x \rightarrow x+dx$ is

$$f(x)dx = \frac{1}{\pi} \frac{\gamma}{x^2 + \gamma^2} dx.$$

Find the probability distribution for the total displacement after N steps. Does it satisfy the central limit theorem? Should it? What are the cumulants of this distribution?

Question 8. Let $y = \pm 1$. Show that

$$P_{1|1}(y, t | y', t') = \frac{1}{2} \left\{ 1 + e^{-2\gamma(t-t')} \right\} \delta_{y,y'} + \frac{1}{2} \left\{ 1 - e^{-2\gamma(t-t')} \right\} \delta_{y,-y'} \quad (284)$$

obeys the Chapman-Kolmogorov equation.

Show that

$$P_1(y, t) = \frac{1}{2} (\delta_{y,1} + \delta_{y,-1}) \quad (285)$$

is a stationary solution. Write $P_{1|1}$ as a 2×2 matrix and formulate the Chapman-Kolmogorov equation as a property of that matrix.

Question 9. This question is about a continuous random walk, also known as a Wiener process.

Show that for $-\infty < y < \infty$ and $t_2 > t_1$ the Chapman-Kolmogorov equation is satisfied for

$$P_{1|1}(y_2, t_2 | y_1, t_1) = \frac{1}{\sqrt{2\pi(t_2 - t_1)}} \exp\left\{-\frac{(y_2 - y_1)^2}{2(t_2 - t_1)}\right\}. \quad (286)$$

Choose $P_1(y_1, 0) = \delta(y_1)$. Show that for $t > 0$

$$P_1(y, t) = \frac{1}{\sqrt{2\pi t}} \exp\left\{-\frac{y^2}{2t}\right\}. \quad (287)$$

Show that $P_1(y, t)$ obeys the diffusion equation

$$\frac{\partial P}{\partial t} = D \frac{\partial^2 P}{\partial y^2} \quad (288)$$

for $D = \frac{1}{2}$. What is the solution for arbitrary $D > 0$?

Question 10. A particle suspended in a fluid undergoes Brownian motion in one dimension with position $x(t)$ and velocity $v(t)$. This motion is modelled by the Langevin equation

$$\frac{dv}{dt} = -\gamma v + \eta(t),$$

where $\eta(t)$ is a Gaussian random variable characterised completely by the averages $\langle \eta(t) \rangle = 0$ and $\langle \eta(t_1)\eta(t_2) \rangle = \Gamma\delta(t_1 - t_2)$. Discuss the physical origin of each of the terms in the Langevin equation.

What is meant by the term *Markov process*? Illustrate your answer by discussing which of the following are Markov processes: (a) $v(t)$ alone; (b) $x(t)$ alone; (c) $v(t)$ and $x(t)$ together.

Show that for $t > 0$

$$x(t) = \frac{v(0)}{\gamma} (1 - e^{-\gamma t}) + \int_0^t dt_1 \int_0^{t_1} dt_2 e^{-\gamma(t_1-t_2)} \eta(t_2)$$

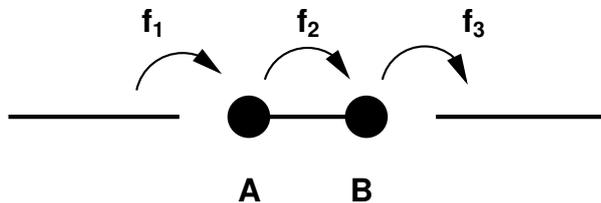
is a solution of the Langevin equation with initial condition $x(0) = 0$. Calculate the average $\langle x(t) v(t) \rangle$ and discuss its limiting behaviour at short and long times.

Question 11. The time evolution of a stochastic system is represented by a master equation of the form

$$\frac{dp_n(t)}{dt} = \sum_m W_{nm} p_m(t).$$

Explain briefly the meaning of this equation and discuss the assumptions on which it is based. What general conditions should the matrix elements W_{nm} satisfy?

A molecule lies between two atomic-scale contacts and conducts charge between them. A simple model of this situation is illustrated below. The model has three states: the molecule may be uncharged, or may carry a single charge at either site A or site B but not both. Charges hop between these sites, and between the sites and the contacts, at the rates indicated in the figure. (For example, a charge at site A has probability f_2 per unit time of hopping to site B .)



Write down a master equation for this model. For the system in equilibrium, calculate the occupation probabilities of the three states, and show that the average number of charges flowing through the molecule per unit time is

$$\frac{f_1 f_2 f_3}{f_1 f_2 + f_1 f_3 + f_2 f_3}.$$

Consider the case $f_1 = f_2 = f_3 \equiv f$. The molecule is uncharged at time $t = 0$. Show that the probability $p(t)$ for it to be uncharged at a later time t is

$$p(t) = \frac{1}{3} + \frac{2}{3} \exp\left(-\frac{3}{2}ft\right) \cos\left(\frac{\sqrt{3}}{2}ft\right).$$