Lecture Notes for the C6 Theory Option

F.H.L. Essler
The Rudolf Peierls Centre for Theoretical Physics
Oxford University, Oxford OX1 3NP, UK

January 30, 2018

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

Contents

I Many-Particle Quantum Mechanics 3

1 Second Quantization 3
- 1.1 Systems of Independent Particles . . . 3
  - 1.1.1 Occupation Number Representation . . . 5
- 1.2 Fock Space . . . 5
  - 1.2.1 Creation and Annihilation Operators . . . 5
  - 1.2.2 Basis of the Fock Space . . . 6
- 1.3 Homework Questions 1-3 . . . 7
  - 1.3.1 Change of Basis . . . 8
- 1.4 Second Quantized Form of Operators . . . 8
  - 1.4.1 Occupation number operators . . . 9
  - 1.4.2 Single-particle operators . . . 9
  - 1.4.3 Two-particle operators . . . 12
- 1.5 Homework Question 4 . . . 14

2 Application I: The Ideal Fermi Gas 14
- 2.1 Quantization in a large, finite volume . . . 15
  - 2.1.1 Ground State . . . 16
  - 2.1.2 Excitations . . . 16
  - 2.1.3 Density Correlations . . . 17
- 2.2 Connection with Quantum Field Theory . . . 19
- 2.3 "Emergent" relativistic description at low energies . . . 20
- 2.4 Homework Questions 5-6 . . . 22

3 Linear Response Theory 23

4 Application II: Weakly Interacting Bosons 25
- 4.1 Ideal Bose Gas . . . 25
- 4.2 Bogoliubov Approximation . . . 26
- 4.3 Bogoliubov Transformation . . . 27
- 4.4 Ground State and Low-lying Excitations . . . 27
- 4.5 Ground state correlation functions . . . 28

4.6 Spontaneous Symmetry Breaking . . . 29
4.7 Depletion of the Condensate . . . 29

5 Application III: Spinwaves in a Ferromagnet 30
5.1 Heisenberg model and spin-rotational SU(2) symmetry . . . 30
5.2 Exact ground states . . . 31
5.3 Spontaneous Symmetry Breaking . . . 32
5.4 Holstein-Primakoff Transformation . . . 32
5.4.1 Heisenberg Antiferromagnet . . . 34
5.5 Homework Questions 7-8 . . . 35

II Phases and Phase Transitions 37

6 The Ising Model 38
6.1 Statistical mechanics of the Ising model . . . 38
6.2 The One-Dimensional Ising Model . . . 38
6.2.1 Transfer matrix approach . . . 39
6.2.2 Averages of observables in the transfer matrix formalism . . . 40
6.3 The Two-Dimensional Ising Model . . . 41
6.3.1 Transfer Matrix Method . . . 41
6.3.2 Spontaneous Symmetry Breaking . . . 43
6.4 Homework Questions 9-11 . . . 44
6.5 Peierls Argument . . . 45
6.6 Mean Field Theory . . . 48
6.7 Solution of the self-consistency equation for $h = 0$ . . . 49
6.8 Vicinity of the Phase Transition . . . 50

7 Critical Behaviour and Universality 50
7.1 Universality . . . 51

8 Landau Theory 52
8.1 Thermodynamic Equilibrium . . . 54
8.2 Beyond the Landau Free Energy . . . 55
8.3 Saddle Point Approximation . . . 56
8.4 Mean Field Exponents . . . 56
8.5 Homework Questions 12-14 . . . 62

9 Other Examples of Phase Transitions 63
9.1 Isotropic-Nematic Transition in Liquid Crystals . . . 63
9.2 Superfluid Transition in Weakly Interacting Bosons . . . 65

10 Regime of validity of the saddle point approximation/mean field theory 67

III Random Systems and Stochastic Processes 69

11 Random Variables 69
11.1 Some Definitions . . . 69
11.2 Discrete-time random walk . . . 70
11.3 The central limit theorem . . . 71

12 Stochastic Processes: Definitions 71

13 Markov Processes 72
13.1 Examples of Markov Processes . . . 73
13.2 Markov Chains . . . 73

14 Brownian Motion 74
14.1 Langevin Equation . . . 75
14.2 Fokker-Planck Equation . . . 76
14.2.1 Moments of the Velocity Difference Probability Distribution . . . 78
14.3 Diffusion Equation . . . 79
14.4 Homework Questions 12-16 . . . 80

Some general remarks:

These notes aim to be self-contained. Homework questions are marked in red, and are placed at appropriate positions in the text, i.e. to work them out you will require only the preceeding material. Passages marked in blue give details on derivations we don't have time to go through in the lectures, or present material that goes beyond the core of the course. In some cases this material will be very useful for particular homework problems. All of the material covered in the course can be found in some form or other in a variety of books. These is no book that covers everything. Some useful references are

- Many-particle QM

  R.P. Feynman, *Statistical Mechanics: A Set of Lectures*, Westview Press.

  A. Altland and B.D. Simons, *Condensed Matter Field Theory*, Cambridge.

- Landau Theory of Phase Transitions

  M. Kardar, *Statistical Physics of Fields*, Cambridge.

Part I
Many-Particle Quantum Mechanics

In the basic QM course you encountered only quantum systems with *very small numbers* of particles. In the harmonic oscillator problem we are dealing with a single QM particle, when solving the hydrogen atom we had one electron and one nucleus. Perhaps the most important field of application of quantum physics is to systems of *many* particles. Examples are the electronic degrees of freedom in solids, superconductors, trapped ultra-cold atomic gases, magnets and so on. The methods you have encountered in the basic QM course are not suitable for studying such problems. In this part of the course we introduce a framework, that will allow us to study the QM of many-particle systems. This new way of looking at things will also reveal very interesting connections to *Quantum Field Theory*.

1 Second Quantization

The formalism we develop in the following is known as *second quantization*.

1.1 Systems of Independent Particles

You already know from second year QM how to solve problems involving *independent particles*

$$H = \sum_{j=1}^{N} H_j \tag{1}$$

where H_j is the Hamiltonian on the j'th particle, e.g.

$$H_j = \frac{\hat{\mathbf{p}}_j^2}{2m} + V(\hat{\mathbf{r}}_j) = -\frac{\hbar^2}{2m}\nabla_j^2 + V(\hat{\mathbf{r}}_j). \tag{2}$$

The key to solving such problems is that $[H_j, H_l] = 0$. We'll now briefly review the necessary steps, switching back and forth quite freely between using states and operators acting on them, and the position representation of the problem (i.e. looking at wave functions).

- Step 1. Solve the single-particle problem

$$H_j|l\rangle = E_l|l\rangle \ . \tag{3}$$

The corresponding wave functions are

$$\phi_l(\mathbf{r}_j) = \langle \mathbf{r}_j|l\rangle. \tag{4}$$

The eigenstates form an orthonormal set

$$\langle l|m\rangle = \delta_{l,m} = \int d^D\mathbf{r}_j \ \phi_l^*(\mathbf{r}_j)\phi_m(\mathbf{r}_j). \tag{5}$$

- Step 2. Form N-particle eigenfunctions as *products*

$$\left(\sum_{j=1}^N H_j\right)\phi_{l_1}(\mathbf{r}_1)\phi_{l_2}(\mathbf{r}_2)\dots\phi_{l_N}(\mathbf{r}_N) = \left(\sum_{j=1}^N E_{l_j}\right)\phi_{l_1}(\mathbf{r}_1)\phi_{l_2}(\mathbf{r}_2)\dots\phi_{l_N}(\mathbf{r}_N) \ . \tag{6}$$

This follows from the fact that in the position representation H_j is a differential operator that acts only on the j'th position $\mathbf{r}_j$. The corresponding eigenstates are tensor products

$$|l_1\rangle \otimes |l_2\rangle \otimes \cdots \otimes |l_N\rangle. \tag{7}$$

- Step 3. Impose the appropriate *exchange symmetry* for indistinguishable particles, e.g.

$$\psi_{l,m}^{(\pm)}(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{\sqrt{2}}\left[\phi_l(\mathbf{r}_1)\phi_m(\mathbf{r}_2) \pm \phi_l(\mathbf{r}_2)\phi_m(\mathbf{r}_1)\right] \ , \quad l \neq m. \tag{8}$$

Generally we require

$$\psi(\dots, \mathbf{r}_i, \dots, \mathbf{r}_j, \dots) = \pm\psi(\dots, \mathbf{r}_j, \dots, \mathbf{r}_i, \dots) \ , \tag{9}$$

where the + sign corresponds to bosons and the − sign to fermions. This is achieved by taking

$$\boxed{\psi_{l_1\dots l_N}(\mathbf{r}_1, \dots, \mathbf{r}_N) = \mathcal{N}\sum_{P\in S_N}(\pm 1)^{|P|}\phi_{l_{P_1}}(\mathbf{r}_1)\dots\phi_{l_{P_N}}(\mathbf{r}_N),} \tag{10}$$

where the sum is over all permutations of $(1, 2, \dots, N)$ and $|P|$ is the number of pair exchanges required to reduce $(P_1, \dots, P_N)$ to $(1, \dots, N)$. The normalization constant $\mathcal{N}$ is

$$\mathcal{N} = \frac{1}{\sqrt{N!n_1!n_2!\dots}} \ , \tag{11}$$

where n_j is the number of times j occurs in the set $\{l_1, \dots, l_N\}$. For fermions the wave functions can be written as *Slater determinants*

$$\psi_{l_1\dots l_N}(\mathbf{r}_1, \dots, \mathbf{r}_N) = \frac{1}{\sqrt{N!}}\det\begin{pmatrix}\phi_{l_1}(\mathbf{r}_1) & \dots & \phi_{l_1}(\mathbf{r}_N)\\ \vdots & & \vdots \\ \phi_{l_N}(\mathbf{r}_1) & \dots & \phi_{l_N}(\mathbf{r}_N)\end{pmatrix}. \tag{12}$$

The states corresponding to (10) are

$$\boxed{|l_1, \dots, l_N\rangle = \mathcal{N}\sum_{P\in S_N}(\pm 1)^{|P|}|l_{P_1}\rangle \otimes \cdots \otimes |l_{P_N}\rangle \ .} \tag{13}$$

1.1.1 Occupation Number Representation

By construction the states have the symmetry

$$|l_{Q_1} \dots l_{Q_N}\rangle = \pm |l_1 \dots l_N\rangle \ , \tag{14}$$

where Q is an arbitrary permutation of $(1, \dots, N)$. As the overall sign of state is irrelevant, we can therefore choose them without loss of generality as

$$|\underbrace{1 \dots 1}_{n_1}\underbrace{2 \dots 2}_{n_2}\underbrace{3 \dots 3}_{n_3} 4 \dots \rangle \equiv |n_1 n_2 n_3 \dots \rangle. \tag{15}$$

In (15) we have as many n_j's as there are single-particle eigenstates, i.e. $\dim H$ [1]. For fermions we have $n_j = 0, 1$ only as a consequence of the *Pauli principle*. The representation (15) is called *occupation number representation*. The n_j's tell us how many particles are in the single-particle state $|j\rangle$. By construction the states $\{|n_1 n_2 n_3 \dots \rangle | \sum_j n_j = N\}$ form an orthonormal basis of our N-particle problem

$$\langle m_1 m_2 m_3 \dots | n_1 n_2 n_3 \dots \rangle = \prod_j \delta_{n_j, m_j} \ , \tag{16}$$

where we have defined $\langle m_1 m_2 m_3 \dots | = | m_1 m_2 m_3 \dots \rangle^\dagger$.

1.2 Fock Space

We now want to allow the particle number to vary. The main reason for doing this is that we will encounter physical problems where particle number is in fact not conserved. Another motivation is that experimental probes like photoemission change particle number, and we want to be able to describe these. The resulting space of states is called *Fock Space*.

1. The state with no particles is called the *vacuum state* and is denoted by $|0\rangle$.
2. N-particle states are $|n_1 n_2 n_3 \dots \rangle$ with $\sum_j n_j = N$.

1.2.1 Creation and Annihilation Operators

Given a basis of our space of states we can define operators by specifying their action on all basis states.

- particle creation operators with quantum number l

$$\boxed{c_l^\dagger |n_1 n_2 \dots \rangle = \begin{cases} 0 & \text{if } n_l = 1 \text{ for fermions} \\ \sqrt{n_l + 1} (\pm 1)^{\sum_{j=1}^{l-1} n_j} |n_1 n_2 \dots n_l + 1 \dots \rangle & \text{else.} \end{cases}} \tag{17}$$

  Here the + (−) sign applies to bosons (fermions).

- particle annihilation operators with quantum number l

$$\boxed{c_l |n_1 n_2 \dots \rangle = \sqrt{n_l} (\pm 1)^{\sum_{j=1}^{l-1} n_j} |n_1 n_2 \dots n_l - 1 \dots \rangle \ .} \tag{18}$$

We note that (18) follows from (17) by

$$\langle m_1 m_2 \dots | c_l^\dagger | n_1 n_2 \dots \rangle^* = \langle n_1 n_2 \dots | c_l | m_1 m_2 \dots \rangle \ . \tag{19}$$

[1] Note that this is different from the particle number N.

The creation and annihilation operators fulfil *canonical (anti)commutation relations*

$$\boxed{[c_l, c_m] = 0 = [c_l^\dagger, c_m^\dagger] \ , \quad [c_l, c_m^\dagger] = \delta_{l,m} \qquad \text{bosons,}} \tag{20}$$

$$\boxed{\{c_l, c_m\} = c_l c_m + c_m c_l = 0 = \{c_l^\dagger, c_m^\dagger\} \ , \quad \{c_l, c_m^\dagger\} = \delta_{l,m} \qquad \text{fermions.}} \tag{21}$$

Exercise 1: Proof of the anticommutations relations

Let us see how to prove these in the fermionic case. For $l < m$ we have

$$\begin{aligned} c_l^\dagger c_m |\dots n_l \dots n_m \dots\rangle &= c_l^\dagger \sqrt{n_m} (-1)^{\sum_{j=1}^{m-1} n_j} |\dots n_l \dots n_m - 1 \dots\rangle \\ &= \sqrt{n_l + 1}\sqrt{n_m} (-1)^{\sum_{j=l}^{m-1} n_j} |\dots n_l + 1 \dots n_m - 1 \dots\rangle. \end{aligned} \tag{22}$$

Similarly we have

$$c_m c_l^\dagger |\dots n_l \dots n_m \dots\rangle = \sqrt{n_l + 1}\sqrt{n_m} (-1)^{1 + \sum_{j=l}^{m-1} n_j} |\dots n_l + 1 \dots n_m - 1 \dots\rangle. \tag{23}$$

This means that for any basis state $|n_1 n_2 \dots\rangle$ we have

$$\{c_l^\dagger, c_m\} |n_1 n_2 \dots\rangle = 0 \ , \qquad \text{if } l > m. \tag{24}$$

This implies that

$$\{c_l^\dagger, c_m\} = 0 \ , \qquad \text{if } l > m. \tag{25}$$

The case $l < m$ works in the same way. This leaves us with the case $l = m$. Here we have

$$c_l^\dagger c_l |\dots n_l \dots n_m \dots\rangle = c_l^\dagger \sqrt{n_l} (-1)^{\sum_{j=1}^{l-1} n_j} |\dots n_l - 1 \dots\rangle = n_l |\dots n_l \dots\rangle. \tag{26}$$

$$\begin{aligned} c_l c_l^\dagger |\dots n_l \dots\rangle &= \begin{cases} c_l \sqrt{n_l + 1} (-1)^{\sum_{j=1}^{l-1} n_j} |\dots n_l + 1 \dots\rangle & \text{if } n_l = 0 \ , \\ 0 & \text{if } n_l = 1 \ , \end{cases} \\ &= \begin{cases} |\dots n_l \dots\rangle & \text{if } n_l = 0 \ , \\ 0 & \text{if } n_l = 1 \ , \end{cases} \end{aligned} \tag{27}$$

Combining these we find that

$$\{c_l^\dagger, c_l\} |\dots n_l \dots\rangle = |\dots n_l \dots\rangle \ , \tag{28}$$

and as the states $|\dots n_l \dots\rangle$ form a basis this implies

$$\{c_l^\dagger, c_l\} = 1. \tag{29}$$

Note that here 1 really means the identity operator $\mathbf{1}$.

1.2.2 Basis of the Fock Space

We are now in a position to write down our Fock space basis in a very convenient way.

- Fock vacuum (state without any particles)

$$|0\rangle. \tag{30}$$

- Single-particle states

$$|0\dots0\underbrace{1}_{l}0\dots\rangle = c_l^\dagger|0\rangle \ . \tag{31}$$

- N-particle states

$$|n_1 n_2 \dots\rangle = \prod_j \frac{1}{\sqrt{n_j!}}\left(c_j^\dagger\right)^{n_j}|0\rangle \ . \tag{32}$$

1.3 Homework Questions 1-3

Question 1. Consider a fermion 'system' with just one single-particle orbital, so that the only states of the system are $|0\rangle$ (unoccupied) and $|1\rangle$ (occupied). Show that we can represent the operators a and $a^\dagger$ by the matrices

$$a^\dagger = \begin{pmatrix} 0 & 0 \\ C & 0 \end{pmatrix}, \quad a = \begin{pmatrix} 0 & C^* \\ 0 & 0 \end{pmatrix}.$$

You can do this by checking the values of aa, $a^\dagger a^\dagger$ and $a^\dagger a + aa^\dagger$. What values may the constant C take?

Question 2. A quantum-mechanical Hamiltonian for a system of an even number N of point unit masses interacting by nearest-neighbour forces in one dimension is given by

$$H = \frac{1}{2}\sum_{r=1}^{N}\left(p_r^2 + (q_{r+1} - q_r)^2\right),$$

where the Hermitian operators q_r, p_r satisfy the commutation relations $[q_r, q_s] = [p_r, p_s] = 0$, $[q_r, p_s] = \mathrm{i}\delta_{rs}$, and where $q_{r+N} = q_r$. New operators Q_k, P_k are defined by

$$q_r = \frac{1}{\sqrt{N}}\sum_k Q_k \mathrm{e}^{\mathrm{i}kr} \qquad \text{and} \qquad p_r = \frac{1}{\sqrt{N}}\sum_k P_k \mathrm{e}^{-\mathrm{i}kr},$$

where $k = 2\pi n/N$ with $n = -N/2+1, \dots, 0, \dots, N/2$.

Show that:

(a) $Q_k = \dfrac{1}{\sqrt{N}}\displaystyle\sum_{s=1}^{N} q_s \mathrm{e}^{-\mathrm{i}ks}$ and $P_k = \frac{1}{\sqrt{N}}\sum_{s=1}^{N} p_s \mathrm{e}^{\mathrm{i}ks}$

(b) $[Q_k, P_{k'}] = \mathrm{i}\delta_{kk'}$

(c) $H = \frac{1}{2}\left(\sum_k P_k P_{-k} + \omega^2 Q_k Q_{-k}\right)$, where $\omega^2 = 2(1-\cos k)$.

Similarly to the treatment of the simple harmonic oscillator in QM I we then define annihilation operators a_k by

$$a_k = \frac{1}{(2\omega_k)^{1/2}}(\omega_k Q_k + \mathrm{i}P_{-k}).$$

Show that the Hermitian conjugate operators are

$$a_k^\dagger = \frac{1}{(2\omega_k)^{1/2}}(\omega_k Q_{-k} - \mathrm{i}P_k),$$

and determine the canonical commutation relations for a_k and $a_p^\dagger$. Construct the Fock space of states and determine the eigenstates and eigenvalues of H.

Question 3. Bosonic creation operators are defined through their action on basis states in the occupation number representation as

$$c_l^\dagger|n_1 n_2 \dots\rangle = \sqrt{n_l + 1}|n_1 n_2 \dots n_l + 1 \dots\rangle \ , \tag{33}$$

a) Deduce from this how bosonic annihilation operators act.
b) Show that the creation and annihilation operators fulfil *canonical commutation relations*

$$[c_l, c_m] = 0 = [c_l^\dagger, c_m^\dagger] \ , \quad [c_l, c_m^\dagger] = \delta_{l,m}. \tag{34}$$

1.3.1 Change of Basis

The Fock space is built from a given basis of single-particle states

$$\boxed{\text{single-particle states } |l\rangle} \longrightarrow \boxed{\text{N-particle states } |n_1 n_2 \dots\rangle} \longrightarrow \boxed{\text{Fock Space}}. \tag{35}$$

You know from second year QM that it is often convenient to switch from one basis to another, e.g. from energy to momentum eigenstates. This is achieved by a unitary transformation

$$\{|l\rangle\} \longrightarrow \{|\alpha\rangle\} \ , \tag{36}$$

where

$$|\alpha\rangle = \sum_l \underbrace{\langle l|\alpha\rangle}_{U_{l\alpha}} \ |l\rangle. \tag{37}$$

By construction

$$\sum_\alpha U_{l\alpha} U^\dagger_{\alpha m} = \sum_\alpha \langle l|\alpha\rangle\langle\alpha|m\rangle = \langle l|m\rangle = \delta_{lm}. \tag{38}$$

We now want to "lift" this unitary transformation to the level of the Fock space. We know that

$$\begin{aligned} |l\rangle &= c_l^\dagger|0\rangle \ , \\ |\alpha\rangle &= d_\alpha^\dagger|0\rangle \ . \end{aligned} \tag{39}$$

On the other hand we have

$$|\alpha\rangle = \sum_l U_{l\alpha}|l\rangle = \sum_l U_{l\alpha} c_l^\dagger|0\rangle. \tag{40}$$

This suggests that we take

$$\boxed{d_\alpha^\dagger = \sum_l U_{l\alpha} c_l^\dagger,} \tag{41}$$

and this indeed reproduces the correct transformation for N-particle states. Taking the hermitian conjugate we obtain the transformation law for annihilation operators

$$\boxed{d_\alpha = \sum_l U^\dagger_{\alpha l} c_l.} \tag{42}$$

We emphasize that these transformation properties are compatible with the (anti)commutation relations (as they must be). For fermions

$$\{d_\alpha, d_\beta^\dagger\} = \sum_{l,m} U^\dagger_{\alpha l} U_{m\beta} \underbrace{\{c_l, c_m^\dagger\}}_{\delta_{l,m}} = \sum_l U^\dagger_{\alpha l} U_{l\beta} = (U^\dagger U)_{\alpha\beta} = \delta_{\alpha,\beta}. \tag{43}$$

1.4 Second Quantized Form of Operators

In the next step we want to know how observables such as H, P, X etc act on the Fock space.

1.4.1 Occupation number operators

These are the simplest hermitian operators we can build from c_l and $c_m^\dagger$. They are defined as

$$\hat{n}_l \equiv c_l^\dagger c_l. \tag{44}$$

From the definition of c_l and $c_l^\dagger$ it follows immediately that

$$\hat{n}_l |n_1 n_2 \dots\rangle = n_l |n_1 n_2 \dots\rangle. \tag{45}$$

1.4.2 Single-particle operators

When acting on N-particle states *single-particle operators* can be written in the form

$$\hat{O} = \sum_j \hat{o}_j \ , \tag{46}$$

where the operator $\hat{o}_j$ acts only on the j'th particle. Examples are kinetic and potential energy operators

$$\hat{T} = \sum_j \frac{\hat{p}_j^2}{2m} \ , \qquad \hat{V} = \sum_j V(\hat{x}_j). \tag{47}$$

In terms of tensor-products (46) means that

$$\hat{O} = \sum_{j=1}^{N} \underbrace{\mathbf{1} \otimes \mathbf{1} \otimes \cdots \otimes \mathbf{1}}_{j-1} \otimes \hat{o} \otimes \mathbf{1} \otimes \cdots \otimes \mathbf{1} \ . \tag{48}$$

We want to represent $\hat{O}$ on the Fock space built from single-particle eigenstates $|\alpha\rangle$. We do this in two steps:

- Step 1: We first represent $\hat{O}$ in a basis of the Fock space built from the eigenstates of $\hat{o}$

  $$\hat{o}|l\rangle = \lambda_l |l\rangle = \lambda_l c_l^\dagger |0\rangle. \tag{49}$$

  Then, when acting on an N-particle state (13), we have

  $$\hat{O}|l_1, l_2, \dots, l_N\rangle = \left[\sum_{j=1}^{N} \lambda_j\right] |l_1, l_2, \dots, l_N\rangle. \tag{50}$$

  This is readily translated into the occupation number representation

  $$\hat{O}|n_1 n_2 \dots\rangle = \left[\sum_k n_k \lambda_k\right] |n_1 n_2 \dots\rangle. \tag{51}$$

  As $|n_1 n_2 \dots\rangle$ constitute a basis, this together with (45) imply that we can represent $\hat{O}$ in the form

  $$\hat{O} = \sum_k \lambda_k \hat{n}_k = \sum_k \lambda_k c_k^\dagger c_k. \tag{52}$$

- Step 2: Now that we have a representation of $\hat{O}$ in the Fock space basis built from the single-particle states $|l\rangle$, we can use a basis transformation to $\{|\alpha\rangle\}$ to obtain a representation in a general basis. Using that $\langle k|\hat{O}|k'\rangle = \delta_{k,k'}\lambda_k$ we can rewrite (52) in the form

  $$\hat{O} = \sum_{k,k'} \langle k'|\hat{O}|k\rangle c_{k'}^\dagger c_k. \tag{53}$$

Then we apply our general rules for a change of single-particle basis of the Fock space

$$c_k^\dagger = \sum_\alpha \langle \alpha | l \rangle d_\alpha^\dagger. \tag{54}$$

Substituting (54) and the analogous relation for annihilation operators into (53) we have

$$\hat{O} = \sum_{\alpha,\beta} \underbrace{\sum_{k'} \left(\langle \alpha | k' \rangle \langle k' | \right)}_{\langle \alpha |} \hat{O} \underbrace{\sum_k |k\rangle \langle k | \beta \rangle}_{|\beta\rangle} d_\alpha^\dagger d_\beta. \tag{55}$$

This gives us the final result

$$\boxed{\hat{O} = \sum_{\alpha,\beta} \langle \alpha | \hat{O} | \beta \rangle \; d_\alpha^\dagger d_\beta.} \tag{56}$$

We now work out a number of explicit examples of Fock space representations for single-particle operators.

1. Momentum Operators $\mathbf{P}$ in the infinite volume:

   (i) Let us first consider $\mathbf{P}$ in the single-particle basis of momentum eigenstates

   $$\hat{\mathbf{P}} |\mathbf{k}\rangle = \mathbf{k} |\mathbf{k}\rangle \; , \quad \langle \mathbf{p} | \mathbf{k} \rangle = (2\pi\hbar)^3 \delta^{(3)}(\mathbf{p} - \mathbf{k}). \tag{57}$$

   Aside 1: Remark

   These are shorthand notations for

   $$\hat{P}_a |k_x, k_y, k_z\rangle = k_a |k_x, k_y, k_z\rangle \; , \quad a = x, y, z. \tag{58}$$

   and

   $$\langle p_x, p_y, p_z | k_x, k_y, k_z \rangle = (2\pi\hbar)^3 \delta(k_x - p_x) \delta(k_y - p_y) \delta(k_z - p_z) \; . \tag{59}$$

   Using our general result for representing single-particle operators in a Fock space built from their eigenstates (52) we have

   $$\hat{\mathbf{P}} = \int \frac{d^3\mathbf{p}}{(2\pi\hbar)^3} \mathbf{p} c^\dagger(\mathbf{p}) c(\mathbf{p}) \; , \qquad [c^\dagger(\mathbf{k}), c(\mathbf{p})\} = (2\pi\hbar)^3 \delta^{(3)}(\mathbf{p} - \mathbf{k}). \tag{60}$$

   Here we have introduced a notation

   $$[c(\mathbf{k}), c^\dagger(\mathbf{p})\} = \begin{cases} c(\mathbf{k}) c^\dagger(\mathbf{p}) - c^\dagger(\mathbf{p}) c(\mathbf{k}) & \text{for bosons} \\ c(\mathbf{k}) c^\dagger(\mathbf{p}) + c^\dagger(\mathbf{p}) c(\mathbf{k}) & \text{for fermions.} \end{cases} \tag{61}$$

   (ii) Next we want to represent $\hat{\mathbf{P}}$ in the single-particle basis of position eigenstates

   $$\hat{\mathbf{X}} |\mathbf{x}\rangle = \mathbf{x} |\mathbf{x}\rangle \; , \quad \langle \mathbf{x} | \mathbf{x}' \rangle = \delta^{(3)}(\mathbf{x} - \mathbf{x}'). \tag{62}$$

   Our general formula (56) gives

   $$\hat{\mathbf{P}} = \int d^3\mathbf{x} d^3\mathbf{x}' \; \langle \mathbf{x}' | \hat{\mathbf{P}} | \mathbf{x} \rangle c^\dagger(\mathbf{x}') c(\mathbf{x}) \; . \tag{63}$$

   We can simplify this by noting that

   $$\langle \mathbf{x}' | \hat{\mathbf{P}} | \mathbf{x} \rangle = -i\hbar \nabla_{\mathbf{x}'} \delta^{(3)}(\mathbf{x} - \mathbf{x}'), \tag{64}$$

   which allows us to eliminate three of the integrals

   $$\hat{\mathbf{P}} = \int d^3\mathbf{x} d^3\mathbf{x}' \left[-i\hbar \nabla_{\mathbf{x}'} \delta^{(3)}(\mathbf{x} - \mathbf{x}') \right] c^\dagger(\mathbf{x}') c(\mathbf{x}) = \int d^3\mathbf{x} c^\dagger(\mathbf{x}) \left(-i\hbar \nabla_{\mathbf{x}} \right) c(\mathbf{x}). \tag{65}$$

2. Single-particle Hamiltonian:

$$H = \sum_{j=1}^{N} \frac{\hat{\mathbf{p}}_j^2}{2m} + V(\hat{\mathbf{x}}_j). \tag{66}$$

(i) Let us first consider H in the single-particle basis of energy eigenstates $H|l\rangle = E_l|l\rangle$, $|l\rangle = c_l^\dagger|0\rangle$. Our result (52) tells us that

$$H = \sum_l E_l c_l^\dagger c_l. \tag{67}$$

(ii) Next we consider the position representation, i.e. we take position eigenstates $|\mathbf{x}\rangle = c^\dagger(\mathbf{x})|0\rangle$ as a basis of single-particle states. Then by (56)

$$H = \int d^3\mathbf{x} d^3\mathbf{x}' \ \langle \mathbf{x}'|H|\mathbf{x}\rangle \ c^\dagger(\mathbf{x}')c(\mathbf{x}). \tag{68}$$

Substituting (66) into (68) and using

$$\langle \mathbf{x}'|V(\hat{\mathbf{x}})|\mathbf{x}\rangle = V(\mathbf{x})\delta^{(3)}(\mathbf{x}-\mathbf{x}') \ , \quad \langle \mathbf{x}'|\hat{\mathbf{p}}^2|\mathbf{x}\rangle = -\hbar^2\nabla^2\delta^{(3)}(\mathbf{x}-\mathbf{x}') \ , \tag{69}$$

we arrive at the position representation

$$\boxed{H = \int d^3\mathbf{x} \ c^\dagger(\mathbf{x}) \left[-\frac{\hbar^2\nabla^2}{2m} + V(\mathbf{x})\right] c(\mathbf{x}).} \tag{70}$$

(iii) Finally we consider the momentum representation, i.e. we take momentum eigenstates $|\mathbf{p}\rangle = c^\dagger(\mathbf{p})|0\rangle$ as a basis of single-particle states. Then by (56)

$$H = \int \frac{d^3\mathbf{p} d^3\mathbf{p}'}{(2\pi\hbar)^6} \ \langle \mathbf{p}'|H|\mathbf{p}\rangle \ c^\dagger(\mathbf{p}')c(\mathbf{p}). \tag{71}$$

Matrix elements of the kinetic energy operator are simple

$$\langle \mathbf{p}'|\hat{\mathbf{p}}^2|\mathbf{p}\rangle = \mathbf{p}^2\langle \mathbf{p}'|\mathbf{p}\rangle = \mathbf{p}^2(2\pi\hbar)^3\delta^{(3)}(\mathbf{p}-\mathbf{p}'). \tag{72}$$

Matrix elements of the potential can be calcuated as follows

$$\begin{aligned} \langle \mathbf{p}'|\hat{V}|\mathbf{p}\rangle &= \int d^3\mathbf{x} d^3\mathbf{x}' \ \langle \mathbf{p}'|\mathbf{x}'\rangle\langle \mathbf{x}'|\hat{V}|\mathbf{x}\rangle\langle \mathbf{x}|\mathbf{p}\rangle = \int d^3\mathbf{x} d^3\mathbf{x}' \ \underbrace{\langle \mathbf{x}'|\hat{V}|\mathbf{x}\rangle}_{V(\mathbf{x})\delta^{(3)}(\mathbf{x}-\mathbf{x}')} \ e^{\frac{i}{\hbar}\mathbf{p}\cdot\mathbf{x} - \frac{i}{\hbar}\mathbf{p}'\cdot\mathbf{x}'} \\ &= \int d^3\mathbf{x} \ V(\mathbf{x}) e^{\frac{i}{\hbar}(\mathbf{p}-\mathbf{p}')\cdot\mathbf{x}} = \tilde{V}(\mathbf{p}-\mathbf{p}'), \end{aligned} \tag{73}$$

where $\tilde{V}(\mathbf{p})$ is essentially the *three-dimensional Fourier transform* of the (ordinary) function $V(\mathbf{x})$. Hence

$$\boxed{H = \int \frac{d^3\mathbf{p}}{(2\pi\hbar)^3} \ \frac{\mathbf{p}^2}{2m} c^\dagger(\mathbf{p})c(\mathbf{p}) + \int \frac{d^3\mathbf{p} d^3\mathbf{p}'}{(2\pi\hbar)^6} \ \tilde{V}(\mathbf{p}-\mathbf{p}')c^\dagger(\mathbf{p}')c(\mathbf{p}).} \tag{74}$$

1.4.3 Two-particle operators

These are operators that act on two particles at a time. A good example is the interaction potential $V(\hat{\mathbf{r}}_1, \hat{\mathbf{r}}_2)$ between two particles at positions $\mathbf{r}_1$ and $\mathbf{r}_2$. For N particles we want to consider

$$\hat{V} = \sum_{i<j}^{N} V(\hat{\mathbf{r}}_i, \hat{\mathbf{r}}_j). \tag{75}$$

In the Fock space basis built from single-particle position eigenstates this is represented as

$$\boxed{\hat{V} = \frac{1}{2}\int d^3\mathbf{r} d^3\mathbf{r}' \; c^\dagger(\mathbf{r})c^\dagger(\mathbf{r}')V(\mathbf{r},\mathbf{r}')c(\mathbf{r}')c(\mathbf{r}).} \tag{76}$$

Note that when writing down the first quantized expression (75), we assumed that the operators acts specifically on states with N particles. On the other hand, (76) acts on the Fock space, i.e. on states where the particle number can take any value. The action of (76) on N-particle states (where N is fixed but arbitrary) is equal to the action of (75).

Aside 2: Derivation of (76)

Let us concentrate on the fermionic case. The bosonic case can be dealt with analogously. We start with our original representation of N-particle states (13)

$$|\mathbf{r}_1, \ldots, \mathbf{r}_N\rangle = \mathcal{N} \sum_{P \in S_N} (-1)^{|P|} |\mathbf{r}_1\rangle \otimes \ldots |\mathbf{r}_N\rangle \; . \tag{77}$$

Then

$$\hat{V}|\mathbf{r}_1, \ldots, \mathbf{r}_N\rangle = \sum_{i<j} V(\mathbf{r}_i, \mathbf{r}_j)|\mathbf{r}_1, \ldots, \mathbf{r}_N\rangle = \frac{1}{2}\sum_{i\neq j} V(\mathbf{r}_i, \mathbf{r}_j)|\mathbf{r}_1, \ldots, \mathbf{r}_N\rangle \; . \tag{78}$$

On the other hand we know that

$$|\mathbf{r}_1, \ldots, \mathbf{r}_N\rangle = \prod_{j=1}^{N} c^\dagger(\mathbf{r}_j)|0\rangle. \tag{79}$$

Now consider

$$c(\mathbf{r})|\mathbf{r}_1, \ldots, \mathbf{r}_N\rangle = c(\mathbf{r}) \prod_{j=1}^{N} c^\dagger(\mathbf{r}_j)|0\rangle = [c(\mathbf{r}), \prod_{j=1}^{N} c^\dagger(\mathbf{r}_j)\}|0\rangle \; , \tag{80}$$

where is the last step we have used that $c(\mathbf{r})|0\rangle = 0$, and $[A, B\}$ is an anticommutator if both A and B involve an odd number of fermions and a commutator otherwise.
In our case we have a commutator for even N and an anticommutator for odd N.
By repeatedly adding and subtracting terms we find that

$$\begin{aligned}[c(\mathbf{r}), \prod_{j=1}^{N} c^\dagger(\mathbf{r}_j)\} \;\; &= \;\; \{c(\mathbf{r}), c^\dagger(\mathbf{r}_1)\} \prod_{j=2}^{N} c^\dagger(\mathbf{r}_j) - c^\dagger(\mathbf{r}_1)\{c(\mathbf{r}), c^\dagger(\mathbf{r}_2)\} \prod_{j=3}^{N} c^\dagger(\mathbf{r}_j) \\ &\quad + \ldots + (-1)^{N-1} \prod_{j=1}^{N-1} c^\dagger(\mathbf{r}_j)\{c(\mathbf{r}), c^\dagger(\mathbf{r}_N)\}.\end{aligned} \tag{81}$$

Using that $\{c(\mathbf{r}), c^\dagger(\mathbf{r}_j)\} = \delta^{(3)}(\mathbf{r}-\mathbf{r}_j)$ we then find

$$c(\mathbf{r})|\mathbf{r}_1,\ldots,\mathbf{r}_N\rangle = \sum_{n=1}^{N}(-1)^{n-1}\delta^{(3)}(\mathbf{r}-\mathbf{r}_n)\prod_{j\neq n}^{N} c^\dagger(\mathbf{r}_j)|0\rangle = \sum_{n=1}^{N}(-1)^{n-1}\delta^{(3)}(\mathbf{r}-\mathbf{r}_n)|\mathbf{r}_1\ldots\overbrace{\mathbf{r}_n}^{\text{missing}}\ldots\mathbf{r}_N\rangle. \tag{82}$$

Hence

$$\underbrace{c^\dagger(\mathbf{r}')c(\mathbf{r}')}_{\text{number op.}}c(\mathbf{r})|\mathbf{r}_1,\ldots,\mathbf{r}_N\rangle = \sum_{n=1}^{N}(-1)^{n-1}\delta^{(3)}(\mathbf{r}-\mathbf{r}_n)\sum_{m\neq n}^{N}\delta^{(3)}(\mathbf{r}'-\mathbf{r}_m)\;|\mathbf{r}_1\ldots\overbrace{\mathbf{r}_n}^{\text{missing}}\ldots\mathbf{r}_N\rangle, \tag{83}$$

and finally

$$c^\dagger(\mathbf{r})c^\dagger(\mathbf{r}')c(\mathbf{r}')c(\mathbf{r})|\mathbf{r}_1,\ldots,\mathbf{r}_N\rangle = \sum_{n=1}^{N}\delta^{(3)}(\mathbf{r}-\mathbf{r}_n)\sum_{m\neq n}^{N}\delta^{(3)}(\mathbf{r}'-\mathbf{r}_m)\;|\mathbf{r}_1\ldots\mathbf{r}_n\ldots\mathbf{r}_N\rangle. \tag{84}$$

This implies that

$$\frac{1}{2}\int d^3\mathbf{r}d^3\mathbf{r}'\;V(\mathbf{r},\mathbf{r}')\;c^\dagger(\mathbf{r})c^\dagger(\mathbf{r}')c(\mathbf{r}')c(\mathbf{r})|\mathbf{r}_1,\ldots,\mathbf{r}_N\rangle = \frac{1}{2}\sum_{n\neq m}V(\mathbf{r}_n,\mathbf{r}_m)|\mathbf{r}_1,\ldots,\mathbf{r}_N\rangle. \tag{85}$$

As $\{|\mathbf{r}_1,\ldots,\mathbf{r}_N\rangle\}$ form a basis, this establishes (76).

Using our formula for basis transformations (41)

$$c^\dagger(\mathbf{r}) = \sum_l \langle l|\mathbf{r}\rangle\; c_l^\dagger, \tag{86}$$

we can transform (76) into a general basis. We have

$$\hat{V} \;=\; \frac{1}{2}\sum_{ll'mm'}\int d^3\mathbf{r}d^3\mathbf{r}'\;V(\mathbf{r},\mathbf{r}')\langle l|\mathbf{r}\rangle\langle l'|\mathbf{r}'\rangle\langle\mathbf{r}'|m'\rangle\langle\mathbf{r}|m\rangle c_l^\dagger c_{l'}^\dagger c_{m'}c_m\;. \tag{87}$$

We can rewrite this by using that the action of $\hat{V}$ on two-particle states is obtained by taking $N=2$ in (75), which tells us that $\hat{V}|\mathbf{r}\rangle\otimes|\mathbf{r}'\rangle = V(\mathbf{r},\mathbf{r}')|\mathbf{r}\rangle\otimes|\mathbf{r}'\rangle$. This implies

$$\begin{aligned} V(\mathbf{r},\mathbf{r}')\langle l|\mathbf{r}\rangle\langle l'|\mathbf{r}'\rangle\langle\mathbf{r}'|m'\rangle\langle\mathbf{r}|m\rangle &= V(\mathbf{r},\mathbf{r}')\left[\langle l|\otimes\langle l'|\right]\left[|\mathbf{r}\rangle\otimes|\mathbf{r}'\rangle|\right]\left[\langle\mathbf{r}|\otimes\langle\mathbf{r}'|\right]\left[|m\rangle\otimes|m'\rangle|\right] \\ &= \left[\langle l|\otimes\langle l'|\right]\hat{V}\left[|\mathbf{r}\rangle\otimes|\mathbf{r}'\rangle|\right]\left[\langle\mathbf{r}|\otimes\langle\mathbf{r}'|\right]\left[|m\rangle\otimes|m'\rangle|\right] \end{aligned} \tag{88}$$

Now we use that

$$\int d^3\mathbf{r}d^3\mathbf{r}'\left[|\mathbf{r}\rangle\otimes|\mathbf{r}'\rangle\right]\left[\langle\mathbf{r}|\otimes\langle\mathbf{r}'|\right] = \mathbf{1} \tag{89}$$

to obtain

$$\hat{V} \;=\; \frac{1}{2}\sum_{l,l',m,m'}\left[\langle l|\otimes\langle l'|\right]|\hat{V}|\left[m\rangle\otimes|m'\rangle\right]c_l^\dagger c_{l'}^\dagger c_{m'}c_m. \tag{90}$$

Finally we can express everything in terms of states with the correct exchange symmetry

$$|mm'\rangle = \frac{1}{\sqrt{2}}\left[|m\rangle\otimes|m'\rangle \pm |m'\rangle\otimes|m\rangle\right] \qquad (m\neq m'). \tag{91}$$

in the form

$$\boxed{\hat{V} = \sum_{(ll'),(mm')} \langle ll'|\hat{V}|mm'\rangle c_l^\dagger c_{l'}^\dagger c_{m'} c_m \ .} \tag{92}$$

Here the sums are over a basis of 2-particle states. In order to see that (90) is equal to (92) observe that

$$\sum_{m,m'} [|m\rangle \otimes |m'\rangle] c_{m'} c_m = \frac{1}{2} \sum_{m,m'} [|m\rangle \otimes |m'\rangle \pm |m'\rangle \otimes |m\rangle] c_{m'} c_m = \frac{1}{\sqrt{2}} \sum_{m,m'} |mm'\rangle \ c_{m'} c_m \tag{93}$$

Here the first equality follows from relabelling summation indices $m \leftrightarrow m'$ and using the (anti)commutation relations between c_m and $c_{m'}$ to bring them back in the right order. The second equality follows from the definition of 2-particle states $|mm'\rangle$. Finally we note that because $|mm'\rangle = \pm|m'm\rangle$ (the minus sign is for fermions) we have

$$\frac{1}{\sqrt{2}} \sum_{m,m'} |mm'\rangle \ c_{m'} c_m = \sqrt{2} \sum_{(mm')} |mm'\rangle \ c_{m'} c_m, \tag{94}$$

where the sum is now over a basis of 2-particle states with the appropriate exchange symmetry. The representation (92) generalizes to arbitrary two-particle operators $\mathcal{O}$.

1.5 Homework Question 4

Question 4. Consider the N-particle interaction potential

$$\hat{V} = \sum_{i<j}^{N} V(\hat{\mathbf{r}}_i, \hat{\mathbf{r}}_j),$$

where $V(\hat{\mathbf{r}}_i, \hat{\mathbf{r}}_j) = V(\hat{\mathbf{r}}_j, \hat{\mathbf{r}}_i)$. Show that in second quantization it is expressed as

$$\hat{V} = \frac{1}{2} \int d^3\mathbf{r} d^3\mathbf{r}' \ V(\mathbf{r}, \mathbf{r}') \ c^\dagger(\mathbf{r}) c^\dagger(\mathbf{r}') c(\mathbf{r}') c(\mathbf{r}).$$

To do so consider the action of $\hat{V}$ on a basis of N-particle position eigenstates

$$|\mathbf{r}_1 \dots \mathbf{r}_N\rangle = \frac{1}{\sqrt{N! n_1! n_2! \dots}} \sum_P (\pm 1)^{|P|} |\mathbf{r}_1\rangle \otimes |\mathbf{r}_2\rangle \otimes |\mathbf{r}_N\rangle = \frac{1}{\sqrt{n_1! n_2! \dots}} \prod_{j=1}^{N} c^\dagger(\mathbf{r}_j)|0\rangle \ ,$$

where n_j is the occupation number of the j^{th} single-particle state. Argue that in an arbitrary basis of single-particle eigenstates $|l\rangle$ $\hat{V}$ can be expressed in the form

$$\hat{V} = \sum_{ll'mm'} \langle ll'|\hat{V}|mm'\rangle c_l^\dagger c_{l'}^\dagger c_{m'} c_m.$$

2 Application I: The Ideal Fermi Gas

Consider an ideal gas of spin-1/2 fermions. The creation operators in the momentum representation (in the infinite volume) are

$$c_\sigma^\dagger(\mathbf{p}) \ , \quad \sigma = \uparrow, \downarrow \ . \tag{95}$$

They fulfil canonical anticommutation relations

$$\{c_\sigma(\mathbf{p}), c_\tau(\mathbf{k})\} = 0 = \{c_\sigma^\dagger(\mathbf{p}), c_\tau^\dagger(\mathbf{k})\} \ , \qquad \{c_\sigma(\mathbf{p}), c_\tau^\dagger(\mathbf{k})\} = \delta_{\sigma,\tau} (2\pi\hbar)^3 \delta^{(3)}(\mathbf{k} - \mathbf{p}). \tag{96}$$

The Hamiltonian, in the grand canonical ensemble, is

$$H - \mu\hat{N} = \int \frac{d^3\mathbf{p}}{(2\pi\hbar)^3} \underbrace{\left[\frac{\mathbf{p}^2}{2m} - \mu\right]}_{\epsilon(\mathbf{p})} \sum_{\sigma=\uparrow,\downarrow} c^\dagger_\sigma(\mathbf{p}) c_\sigma(\mathbf{p}). \tag{97}$$

Here $\mu > 0$ is the chemical potential. As $c^\dagger_\sigma(\mathbf{p})c_\sigma(\mathbf{p}) = \hat{n}_\sigma(\mathbf{p})$ is the number operator for spin-σ fermions with momentum $\mathbf{p}$, we can easily deduce the action of the Hamiltonian on states in the Fock space:

$$\begin{aligned}
\left[H - \mu\hat{N}\right]|0\rangle &= 0\ , \\
\left[H - \mu\hat{N}\right] c^\dagger_\sigma(\mathbf{p})|0\rangle &= \epsilon(\mathbf{p})\ c^\dagger_\sigma(\mathbf{p})|0\rangle\ , \\
\left[H - \mu\hat{N}\right] \prod_{j=1}^{n} c^\dagger_{\sigma_j}(\mathbf{p}_j)|0\rangle &= \left[\sum_{k=1}^{n} \epsilon(\mathbf{p}_k)\right] \prod_{j=1}^{n} c^\dagger_{\sigma_j}(\mathbf{p}_j)|0\rangle\ .
\end{aligned} \tag{98}$$

2.1 Quantization in a large, finite volume

In order to construct the ground state and low-lying excitations, it is convenient to work with a discrete set of momenta. This is achieved by considering the gas in a large, periodic box of linear size L. Momentum eigenstates are obtained by solving the eigenvalue equation e.g. in the position representation

$$\hat{\mathbf{p}}\psi_{\mathbf{k}}(\mathbf{r}) = -i\hbar\nabla\psi_{\mathbf{k}}(\mathbf{r}) = \mathbf{k}\psi_{\mathbf{k}}(\mathbf{r}). \tag{99}$$

The solutions are plane waves

$$\psi_{\mathbf{k}}(\mathbf{r}) = e^{\frac{i}{\hbar}\mathbf{k}\cdot\mathbf{r}}. \tag{100}$$

Imposing periodic boundary conditions ($\mathbf{e}_a$ is the unit vector in the a direction)

$$\psi_{\mathbf{k}}(\mathbf{r} + L\mathbf{e}_a) = \psi_{\mathbf{k}}(\mathbf{r}) \quad \text{for } a = x, y, z, \tag{101}$$

gives quantization conditions for the momenta $\mathbf{k}$

$$e^{\frac{i}{\hbar}Lk_a} = 1 \Rightarrow k_a = \frac{2\pi\hbar n_a}{L}\ , \quad a = x, y, z. \tag{102}$$

To summarize, in a large, periodic box the momenta are quantized as

$$\mathbf{k} = \frac{2\pi\hbar}{L}\begin{pmatrix} n_x \\ n_y \\ n_z \end{pmatrix} \tag{103}$$

Importantly, we can now normalize the eigenstates to 1, i.e.

$$\psi_{\mathbf{k}}(\mathbf{r}) = \frac{1}{L^{\frac{3}{2}}} e^{\frac{i}{\hbar}\mathbf{k}\cdot\mathbf{r}}. \tag{104}$$

Hence

$$\langle\mathbf{k}|\mathbf{k}'\rangle = \int d^3\mathbf{r}\psi^*_{\mathbf{k}}(\mathbf{r})\psi_{\mathbf{k}'}(\mathbf{r}) = \delta_{\mathbf{k},\mathbf{k}'}. \tag{105}$$

As a consequence of the different normalization of single-particle states, the anticommutation relations of creation/annihilation operators are changed and now read

$$\{c_\sigma(\mathbf{p}), c_\tau(\mathbf{k})\} = 0 = \{c^\dagger_\sigma(\mathbf{p}), c^\dagger_\tau(\mathbf{k})\}\ , \qquad \{c^\dagger_\sigma(\mathbf{p}), c_\tau(\mathbf{k})\} = \delta_{\sigma,\tau}\delta_{\mathbf{k},\mathbf{p}}. \tag{106}$$

The Hamiltonian is

$$\boxed{H - \mu\hat{N} = \sum_{\mathbf{p}} \epsilon(\mathbf{p}) \sum_{\sigma=\uparrow,\downarrow} c^\dagger_\sigma(\mathbf{p}) c_\sigma(\mathbf{p}).} \tag{107}$$

We define a *Fermi momentum* by

$$\frac{p_F^2}{2m} = \mu. \tag{108}$$

2.1.1 Ground State

Then the lowest energy state is obtained by *filling all negative energy single-particle states*, i.e.

$$\boxed{|\mathrm{GS}\rangle = \prod_{|\mathbf{p}|<p_F,\sigma} c^\dagger_\sigma(\mathbf{p})|0\rangle.} \tag{109}$$

The ground state energy is

$$E_{\mathrm{GS}} = \sum_\sigma \sum_{|\mathbf{p}|<p_F} \epsilon(\mathbf{p}). \tag{110}$$

This is extensive (proportional to the volume) as expected. You can see the advantage of working in a finite volume: the product in (109) involves only a finite number of factors and the ground state energy is finite. The ground state momentum is

$$P_{\mathrm{GS}} = \sum_\sigma \sum_{|\mathbf{p}|<p_F} \mathbf{p} = 0. \tag{111}$$

The ground state momentum is zero, because is a state with momentum $\mathbf{p}$ contributes to the sum, then so does the state with momentum $-\mathbf{p}$.

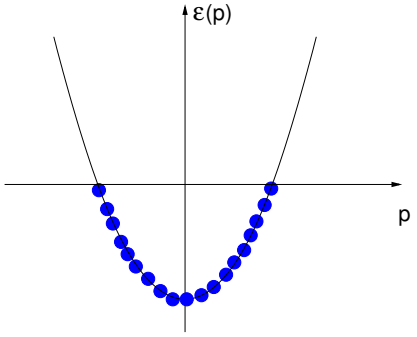


Figure 1: Ground state in the 1 dimensional case. Blue circles correspond to "filled" single-particle states.

2.1.2 Excitations

- Particle excitations

$$c^\dagger_\sigma(\mathbf{k})|\mathrm{GS}\rangle \qquad \text{with } |\mathbf{k}| > p_F. \tag{112}$$

  Their energies and momenta are

$$E = E_{\mathrm{GS}} + \epsilon(\mathbf{k}) > E_{\mathrm{GS}}\ , \quad \mathbf{P} = \mathbf{k}. \tag{113}$$

- Hole excitations

$$c_\sigma(\mathbf{k})|\mathrm{GS}\rangle \qquad \text{with } |\mathbf{k}| < p_F. \tag{114}$$

  Their energies and momenta are

$$E = E_{\mathrm{GS}} - \epsilon(\mathbf{k}) > E_{\mathrm{GS}}\ , \quad \mathbf{P} = -\mathbf{k}. \tag{115}$$

- Particle-hole excitations

$$c^\dagger_\sigma(\mathbf{k})c_\tau(\mathbf{p})|\mathrm{GS}\rangle \qquad \text{with } |\mathbf{k}| > p_F > |\mathbf{p}|. \tag{116}$$

  Their energies and momenta are

$$E = E_{\mathrm{GS}} + \epsilon(\mathbf{k}) - \epsilon(\mathbf{p}) > E_{\mathrm{GS}}\ , \quad \mathbf{P} = \mathbf{k} - \mathbf{p}. \tag{117}$$

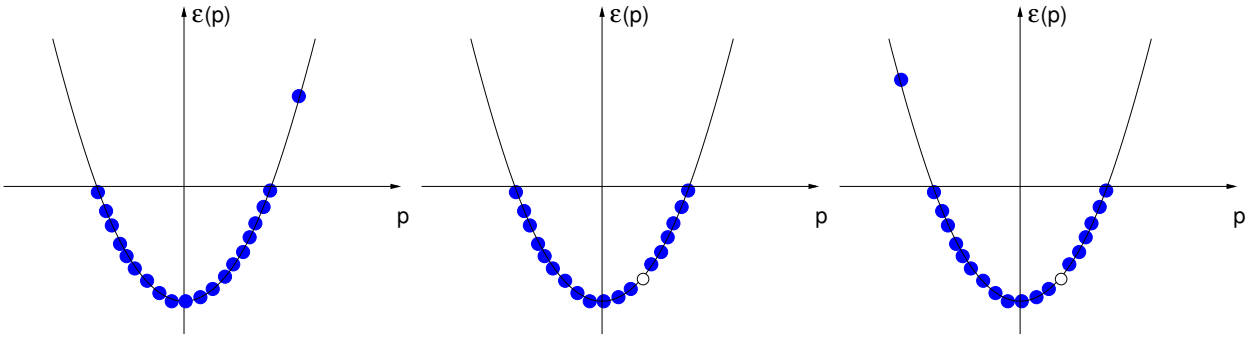


Figure 2: Some simple excited states: (a) particle (b) hole (c) particle-hole.

2.1.3 Density Correlations

Consider the single-particle operator

$$o = |\mathbf{r}\rangle\langle\mathbf{r}| \tag{118}$$

It represents the particle density at position $|\mathbf{r}\rangle$ as can be seen by acting on position eigenstates. In second quantization it is

$$\rho(\mathbf{r}) = \sum_\sigma \int d^3\mathbf{r}' d^3\mathbf{r}'' \ \langle\mathbf{r}'|o|\mathbf{r}''\rangle \ c_\sigma^\dagger(\mathbf{r}')c_\sigma(\mathbf{r}'') = \sum_\sigma c_\sigma^\dagger(\mathbf{r})c_\sigma(\mathbf{r}). \tag{119}$$

1. One-point function.

   We now want to determine the expectation value of this operator in the ground state

   $$\langle \mathrm{GS}|\rho(\mathbf{r})|\mathrm{GS}\rangle = \sum_\sigma \langle \mathrm{GS}|c_\sigma^\dagger(\mathbf{r})c_\sigma(\mathbf{r})|\mathrm{GS}\rangle. \tag{120}$$

   A crucial observation is that the ground state has a simple description in terms of the Fock space built from momentum eigenstates. Hence what we want to do is to work out the momentum representation of $\rho(\mathbf{r})$. We know from our general formula (42) that

   $$c_\sigma(\mathbf{r}) = \sum_{\mathbf{p}} \underbrace{\langle\mathbf{r}|\mathbf{p}\rangle}_{L^{-3/2}e^{\frac{i}{\hbar}\mathbf{p}\cdot\mathbf{r}}} c_\sigma(\mathbf{p}). \tag{121}$$

   Substituting this as well as the analogous expression for the creation operator into (120), we obtain

   $$\langle \mathrm{GS}|\rho(\mathbf{r})|\mathrm{GS}\rangle = \sum_\sigma \frac{1}{L^3} \sum_{\mathbf{p},\mathbf{p}'} e^{\frac{i}{\hbar}(\mathbf{p}-\mathbf{p}')\cdot\mathbf{r}} \langle \mathrm{GS}|c_\sigma^\dagger(\mathbf{p}')c_\sigma(\mathbf{p})|\mathrm{GS}\rangle. \tag{122}$$

   For the expectation value $\langle \mathrm{GS}|c_\sigma^\dagger(\mathbf{p}')c_\sigma(\mathbf{p})|\mathrm{GS}\rangle$ to be non-zero, we must have that $c_\sigma^\dagger(\mathbf{p}')c_\sigma(\mathbf{p})|\mathrm{GS}\rangle$ reproduces $|\mathrm{GS}\rangle$ itself. The only way this is possible is if $|\mathbf{p}| < p_F$ (so that the c pokes a hole in the Fermi sea) and $\mathbf{p}' = \mathbf{p}$ (so that the $c^\dagger$ precisely fills the hole made by the c). By this reasoning we have

   $$\langle \mathrm{GS}|c_\sigma^\dagger(\mathbf{p}')c_\tau(\mathbf{p})|\mathrm{GS}\rangle = \delta_{\sigma,\tau}\delta_{\mathbf{p},\mathbf{p}'}\theta(p_F - |\mathbf{p}'|). \tag{123}$$

   Similarly we can show that

   $$\begin{aligned} \langle \mathrm{GS}|c_\sigma(\mathbf{p}')c_\tau^\dagger(\mathbf{p})|\mathrm{GS}\rangle &= \delta_{\sigma,\tau}\delta_{\mathbf{p},\mathbf{p}'}\theta(|\mathbf{p}| - p_F), \\ \langle \mathrm{GS}|c_\sigma(\mathbf{p}')c_\tau(\mathbf{p})|\mathrm{GS}\rangle &= 0 = \langle \mathrm{GS}|c_\sigma^\dagger(\mathbf{p}')c_\tau^\dagger(\mathbf{p})|\mathrm{GS}\rangle. \end{aligned} \tag{124}$$

   Substituting (123) back into (122) we find

   $$\langle \mathrm{GS}|\rho(\mathbf{r})|\mathrm{GS}\rangle = \sum_\sigma \frac{1}{L^3} \sum_{\mathbf{p},\mathbf{p}'} e^{\frac{i}{\hbar}(\mathbf{p}-\mathbf{p}')\cdot\mathbf{r}} \delta_{\mathbf{p},\mathbf{p}'} \ \theta(p_F - |\mathbf{p}|) = \underbrace{2}_{\text{spin}} \frac{1}{L^3} \sum_{\mathbf{p}} \theta(p_F - |\mathbf{p}|) = \frac{N}{L}. \tag{125}$$

So our expectation value gives precisely the particle density. This is expected because our system is translationally invariant and therefore $\langle \mathrm{GS}|\rho(\mathbf{r})|\mathrm{GS}\rangle$ cannot depend on $\mathbf{r}$.

2. Two-point function.

Next we want to determine the two-point function

$$\langle \mathrm{GS}|\rho(\mathbf{r})\rho(\mathbf{r}')|\mathrm{GS}\rangle = \sum_{\sigma,\tau} \frac{1}{L^6} \sum_{\mathbf{p},\mathbf{p}'} \sum_{\mathbf{k},\mathbf{k}'} e^{\frac{i}{\hbar}(\mathbf{p}-\mathbf{p}')\cdot\mathbf{r}} e^{\frac{i}{\hbar}(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}'} \langle \mathrm{GS}|c_\sigma^\dagger(\mathbf{p}')c_\sigma(\mathbf{p})c_\tau^\dagger(\mathbf{k}')c_\tau(\mathbf{k})|\mathrm{GS}\rangle. \tag{126}$$

The difficulty in calculating the ground state expectation value is that the $c_\sigma(\mathbf{p})$ ($c_\sigma^\dagger(\mathbf{p})$) annihilate the ground state only for $|\mathbf{p}| > p_F$ ($|\mathbf{p}| < p_F$). A way around this is to define new annihilation and creation operators by

$$d_\sigma(\mathbf{p}) = \theta(|\mathbf{p}| - p_F)c_\sigma(\mathbf{p}) + \theta(p_F - |\mathbf{p}|)c_\sigma^\dagger(\mathbf{p}) \ , \quad d_\sigma^\dagger(\mathbf{p}) = \big(d_\sigma(\mathbf{p})\big)^\dagger. \tag{127}$$

These fulfil canonical anticommutation relations

$$\{d_\sigma(\mathbf{p}), d_\tau^\dagger(\mathbf{k})\} = \delta_{\sigma,\tau}\delta_{\mathbf{k},\mathbf{p}} \ , \quad \{d_\sigma(\mathbf{p}), d_\tau(\mathbf{k})\} = 0 \ , \tag{128}$$

and by construction we have

$$\langle \mathrm{GS}|c_\sigma^\dagger(\mathbf{p}) = 0 = c_\sigma(\mathbf{p})|\mathrm{GS}\rangle. \tag{129}$$

This is an example of a so-called *particle-hole transformation*. We may now use the inverse transformation

$$c_\sigma(\mathbf{p}) = \theta(|\mathbf{p}| - p_F)d_\sigma(\mathbf{p}) + \theta(p_F - |\mathbf{p}|)d_\sigma^\dagger(\mathbf{p}) \tag{130}$$

to replace the creation and annihilation operators in (126) by the $d_\sigma(\mathbf{p})$, $d_\sigma^\dagger(\mathbf{p})$. After that we simply apply the anticommutation relations (128) to move all annihilation operators to the right, and finally use (129). The result is

$$\begin{aligned} \langle \mathrm{GS}|c_\sigma^\dagger(\mathbf{p}')c_\sigma(\mathbf{p})c_\tau^\dagger(\mathbf{k}')c_\tau(\mathbf{k})|\mathrm{GS}\rangle &= \delta_{\mathbf{k},\mathbf{k}'}\delta_{\mathbf{p},\mathbf{p}'}\theta(p_F - |\mathbf{p}|)\theta(p_F - |\mathbf{k}|) \\ &\quad + \delta_{\sigma,\tau}\delta_{\mathbf{p},\mathbf{k}'}\delta_{\mathbf{k},\mathbf{p}'}\theta(|\mathbf{k}'| - p_F)\theta(p_F - |\mathbf{k}|). \end{aligned} \tag{131}$$

Observe that by virtue of (123) and (124) this can be rewritten in the form

$$\langle \mathrm{GS}|c_\sigma^\dagger(\mathbf{p}')c_\sigma(\mathbf{p})|\mathrm{GS}\rangle\langle \mathrm{GS}|c_\tau^\dagger(\mathbf{k}')c_\tau(\mathbf{k})|\mathrm{GS}\rangle + \langle \mathrm{GS}|c_\sigma^\dagger(\mathbf{p}')c_\tau(\mathbf{k})|\mathrm{GS}\rangle\langle \mathrm{GS}|c_\sigma(\mathbf{p})c_\tau^\dagger(\mathbf{k}')|\mathrm{GS}\rangle \ . \tag{132}$$

The fact that the 4-point function (131) can be written as a sum over products of two-point functions is a reflection of *Wick's theorem* for noninteracting spin-1/2 fermions. This is not part of the syllabus and we won't dwell on it, but apart from extra minus signs, this says that $2n$-point functions are given by the sum over all possible "pairings", giving rise to a product of two-point functions. In our particular case this gives

$$\begin{aligned} \langle c_\sigma^\dagger(\mathbf{p}')c_\sigma(\mathbf{p})c_\tau^\dagger(\mathbf{k}')c_\tau(\mathbf{k})\rangle &= \langle c_\sigma^\dagger(\mathbf{p}')c_\sigma(\mathbf{p})\rangle\langle c_\tau^\dagger(\mathbf{k}')c_\tau(\mathbf{k})\rangle - \langle c_\sigma^\dagger(\mathbf{p}')c_\tau^\dagger(\mathbf{k}')\rangle\langle c_\sigma(\mathbf{p})c_\tau(\mathbf{k})\rangle \\ &+ \langle c_\sigma^\dagger(\mathbf{p}')c_\tau(\mathbf{k})\rangle\langle c_\sigma(\mathbf{p})c_\tau^\dagger(\mathbf{k}')\rangle, \end{aligned} \tag{133}$$

and using that the two point function of two creation or two annihilation operators is zero we obtain (132). Substituting (131) back in to (126) gives

$$\begin{aligned} \langle \mathrm{GS}|\rho(\mathbf{r})\rho(\mathbf{r}')|\mathrm{GS}\rangle &= \sum_{\sigma,\sigma'} \frac{1}{L^6} \sum_{\mathbf{k},\mathbf{p}} \theta(p_F - |\mathbf{k}|)\theta(p_F - |\mathbf{p}|) \\ &+ \sum_{\sigma} \frac{1}{L^6} \sum_{\mathbf{k},\mathbf{k}'} \theta(|\mathbf{k}| - p_F)\theta(p_F - |\mathbf{k}'|)e^{\frac{i}{\hbar}(\mathbf{k}-\mathbf{k}')\cdot(\mathbf{r}-\mathbf{r}')} \\ &= \big(\langle \mathrm{GS}|\rho(\mathbf{r})|\mathrm{GS}\rangle\big)^2 + \frac{2}{L^3} \sum_{|\mathbf{k}|>p_F} e^{\frac{i}{\hbar}\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')} \frac{1}{L^3} \sum_{|\mathbf{k}'|<p_F} e^{-\frac{i}{\hbar}\mathbf{k}'\cdot(\mathbf{r}-\mathbf{r}')}. \end{aligned} \tag{134}$$

In the limit $L \to \infty$ we can simplify this expression further.

Aside 3: Evaluating k-sums

Evaluting the $\mathbf{k}$ sums for large L: The idea is to turn sums into integrals

$$\frac{1}{L^3}\sum_{|\mathbf{k}|<p_F} e^{\frac{i}{\hbar}\mathbf{k}\cdot\mathbf{R}} \rightarrow \int \frac{d^3\mathbf{k}}{(2\pi\hbar)^3}\theta(p_F-|\mathbf{k}|)e^{\frac{i}{\hbar}\mathbf{k}\cdot\mathbf{R}} = \int_0^\infty dp p^2 \int_0^\pi d\vartheta \sin\vartheta \int_0^{2\pi} d\varphi \, \frac{\theta(p_F-\hbar p)}{(2\pi)^3} e^{ip|\mathbf{R}|\cos\vartheta}$$
$$= \int_0^{p_F/\hbar} \frac{dp}{(2\pi)^2}\frac{2p\sin(p|\mathbf{R}|)}{|\mathbf{R}|} = \frac{\sin(p_F|\mathbf{R}|)-p_F|\mathbf{R}|\cos(p_F|\mathbf{R}|)}{2\pi^2|\mathbf{R}|^3} \equiv h(|\mathbf{R}|). \tag{135}$$

Here we have introduced spherical polar coordinates such that the z-axis of our co-ordinate system is along the $\mathbf{R}$ direction, and

$$\begin{aligned} k_x &= \hbar p \sin\vartheta\cos\varphi \ , \\ k_y &= \hbar p \sin\vartheta\sin\varphi \ , \\ k_z &= \hbar p\cos\vartheta. \end{aligned} \tag{136}$$

The other sum works similarly

$$\frac{1}{L^3}\sum_{|\mathbf{k}|>p_F} e^{\frac{i}{\hbar}\mathbf{k}\cdot\mathbf{R}} = \frac{1}{L^3}\sum_{\mathbf{k}} e^{\frac{i}{\hbar}\mathbf{k}\cdot\mathbf{R}} - \frac{1}{L^3}\sum_{|\mathbf{k}|<p_F} e^{\frac{i}{\hbar}\mathbf{k}\cdot\mathbf{R}}. \tag{137}$$

The second part is evaluated above, while the first part is

$$\frac{1}{L^3}\sum_{\mathbf{k}} e^{\frac{i}{\hbar}\mathbf{k}\cdot\mathbf{R}} = \delta^{(3)}(\mathbf{R}). \tag{138}$$

The equality can be proved by multiplying both sides by a test-function $f(\mathbf{R})$ and then integrating over $\mathbf{R}$:

$$\int d^3\mathbf{R}\frac{1}{L^3}\sum_{\mathbf{k}} e^{\frac{i}{\hbar}\mathbf{k}\cdot\mathbf{R}} f(\mathbf{R}) = \frac{1}{L^3}\sum_{\mathbf{k}}\int d^3\mathbf{R} e^{\frac{i}{\hbar}\mathbf{k}\cdot\mathbf{R}} f(\mathbf{R}) = \frac{1}{L^3}\sum_{\mathbf{k}} \tilde{f}_{\mathbf{k}} = f(\mathbf{0}). \tag{139}$$

Here we have used standard definitions for Fourier series, cf Riley/Hobson/Bence 12.7.

Using these simplifications for large L we arrive at our final answer

$$\boxed{\langle \mathrm{GS}|\rho(\mathbf{r})\rho(\mathbf{r}')|\mathrm{GS}\rangle = \langle \mathrm{GS}|\rho(\mathbf{r})|\mathrm{GS}\rangle^2 + \langle \mathrm{GS}|\rho(\mathbf{r})|\mathrm{GS}\rangle\delta^{(3)}(\mathbf{r}-\mathbf{r}') - 2\left[h(|\mathbf{r}-\mathbf{r}'|)\right]^2.} \tag{140}$$

The first two terms are the same as for a *classical* ideal gas, while the third contribution is due to the *fermionic statistics* (Pauli exclusion: "fermions don't like to be close to one another").

2.2 Connection with Quantum Field Theory

Using our general result (70) we can express the Hamiltonian of the free Fermi gas in the position representation

$$H = \sum_\sigma \int d^3\mathbf{x}\ c_\sigma^\dagger(\mathbf{x})\left[-\frac{\hbar^2\nabla^2}{2m} - \mu\right] c_\sigma(\mathbf{x}). \tag{141}$$

Aside 4: Schrödinger vs Heisenberg pictures and Heisenberg equation of motion

Let us consider time evolution in a system with time-independent Hamiltonian. In the *Schrödinger picture* one takes the states to be time-dependent and operators to be time-independent. Time evolution is governed by the time dependent Schrödinger equation

$$i\hbar\frac{d|\psi(t)\rangle}{dt} = H|\psi(t)\rangle \ . \tag{142}$$

This is formally solved by

$$|\psi(t)\rangle = e^{-\frac{i}{\hbar}Ht}|\psi(0)\rangle \ . \tag{143}$$

Measurable quantities are related to matrix elements of observables $\mathcal{O}$

$$\langle\phi(t)|\mathcal{O}|\psi(t)\rangle = \langle\phi(0)|\underbrace{e^{\frac{i}{\hbar}Ht}\mathcal{O}e^{-\frac{i}{\hbar}Ht}}_{\mathcal{O}(t)}|\psi(0)\rangle \tag{144}$$

We see that we can calculate these in an alternative way: we can fix a basis of states at $t = 0$ and calculate all matrix elements in this basis, while time evolution is now moved to the operators. This formulation of Quantum mechanics is known as *Heisenberg picture*. For a time-independent Hamiltonian the time evolution of operators is then governed by the *Heisenberg equation of motion*

$$\frac{d\mathcal{O}}{dt} = \frac{i}{\hbar}[H, \mathcal{O}(t)] \ . \tag{145}$$

The Heisenberg equations of motion for $c_\sigma(\mathbf{x}, t)$ are

$$\frac{\partial c_\sigma(\mathbf{x}, t)}{\partial t} = -\frac{i}{\hbar}\left[-\frac{\hbar^2\nabla^2}{2m} - \mu\right] c_\sigma(\mathbf{x}, t). \tag{146}$$

These correspond to the Euler-Lagrange equations

$$\partial_t \frac{\delta L}{\delta\partial_t c^\dagger_\sigma(\mathbf{x}, t)} = -\sum_a \partial_a \frac{\delta L}{\delta\partial_a c^\dagger_\sigma(\mathbf{x}, t)} + \frac{\delta L}{\delta c^\dagger_\sigma(\mathbf{x}, t)} \tag{147}$$

obtained from the following Lagrangian

$$L = \sum_\sigma \int d^3\mathbf{x}\ c^\dagger_\sigma(\mathbf{x}, t)\left[i\hbar\frac{\partial}{\partial t} + \frac{\hbar^2\nabla^2}{2m} + \mu\right] c_\sigma(\mathbf{x}, t) \ , \tag{148}$$

which describes a non-relativistic Quantum Field Theory for the fermionic fields $c_\sigma(\mathbf{x})$.

2.3 "Emergent" relativistic description at low energies

Let us now for simplicity consider the case $D = 1$. We have seen above that the low-energy degrees of freedom "live" in the vicinities of the Fermi momenta $\pm p_F$. The expansion of $c_\sigma(x)$ in terms of momentum modes reads

$$c_\sigma(x) = \frac{1}{\sqrt{L}}\sum_p c_\sigma(p)\ e^{\frac{i}{\hbar}px} \ . \tag{149}$$

Let us now imagine that we probe our Fermi system only in ways that involve very small energy transfers to the system. This kind of situation is in fact often encountered when doing experiments on solids. Such experimental probes are sensitive only to states that have an energy close to that of the ground state. Let us thus "project" our theory to low-energy degrees of freedom only: this amounts to retain only momentum modes in small intervals around the Fermi points

$$|p \pm p_F| < \Lambda \ll p_F. \tag{150}$$

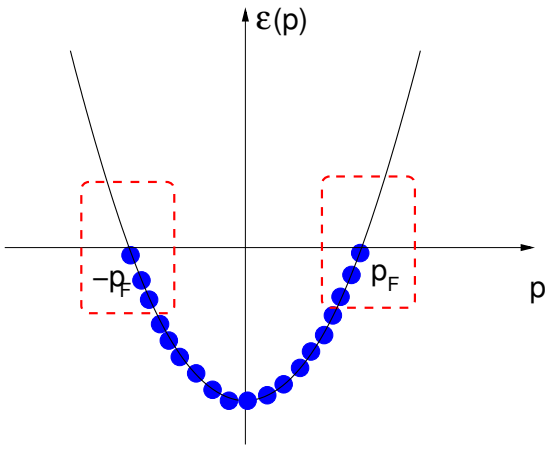


Figure 3: The low-energy degrees of freedom (particle/hole excitations with small excitation energies) involve only momenta close to the Fermi points $\pm p_F$.

Here Λ is a momentum *cutoff*. This gives

$$c_\sigma(x) \longrightarrow e^{\frac{i}{\hbar}p_F x} \underbrace{\frac{1}{\sqrt{L}} \sum_{|p|<\Lambda} c_\sigma(p_F+p)\ e^{\frac{i}{\hbar}px}}_{R_\sigma(x)} + e^{-\frac{i}{\hbar}p_F x} \underbrace{\frac{1}{\sqrt{L}} \sum_{|p|<\Lambda} c_\sigma(-p_F+p)\ e^{\frac{i}{\hbar}px}}_{L_\sigma(x)} . \tag{151}$$

The low-energy projection of our Hamiltonian is

$$H-\mu N \longrightarrow H_{\text{low}} = \sum_{|p|<\Lambda,\sigma} \epsilon(p_F+p)c_\sigma^\dagger(p_F+p)c_\sigma(p_F+p) + \epsilon(-p_F+p)c_\sigma^\dagger(-p_F+p)c_\sigma(-p_F+p)\ , \tag{152}$$

where

$$\epsilon(\pm p_F+p) = \frac{(\pm p_F+p)^2}{2m} - \mu = \pm\frac{pp_F}{m}\left(1\pm\frac{p}{2p_F}\right) \approx \pm\frac{pp_F}{m} \equiv \pm v_F p. \tag{153}$$

Here v_F is called *Fermi velocity*. Neglecting the small correction terms to the linear dispersion we have

$$\begin{aligned} H_{\text{low}} &\approx \sum_{|p|<\Lambda,\sigma} v_F p\left[c_\sigma^\dagger(p_F+p)c_\sigma(p_F+p) - c_\sigma^\dagger(-p_F+p)c_\sigma(-p_F+p)\right] \\ &= \sum_\sigma -i\hbar v_F \int dx \left[R_\sigma^\dagger(x)\partial_x R_\sigma(x) - L_\sigma^\dagger(x)\partial_x L_\sigma(x)\right]. \end{aligned} \tag{154}$$

This Hamiltonian is closely related to the 1+1 dimensional Dirac equation. The Lagrangian density of the Dirac field is

$$\mathcal{L}_D = i\hbar\bar\Psi\left[\gamma^0\partial_t - c\gamma^1\partial_1\right]\Psi - mc^2\bar\Psi\Psi\ , \tag{155}$$

Here c is the speed of light, $\gamma^0=\sigma^x$ and $\gamma^1=-i\sigma^y$ are gamma matrices in two dimensions, $\bar\Psi=\Psi^\dagger\gamma^0$ and Ψ itself is a two dimensional spinor

$$\Psi(x) = \begin{pmatrix} R(x) \\ L(x) \end{pmatrix}. \tag{156}$$

Writing (155) out in components we have

$$\mathcal{L}_D = i\hbar\left[R^\dagger\partial_t R + L^\dagger\partial_t L\right] + ic\left[R^\dagger\partial_x R - L^\dagger\partial_x L\right] - mc^2\left[R^\dagger L + L^\dagger R\right]. \tag{157}$$

The corresponding Hamiltonian density

$$\mathcal{H}_D = -i\hbar c\left[R^\dagger\partial_x R - L^\dagger\partial_x L\right] + mc^2\left[R^\dagger L + L^\dagger R\right]. \tag{158}$$

Comparing (154) to (158) we conclude that at low energies the Fermi gas is described by two copies (one for each spin projection) of a massless Dirac field. The resulting theory is relativistic apart from the speed of light being replaced by the Fermi velocity v_F. This symmetry was not present in our original problem, but *emerged* in the vicinity of a non-trivial ground state. Such emergent descriptions of the low energy degrees of freedom in terms of relativistic QFTs are fairly common in condensed matter physics. Usually the QFTs are however strongly interacting!

2.4 Homework Questions 5-6

Question 5. Consider a system of fermions moving freely on a one-dimensional ring of length L, i.e. periodic boundary conditions are applied between $x = 0$ and $x = L$. The fermions are all in the same spin state, so that spin quantum numbers may be omitted. Fermion creation and annihilation operators at the point x are denoted by $\psi^\dagger(x)$ and $\psi(x)$.
a) Write down the complete set of anticommutation relation satisfied by $\psi^\dagger(x_1)$ and $\psi(x_2)$.
b) Write down the wave-functions of single-particle momentum eigenstates (make sure to take the boundary conditions into account!). What are the allowed values of momentum? Using this result, derive an expression for the momentum space creation and annihilation operators $\Psi_p^\dagger$ and Ψ_p in terms of $\psi^\dagger(x)$ and $\psi(x)$ (hint: use the general result for basis transformation obtained in the lecture notes).
c) Starting with your expression for the anticommutator $\{\psi^\dagger(x_1), \psi(x_2)\}$, evaluate $\{\Psi_p^\dagger, \Psi_q\}$.
d) Derive an expression for $\psi(x)$ in terms of Ψ_k.
e) The density operator $\rho(x)$ is defined by $\rho(x) = \psi^\dagger(x)\,\psi(x)$. The number operator is

$$N = \int_0^L dx\, \rho(x)\,.$$

Express $\rho(x)$ in terms of $\Psi_p^\dagger$ and Ψ_q, and show from this that

$$N = \sum_k \Psi_k^\dagger \Psi_k\,.$$

Let $|0\rangle$ be the vacuum state (containing no particles) and define $|\phi\rangle$ by

$$|\phi\rangle = A \prod_k (u_k + v_k \Psi_k^\dagger)|0\rangle,$$

where u_k and v_k are complex numbers depending on the label k, and A is a normalisation constant.

Evaluate (i) $|A|^2$, (ii) $\langle\phi|N|\phi\rangle$, and (iii) $\langle\phi|N^2|\phi\rangle$. Under what conditions is $|\phi\rangle$ an eigenstate of particle number?

Question 6. Consider a system of fermions in which the functions $\varphi_\ell(x)$, $\ell = 1, 2 \ldots N$, form a complete orthonormal basis for single particle wavefunctions.
a) Explain how Slater determinants may be used to construct a complete orthonormal basis for n-particle states with $n = 2, 3 \ldots N$. Calculate the normalisation constant for such a Slater determinant at a general value of n. How many independent n-particle states are there for each n?
b) Let $C_\ell^\dagger$ and C_ℓ be fermion creation and destruction operators which satisfy the usual anticommutation relations. The quantities a_k are defined by

$$a_k = \sum_{\ell=1}^{N} U_{k\ell} C_\ell,$$

where $U_{k\ell}$ are elements of an $N \times N$ matrix, U. Write down an expression for $a_k^\dagger$. Find the condition which must be satisfied by the matrix U in order that the operators $a_k^\dagger$ and a_k also satisfy fermion anticommutation relations.

c) Non-interacting spinless fermions move in one dimension in an infinite square-well potential, with position coordinate $0 \leq x \leq L$. The normalised single particle energy eigenstates are

$$\varphi_\ell(x) = \left(\frac{2}{L}\right)^{1/2} \sin\left(\frac{\ell \pi x}{L}\right) ,$$

and the corresponding fermion creation operator is $C_\ell^\dagger$.

Write down expressions for $C^\dagger(x)$, the fermion creation operator at the point x, and for $\rho(x)$, the particle density operator, in terms of $C_\ell^\dagger$, C_ℓ and $\varphi_\ell(x)$.

d) What is the ground state expectation value $\langle \rho(x) \rangle$ in a system of n fermions?

In the limit $n \to \infty$, $L \to \infty$, taken at fixed average density $\rho_0 = n/L$, show that

$$\langle \rho(x) \rangle = \rho_0 \left[1 - \frac{\sin 2\pi\rho_0 x}{2\pi\rho_0 x}\right] .$$

Sketch this function and comment briefly on its behaviour for $x \to 0$ and $x \to \infty$.

3 Linear Response Theory

Now is a good time to address the question what quantities are of experimental interest.To that end, let us consider a quantum system described by a time-independent Hamiltonian H_0 that initially is in thermal equilibrium at a temperature T. Expectation values of physical observables are then given by

$$\langle \mathcal{O} \rangle_\beta \equiv \frac{1}{Z} \mathrm{Tr}\left[e^{-\beta H_0} \mathcal{O}\right] , \tag{159}$$

where $\beta = 1/(k_B T)$ and Z is the *partition function*

$$Z = \mathrm{Tr}\left[e^{-\beta H_0}\right] . \tag{160}$$

In terms of energy eigenstates $H_0|n\rangle = E_n|n\rangle$ we have

$$\langle \mathcal{O} \rangle_\beta = \frac{1}{Z} \sum_n e^{-\beta E_n} \langle n|\mathcal{O}|n\rangle . \tag{161}$$

Suppose now that we apply a very weak time-dependent external perturbation $\hat{V}(t)$ to our system that drives it out of thermal equilibrium. The time evolution of the system is then governed by the Hamiltonian

$$H(t) = H_0 + V(t) . \tag{162}$$

The time evolution of states is given by the time-dependent Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = H(t)|\psi(t)\rangle . \tag{163}$$

As the perturbation is very weak it is useful to separate its effects on the time evolution from those of H_0. This gives rise to the so-called *interaction picture*. We define interaction picture operators by

$$A_{\mathrm{int}}(t) \equiv e^{\frac{i}{\hbar} H_0 t} A e^{-\frac{i}{\hbar} H_0 t} . \tag{164}$$

We then split the time evolution of states into the contribution from H_0 and the remainder as follows

$$|\psi(t)\rangle = e^{-\frac{i}{\hbar} H_0 t} U(t, t_0) e^{\frac{i}{\hbar} H_0 t_0} |\psi(t_0)\rangle . \tag{165}$$

Substituting this back into (163) and using that the state is arbitrary then gives an equation for $U(t,t_0)$

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}U(t,t_0) &= e^{\frac{i}{\hbar}H_0 t}V(t)e^{-\frac{i}{\hbar}H_0 t}\ U(t,t_0) \\ &= V_{\rm int}(t)\ U(t,t_0)\ . \end{aligned} \tag{166}$$

The solution of this equation under the conditions that $U(t_0,t_0)=1$ is

$$\begin{aligned} U(t,t_0) &= 1-\frac{i}{\hbar}\int_{t_0}^{t}dt'\ V_{\rm int}(t') - \frac{1}{\hbar^2}\int_{t_0}^{t}dt'\int_{t_0}^{t'}dt''\ V_{\rm int}(t')V_{\rm int}(t'')+\dots \\ &= \mathcal{T}\exp\left(\frac{-i}{\hbar}\int_{t_0}^{t}dt'\ V_{\rm int}(t')\right)\ , \end{aligned} \tag{167}$$

where we have defined a *time ordering operation* by

$$\mathcal{T}\ A(t)B(t') = \begin{cases} A(t)B(t') & \text{if } t>t' \\ B(t')A(t) & \text{if } t<t' \end{cases}\ . \tag{168}$$

Let us now apply this formalism to the expectation value of an observable $\mathcal{O}$ at time $t>t_0$, where we imagine that our perturbation vanishes for earlier times. Then

$$\begin{aligned} \langle\mathcal{O}(t)\rangle_\beta &\equiv \frac{1}{Z}\sum_n e^{-\beta E_n}\langle n(t)|\mathcal{O}|n(t)\rangle \\ &= \frac{1}{Z}\sum_n e^{-\beta E_n}\langle n|U^\dagger(t,t_0)\ \mathcal{O}_{\rm int}(t)\ U(t,t_0)|n\rangle \\ &= \frac{1}{Z}\sum_n e^{-\beta E_n}\left\{\langle n|\mathcal{O}_{\rm int}(t)|n\rangle - \frac{i}{\hbar}\int_{t_0}^{t}dt'\langle n|[\mathcal{O}_{\rm int}(t),V_{\rm int}(t')]|n\rangle+\dots\right\} \\ &\approx \langle\mathcal{O}\rangle_\beta - \frac{i}{\hbar}\int_{t_0}^{t}dt'\langle[\mathcal{O}_{\rm int}(t),V_{\rm int}(t')]\rangle_\beta\ . \end{aligned} \tag{169}$$

Here we have used that $e^{-\frac{i}{\hbar}H_0 t_0}|n\rangle = e^{-\frac{i}{\hbar}E_0 t_0}|n\rangle$. Restricting our discussion to the case $V(t)=f(t)V$ where $f(t)$ is a time dependent function and taking $t_0\to-\infty$ we obtain

$$\langle\mathcal{O}(t)\rangle_\beta - \langle\mathcal{O}\rangle_\beta \approx \int_{-\infty}^{\infty}dt'\ \chi_{\mathcal{O},V}(t,t')\ f(t')\ , \tag{170}$$

where $\chi_{\mathcal{O},V}$ is a *retarded correlation function* of the operators $\mathcal{O}$ and V

$$\chi_{\mathcal{O},V}(t,t') = -\frac{i}{\hbar}\theta(t-t')\ \langle[\mathcal{O}_{\rm int}(t),V_{\rm int}(t')]\rangle_\beta\ . \tag{171}$$

It is easy to see (by writing out the average in a basis of eigenstates of H_0 and using (164)) that this correlation function depends only on the time difference $t-t'$.

Let us now look at an example of this linear response formalism. Let us take the Fermi gas as our system and perturb it by a time-dependent potential of the form

$$V = \int d^3\mathbf{r}\ \rho(\mathbf{r})\ \phi_{\rm ext}(\mathbf{r},t)\ , \tag{172}$$

where $\rho(\mathbf{r})$ is the particle density operator introduced previously. Imposing such a potential is readily done experimentally. If we then measure the density variation induced by the external potential we find according to our calculation above

$$\langle\rho(\mathbf{r},t)\rangle_\beta - \langle\rho(\mathbf{r})\rangle_\beta \approx \int_{-\infty}^{\infty}dt'\int d^3\mathbf{r}'\ \phi_{\rm ext}(\mathbf{r},t')\ \chi_{\rho\rho}(\mathbf{r},t;\mathbf{r}',t'). \tag{173}$$

Here $\chi_{\rho\rho}$ is the *retarded density-density correlation function*

$$\chi_{\rho\rho}(\mathbf{r},t;\mathbf{r}',t') = -\frac{i}{\hbar}\theta(t-t')\ \langle[\rho_{\rm int}(\mathbf{r},t),\rho_{\rm int}(\mathbf{r}',t')]\rangle_\beta\ . \tag{174}$$

This is a particular example of a *response function*. These are the quantities that are measured in many kinds of experiments.

4 Application II: Weakly Interacting Bosons

As you know from Statistical Mechanics, the ideal Bose gas displays the very interesting phenomenon of *Bose condensation*. This has been observed in systems of trapped Rb atoms and led to the award of the Nobel prize in 2001 to Ketterle, Cornell and Wiemann. The atoms in these experiments are bosonic, but the atom-atom interactions are not zero. We now want to understand the effects of interactions in the framework of a microscopic theory. The kinetic energy operator is expressed in terms of creation/annihilation operators single-particle momentum eigenstates as

$$\hat{T} = \sum_{\mathbf{p}} \frac{\mathbf{p}^2}{2m} c^\dagger(\mathbf{p})c(\mathbf{p}). \tag{175}$$

Here we have assumed that our system in enclosed in a large, periodic box of linear dimension L. The boson-boson interaction is most easily expressed in position space

$$\hat{V} = \frac{1}{2}\int d^3\mathbf{r}d^3\mathbf{r}'\ c^\dagger(\mathbf{r})c^\dagger(\mathbf{r}')V(\mathbf{r},\mathbf{r}')c(\mathbf{r}')c(\mathbf{r}) \tag{176}$$

A good model for the potential $V(\mathbf{r},\mathbf{r}')$ is to take it of the form

$$V(\mathbf{r},\mathbf{r}') = U\delta^{(3)}(\mathbf{r}-\mathbf{r}'), \tag{177}$$

i.e. bosons interact only if they occupy the same point in space. Changing to the momentum space description

$$c(\mathbf{r}) = \frac{1}{L^{3/2}}\sum_{\mathbf{p}} e^{\frac{i}{\hbar}\mathbf{p}\cdot\mathbf{r}}c(\mathbf{p}), \tag{178}$$

we have

$$\hat{V} = \frac{U}{2L^3}\sum_{\mathbf{p_1},\mathbf{p_2},\mathbf{p_3}} c^\dagger(\mathbf{p_1})c^\dagger(\mathbf{p_2})c(\mathbf{p_3})c(\mathbf{p_1}+\mathbf{p_2}-\mathbf{p_3}). \tag{179}$$

4.1 Ideal Bose Gas

For $U = 0$ we are dealing with an ideal Bose gas and we know that the ground state is a *condensate*: all particles occupy the lowest-energy single-particle state, i.e. the zero-momentum state

$$|\mathrm{GS}\rangle_0 = \frac{1}{\sqrt{N!}}\Big(c^\dagger(\mathbf{p}=\mathbf{0})\Big)^N|0\rangle. \tag{180}$$

So $\mathbf{p} = 0$ is special, and in particular we have

$${}_0\langle\mathrm{GS}|c^\dagger(\mathbf{p}=\mathbf{0})c(\mathbf{p}=\mathbf{0})|\mathrm{GS}\rangle_0 = N. \tag{181}$$

4.2 Bogoliubov Approximation

For small $U > 0$ we expect the Bose-Einstein condensate to persist, i.e. we expect

$$\langle \mathrm{GS}|c^\dagger(\mathbf{p}=\mathbf{0})c(\mathbf{p}=\mathbf{0})|\mathrm{GS}\rangle = N_0 \gg 1. \tag{182}$$

However,

$$[c^\dagger(\mathbf{0})c(\mathbf{0}), \hat{V}] \neq 0, \tag{183}$$

so that the number of $\mathbf{p}=\mathbf{0}$ bosons is not conserved, and the ground state $|\mathrm{GS}\rangle$ will be a superposition of states with different numbers of $\mathbf{p}=\mathbf{0}$ bosons. However, for the ground state and low-lying excited states we will have

$$\langle \Psi|c^\dagger(\mathbf{0})c(\mathbf{0})|\Psi\rangle \simeq N_0 \ , \tag{184}$$

where N_0, crucially, is a very large number. The *Bogoliubov approximation* states that, *when acting on the ground state or low-lying excited states*, we in fact have

$$\boxed{c^\dagger(\mathbf{0}) \simeq \sqrt{N_0} \ , \qquad c(\mathbf{0}) \simeq \sqrt{N_0} \ ,} \tag{185}$$

i.e. creation and annihilation operators are approximately diagonal. This is a much stronger statement than (184), and at first sight looks rather strange. It amounts to making an ansatz for low-energy states $|\psi\rangle$ that fulfils

$$\langle \psi'|c(\mathbf{0})|\psi\rangle = \sqrt{N_0}\langle\psi'|\psi\rangle + \dots \tag{186}$$

where the dots denote terms that are small compared to $\sqrt{N_0}$. We'll return to what this implies for the structure of $|\psi\rangle$ a little later. Using (185) we may expand H in inverse powers of N_0

$$\begin{aligned} H &= \sum_{\mathbf{p}} \frac{\mathbf{p}^2}{2m} c^\dagger(\mathbf{p})c(\mathbf{p}) \\ &\quad + \frac{U}{2L^3}N_0^2 + \frac{UN_0}{2L^3}\sum_{\mathbf{k}\neq\mathbf{0}} 2c^\dagger(\mathbf{k})c(\mathbf{k}) + 2c^\dagger(-\mathbf{k})c(-\mathbf{k}) + c^\dagger(\mathbf{k})c^\dagger(-\mathbf{k}) + c(-\mathbf{k})c(\mathbf{k}) \\ &\quad + \dots \end{aligned} \tag{187}$$

Note that there is no term that goes as $N_0^{\frac{3}{2}}$ because setting the momentum of three of the creation/annihilation operators in (179) to zero forces the last momentum to be zero as well. Now use that

$$N_0 = c^\dagger(\mathbf{0})c(\mathbf{0}) = N - \sum_{\mathbf{p}\neq\mathbf{0}} c^\dagger(\mathbf{p})c(\mathbf{p}), \tag{188}$$

where N is the (conserved) total number of bosons, and define

$$\rho = \frac{N}{L^3} = \text{density of particles.} \tag{189}$$

Then our Hamiltonian becomes

$$\boxed{H = \frac{U\rho}{2}N + \sum_{\mathbf{p}\neq\mathbf{0}} \underbrace{\left[\frac{\mathbf{p}^2}{2m} + U\rho\right]}_{\epsilon(\mathbf{p})} c^\dagger(\mathbf{p})c(\mathbf{p}) + \frac{U\rho}{2}\left[c^\dagger(\mathbf{p})c^\dagger(-\mathbf{p}) + c(-\mathbf{p})c(\mathbf{p})\right] + \dots} \tag{190}$$

The Bogoliubov approximation has reduced the complicated four-boson interaction to two-boson terms. The price we pay is that we have to deal with the "pairing"-terms quadratic in creation/annihilation operators.

4.3 Bogoliubov Transformation

Consider the creation/annihilation operators defined by

$$\begin{pmatrix} b(\mathbf{p}) \\ b^\dagger(-\mathbf{p}) \end{pmatrix} = \begin{pmatrix} \cosh(\theta_\mathbf{p}) & \sinh(\theta_\mathbf{p}) \\ \sinh(\theta_\mathbf{p}) & \cosh(\theta_\mathbf{p}) \end{pmatrix} \begin{pmatrix} c(\mathbf{p}) \\ c^\dagger(-\mathbf{p}). \end{pmatrix} \tag{191}$$

It is easily checked that for any choice of Bogoliubov angle $\theta_\mathbf{p} = \theta_{-\mathbf{p}}$

$$[b(\mathbf{p}), b(\mathbf{q})] = 0 = [b^\dagger(\mathbf{p}), b^\dagger(\mathbf{q})] \ , \qquad [b(\mathbf{p}), b^\dagger(\mathbf{q})] = \delta_{\mathbf{p},\mathbf{q}}. \tag{192}$$

In terms of the Bogoliubov bosons the Hamiltonian becomes

$$\begin{aligned} H &= \text{const} + \frac{1}{2}\sum_{\mathbf{p}\neq\mathbf{0}} \left[\left(\frac{\mathbf{p}^2}{2m} + U\rho \right) \cosh(2\theta_\mathbf{p}) - U\rho \sinh(2\theta_\mathbf{p}) \right] \left[b^\dagger(\mathbf{p})b(\mathbf{p}) + b^\dagger(-\mathbf{p})b(-\mathbf{p}) \right] \\ &\quad - \left[\left(\frac{\mathbf{p}^2}{2m} + U\rho \right) \sinh(2\theta_\mathbf{p}) - U\rho \cosh(2\theta_\mathbf{p}) \right] \left[b^\dagger(\mathbf{p})b^\dagger(-\mathbf{p}) + b(-\mathbf{p})b(\mathbf{p}) \right] + \ldots \end{aligned} \tag{193}$$

Now we choose

$$\tanh(2\theta_\mathbf{p}) = \frac{U\rho}{\frac{\mathbf{p}^2}{2m} + U\rho}, \tag{194}$$

as this removes the $b^\dagger b^\dagger + bb$ terms, and leaves us with a *diagonal* Hamiltonian

$$\boxed{H = \text{const} + \sum_{\mathbf{p}\neq\mathbf{0}} E(\mathbf{p}) b^\dagger(\mathbf{p}) b(\mathbf{p}) + \ldots} \tag{195}$$

where

$$E(\mathbf{p}) = \sqrt{\left(\frac{\mathbf{p}^2}{2m} + U\rho \right)^2 - (U\rho)^2} \ . \tag{196}$$

We note that

$$E(\mathbf{p}) \longrightarrow \frac{\mathbf{p}^2}{2m} \qquad \text{for } |\mathbf{p}| \to \infty, \tag{197}$$

which tells us that at high momenta (and hence high energies) we recover the quadratic dispersion. In this limit $\theta_\mathbf{p} \to 0$, so that the Bogoliubov bosons reduce to the "physical" bosons we started with. On the other hand

$$E(\mathbf{p}) \longrightarrow \sqrt{\frac{U\rho}{m}} |\mathbf{p}| \qquad \text{for } |\mathbf{p}| \to 0. \tag{198}$$

So here we have a *linear* dispersion.

4.4 Ground State and Low-lying Excitations

We note that the Hamiltonian (195) involves only creation/annihilation operators with $\mathbf{p} \neq 0$. Formally, we can define zero-momentum Bogoliubov bosons as simply being equal to the original ones

$$b(\mathbf{0}) = c(\mathbf{0}) \ . \tag{199}$$

Let us now define the Bogoliubov vacuum state $|\tilde{0}\rangle$ by

$$b(\mathbf{p})|\tilde{0}\rangle = 0 \ . \tag{200}$$

Clearly, for $\mathbf{p} \neq 0$ we have $E(\mathbf{p}) > 0$, and hence no Bogoliubov quasiparticles will be present in the ground state. On the other hand, a basic assumption we made was that

$$\langle \text{GS}|b(\mathbf{0})|\text{GS}\rangle \simeq \sqrt{N_0}. \tag{201}$$

In order to get an idea what this implies for the structure of the ground state, let us express it in the general form

$$|\mathrm{GS}\rangle = \sum_{n=0}^{\infty} \alpha_n \big(b^\dagger(\mathbf{0})\big)^n |\tilde{0}\rangle \ . \tag{202}$$

Eqn (201) then implies that

$$\alpha_{n+1} \simeq \frac{\sqrt{N_0}}{n+1}\alpha_n \ . \tag{203}$$

Replacing this approximate relation by an equality leads to a *coherent state*

$$|\mathrm{GS}\rangle = e^{-N_0/2} e^{\sqrt{N_0}b^\dagger(\mathbf{0})}|\tilde{0}\rangle. \tag{204}$$

Low-lying excited states can now be obtained by creating Bogoliubov quasipartices, e.g.

$$b^\dagger(\mathbf{q})|\mathrm{GS}\rangle, \tag{205}$$

is a particle-excitation with energy $E(\mathbf{q}) > 0$. Note that the eigenstates of the Hamiltonian are obtained by acting with the Bogoliubov creation operators on the ground state and that the "elementary excitations" are therefore the Bogoliubov bosons rather than the original bosons. The way to understand this is to note that the ground state is a rather complicated many-boson state, and excitations involve the *collective* motion of all these bosons. Inspection of the Bogoliubov angle (194) shows that $\theta_{\mathbf{p}}$ goes to zero for large momenta $|\mathbf{p}|$, and concomitantly $b^\dagger(\mathbf{p}) \approx c^\dagger(\mathbf{p})$, *cf.* (191). In other words, at high energies the Bogoliubov bosons essentially reduce to the original ones, which is also reflected in the dispersion relation (197).

4.5 Ground state correlation functions

We are now in a position to work out correlation functions in the ground state such as

$$\langle \mathrm{GS}|c^\dagger(\mathbf{p})c(\mathbf{q})|\mathrm{GS}\rangle \ , \quad \mathbf{p},\mathbf{q} \neq \mathbf{0}. \tag{206}$$

Inverting the Bogoliubov transformation (191) we have

$$\begin{aligned} c^\dagger(\mathbf{p}) &= \cosh(\theta_{\mathbf{p}})b^\dagger(\mathbf{p}) - \sinh(\theta_{\mathbf{p}})b(-\mathbf{p}) \ , \\ c(\mathbf{q}) &= \cosh(\theta_{\mathbf{q}})b(\mathbf{q}) - \sinh(\theta_{\mathbf{q}})b^\dagger(-\mathbf{q}) \ . \end{aligned} \tag{207}$$

Using that

$$\langle \mathrm{GS}|b^\dagger(\mathbf{p}) = 0 = b(\mathbf{q})|\mathrm{GS}\rangle, \tag{208}$$

we find that

$$\begin{aligned} \langle \mathrm{GS}|c^\dagger(\mathbf{p})c(\mathbf{q})|\mathrm{GS}\rangle &= \sinh(\theta_{\mathbf{q}})\sinh(\theta_{\mathbf{q}})\langle \mathrm{GS}|b(-\mathbf{p})b^\dagger(-\mathbf{q})|\mathrm{GS}\rangle \\ &= \sinh^2(\theta_{\mathbf{p}})\delta_{\mathbf{q},\mathbf{p}} \quad (\mathbf{p},\mathbf{q} \neq \mathbf{0}). \end{aligned} \tag{209}$$

This tells us that, in contrast to the ideal Bose gas, in the ground state of the interacting Bose gas we have a finite density of bosons with non-zero momentum

$$\langle \mathrm{GS}|c^\dagger(\mathbf{p})c(\mathbf{p})|\mathrm{GS}\rangle = \sinh^2(\theta_{\mathbf{p}}) \ . \tag{210}$$

Another feature of the ground state is that the two-point function of two annihilation/creation operators is non-zero

$$\langle \mathrm{GS}|c(\mathbf{p})c(\mathbf{q})|\mathrm{GS}\rangle = \langle \mathrm{GS}|c^\dagger(\mathbf{q})c^\dagger(\mathbf{p})|\mathrm{GS}\rangle = -\cosh(\theta_{\mathbf{p}})\sinh(\theta_{\mathbf{q}})\delta_{\mathbf{p},-\mathbf{q}}. \tag{211}$$

4.6 Spontaneous Symmetry Breaking

Eqns (211) imply that boson number is not a good quantum number in the ground state. More formally, we say that *the ground state spontaneously breaks the U(1) symmetry of the Hamiltonian* $H = \hat{T} + \hat{V}$. Let us explain that statement. The Hamiltonian is invariant under the symmetry operation

$$\begin{aligned} \hat{U}c(\mathbf{p})\hat{U}^\dagger &= e^{i\phi}c(\mathbf{p}) \ , \quad \phi \in \mathbb{R} \ , \\ \hat{U}c^\dagger(\mathbf{p})\hat{U}^\dagger &= e^{-i\phi}c^\dagger(\mathbf{p}) \ , \end{aligned} \tag{212}$$

i.e.

$$\hat{U}H\hat{U}^\dagger = H. \tag{213}$$

The reason for this is that all terms in H involve the same number of creation as annihilation operators, and the total particle number is therefore conserved. This is referred to as a global U(1) symmetry (as the transformations (212) form a group called U(1)). Let us now investigate how ground state expectation values transform. We have

$$\langle \mathrm{GS}|c(\mathbf{p})c(\mathbf{q})|\mathrm{GS}\rangle = \langle \mathrm{GS}|\hat{U}^\dagger\hat{U}c(\mathbf{p})\hat{U}^\dagger\hat{U}c(\mathbf{q})\hat{U}^\dagger\hat{U}|\mathrm{GS}\rangle = e^{2i\phi}\langle \mathrm{GS}|\hat{U}^\dagger c(\mathbf{p})c(\mathbf{q})\hat{U}|\mathrm{GS}\rangle \ . \tag{214}$$

If the ground state were invariant under the symmetry, we would have $\hat{U}|\mathrm{GS}\rangle = |\mathrm{GS}\rangle$. Eqn (214) would then imply that $\langle \mathrm{GS}|c(\mathbf{p})c(\mathbf{q})|\mathrm{GS}\rangle = 0$. Reversing the argument, we see that a non-zero value of the expectation value (211) implies that the ground state *cannot be* invariant under the U(1) symmetry, and in fact *"breaks it spontaneously"*.

4.7 Depletion of the Condensate

We started out by asserting that for small interactions $U > 0$ we retain a Bose-Einstein condensate, i.e. the consensate fraction N_0/N remains large. We can now check that this assumption is *self-consistent*. We have

$$N_0 = N - \sum_{\mathbf{p}\neq\mathbf{0}} c^\dagger(\mathbf{p})c(\mathbf{p}). \tag{215}$$

Thus in the ground state

$$\frac{N_0}{N} = 1 - \frac{1}{N}\sum_{\mathbf{p}\neq\mathbf{0}}\langle \mathrm{GS}|c^\dagger(\mathbf{p})c(\mathbf{p})|\mathrm{GS}\rangle = 1 - \frac{1}{N}\sum_{\mathbf{p}\neq\mathbf{0}}\sinh^2(\theta_\mathbf{p}), \tag{216}$$

where we have used (209). This equals

$$\frac{N_0}{N} = 1 - \frac{1}{2N}\sum_{\mathbf{p}\neq\mathbf{0}}\left[\frac{1}{\sqrt{1-\tanh^2(2\theta_\mathbf{p})}} - 1\right] = 1 - \frac{1}{2N}\sum_{\mathbf{p}\neq\mathbf{0}}\left[\frac{1}{\sqrt{1-\left[\frac{U\rho}{\epsilon(\mathbf{p})}\right]^2}} - 1\right]. \tag{217}$$

We again turn this into an integral and evaluate it in spherical polar coordinates, which gives

$$\frac{N_0}{N} \approx 1 - \frac{2\pi}{\rho}\int_0^\infty \frac{dp}{(2\pi\hbar)^3}p^2\left[\frac{1}{\sqrt{1-\left[\frac{U\rho}{\epsilon(\mathbf{p})}\right]^2}} - 1\right]. \tag{218}$$

By means of the substitution $p = \sqrt{2mU\rho}z$ we can see that the integral is proportional to $U^{3/2}$ and thus indeed small for small U.

5 Application III: Spinwaves in a Ferromagnet

Consider the following model of a magnetic insulator: at each site $\mathbf{r}$ of a D-dimensional with N sites lattice we have a magnetic moment. In QM such magnetic moments are described by three spin-operators

$$S^\alpha_{\mathbf{r}} \ , \quad \alpha = x, y, z \ , \tag{219}$$

which fulfil the angular momentum commutation relations

$$[S^\alpha_{\mathbf{r}}, S^\beta_{\mathbf{r}'}] = \delta_{\mathbf{r},\mathbf{r}'} i\epsilon_{\alpha\beta\gamma} S^\gamma_{\mathbf{r}}. \tag{220}$$

We will assume that the spin are large in the sense that

$$\mathbf{S}^2_{\mathbf{r}} = \sum_\alpha \left(S^\alpha_{\mathbf{r}}\right)^2 = s(s+1) \gg 1. \tag{221}$$

Let us begin by constructing a basis of quantum mechanical states. At each site we have $2s+1$ eigenstates of $S^z_{\mathbf{r}}$

$$S^z_{\mathbf{r}}|m\rangle_{\mathbf{r}} = m|m\rangle_{\mathbf{r}}, \quad m = s, s-1, \ldots, -s. \tag{222}$$

They can be constructed from $|s\rangle_{\mathbf{r}}$ using spin lowering operators $S^-_{\mathbf{r}} = S^x_{\mathbf{r}} - iS^y_{\mathbf{r}}$

$$|s-n\rangle_{\mathbf{r}} = \frac{1}{\mathcal{N}_n}\left(S^-_{\mathbf{r}}\right)^n |s\rangle_{\mathbf{r}} \ , \quad n = 0, 1, \ldots, 2s, \tag{223}$$

where $\mathcal{N}_n$ are normalization constants. A basis of states is then given by

$$\prod_{\mathbf{r}} |s_{\mathbf{r}}\rangle_{\mathbf{r}} \ , \quad -s \le s_{\mathbf{r}} \le s \text{ spin on site } \mathbf{r}. \tag{224}$$

5.1 Heisenberg model and spin-rotational SU(2) symmetry

An appropriate Hamiltonian for a ferromagnetic insulator was derived by Heisenberg

$$\boxed{H = -J \sum_{\langle \mathbf{r},\mathbf{r}'\rangle} \mathbf{S}_{\mathbf{r}} \cdot \mathbf{S}_{\mathbf{r}'}.} \tag{225}$$

Here $\langle \mathbf{r}, \mathbf{r}'\rangle$ denote nearest-neighbour pairs of spins and we will assume that $J > 0$. The model (225) is known as the ferromagnetic *Heisenberg model*. You can check that the Hamiltonian (225) commutes with the three total spin operators

$$[H, S^\alpha] = 0 \ , \quad S^\alpha = \sum_{\mathbf{r}} S^\alpha_{\mathbf{r}}. \tag{226}$$

These imply that the Hamiltonian is invariant under general rotations (in spin space)

$$e^{i\boldsymbol{\alpha}\cdot\mathbf{S}} H e^{-i\boldsymbol{\alpha}\cdot\mathbf{S}} = H. \tag{227}$$

The transformations (227) form a group known as SU(2), and the Heisenberg Hamiltonian (225) is invariant under them.

5.2 Exact ground states

One ground state of H is given by

$$\boxed{|\mathrm{GS}\rangle = \prod_{\mathbf{r}} |s\rangle_{\mathbf{r}}.} \tag{228}$$

Its energy is

$$H|\mathrm{GS}\rangle = -J \sum_{\langle \mathbf{r},\mathbf{r}'\rangle} s^2 |\mathrm{GS}\rangle = -Js^2 N_B |\mathrm{GS}\rangle, \tag{229}$$

where N_B is the total number of bonds in our lattice. The total spin lowering operator $S^- = \sum_{\mathbf{r}} S^-_{\mathbf{r}}$ commutes with H by virtue of (226) and hence

$$|\mathrm{GS}, n\rangle = \frac{1}{N_n} \left(S^-\right)^n |\mathrm{GS}\rangle \ , \quad 0 \leq n \leq 2sN \tag{230}$$

are ground states as well (as they have the same energy). Here N_n is a normalization.

Aside 5: Proof that $|\mathrm{GS}\rangle$ is a ground state

We note that the spin-spin interaction can be written in the form

$$2\mathbf{S}_{\mathbf{r}} \cdot \mathbf{S}_{\mathbf{r}'} = (\mathbf{S}_{\mathbf{r}} + \mathbf{S}_{\mathbf{r}'})^2 - \mathbf{S}^2_{\mathbf{r}} - \mathbf{S}^2_{\mathbf{r}'} = \mathbf{J}^2 - 2s(s+1). \tag{231}$$

Here $\mathbf{J}^2$ is the total angular momentum squared. Its eigenvalues follow from the theory of adding angular momenta to be

$$\mathbf{J}^2 |j, m\rangle = j(j+1)|j, m\rangle \ , \quad j = 2s, 2s-1, \ldots, 1, 0. \tag{232}$$

This tells us that the maximal eigenvalue of $\mathbf{J}^2$ is $2s(2s+1)$, and by expanding $|\psi\rangle$ in a basis of eigenstates of $\mathbf{J}^2$ we can easily show that

$$\begin{aligned} \langle\psi|\mathbf{J}^2|\psi\rangle &= \sum_{j,m,j',m'} \langle\psi|j,m\rangle\langle j,m|\mathbf{J}^2|j',m'\rangle\langle j',m'|\psi\rangle \\ &= \sum_{j,m} |\langle\psi|j,m\rangle|^2 j(j+1) \leq 2s(2s+1) \sum_{j,m} |\langle\psi|j,m\rangle|^2 = 2s(2s+1). \end{aligned} \tag{233}$$

This tells us that

$$\langle\psi|\mathbf{S}_{\mathbf{r}} \cdot \mathbf{S}_{\mathbf{r}'}|\psi\rangle \leq s^2. \tag{234}$$

This provides us with a bound on the eigenvalues of the Hamiltonian, as

$$\langle\psi|H|\psi|\rangle \geq -J \sum_{\langle \mathbf{r},\mathbf{r}'\rangle} s^2 = -Js^2 Nz. \tag{235}$$

The state we have constructed saturates this bound, so must be a ground state.

Let us now see how the SU(2) symmetry is reflected in expectation values of operators $\mathcal{O}$. At finite temperature we have

$$\langle\mathcal{O}\rangle_\beta = \frac{1}{Z(\beta)} \mathrm{Tr}\left[e^{-\beta H}\mathcal{O}\right], \tag{236}$$

where $Z(\beta) = \mathrm{Tr}[e^{-\beta H}]$ is the partition function and $\beta = 1/k_B T$. In the $T \to 0$ limit we have

$$\langle\mathcal{O}\rangle_\infty = \frac{1}{2sN+1} \sum_{n=0}^{2sN} \langle\mathrm{GS}, \mathrm{n}|\mathcal{O}|\mathrm{GS}, \mathrm{n}\rangle, \tag{237}$$

i.e. we average over all ground states. The thermal average, as well as its $T = 0$ limit, are invariant under rotations in spin space. Indeed, under a rotation in spin space we have

$$\langle e^{i\boldsymbol{\alpha}\cdot\mathbf{S}}\mathcal{O}e^{-i\boldsymbol{\alpha}\cdot\mathbf{S}}\rangle_\beta = \frac{1}{Z(\beta)}\mathrm{Tr}\left[e^{-\beta H}e^{i\boldsymbol{\alpha}\cdot\mathbf{S}}\mathcal{O}e^{-i\boldsymbol{\alpha}\cdot\mathbf{S}}\right] \tag{238}$$

where $\mathbf{S} = \sum_{\mathbf{r}}\mathbf{S}_{\mathbf{r}}$ are the global spin operators. Using the cyclicity of the trace and the fact that H commutes with the global spin operators, we see that this equals $\langle\mathcal{O}\rangle_\beta$. If we choose as our operator $\mathcal{O}$ any of the global spin operators, and consider a rotation by π around one of the orthogonal axes, we see that the magnetization always vanishes

$$\langle S^\alpha\rangle_\beta = 0\ , \quad \alpha = x, y, z. \tag{239}$$

Physically this is what one would expect for a system that is spin rotationally invariant, i.e. looks the same in any direction in spin space.

5.3 Spontaneous Symmetry Breaking

In a real system, the $2sN + 1$-fold ground state degeneracy is usually broken through imperfections. In practice the details of these imperfections are not important, the only thing that matters is that the symmetry gets broken. To keep things simple, one retains the spin-rotationally symmetric Hamiltonian, and says that the ground state breaks the symmetry "spontaneously".

A convenient mathematical description of this effect is as follows. Imagine adding an infinitesimal magnetic field $-\epsilon\sum_{\mathbf{r}}S^z_{\mathbf{r}}$ to the Hamiltonian. This will break the symmetry and hence the degeneracy of the ground states, which now will have energies

$$E_{\mathrm{GS,n}} = -Js^2N_B - \epsilon(sN - n). \tag{240}$$

Now consider the sequence of limits

$$\lim_{\epsilon\to 0}\lim_{N\to\infty}[E_{\mathrm{GS,n}} - E_{\mathrm{GS},0}] = \begin{cases} 0 & \text{if } \lim_{N\to\infty}\frac{n}{N} = 0\ , \\ \infty & \text{else.} \end{cases} \tag{241}$$

This means that if we *define* the thermodynamic limit in the above way, then the only surviving ground states will have magnetization per site s, i.e. contain only a non-extensive number of spin flips. In all of these remaining ground states the spin rotational symmetry has been broken. As we have taken $\epsilon \to 0$ our Hamiltonian is again SU(2) symmetric, but the remaining ground states "spontaneously" break this symmetry.

5.4 Holstein-Primakoff Transformation

We succeeded in finding the ground states of H because of their simple structure. For more general spin Hamiltonians, or even the Hamiltonian (225) with negative value of J, this will no longer work and we need a more general, but approximate way of dealing with such problems. This is provided by *(linear) spinwave theory.*

As shown by Holstein and Primakoff, spin operators can be represented in terms of bosonic creation and annihilation operators as follows:

$$S^z_{\mathbf{r}} = s - a^\dagger_{\mathbf{r}}a_{\mathbf{r}}\ , \quad S^+_{\mathbf{r}} = S^x_{\mathbf{r}} + iS^y_{\mathbf{r}} = \sqrt{2s}\sqrt{1 - \frac{a^\dagger_{\mathbf{r}}a_{\mathbf{r}}}{2s}}a_{\mathbf{r}}\ . \tag{242}$$

You can check that the bosonic commutation relations

$$[a_{\mathbf{r}}, a^\dagger_{\mathbf{r}'}] = \delta_{\mathbf{r},\mathbf{r}'} \tag{243}$$

imply that

$$[S^\alpha_{\mathbf{r}}, S^\beta_{\mathbf{r}'}] = \delta_{\mathbf{r},\mathbf{r}'} i\epsilon_{\alpha\beta\gamma} S^\gamma_{\mathbf{r}}. \tag{244}$$

However, there is a caveat: the spaces of QM states are different! At site $\mathbf{r}$ we have

$$\left(S_{\mathbf{r}}\right)^n |s\rangle_{\mathbf{r}} \ , \quad n = 0, \ldots, 2s \tag{245}$$

for spins, but for bosons there are infinitely many states

$$\left(a^\dagger_{\mathbf{r}}\right)^n |0\rangle_{\mathbf{r}} \ , \quad n = 0, \ldots, \infty. \tag{246}$$

To make things match, we must impose a *constraint*, that there are at most 2s bosons per site. Now we take advantage of the fact that we have assumed s to be large: in the ground state there are no bosons present, because

$$\langle \mathrm{GS}|s - a^\dagger_{\mathbf{r}} a_{\mathbf{r}}|\mathrm{GS}\rangle = \langle \mathrm{GS}|S^z_{\mathbf{r}}|\mathrm{GS}\rangle = s = \tag{247}$$

Low-lying excited states will only have a few bosons, so for large enough s we don't have to worry about the constraint. Using the Holstein-Primakoff transformation, we can rewrite H in a $1/s$ *expansion*

$$\boxed{H = -J \sum_{\langle \mathbf{r},\mathbf{r}'\rangle} s^2 - s\left[a^\dagger_{\mathbf{r}} a_{\mathbf{r}} + a^\dagger_{\mathbf{r}'} a_{\mathbf{r}'} - a^\dagger_{\mathbf{r}} a_{\mathbf{r}'} - a^\dagger_{\mathbf{r}'} a_{\mathbf{r}}\right] + \ldots} \tag{248}$$

Here the dots indicate terms proportional to s^0, s^{-1}, etc. Once again using that s is large, we drop these terms (for the time being). We then can diagonalize H by going to momentum space

$$a_{\mathbf{r}} = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}} a(\mathbf{k}) \ , \qquad [a(\mathbf{k}), a^\dagger(\mathbf{p})] = \delta_{\mathbf{k},\mathbf{p}} \ , \tag{249}$$

Exercise 2: Quantization conditions for lattice "momenta"

As we are dealing with a lattice model translational invariance is now reduced to translations by multiples of the lattice spacing in the various crystallographic directions. On a hypercubic lattice with periodic boundary conditions we have

$$a(\mathbf{r} + L\mathbf{e}_a) = a(\mathbf{r}), \tag{250}$$

where $\mathbf{e}_a$ denotes the unit vector in the a direction and L is the linear size of our system (so $N = L^D$). In D=1 periodic boundary conditions imply that our spin model lives on a ring, and (250) then simply means that if we go around the ring once we return precisely to where we started. Out goal is to define a set of N linearly independent operators

$$a(\mathbf{k}) = \frac{1}{\sqrt{N}} \sum_{\mathbf{r}} e^{-i\mathbf{k}\cdot\mathbf{r}} a_{\mathbf{r}} \ . \tag{251}$$

Using (250) we conclude that

$$a(\mathbf{k}) = \frac{1}{\sqrt{N}} \sum_{\mathbf{r}} e^{-i\mathbf{k}\cdot\mathbf{r}} a_{\mathbf{r}+L\mathbf{e}_a} = \frac{1}{\sqrt{N}} \sum_{\mathbf{r}} e^{-i\mathbf{k}\cdot(\mathbf{r}-L\mathbf{e}_a)} a_{\mathbf{r}}, \tag{252}$$

which leads to the conditions

$$e^{iL\mathbf{k}\cdot\mathbf{e}_a} = 1. \tag{253}$$

A solution to these conditions that provides us with linearly independent $a(\mathbf{k})$'s is

$$k_a = \frac{2\pi}{L} n_a \ , \quad n_a = 1, 2, \ldots, L. \tag{254}$$

Let us now work out the commutation relations

$$\begin{aligned} [a(\mathbf{k}), a^\dagger(\mathbf{p})] &= \frac{1}{N}\sum_{\mathbf{r},\mathbf{r}'} e^{-i\mathbf{k}\cdot\mathbf{r}} e^{i\mathbf{p}\cdot\mathbf{r}'} \underbrace{[a(\mathbf{r}), a^\dagger(\mathbf{r}')]}_{\delta_{\mathbf{r},\mathbf{r}'}} \\ &= \frac{1}{N}\sum_{\mathbf{r}} e^{-i(\mathbf{k}-\mathbf{p})\cdot\mathbf{r}} = \frac{1}{N}\prod_{a=1}^{D}\underbrace{\sum_{r_a=1}^{L} e^{-i(k_a-p_a)r_a}}_{L\delta_{k_a,p_a}} = \delta_{\mathbf{k},\mathbf{p}}. \end{aligned} \tag{255}$$

In terms of creation and annihilation operators in momentum space the Hamiltonian becomes

$$\boxed{H = -Js^2Nz + \sum_{\mathbf{q}} \epsilon(\mathbf{q}) a^\dagger(\mathbf{q}) a(\mathbf{q}) + \ldots} \tag{256}$$

For a simple cubic lattice the energy is

$$\epsilon(\mathbf{q}) = 2Js\left[3 - \cos q_x - \cos q_x - \cos q_z\right]. \tag{257}$$

For small wave numbers this is quadratic

$$\epsilon(\mathbf{q}) \approx Js|\mathbf{q}|^2 \qquad \text{for } |\mathbf{q}| \to 0. \tag{258}$$

In the context of spontaneous symmetry breaking these gapless excitations are known as *Goldstone modes*.

Let us now revisit the logic underlying our $1/s$ expansion. For things to be consistent, we require that the terms of order s in (256) provide only a small correction to the leading contribution proportional to s^2. This will be the case as long as we are interested only is states $|\Psi\rangle$ such that

$$\langle\Psi|a^\dagger(\mathbf{q})a(\mathbf{q})|\Psi\rangle \ll s. \tag{259}$$

This condition is certainly fulfilled for the ground state and low-lying excited states.

5.4.1 Heisenberg Antiferromagnet

Another example to which spinwave theory can be fruitfully applied is the model

$$H = J\sum_{\langle\mathbf{r},\mathbf{r}'\rangle} \mathbf{S}_{\mathbf{r}}\cdot\mathbf{S}_{\mathbf{r}'}, \tag{260}$$

where $\langle\mathbf{r},\mathbf{r}'\rangle$ denote nearest-neighbour pairs of spins on a simple cubic lattice and $J > 0$. Compared to (225) all we have done is to switch the overall sign of H, but this has important consequences. In particular, it is no longer possible to obtain an exact ground state for the model. Instead, we start by considering our spins to be *classical*. This is justified if we are interested only in states with large spin quantum numbers. We will assume this to be the case and check the self-consistency of our assumption later. In the classical limit we can think of the spins as three-dimensional vectors. The lowest energy configuration is then one, where all neighbouring spins point in opposite directions, i.e. along the three cystal axes the spin configuration

looks like ↑↓↑↓↑↓ .... This is known as a *Néel state*. It is convenient to subdivide our lattice into two sublattices: on sublattice A all spins point in the same direction, while on sublattice B all spins point in the opposite direction. Like the ferromagnet, the model (260) has a global spin-rotational symmetry, that will be spontaneously broken in the ground state. By choosing our spin quantization axis appropriately, the classical ground state can then be written in the form

$$\prod_{\mathbf{r}\in A} |s\rangle_{\mathbf{r}} \prod_{\mathbf{r}'\in B} |-s\rangle_{\mathbf{r}'} \tag{261}$$

The idea is now to map this state to a ferromagnetic one, by inverting the spin quantization axis in the B sublattice. After that we can employ the Holstein-Primakoff transformation to carry out a $1/s$ expansion. As a result of the rotation of spin quatization axis on the B sublattice, the part of the Hamiltonian of order s now contains terms involving two annihilation or two creation operators. Diagonalizing the Hamiltonian then requires a Bogoliubov transformation.

5.5 Homework Questions 7-8

Question 7. A magnetic system consists of two types of Heisenberg spin $\mathbf{S}^A$ and $\mathbf{S}^B$ located respectively on the two inter-penetrating sublattices of an NaCl crystal structure (i.e. a simple cubic structure with alternate A and B in any Cartesian direction). Its Hamiltonian is

$$H = J \sum_{i,j} \mathbf{S}_i^A \cdot \mathbf{S}_j^B$$

where the i, j are nearest neighbours, respectively on the A and B sublattices. J is positive. Show that the classical ground state has all the A spins ferromagnetically aligned in one direction and all the B spins ferromagnetically aligned in the opposite direction. Assume the quantum mechanical ground state is well approximated by the classical one. To a first approximation the spin operators are given in terms of boson operators a, b by

A sublattice	B sublattice
$S_i^z = S^A - a_i^\dagger a_i$	$S_j^z = -S^B + b_j^\dagger b_j$
$S_i^+ \equiv S_i^x + \mathrm{i}S_i^y \simeq (2S^A)^{1/2} a_i$	$S_j^+ \equiv S_j^x + \mathrm{i}S_j^y \simeq (2S^B)^{1/2} b_j^\dagger$
$S_i^- \equiv S_i^x - \mathrm{i}S_i^y \simeq (2S^A)^{1/2} a_i^\dagger$	$S_j^- \equiv S_j^x - \mathrm{i}S_j^y \simeq (2S^B)^{1/2} b_j$

Discuss the validity of this approximation. Use these relations to express the Hamiltonian in terms of the boson operators to quadratic order.

Transforming to crystal momentum space using (with N the number of sites on one sublattice)

$$a_i = N^{-1/2} \sum_{\mathbf{k}} \mathrm{e}^{-\mathrm{i}\mathbf{k}\cdot\mathbf{r}_i} a_{\mathbf{k}}, \quad b_j = N^{-1/2} \sum_{\mathbf{k}} \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}_j} b_{\mathbf{k}}$$

show that your result can be expressed in the form

$$H = E_0 + \sum_{\mathbf{k}} \left[A_{\mathbf{k}} a_{\mathbf{k}}^\dagger a_{\mathbf{k}} + B_{\mathbf{k}} b_{\mathbf{k}}^\dagger b_{\mathbf{k}} + C_{\mathbf{k}} (a_{\mathbf{k}}^\dagger b_{\mathbf{k}}^\dagger + b_{\mathbf{k}} a_{\mathbf{k}}) \right]$$

and determine the coefficients. Hence calculate the spectrum of excitations at low momenta. Consider both the cases with $S^A = S^B$ and $S^A \neq S^B$ and comment on your results.

Question 8. (optional) Consider the ideal Fermi gas at finite density N/V in a periodic 3-dimensional box of length L.
(a) Give an expression of the ground state in terms of creation operators for momentum eigenstates.
(b) Calculate the **single-particle Green's function**

$$\begin{aligned} G_{\sigma\tau}(\omega, \mathbf{q}) &= \int dt\ e^{i\omega(t-t')} \int d^3\mathbf{r}\ e^{-i\mathbf{q}\cdot(\mathbf{r}-\mathbf{r}')} G_{\sigma\tau}(t, \mathbf{r}; t', \mathbf{r}')\ , \\ G_{\sigma\tau}(t, \mathbf{r}; t', \mathbf{r}') &= -i\langle GS | T c_\sigma(\mathbf{r}, t)\ c_\tau^\dagger(\mathbf{r}', t') | GS \rangle, \end{aligned} \tag{262}$$

where T denotes time-ordering (i.e. $T\mathcal{O}(t_1)\mathcal{O}(t_2) = \theta(t_1 - t_2)\mathcal{O}(t_1)\mathcal{O}(t_2) - \theta(t_2 - t_1)\mathcal{O}(t_2)\mathcal{O}(t_1)$ for fermionic operators), and

$$c_\sigma(\mathbf{r}, t) \equiv e^{\frac{i}{\hbar}Ht} c_\sigma(\mathbf{r}) e^{-\frac{i}{\hbar}Ht}. \tag{263}$$

First express the creation/annihilation operators $c^\dagger_\sigma(\mathbf{r}, t)$, $c_\sigma(\mathbf{r}, t)$ in terms of creation/annihilation operators in momentum space $c^\dagger_\sigma(\mathbf{p}, t)$, $c_\sigma(\mathbf{p}, t)$. Then show that for annihilation operators in momentum space we have

$$c_\sigma(\mathbf{p}, t) \equiv e^{\frac{i}{\hbar}Ht} c_\sigma(\mathbf{p}) e^{-\frac{i}{\hbar}Ht} = c_\sigma(\mathbf{p}) e^{-\frac{i}{\hbar}t\epsilon(\mathbf{p})} \ , \tag{264}$$

where $\epsilon(\mathbf{p}) = \mathbf{p}^2/2m - \mu$. Use this to show that

$$c_\sigma(\mathbf{r}, t) = \frac{1}{L^{3/2}} \sum_{\mathbf{p}} e^{-\frac{i}{\hbar}t\epsilon(\mathbf{p}) + i\mathbf{p}\cdot\mathbf{r}} \ c_\sigma(\mathbf{p}). \tag{265}$$

Now insert (265) into (262) and evaluate the ground state expectation value to obtain an integral representation for $G_{\sigma\tau}(t, \mathbf{r}; t', \mathbf{r}')$. Why does the Green's function only depend on $t - t'$ and $\mathbf{r} - \mathbf{r}'$? Finally, calculate $G_{\sigma\tau}(\omega, \mathbf{q})$.

Note: the imaginary part of the single-particle Green's function is (approximately) measured by angle resolved photoemission (ARPES) experiments.

Part II
Phases and 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, a cartoon of which is shown in Fig. 4.

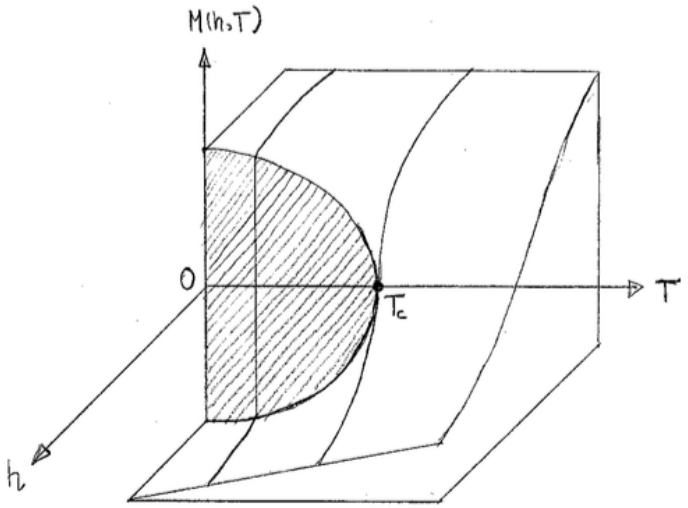


Figure 4: Equilibrium phase diagram for the paramagnetic to ferromagnetic transition. The *magnetization* $M(h,T)$ jumps when crossing zero for $T < T_c$. When lowering the temperature at $h = 0^+$ a *spontaneous magnetization* develops at a critical temperature $T = T_c$ and grows with decreasing temperature.

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

$$Z = \sum_{\text{configurations C}} e^{-E(C)/k_B T} \tag{266}$$

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. 4 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. 4 is second order.

6 The Ising Model

The two-dimensional square-lattice Ising model is one of the simplest statistical models to show a phase transition and as a result has played a hugely important role in the development of the theory of critical phenomena.

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 $\{\sigma_1, \sigma_2, \ldots, \sigma_N\}$ specifies a *configuration*. The corresponding energy is taken to be of the form

$$E(\{\sigma_j\}) = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j - h \sum_{j=1}^{N} \sigma_j \ , \tag{267}$$

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.

6.1 Statistical mechanics of the Ising model

The partition function of the model is

$$Z = \sum_{\sigma_1 = \pm 1} \sum_{\sigma_2 = \pm 1} \cdots \sum_{\sigma_N = \pm 1} e^{-\beta E(\{\sigma_j\})}. \tag{268}$$

The *magnetization per site* is given by

$$m(h) = \frac{1}{N} \langle \sum_{j=1}^{N} \sigma_j \rangle_\beta = \frac{1}{N\beta} \frac{\partial}{\partial h} \ln(Z). \tag{269}$$

The *magnetic susceptibility* is defined as

$$\chi(h) = \frac{\partial m(h)}{\partial h} = \frac{1}{N\beta} \frac{\partial^2}{\partial h^2} \ln(Z). \tag{270}$$

Substituting the expression (268) for the partition function and then carrying out the derivatives it can be expressed in the form

$$\chi(h) = \frac{\beta}{N} \sum_{l,m=1}^{N} \langle \sigma_l \sigma_m \rangle_\beta - \langle \sigma_l \rangle_\beta \langle \sigma_m \rangle_\beta. \tag{271}$$

6.2 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

$$E = \sum_{j=1}^{N} \left[-J \sigma_j \sigma_{j+1} - \frac{h}{2} (\sigma_j + \sigma_{j+1}) \right] \equiv \sum_{j=1}^{N} E(\sigma_j, \sigma_{j+1}), \tag{272}$$

where we have defined

$$\sigma_{N+1} = \sigma_1. \tag{273}$$

The partition function can be expressed in the form

$$Z = \sum_{\sigma_1,\ldots,\sigma_N} \prod_{j=1}^{N} e^{-\beta E(\sigma_j,\sigma_{j+1})}. \tag{274}$$

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

6.2.1 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

$$T_{\sigma\sigma'} = e^{-\beta E(\sigma,\sigma')}. \tag{275}$$

Its explicit form is

$$T = \begin{pmatrix} T_{++} & T_{+-} \\ T_{-+} & T_{--} \end{pmatrix} = \begin{pmatrix} e^{\beta(J+h)} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta(J-h)} \end{pmatrix}. \tag{276}$$

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

$$Z = \sum_{\sigma_1,\ldots,\sigma_N} T_{\sigma_1\sigma_2} T_{\sigma_2\sigma_3} \ldots T_{\sigma_{N-1}\sigma_N} T_{\sigma_N\sigma_1} \tag{277}$$

As desired, this has the structure of a matrix multiplication

$$\boxed{Z = \mathrm{Tr}\left(T^N\right).} \tag{278}$$

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

$$U^\dagger T U = \begin{pmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{pmatrix}, \tag{279}$$

where U is a unitary matrix and

$$\lambda_\pm = e^{\beta J}\cosh(\beta h) \pm \sqrt{e^{2\beta J}\sinh^2(\beta h) + e^{-2\beta J}}. \tag{280}$$

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

$$Z = \mathrm{Tr}\left(UU^\dagger T^N\right) = \mathrm{Tr}\left(U^\dagger T^N U\right) = \mathrm{Tr}\left([U^\dagger T U]^N\right) = \mathrm{Tr}\begin{pmatrix} \lambda_+^N & 0 \\ 0 & \lambda_-^N \end{pmatrix} = \lambda_+^N + \lambda_-^N. \tag{281}$$

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

$$Z = \lambda_+^N \left(1 + \left[\frac{\lambda_-}{\lambda_+}\right]^N\right) = \lambda_+^N \left(1 + e^{-N\ln(\lambda_+/\lambda_-)}\right). \tag{282}$$

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

$$\boxed{Z \simeq \lambda_+^N.} \tag{283}$$

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

$$m(h) = \frac{1}{N\beta}\frac{\partial}{\partial h}\ln(Z). \tag{284}$$

In Fig. 5 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 $h \to 0$, 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.

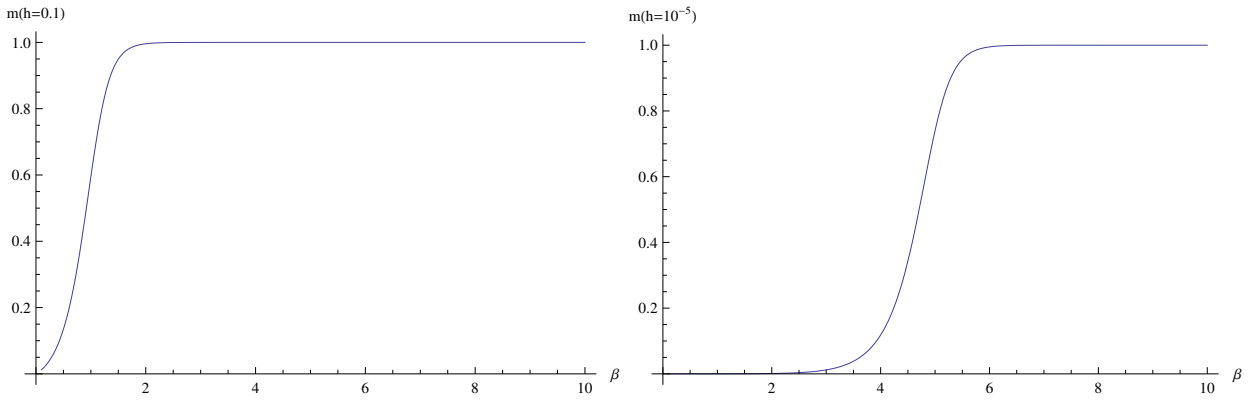


Figure 5: Magnetization per site as a function of inverse temperature for two values of applied magnetic field. We see that when we reduce the magnetic field, the temperature region in which the magnetization is essentially zero grows.

6.2.2 Averages of observables in the transfer matrix formalism

The average magnetization at site j is

$$\langle \sigma_j \rangle_\beta = \frac{1}{Z} \sum_{\sigma_1,\ldots,\sigma_N} \sigma_j e^{-\beta E(\{\sigma_j\})}. \tag{285}$$

We can express this in terms of the transfer matrix as

$$\langle \sigma_j \rangle_\beta = \frac{1}{Z} \sum_{\sigma_1,\ldots,\sigma_N} T_{\sigma_1\sigma_2} T_{\sigma_2\sigma_3} \ldots T_{\sigma_{j-1}\sigma_j} \sigma_j T_{\sigma_j\sigma_{j+1}} \ldots T_{\sigma_N\sigma_1}. \tag{286}$$

Using that

$$\left(T\sigma^z\right)_{\sigma_{j-1}\sigma_j} = T_{\sigma_{j-1}\sigma_j}\sigma_j \ , \tag{287}$$

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

$$\langle \sigma_j \rangle_\beta = \frac{1}{Z} \mathrm{Tr}\left[T^{j-1}\sigma^z T^{N-j+1}\right] = \frac{1}{Z}\mathrm{Tr}\left[T^N \sigma^z\right]. \tag{288}$$

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

$$\langle \sigma_j \rangle_\beta = \frac{1}{Z}\mathrm{Tr}\left[U^\dagger T^N U U^\dagger \sigma^z U\right] = \frac{1}{Z}\mathrm{Tr}\left[\begin{pmatrix} \lambda_+^N & 0 \\ 0 & \lambda_-^N \end{pmatrix} U^\dagger \sigma^z U\right] \ . \tag{289}$$

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

$$T|\pm\rangle = \lambda_\pm |\pm\rangle \tag{290}$$

as

$$U = (|+\rangle, |-\rangle). \tag{291}$$

For $h = 0$ we have

$$U|_{h=0} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{292}$$

This gives

$$\langle \sigma_j \rangle_\beta \Big|_{h=0} = 0. \tag{293}$$

For general h the expression is more complicated

$$U = \begin{pmatrix} \frac{\alpha_+}{\sqrt{1+\alpha_+^2}} & \frac{\alpha_-}{\sqrt{1+\alpha_-^2}} \\ \frac{1}{\sqrt{1+\alpha_+^2}} & \frac{1}{\sqrt{1+\alpha_-^2}} \end{pmatrix} , \quad \alpha_\pm = \sqrt{1+e^{4\beta J}\sinh^2(\beta h)} \pm e^{2\beta J}\sinh(\beta h). \tag{294}$$

The magnetization per site in the thermodynamic limit is then

$$\lim_{N\to\infty}\langle\sigma_j\rangle_\beta = \lim_{N\to\infty}\frac{\left(\frac{\alpha_+^2-1}{\alpha_+^2+1}\right)\lambda_+^N + \left(\frac{\alpha_-^2-1}{\alpha_-^2+1}\right)\lambda_-^N}{\lambda_+^N+\lambda_-^N} = \left(\frac{\alpha_+^2-1}{\alpha_+^2+1}\right). \tag{295}$$

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

$$\boxed{\lim_{h\to 0}\lim_{N\to\infty}\langle\sigma_j\rangle_\beta = 0 , \quad \beta<\infty.} \tag{296}$$

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

$$\langle\sigma_j\sigma_{j+r}\rangle_\beta = \frac{1}{Z}\mathrm{Tr}\left[T^{j-1}\sigma^z T^r \sigma^z T^{N+1-j-r}\right] = \frac{1}{Z}\mathrm{Tr}\left[U^\dagger\sigma^z U\begin{pmatrix}\lambda_+^r & 0\\ 0 & \lambda_-^r\end{pmatrix}U^\dagger\sigma^z U\begin{pmatrix}\lambda_+^{N-r} & 0\\ 0 & \lambda_-^{N-r}\end{pmatrix}\right]. \tag{297}$$

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

$$\langle\sigma_j\sigma_{j+r}\rangle_\beta\Big|_{h=0} = \frac{\lambda_+^{N-r}\lambda_-^r + \lambda_-^{N-r}\lambda_+^r}{\lambda_+^N+\lambda_-^N} \approx \left[\frac{\lambda_-}{\lambda_+}\right]^r = e^{-r/\xi}. \tag{298}$$

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

$$\xi = \frac{1}{\ln\coth(\beta J)}. \tag{299}$$

6.3 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

$$\sigma_{j,k} = \pm 1 , \quad j,k = 1,\ldots,N. \tag{300}$$

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

$$E(\{\sigma_{k,j}\}) = -J\sum_{j,k}\sigma_{k,j}\sigma_{k,j+1} + \sigma_{k,j}\sigma_{k+1,j}. \tag{301}$$

6.3.1 Transfer Matrix Method

The partition function is given by

$$Z = \sum_{\{\sigma_{j,k}\}} e^{-\beta E(\{\sigma_{k,j}\})}. \tag{302}$$

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

$$Z = \sum_{\{\sigma_{j,k}\}} e^{-\beta\sum_{k=1}^N E(k;k+1)}, \tag{303}$$

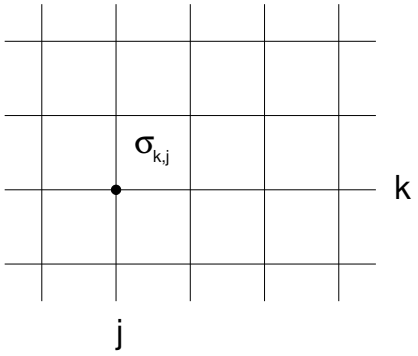


Figure 6: Ising model on the square lattice.

where

$$E(k;k+1) = -J\sum_{j=1}^{N}\sigma_{k,j}\sigma_{k+1,j} + \frac{1}{2}\left[\sigma_{k,j}\sigma_{k,j+1} + \sigma_{k+1,j}\sigma_{k+1,j+1}\right]. \tag{304}$$

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

$$|\boldsymbol{s}\rangle \tag{305}$$

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

$$\langle \boldsymbol{t}|\boldsymbol{s}\rangle = \prod_{j=1}^{N}\delta_{t_j,s_j}. \tag{306}$$

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

$$I = \sum_{\boldsymbol{s}}|\boldsymbol{s}\rangle\langle\boldsymbol{s}|. \tag{307}$$

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

$$\langle\boldsymbol{\sigma}_k|T|\boldsymbol{\sigma}_{k+1}\rangle = e^{-\beta E(k;k+1)}. \tag{308}$$

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

$$Z = \sum_{\boldsymbol{\sigma}_1}\sum_{\boldsymbol{\sigma}_2}\cdots\sum_{\boldsymbol{\sigma}_N}\langle\boldsymbol{\sigma}_1|T|\boldsymbol{\sigma}_2\rangle\langle\boldsymbol{\sigma}_2|T|\boldsymbol{\sigma}_3\rangle\ldots\langle\boldsymbol{\sigma}_{N-1}|T|\boldsymbol{\sigma}_N\rangle\langle\boldsymbol{\sigma}_N|T|\boldsymbol{\sigma}_1\rangle \tag{309}$$

We now may use (307) to carry out the sums over spins, which gives

$$\boxed{Z = \mathrm{Tr}\left[T^N\right]}, \tag{310}$$

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

$$Z = \sum_{j=1}^{2^N}\lambda_j^N\ , \tag{311}$$

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

$$F = -k_B T \ln(Z) = -k_B T \ln\left[\lambda_{\max}^N \sum_{j=1}^{2^N} \left(\frac{\lambda_j}{\lambda_{\max}}\right)^N\right] = -k_B T N \ln(\lambda_{\max}) - k_B T \ln\left[\sum_j \left(\frac{\lambda_j}{\lambda_{\max}}\right)^N\right], \quad (312)$$

where $\lambda_{\max}$ is the largest eigenvalue of T, which we assume to be unique. As $|\lambda_j/\lambda_{\max}| < 1$, the second contribution in (312) is bounded by $-k_B T N \ln(2)$, and we see that in the thermodynamic limit the free energy per site is

$$\boxed{f = \lim_{N\to\infty} \frac{F}{N^2} = \lim_{N\to\infty} -\frac{k_B T}{N} \ln(\lambda_{\max}).} \quad (313)$$

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 (302) 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.

6.3.2 Spontaneous Symmetry Breaking

Surprisingly, the transfer matrix of the 2D Ising model can be diagonalized exactly. Unfortunately we don't have the time do 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.7. At low temperatures $T < T_c$

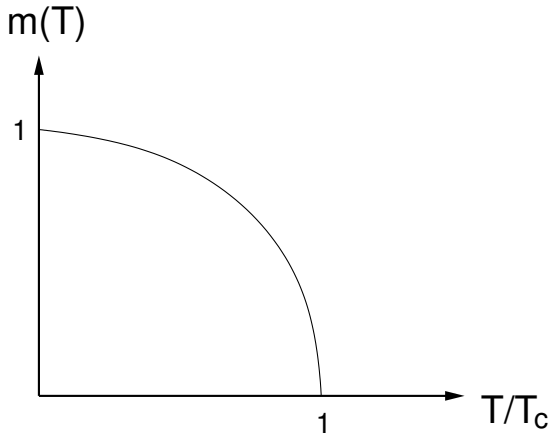


Figure 7: Phase Transition in the square lattice Ising model.

there is a non-zero magnetization per site, even though we did not apply a magnetic field. This is surprising, because our energy (301) is unchanged if we flip all spins

$$\sigma_{j,k} \to -\sigma_{j,k}. \quad (314)$$

The operation (314) is a discrete (two-fold) symmetry of the Ising model. Because we have translational invariance, the magnetization per site is

$$m = \langle \sigma_{j,k} \rangle_\beta. \quad (315)$$

Hence a non-zero value of m signifies the *spontaneous breaking* of the discrete symmetry (314). 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

$$\delta E = -\epsilon \sum_{j,k} \sigma_{j,k} \tag{316}$$

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

$$E_\pm = -JN_B \mp \epsilon N \ , \tag{317}$$

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

$$f(T) \equiv \lim_{\epsilon \to 0} \lim_{N \to \infty} \frac{-k_B T \ln(Z)}{N^2}. \tag{318}$$

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

$$\frac{Z_-}{Z_+} = e^{-2\epsilon N/k_B T} \ . \tag{319}$$

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

$$\boxed{\lim_{\epsilon \to 0} \lim_{N \to \infty} Z \neq \lim_{N \to \infty} \lim_{\epsilon \to 0} Z.} \tag{320}$$

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.

6.4 Homework Questions 9-11

Question 9. 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

$$\mathcal{H} = -J \sum_{\langle ij \rangle} n_i n_j - \mu \sum_i n_i \, , \tag{321}$$

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

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

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

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

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

$$H = -J \sum_{i=0}^{L-1} \delta_{\sigma_i, \sigma_{i+1}}, \tag{322}$$

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 (322). 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 $L \to \infty$

$$f = -k_B T \ln \left(3 + e^{\beta J} - 1 \right). \tag{323}$$

(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

$$\langle \delta_{\sigma_\ell, 1} \rangle = \frac{1}{3} + \frac{2}{3} e^{-\ell/\xi}. \tag{324}$$

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

Question 11. 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. 8. The energy is

$$E = -J \sum_{j=1}^{N-1} \sigma_j \sigma_{j+1} + 2h \sum_{j=1}^{(N-1)/2} \sigma_{2j}. \tag{325}$$

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

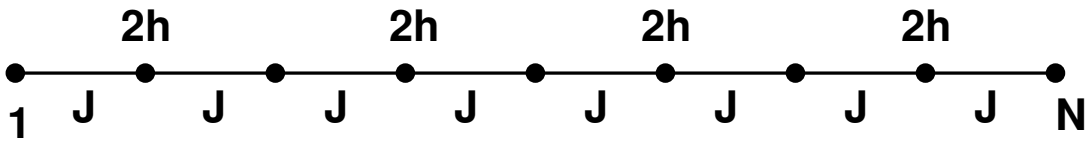


Figure 8: Open Ising chain with magnetic field applied to all even sites.

$$Z = \langle u | T^{(N-1)/2} | v \rangle \ , \tag{326}$$

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 $h = 0$.

6.5 Peierls Argument

The 2D Ising model can be solved exactly by the transfer matrix method introduced above. Interestingly, it is possible to establish the existence of a finite-temperature phase transition in the model without solving it. This was pioneered by Sir Rudolf Peierls, who was the Wykeham Professor of Theoretical Physics at

Oxford for many years and during his career made numerous ground breaking contributions in particular to condensed matter physics.

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.

Consider the Ising model on the square lattice with boundary conditions such that all spins on the boundary are up, i.e. take the value +1. You can think of these boundary conditions as a symmetry breaking field. The bulk magnetic field is taken to be zero. Configurations look like the one shown in Fig. 9, and can be characterized by *domains walls*. These are lines separating + and − spins such that

1. The + (−) spins lie always to the left (right) of the wall.
2. Where ambiguities remain, the wall is taken to bend to the right.
3. The *length* of the wall is defined as the number of lattice spacings it traverses.

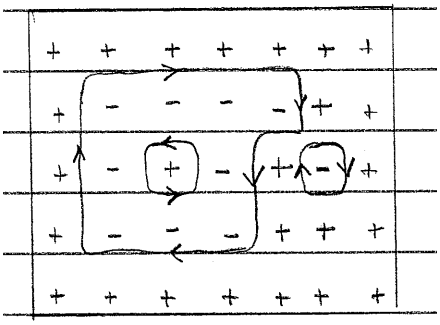


Figure 9: A configuration of spins and the corresponding domain walls.

A wall of length b encloses at most $b^2/16$ spins. The total number of domain walls of length b, $m(b)$, is bounded by

$$\boxed{m(b) \leq 4N_t 3^{b-1},} \tag{327}$$

where N_t is the total number of sites. This can be seens as follows:

- the first link can go into less than $4N_t$ positions (starting at an arbitrary site and going in any of the four possible directions).
- subsequent links have at most 3 possible directions each.

Let us denote the i'th domain wall of length b by (b, i). Next consider a particular configuration $\boldsymbol{\sigma} = \{\sigma_{j,k}\}$ of spins on the lattice, and define

$$X_{\boldsymbol{\sigma}}(b,i) = \begin{cases} 1 & \text{if (b,i) occurs in } \boldsymbol{\sigma} \\ 0 & \text{else} \end{cases} \tag{328}$$

Then the total number of − spins in $\boldsymbol{\sigma}$ is bounded by

$$N_- \leq \sum_b \frac{b^2}{16} \sum_{i=1}^{m(b)} X_{\boldsymbol{\sigma}}(b,i), \tag{329}$$

because each spin is enclosed by at least one domain wall due to our choice of boundary conditions. Taking thermal averages, we have

$$\langle N_- \rangle_\beta \leq \sum_b \frac{b^2}{16} \sum_{i=1}^{m(b)} \langle X_{\boldsymbol{\sigma}}(b,i) \rangle_\beta. \tag{330}$$

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

$$\langle X_{\boldsymbol{\sigma}}(b,i) \rangle_\beta = \frac{1}{Z} {\sum_\sigma}' e^{-\beta E(\boldsymbol{\sigma})}, \tag{331}$$

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

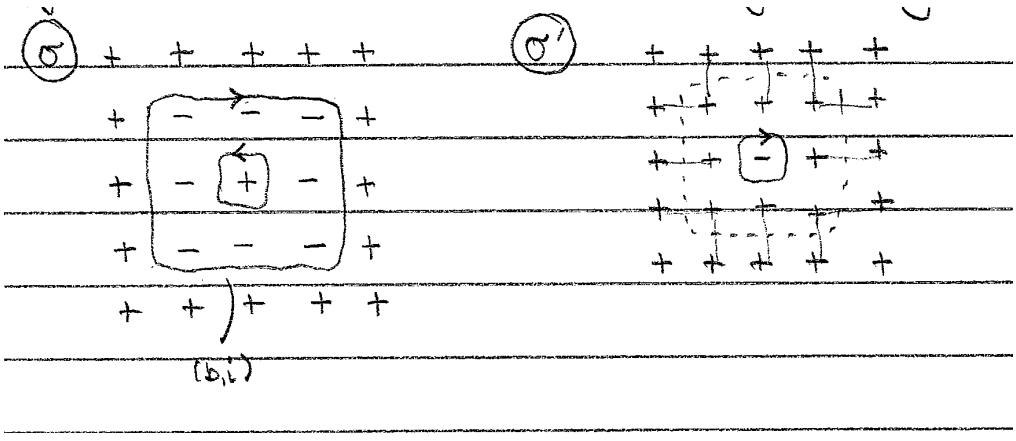


Figure 10: Configurations $\boldsymbol{\sigma}$ and $\boldsymbol{\sigma}'$ related by reversing all spins inside the domain wall (b,i). Shown are all the bonds whose energies have been changed from $-J$ to J.

are related by

$$E(\boldsymbol{\sigma}) = E(\boldsymbol{\sigma}') + 2bJ. \tag{332}$$

This can be used to obtain a bound on Z

$$Z \geq {\sum_{\boldsymbol{\sigma}'}}' e^{-\beta E(\boldsymbol{\sigma}')} \geq {\sum_{\boldsymbol{\sigma}}}' e^{-\beta E(\boldsymbol{\sigma})} e^{2bJ\beta}, \tag{333}$$

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

$$\langle X_{\boldsymbol{\sigma}}(b,i) \rangle_\beta = \frac{1}{Z} {\sum_\sigma}' e^{-\beta E(\boldsymbol{\sigma})} \leq e^{-2bJ\beta}. \tag{334}$$

Now we put everything together

$$\langle N_- \rangle_\beta \;\leq\; \sum_b \frac{b^2}{16} \sum_{i=1}^{m(b)} e^{-2\beta J b} \leq \sum_b \frac{b^2}{16} 4 N_t 3^{b-1} e^{-2\beta J b} = \frac{N_t}{12} \sum_{b=4,6,8,\ldots} b^2 \left[3e^{-2\beta J}\right]^b. \tag{335}$$

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

$$\boxed{\langle N_- \rangle_\beta \leq 108 N_t e^{-8\beta J}.} \tag{336}$$

So, at low temperatures we have

$$\boxed{\frac{\langle N_- \rangle_\beta}{N_t} \ll \frac{1}{2}.} \tag{337}$$

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

6.6 Mean Field Theory

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

$$E = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j - h \sum_j \sigma_j. \tag{338}$$

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

$$m = \frac{1}{N} \sum_{j=1}^{N} \langle \sigma_j \rangle_\beta. \tag{339}$$

We now rewrite the energy using

$$\sigma_i = m + (\sigma_i - m). \tag{340}$$

In particular we have

$$\sigma_i \sigma_j = m^2 + m(\sigma_j - m) + m(\sigma_i - m) + (\sigma_i - m)(\sigma_j - m). \tag{341}$$

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

$$\boxed{E_{\rm MF} = -J \sum_{\langle ij \rangle} -m^2 + m(\sigma_i + \sigma_j) - h \sum_j \sigma_j.} \tag{342}$$

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 (342) further by noting that

$$\begin{aligned} -J \sum_{\langle ij \rangle} -m^2 &= J m^2 \frac{Nz}{2}\,, \\ \sum_{\langle ij \rangle} \sigma_i + \sigma_j &= z \sum_j \sigma_j. \end{aligned} \tag{343}$$

The mean-field energy then becomes

$$\boxed{E_{\rm MF} = \frac{JNz}{2} m^2 - (Jzm + h) \sum_{j=1}^{N} \sigma_j.} \tag{344}$$

The partition function in the mean-field approximation is

$$\begin{aligned} Z_{\rm MF} = \sum_{\{\sigma_j\}} e^{-\beta E_{\rm MF}} &= e^{-\frac{NJz\beta m^2}{2}} \sum_{\sigma_1} \cdots \sum_{\sigma_N} \prod_{j=1}^{N} e^{\beta(Jzm+h)\sigma_j} \\ &= e^{-\frac{NJz\beta m^2}{2}} \left[\sum_{\sigma_1} e^{\beta(Jzm+h)\sigma_1}\right] \cdots \left[\sum_{\sigma_N} e^{\beta(Jzm+h)\sigma_N}\right] \\ &= e^{-\frac{NJz\beta m^2}{2}} \left[2\cosh(Jzm\beta + h\beta)\right]^N . \end{aligned} \tag{345}$$

The magnetization per site is

$$m = \frac{1}{N}\sum_{j=1}^{N} \langle \sigma_j \rangle_\beta = \langle \sigma_N \rangle_\beta, \tag{346}$$

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

$$\boxed{m = \langle \sigma_j \rangle_\beta = \frac{1}{Z_{\rm MF}} e^{-\frac{NJz\beta m^2}{2}} \sum_{\sigma_1} \cdots \sum_{\sigma_N} \sigma_N \prod_{j=1}^{N} e^{\beta(Jzm+h)\sigma_j} = \tanh(Jzm\beta + \beta h).} \tag{347}$$

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

6.7 Solution of the self-consistency equation for $h = 0$

For zero field the self-consistency equation reads

$$m = \tanh(Jzm\beta). \tag{348}$$

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

$$m = \begin{cases} 0 & \text{if } Jz\beta < 1 \\ \pm m_0, 0 & \text{if } Jz\beta > 1 \end{cases}. \tag{349}$$

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

$$f_{\rm MF} = -\frac{1}{\beta N}\ln(Z_{\rm MF}) = \frac{Jzm^2}{2} - \frac{1}{\beta}\ln[2\cosh(Jzm\beta)]. \tag{350}$$

We have

$$\frac{\partial^2 f_{\rm MF}}{\partial m^2} = Jz\left[1 - \frac{Jz\beta}{\cosh(Jzm\beta)}\right]. \tag{351}$$

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

$$\boxed{m = \begin{cases} 0 & \text{if } T > T_c \\ \pm m_0 & \text{if } T < T_c \end{cases},} \tag{352}$$

where the transition temperature is

$$T_c = \frac{Jz}{k_B}. \tag{353}$$

6.8 Vicinity of the Phase Transition

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

$$t = \frac{T - T_c}{T_c}. \tag{354}$$

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

1. Magnetization per site

$$\boxed{m\Big|_{h=0} \simeq \begin{cases} 0 & \text{if } T > T_c, \\ \sqrt{-3t} & \text{if } T < T_c. \end{cases}} \tag{355}$$

2. Magnetic susceptibility in zero field

$$\boxed{\chi = \frac{\partial m}{\partial h}\Big|_{h=0} \simeq \begin{cases} \frac{1}{k_B T_c} t^{-1} & \text{if } T > T_c, \\ \frac{1}{2k_B T_c}(-t)^{-1} & \text{if } T < T_c, \end{cases}} \tag{356}$$

3. Free energy per site and heat capacity per volume

$$\boxed{f_{\mathrm{MF}}\Big|_{h=0} \simeq -k_B T \ln 2 + \begin{cases} 0 & \text{if } T > T_c \\ -\frac{3k_B T_c}{4} t^2 & \text{if } T < T_c. \end{cases}} \tag{357}$$

$$\boxed{\frac{C}{V} = -T\frac{\partial^2 f_{\mathrm{MF}}}{\partial T^2}\Big|_{h=0} \simeq \begin{cases} 0 & \text{if } T > T_c \\ \frac{3k_B}{2} & \text{if } T < T_c. \end{cases}} \tag{358}$$

7 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

$$\boxed{m(T) = \lim_{h \to +0} \lim_{V \to \infty} \frac{1}{V} M(h, T).} \tag{359}$$

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

$$\boxed{m(T) \sim \begin{cases} 0 & \text{if } T > T_c \\ |t|^\beta & \text{if } T < T_c \end{cases} \qquad t = \frac{T - T_c}{T_c}.} \tag{360}$$

   β is a *critical exponent*.

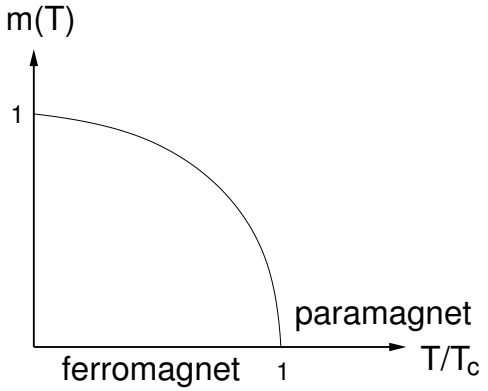


Figure 11: Order parameter for the paramagnet to ferromagnet transition in zero field.

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 $\gamma_\pm$. For our magnet

$$\boxed{\chi_\pm(T) = \frac{\partial}{\partial h}\Big|_{h=0} \lim_{V\to\infty} \frac{1}{V} M(h,T) \sim |t|^{-\gamma_\pm}.} \tag{361}$$

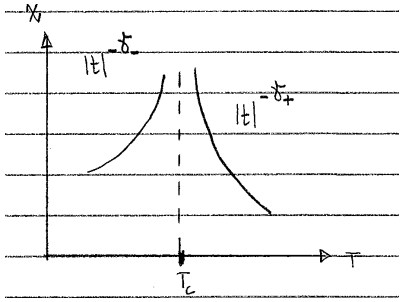


Figure 12: Critical behaviour of the magnetic susceptibility.

3. Heat Capacity

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

$$\boxed{C(T) = -T\frac{\partial^2 F}{\partial T^2} \sim \begin{cases} A_+|t|^{-\alpha_+} & \text{if } T > T_c \\ A_-|t|^{-\alpha_-} & \text{if } T < T_c. \end{cases}} \tag{362}$$

Depending on the signs of $\alpha_\pm$ this may or may not be singular, see Fig. 13.

7.1 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

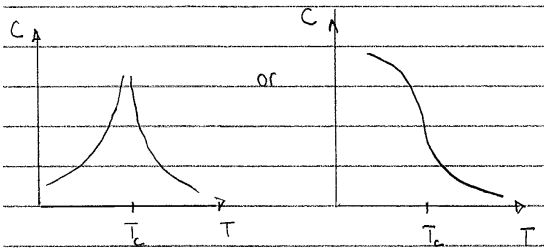


Figure 13: Critical behaviour of the heat capacity.

the same critical behaviour!

8 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 simplest case of a paramagnet to ferromagnet transition the order parameter is the magnetization per site, i.e. a real number. What we mean by simplest case here is that at a microscopic level the magnetic moments are like Ising spins, i.e. point either in a particular direction or in the opposite direction. The situation where the magnetic moments are modelled by vectors that can point in any direction can also be treated by the method we will now describe.

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

$$\boxed{\beta F = \int d^D\mathbf{r} f[\Phi(\mathbf{r})].} \tag{363}$$

4. By construction of the order parameter(s), the latter is (are) *small* close to our critical point. This allows us to expand $f[\Phi(\mathbf{r})]$ is 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

$$\boxed{f[\phi] \simeq \text{const} - h\phi + \alpha_2\phi^2 + \frac{1}{2}|\nabla\phi|^2 + \alpha_3\phi^3 + \alpha_4\phi^4 + \dots} \tag{364}$$

where the coefficient of the gradient term is fixed by convention to be 1/2 (basically this amounts to a rescaling of the order parameter by a constant and using the re-scaled quantity as our order parameter field). This makes ϕ in general dimensionful

$$\dim[\phi(\mathbf{r})] = (\text{length})^{1-D/2}. \tag{365}$$

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 the only give boundary contributions to F. *The coefficients α_n are a priori all functions of temperature and magnetic field.*

5. In translationally invariant systems the free energy is minimized by **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 $V(\phi(\mathbf{r})) = -h\phi + \alpha_2\phi^2 + \alpha_3\phi^3 + \alpha_4\phi^4$ 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 critial point to constrain the α_j.

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

   $$\boxed{\alpha_4 > 0.} \tag{366}$$

   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.

   $$\phi \to -\phi, \tag{367}$$

   then the free energy must respect this symmetry. A ferromagnet has the symmetry (367) 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. 14. We see that the 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 $T - T_c$)

   $$\boxed{\alpha_2(t) = At + \mathcal{O}(t^2)\ , \qquad t = \frac{T - T_c}{T_c}\ ,\ A > 0.} \tag{368}$$

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

   $$\alpha_4(t) = \alpha_4(0) + \mathcal{O}(t). \tag{369}$$

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 **r**-independent, so that we merely need to scrutinize the potential $V(\phi)$ to understand the nature of the phase transition. In Fig. 15 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.

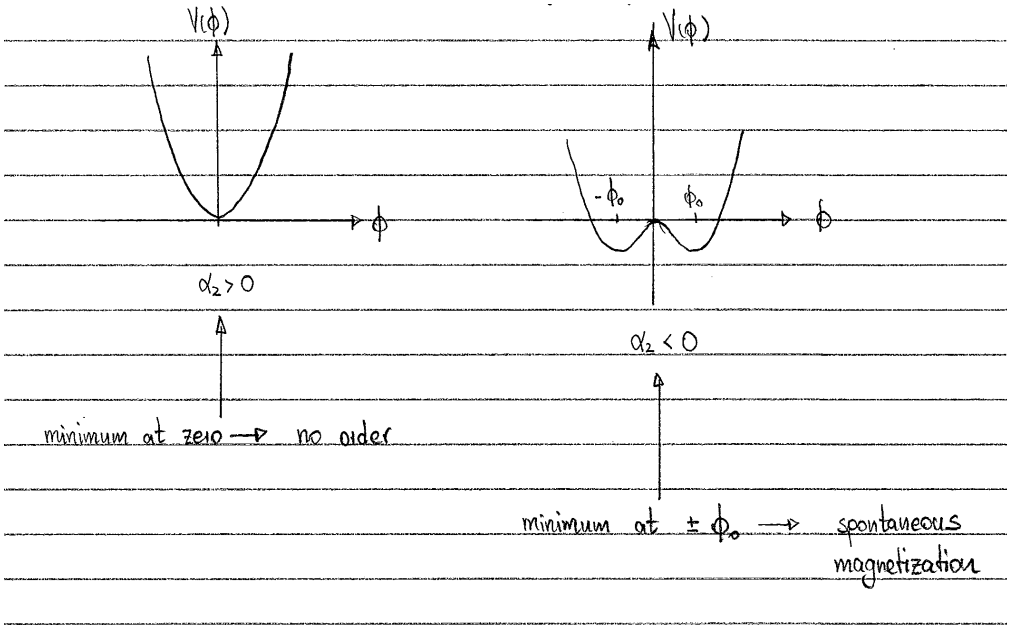


Figure 14: Minima of the potential $V(\phi)$. For $\alpha_2 > 0$ the minima occurs at $\phi = 0$, so there is no ferromagnetic order. For $\alpha_2 < 0$ there are two minima at $\phi = \pm\phi_0$, corresponding to the emergence of a spontaneous magnetization.

8.1 Thermodynamic Equilibrium

The state of thermodynamic equilibrium is obtained by minimizing the free energy

$$\beta F = \int d^D\mathbf{r} \left[-h\phi + \alpha_2\phi^2 + \alpha_3\phi^3 + \frac{1}{2}|\nabla\phi|^2 + \alpha_4\phi^4 \right], \tag{370}$$

where $\alpha_3(h,t)$ must vanish at zero field $\alpha_3(0,t) = 0$. 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

$$\frac{\delta F}{\delta\phi(\mathbf{r})} = 0. \tag{371}$$

The resulting nonlinear differential equation is

$$\boxed{-\nabla^2\phi(\mathbf{r}) + 2\alpha_2\phi(\mathbf{r}) + 3\alpha_3\phi^2(\mathbf{r}) + 4\alpha_4\phi^3(\mathbf{r}) - h = 0.} \tag{372}$$

It is easy to see (see the discussion above) that the solutions to (372) 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

$$\boxed{\phi(\mathbf{r}) = \begin{cases} 0 & \text{if } \alpha_2 > 0 \leftrightarrow T > T_c \\ \pm\phi_0 = \pm\sqrt{-\frac{\alpha_2}{2\alpha_4}} & \text{if } \alpha_2 < 0 \leftrightarrow T < T_c \end{cases}.} \tag{373}$$

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.

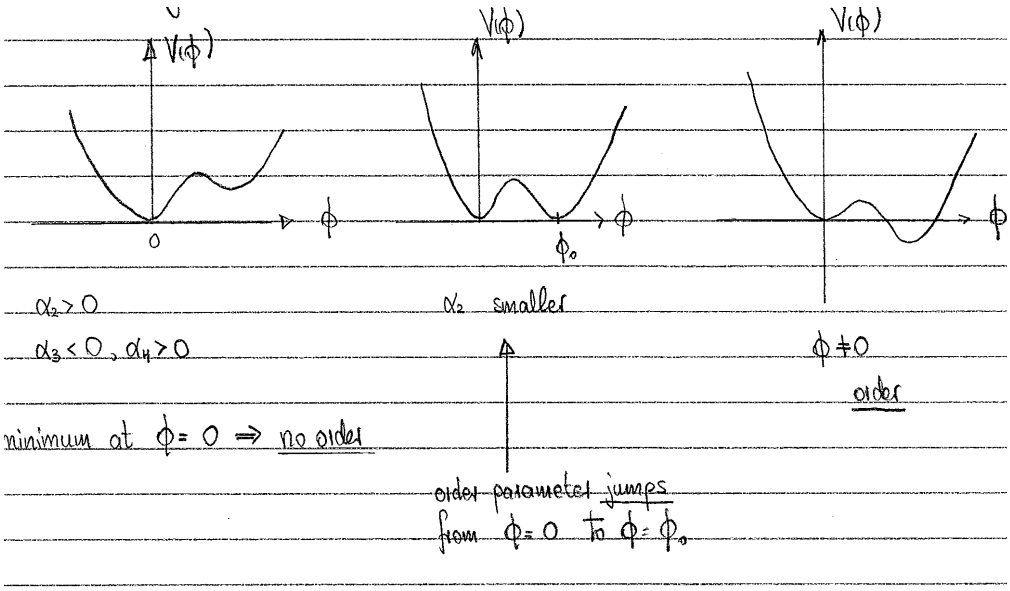


Figure 15: Minima of the potential $V(\phi)$ for $\alpha_3 < 0$. Decreasing the value of α_2 leads to a discontinuous jump in the order parameter at some critical value. The transition is therefore first order.

8.2 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

$$Z_{\text{micro}} = \sum_{\text{configurations C}} e^{-\beta E(C)} \ . \tag{374}$$

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

$$Z \longrightarrow \sum_{\substack{\text{order parameter} \\ \text{configurations}}} e^{-\beta \mathcal{H}} \ . \tag{375}$$

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

$$\beta\mathcal{H} = \int d^D\mathbf{r} \ f[\phi(\mathbf{r})] \ , \tag{376}$$

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 (375) is really the *functional integral*

$$Z = \int \mathcal{D}\phi(\mathbf{r}) \ e^{-\beta\mathcal{H}}. \tag{377}$$

The latter is defined as follows:

- We first discretize our D-dimensional spatial variable

$$\mathbf{r} \longrightarrow a_0 \mathbf{n} = a_0(n_1, \ldots, n_D), \tag{378}$$

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

$$\begin{aligned} \phi(\mathbf{r}) &\longrightarrow \phi_{\mathbf{n}} \ , \\ \partial_{r_j}\Phi(\mathbf{r}) &\longrightarrow \frac{\phi_{\mathbf{n}+\mathbf{e}_j} - \phi_{\mathbf{n}}}{a_0}, \end{aligned} \tag{379}$$

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

- The Landau-Ginzburg Hamiltonian is discretized as

$$\beta\mathcal{H} \longrightarrow \sum_{\mathbf{n}} f[\phi_{\mathbf{n}}] a_0^D. \tag{380}$$

- The functional integral is then defined as follows

$$\boxed{Z = \int \mathcal{D}\phi(\mathbf{r}) \ e^{-\beta\mathcal{H}} \equiv \lim_{N\to\infty} \int \prod_{\mathbf{n}} d\phi_{\mathbf{n}} e^{-\sum_{\mathbf{m}} f[\phi_{\mathbf{m}}] a_0^D}.} \tag{381}$$

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

$$\langle \phi(\mathbf{r})\phi(\mathbf{0})\rangle_\beta \equiv \frac{1}{Z} \int \mathcal{D}\phi \ \phi(\mathbf{r})\phi(\mathbf{0}) \ e^{-\beta\mathcal{H}}. \tag{382}$$

8.3 Saddle Point Approximation

The Landau-Ginzburg field theory

$$Z = \int \mathcal{D}\phi(\mathbf{r}) \ e^{-\int d^D\mathbf{r} f[\phi(\mathbf{r})]}, \tag{383}$$

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

$$\beta\mathcal{H} = \int d^D\mathbf{r} f[\phi(\mathbf{r})]. \tag{384}$$

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.

8.4 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

$$\boxed{\phi_0 = \pm\sqrt{\frac{A}{2\alpha_4}}|t|^{\frac{1}{2}}.} \tag{385}$$

  This gives the critical exponent

$$\boxed{\beta = \frac{1}{2}.} \tag{386}$$

- Magnetic susceptibility.

  Differentiating (372) with respect to h gives for $\mathbf{r}$-independent solutions

$$2\alpha_2\frac{\partial\phi}{\partial h} + 6\alpha_3\frac{\partial\phi}{\partial h}\phi + 12\alpha_4\frac{\partial\phi}{\partial h}\phi^2 = 1 - 2\frac{\partial\alpha_2}{\partial h}\phi - 3\frac{\partial\alpha_3}{\partial h}\phi^2 - 4\frac{\partial\alpha_4}{\partial h}\phi^3. \tag{387}$$

  On the right hand side we only need to retain the first term as the others are small close to the transition. The zero-field susceptibility is then

$$\chi = \frac{\partial\phi}{\partial h}\Big|_{h=0} \approx \frac{1}{2\alpha_2 + 12\alpha_4\phi^2}, \tag{388}$$

  where we have used that $\alpha_3(0,t) = 0$. Using that $\alpha_2 = At$ and $\phi_0^2 = A|t|/2\alpha_4$ this becomes

$$\boxed{\chi = \begin{cases} \frac{1}{2At} & \text{if } t > 0, \\ \frac{1}{4A|t|} & \text{if } t < 0. \end{cases}} \tag{389}$$

  This gives the critical exponents

$$\boxed{\gamma_\pm = 1.} \tag{390}$$

- Heat capacity.

  The heat capacity is defined by

$$C(T, h = 0) = -T\frac{\partial^2 F}{\partial T^2}. \tag{391}$$

  The saddle point contribution to the free energy is

$$\frac{F}{V} \sim \begin{cases} 0 & \text{if } t > 0, \\ -\frac{A^2 k_B T_c t^2}{4\alpha_4} & \text{if } t < 0, \end{cases} \tag{392}$$

  giving

$$\boxed{C(T, h = 0) \sim \begin{cases} 0 & \text{if } t > 0, \\ \frac{A^2 k_B V}{2\alpha_4} & \text{if } t < 0, \end{cases}} \tag{393}$$

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

$$\boxed{\alpha = 0.} \tag{394}$$

- 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

$$\langle\phi(\mathbf{r})\phi(\mathbf{0})\rangle_\beta - \langle\phi(\mathbf{r})\rangle_\beta\langle\phi(\mathbf{0})\rangle_\beta = e^{-|\mathbf{r}|/\xi} , \qquad |\mathbf{r}| \to \infty. \tag{395}$$

  The correlation length ξ diverges when the critical point is approached

$$\boxed{\xi \sim |t|^{-\nu}.} \tag{396}$$

  The relation (396) 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

$$\langle\phi(\mathbf{r})\phi(\mathbf{0})\rangle_\beta \approx \frac{1}{Z}\int \mathcal{D}\phi\ \phi(\mathbf{r})\phi(\mathbf{0}) \exp\left(-\int d^D\mathbf{r}' \left[\frac{1}{2}|\nabla\phi(\mathbf{r}')|^2 + \alpha_2\phi^2(\mathbf{r}')\right]\right). \tag{397}$$

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

  The two-point function (397) can be calculated using a *generating functional*

$$Z[h] = \int \mathcal{D}\phi\ \exp\left(-\int d^D\mathbf{r}' \left[\frac{1}{2}|\nabla\phi(\mathbf{r}')|^2 + \alpha_2\phi^2(\mathbf{r}') - h(\mathbf{r}')\phi(\mathbf{r}')\right]\right). \tag{398}$$

  In order to work out how $Z[h]$ is related to the correlation functions we are interested in we need to know how to take "functional derivatives" of $Z[h]$ with respect to "sources" $h(\mathbf{r})$. The necessary technology is summarized below.

Aside 6: Functionals and functional derivatives

FUNCTIONALS

What is a *functional*? You all know that a real function can be viewed as a *map* from e.g. an interval $[a, b]$ to the real numbers

$$f : [a, b] \to \mathbb{R} , \quad x \to f(x). \tag{399}$$

A functional is similar to a function in that it maps all elements in a certain domain to real numbers, however, the nature of its domain is very different. Instead of acting on all points of an interval or some other subset of the real numbers, the domain of functionals consists of (suitably chosen) classes of functions. In other words, given some class $\{f\}$ of functions, a functional F is a map

$$F : \{f\} \to \mathbb{R} , \quad f \to F[f]. \tag{400}$$

We now consider two specific examples of functionals.

1. The distance between two points. A very simple functional F consists of the map which assigns to all paths between two fixed points the length of the path. To write this functional explicitly, let us consider a simple two-dimensional situation in the (x, y) plane and choose

two points (x_1, y_1) and (x_2, y_2). We consider the set of paths that do not turn back, i.e. paths along which x increases monotonically as we go from (x_1, y_1) to (x_2, y_2). These can be described by the set of functions $\{f\}$ on the interval $[x_1, x_2]$ satisfying $f(x_1) = y_1$ and $f(x_2) = y_2$. The length of a path is then given by the well-known expression

$$F[f(x)] = \int_{x_1}^{x_2} dx' \, \sqrt{1 + f'(x')^2} \, . \tag{401}$$

2. *Action Functionals.* These are very important in Physics. Let us recall their definition in the context of classical mechanics. Start with n generalised coordinates $\mathbf{q}(t) = (q_1(t), \ldots, q_n(t))$ and a Lagrangian $L = L(\mathbf{q}, \dot{\mathbf{q}})$. Then, the action functional $S[\mathbf{q}]$ is defined by

$$S[\mathbf{q}] = \int_{t_1}^{t_2} dt \, L(\mathbf{q}(t), \dot{\mathbf{q}}(t)) \, . \tag{402}$$

It depends on classical paths $\mathbf{q}(t)$ between times t_1 and t_2 satisfying the boundary conditions $\mathbf{q}(t_1) = \mathbf{q}_1$ and $\mathbf{q}(t_2) = \mathbf{q}_2$.

FUNCTIONAL DIFFERENTIATION

In both the examples given above a very natural question to ask is what function *extremizes* the functional. In the first example this corresponds to wanting to know the path that minimizes the distance between two points. In the second example the extremum of the action functional gives the solutions to the classical equations of motion. This is known as *Hamilton's principle.* In order to figure out what function extremizes the functional it is very useful to generalize the notion of a derivative. For our purposes we define the *functional derivative* by

$$\boxed{\frac{\delta F[f(x)]}{\delta f(y)} = \lim_{\epsilon \to 0} \frac{F[f(x) + \epsilon \delta(x - y)] - F[f(x)]}{\epsilon} \, .} \tag{403}$$

Here, as usual, we should think of the δ-function as being defined as the limit of a test function, e.g.

$$\delta(x) = \lim_{a \to 0} \frac{1}{\sqrt{\pi} a} e^{-x^2/a^2} \, , \tag{404}$$

and take the limit $a \to 0$ only in the end (after commuting the limit with all other operations such as the $\lim_{\epsilon \to 0}$ in (403)). Importantly, the derivative defined in this way is a linear operation which satisfies the product and chain rules of ordinary differentiation and commutes with ordinary integrals and derivatives. Let us see how functional differentiation works for our two examples.

1. The distance between two points. In analogy with finding stationary points of functions we want to extremize (401) by setting its functional derivative equal to zero

$$0 = \frac{\delta F[f(x)]}{\delta f(y)} \, . \tag{405}$$

We first do the calculation by using the definition (403).

$$\frac{\delta F[f(x)]}{\delta f(y)} = \lim_{\epsilon \to 0} \int_{x_1}^{x_2} dx' \, \frac{\sqrt{1 + [f'(x') + \epsilon \delta'(x' - y)]^2} - \sqrt{1 + [f'(x')]^2}}{\epsilon} . \tag{406}$$

The Taylor expansion of the square root is $\sqrt{1 + 2\epsilon} = 1 + \epsilon + \ldots$, which gives

$$\sqrt{1 + [f'(x') + \epsilon \delta'(x' - y)]^2} \;\; = \;\; \sqrt{1 + [f'(x')]^2} + \frac{\epsilon f'(x') \delta'(x' - y)}{\sqrt{1 + [f'(x')]^2}} + \mathcal{O}(\epsilon^2) \, , \tag{407}$$

where $\delta'(x)$ is the derivative of the delta-function and $\mathcal{O}(\epsilon^2)$ denote terms proportional to ϵ^2. Substituting this back into (406) we have [a]

$$\frac{\delta F[f(x)]}{\delta f(y)} = \int_{x_1}^{x_2} dx' \, \frac{\delta'(x'-y) f'(x')}{\sqrt{1+[f'(x')]^2}} = -\frac{d}{dy} \frac{f'(y)}{\sqrt{1+[f'(y)]^2}}. \tag{409}$$

The solution to (405) is thus

$$f'(y) = \text{const}, \tag{410}$$

which describes a straight line. In practice we don't really go back to the definition of the functional derivative any more than we use the definition of an ordinary derivative to work it out, but proceed as follows.

- We first interchange the functional derivative and the integration

$$\frac{\delta F[f(x)]}{\delta f(y)} = \int_{x_1}^{x_2} dx' \, \frac{\delta}{\delta f(y)} \sqrt{1+[f'(x')]^2}. \tag{411}$$

- Next we use the chain rule

$$\frac{\delta \sqrt{1+f'(x')^2}}{\delta f(y)} = \frac{1}{2\sqrt{1+f'(x')^2}} \frac{\delta(1+f'(x')^2)}{\delta f(y)} = \frac{f'(x')}{\sqrt{1+f'(x')^2}} \frac{\delta f'(x')}{\delta f(y)}. \tag{412}$$

- Finally we interchange the functional and the ordinary derivative

$$\frac{\delta f'(x')}{\delta f(y)} = \frac{d}{dx'} \frac{\delta f(x')}{\delta f(y)} = \frac{d}{dx'} \delta(x'-y) \ . \tag{413}$$

  The last identity follows from our definition (403).

Now we can put everything together and arrive at the same answer (409).

[a] In the last step we have used

$$\int_a^b dx' \delta'(x'-y) g(x') = -g'(y) \ , \tag{408}$$

which can be proved by "integration by parts".

Exercise 3

We now want to try out the ideas introduced above on our second example and extremize the classical action (402) in order to obtain the classical equations of motion. We first interchange functional derivative and integration and then use the chain rule to obtain

$$\frac{\delta S[\mathbf{q}]}{\delta q_i(t)} = \frac{\delta}{\delta q_i(t)} \int_{t_1}^{t_2} d\tilde{t} \, L(\mathbf{q}(\tilde{t}), \dot{\mathbf{q}}(\tilde{t})) \tag{414}$$

$$= \int_{t_1}^{t_2} d\tilde{t} \left[\frac{\partial L}{\partial q_j}(\mathbf{q}, \dot{\mathbf{q}}) \frac{\delta q_j(\tilde{t})}{\delta q_i(t)} + \frac{\partial L}{\partial \dot{q}_j}(\mathbf{q}, \dot{\mathbf{q}}) \frac{\delta \dot{q}_j(\tilde{t})}{\delta q_i(t)} \right] \tag{415}$$

$$\tag{416}$$

We now use that $\frac{\delta \dot{q}_j(\tilde{t})}{\delta q_i(t)} = \frac{d}{d\tilde{t}} \frac{\delta q_j(\tilde{t})}{\delta q_i(t)}$ and integrate by parts with respect to $\tilde{t}$

$$\frac{\delta S[\mathbf{q}]}{\delta q_i(t)} = \int_{t_1}^{t_2} d\tilde{t} \left[\frac{\partial L}{\partial q_j}(\mathbf{q}, \dot{\mathbf{q}}) - \frac{d}{d\tilde{t}} \frac{\partial L}{\partial \dot{q}_j}(\mathbf{q}, \dot{\mathbf{q}}) \right] \frac{\delta q_j(\tilde{t})}{\delta q_i(t)} \tag{417}$$

$$= \int_{t_1}^{t_2} d\tilde{t} \left[\frac{\partial L}{\partial q_j}(\mathbf{q}, \dot{\mathbf{q}}) - \frac{d}{d\tilde{t}} \frac{\partial L}{\partial \dot{q}_j}(\mathbf{q}, \dot{\mathbf{q}}) \right] \delta_{ij} \delta(\tilde{t} - t) = \frac{\partial L}{\partial q_i}(\mathbf{q}, \dot{\mathbf{q}}) - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i}(\mathbf{q}, \dot{\mathbf{q}}) \ . \tag{418}$$

In the second last step we have used

$$\frac{\delta q_j(\tilde{t})}{\delta q_i(t)} = \delta_{ij} \delta(\tilde{t} - t) \ , \tag{419}$$

which follows straightforwardly from our general definition (403). Thus we conclude that the extrema of the classical action are given by paths that fulfil the equations of motion

$$\boxed{\frac{\partial L}{\partial q_i}(\mathbf{q}, \dot{\mathbf{q}}) - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i}(\mathbf{q}, \dot{\mathbf{q}}) = 0.} \tag{420}$$

Nice.

Taking functional derivatives of $Z[h]$ we find that

$$\langle \phi(\mathbf{r}) \phi(\mathbf{0}) \rangle_\beta = \frac{\delta}{\delta h(\mathbf{r})} \frac{\delta}{\delta h(\mathbf{0})} \Big|_{h=0} \ln(Z[h]). \tag{421}$$

We calculate the generating functional by going to Fourier space

$$\phi(\mathbf{r}) = \int \frac{d^D \mathbf{p}}{(2\pi)^D} e^{-i\mathbf{p}\cdot\mathbf{r}} \tilde{\phi}(\mathbf{p}), \quad h(\mathbf{r}) = \int \frac{d^D \mathbf{p}}{(2\pi)^D} e^{-i\mathbf{p}\cdot\mathbf{r}} \tilde{h}(\mathbf{p}). \tag{422}$$

This gives

$$\beta \mathcal{H} = \int \frac{d^D \mathbf{p}}{(2\pi)^D} \left[\left(\frac{\mathbf{p}^2}{2} + \alpha_2 \right) \tilde{\phi}(\mathbf{p}) \tilde{\phi}(-\mathbf{p}) - h(\mathbf{p}) \tilde{\phi}(-\mathbf{p}) \right]. \tag{423}$$

Next we "complete the square" by changing variables to

$$\tilde{\varphi}(\mathbf{p}) = \tilde{\phi}(\mathbf{p}) - \frac{\tilde{h}(\mathbf{p})}{\mathbf{p}^2 + 2\alpha_2}. \tag{424}$$

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

$$Z[h] = \int \mathcal{D}\tilde{\varphi} \exp\left(- \int \frac{d^D \mathbf{p}}{(2\pi)^D} \left[\frac{\mathbf{p}^2}{2} + \alpha_2 \right] \tilde{\varphi}(\mathbf{p}) \tilde{\varphi}(-\mathbf{p}) \right) \exp\left(\frac{1}{2} \int \frac{d^D \mathbf{p}}{(2\pi)^D} \frac{\tilde{h}(\mathbf{p}) \tilde{h}(-\mathbf{p})}{\mathbf{p}^2 + 2\alpha_2} \right). \tag{425}$$

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

$$Z[h] = \mathcal{N} \exp\left(\frac{1}{2} \int d^D \mathbf{r} d^D \mathbf{r}' h(\mathbf{r}) G(\mathbf{r} - \mathbf{r}') h(\mathbf{r}') \right) \ ,$$

$$G(\mathbf{r}) = \int \frac{d^D \mathbf{p}}{(2\pi)^D} \frac{e^{-i\mathbf{p}\cdot\mathbf{r}}}{\mathbf{p}^2 + 2\alpha_2}. \tag{426}$$

Taking functional derivatives we have

$$\langle \phi(\mathbf{r}) \phi(\mathbf{0}) \rangle_\beta = G(\mathbf{r}) \sim (2\alpha_2)^{\frac{D-3}{4}} \frac{e^{-|\mathbf{r}|\sqrt{2\alpha_2}}}{|\mathbf{r}|^{\frac{D-1}{2}}} \ , \qquad |\mathbf{r}| \to \infty. \tag{427}$$

This gives the correlation length

$$\xi = \frac{1}{\sqrt{2\alpha_2}} \sim \frac{1}{t^{\frac{1}{2}}}, \tag{428}$$

and thus the critical exponent

$$\boxed{\nu = \frac{1}{2}.} \tag{429}$$

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

$$\beta\mathcal{H} \;=\; \frac{1}{2}\int \frac{d^D\mathbf{p}}{(2\pi)^D}\left[\left(\mathbf{p}^2 + 2\alpha_2\right)\tilde{\phi}(\mathbf{p})\tilde{\phi}(-\mathbf{p}) + \text{quartic}\right]. \tag{430}$$

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

$$\langle\tilde{\phi}(\mathbf{p})\tilde{\phi}(\mathbf{q})\rangle_\beta = \frac{(2\pi)^D\delta^D(\mathbf{p}-\mathbf{q})}{\mathbf{p}^2 + 2\alpha_2}. \tag{431}$$

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

$$V(\phi) = \alpha_2\phi^2 + \alpha_4\phi^4 \simeq \alpha_2\phi_0^2 + \alpha_4\phi_0^4 + (\alpha_2 + 6\alpha_4\phi_0^2)(\phi-\phi_0)^2 + \ldots \tag{432}$$

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

$$f[\delta\phi(\mathbf{r})] \simeq \frac{1}{2}|\nabla\delta\phi|^2 + \tilde{\alpha}_2\delta\phi^2. \tag{433}$$

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*

$$\langle\delta\phi(\mathbf{r})\delta\phi(\mathbf{0})\rangle_\beta = \langle\phi(\mathbf{r})\phi(\mathbf{0})\rangle_\beta - \langle\phi(\mathbf{r})\rangle_\beta\langle\phi(\mathbf{0})\rangle_\beta \sim \frac{e^{-|\mathbf{r}|\sqrt{-4\alpha_2}}}{|\mathbf{r}|^{\frac{D-1}{2}}}\;, \quad |\mathbf{r}|\to\infty. \tag{434}$$

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

8.5 Homework Questions 12-14

Question 12. 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.

Question 13.
(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

$$F = \int d^3\mathbf{r} \left[\frac{1}{2}|\nabla\phi(\mathbf{r})|^2 - h\phi(\mathbf{r}) + \alpha_2\phi^2(\mathbf{r}) + \alpha_3\phi^3(\mathbf{r}) + \alpha_4\phi^4(\mathbf{r}) \right]. \tag{435}$$

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

$$\alpha_2 = At\ , \quad t = \frac{T - T_c}{T_c} \text{ and } A > 0. \tag{436}$$

(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

$$E = \frac{1}{2}\left[\frac{d\phi(x)}{dx}\right]^2 - \alpha_2\phi^2 - \alpha_4\phi^4 \tag{437}$$

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

$$\lim_{x\to\infty} \phi(x) = \phi_1\ , \quad \lim_{x\to-\infty} \phi(x) = \phi_2, \tag{438}$$

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

Question 14. A system with a real, two-component order parameter $(\phi_1(\mathbf{r}), \phi_2(\mathbf{r}))$ has a free energy

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

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?

9 Other Examples of Phase Transitions

9.1 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. 16. What is the order parameter characterizing this

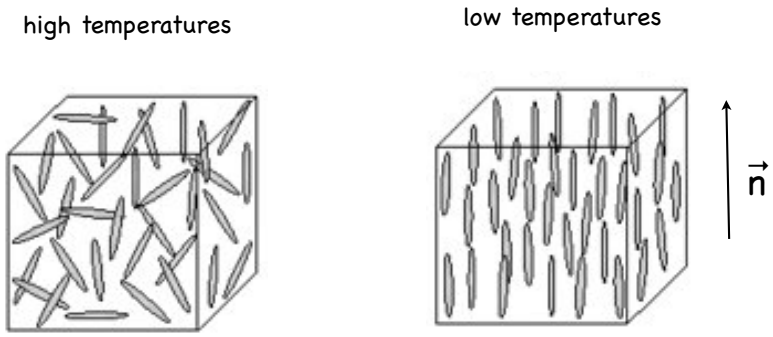


Figure 16: At low temperatures the rod-like molecules spontaneously align along an axis $\vec{n}$.

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

$$\vec{n}(\mathbf{r}) \longrightarrow -\vec{n}(\mathbf{r}). \tag{439}$$

So how about something quadratic like

$$\langle n_i(\mathbf{r}) n_j(\mathbf{r})\rangle_\beta. \tag{440}$$

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

$$\vec{n} = \begin{pmatrix} \sin\theta\cos\varphi \\ \sin\theta\sin\varphi \\ \cos\theta \end{pmatrix}. \tag{441}$$

Then averaging over all possible orientations gives

$$\langle\langle n_i n_j\rangle\rangle = \frac{1}{4\pi}\int_0^\pi d\theta \int_0^{2\pi} d\varphi \sin\theta n_i n_j = \frac{1}{3}\delta_{i,j} \neq 0. \tag{442}$$

This consideration suggests to try

$$\boxed{Q_{ij} = \langle n_i n_j\rangle_\beta - \frac{1}{3}\delta_{i,j}} \tag{443}$$

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

$$Q = \begin{pmatrix} -\frac{1}{3} & 0 & 0 \\ 0 & -\frac{1}{3} & 0 \\ 0 & 0 & \frac{2}{3} \end{pmatrix}. \tag{444}$$

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

$$Q(\mathbf{r}) \longrightarrow RQ(\mathbf{r})R^T. \tag{445}$$

This suggests the following expansion for the free energy density

$$\boxed{f[Q(\mathbf{r})] = \frac{1}{2}|\nabla Q|^2 + \alpha_2 \mathrm{Tr}[Q^2] + \alpha_3 \mathrm{Tr}[Q^3] + \alpha_4 \left(\mathrm{Tr}[Q^2]\right)^2 + \dots} \tag{446}$$

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.

9.2 Superfluid Transition in Weakly Interacting Bosons

Let us recall the second quantized Hamiltonian for weakly repulsive bosons

$$H = \int d^4\mathbf{r} \left[c^\dagger(\mathbf{r}) \left(-\frac{\nabla^2}{2m} \right) c(\mathbf{r}) + \frac{U}{2} c^\dagger(\mathbf{r}) c^\dagger(\mathbf{r}) c(\mathbf{r}) c(\mathbf{r}) \right]. \tag{447}$$

Let us recall that the Hamiltonian (447) exhibits a U(1) symmetry, i.e. it is invariant under the transformation

$$c(\mathbf{r}) \longrightarrow e^{i\theta} c(\mathbf{r}). \tag{448}$$

This expresses the fact that the Hamiltonian conserves the total number of bosons. At temperatures below T_c the symmetry (448) gets spontaneously broken and in the superfluid phase we have macroscopic occupation of the zero momentum single-particle state

$$\langle c(\mathbf{p}=0) \rangle_\beta = \sqrt{N_0}. \tag{449}$$

This implies that at low temperatures in a spatially homogeneous system we have

$$\langle c(\mathbf{r}) \rangle_{T<T_c} \neq 0. \tag{450}$$

These observations suggest that we should try

$$\psi(\mathbf{r}) = \langle c(\mathbf{r}) \rangle_\beta \tag{451}$$

as our order parameter. Importantly this order parameter is *complex valued*. Our Landau free energy must respect the resulting symmetry of the order parameter

$$\psi(\mathbf{r}) \longrightarrow e^{i\theta} \psi(\mathbf{r}) \ , \tag{452}$$

while the symmetry gets spontaneously broken in the ordered phase. Using our usual arguments for writing down the Landau free energy, we arrive at the following expansion in powers of the order parameter

$$\boxed{F[\psi] = \int d^D\mathbf{r} \left[\frac{1}{2m} |\nabla \psi|^2 + \alpha_2 |\psi(\mathbf{r})|^2 + \alpha_4 |\psi(\mathbf{r})|^4 + \dots \right].} \tag{453}$$

In analogy to the case of a real scalar order parameter considered above we require $\alpha_4 > 0$ for stability, while $\alpha_2 = At$ is proportional to the reduced temperature.

Exercise 4: Deriving (453)

It is useful to express the complex field ψ in terms of its real and imaginary part

$$\psi(\mathbf{r}) = \psi_1(\mathbf{r}) + i\psi_2(\mathbf{r}) \ . \tag{454}$$

Under U(1) transformations (452) we have

$$\begin{aligned} \psi_1(\mathbf{r}) &\longrightarrow \cos(\theta)\psi_1(\mathbf{r}) - \sin(\theta)\psi_2(\mathbf{r}) \ , \\ \psi_2(\mathbf{r}) &\longrightarrow \sin(\theta)\psi_1(\mathbf{r}) + \cos(\theta)\psi_2(\mathbf{r}) \ . \end{aligned} \tag{455}$$

Expanding in powers of $\psi_{1,2}$ and only retaining terms that respect the U(1) symmetry we arrive at (453).

In order to proceed it is useful to represent the order parameter in the form

$$\psi(\mathbf{r}) = \rho(\mathbf{r}) \ e^{i\theta(\mathbf{r})} \ , \tag{456}$$

where both ρ and θ are real. Substituting (456) into (453) gives an expansion of the form

$$F[\rho,\theta] = \int d^D\mathbf{r} \left[\frac{1}{2m}|\nabla\rho|^2 + \alpha_2\rho^2(\mathbf{r}) + \alpha_4\rho^4(\mathbf{r}) + \frac{1}{2m}\rho^2(\mathbf{r})|\nabla\theta|^2 + \dots \right] . \tag{457}$$

In a homogeneous system this free energy is minimized by

$$\rho(\mathbf{r}) = \rho_0 = \pm\sqrt{-\frac{\alpha_2}{2\alpha_4}}. \tag{458}$$

Substituting this back into (457) we obtain (at leading order in the expansion) a Gaussian theory that describes the fluctuations around the ordered state

$$F[\delta\rho,\theta] = \int d^D\mathbf{r} \left[\frac{1}{2m}|\nabla\delta\rho|^2 + (\alpha_2 + 6\alpha_4\rho_0^2)\rho^2(\mathbf{r}) + \frac{\rho_0^2}{2m}|\nabla\theta|^2 + \dots \right] . \tag{459}$$

We may now proceed as in our calculating of the correlation length exponent for a real scalar order parameter theory in order to determine the two point function. The result is

$$\langle \delta\rho(\mathbf{r}) \ \delta\rho(\mathbf{0}) \rangle_\beta \ \propto \ \int \frac{d^D\mathbf{p}}{(2\pi)^D} \frac{e^{-i\mathbf{p}\cdot\mathbf{r}}}{\frac{\mathbf{p}^2}{m} + a} \ , \tag{460}$$

where

$$a = 2\alpha_2 + 12\alpha_4\rho_0^2 \ . \tag{461}$$

When considering the two point function of θ we face a complication in that (by construction) the field is defined only modulo 2π. This makes calculations more complicated unless one is in a situation where the fluctuations of θ around some ordered value, which we will take to be zero, are small. In that case the system does not explore configurations for which it becomes important that θ is defined only modulo 2π. Under these assumptions we may again follow our previous calculations to obtain

$$\langle \theta(\mathbf{r}) \ \theta(\mathbf{0}) \rangle_\beta \propto \int \frac{d^D\mathbf{p}}{(2\pi)^D} \frac{e^{-i\mathbf{p}\cdot\mathbf{r}}}{\mathbf{p}^2} \ . \tag{462}$$

This integral is divergent at small momenta for $D < 2$, while for large D we have

$$\langle \theta(\mathbf{r}) \ \theta(\mathbf{0}) \rangle_\beta \propto \frac{1}{|\mathbf{r}|^{D-2}} \ , \quad D > 2. \tag{463}$$

The divergence for $D < 2$ invalidates the assumptions we have made about the fluctuations of θ being small. However, the fact that fluctuations diverge in low dimensions is in fact correct. What happens is that phase fluctuations are so strong, that it becomes impossible to have a globally nonzero order parameter $\langle \psi(\mathbf{r}) \rangle_\beta$.

10 Regime of validity of the saddle point approximation/mean field theory

An important question is how reliable the saddle-point approximation we have used to analyze path integrals like (381) is. A useful consistency check is obtained by comparing the variance of the order parameter to its square when averaged over a spatial region with a radius set by the correlation length ξ. The variance gives a measure of the size of the fluctuations around the average and for a translationally invariant system is given by

$$\begin{aligned}\langle(\delta\phi)^2\rangle_{R_\xi} &= \Big\langle\Big[\int_{R_\xi} d^D\mathbf{r}\ \big(\underbrace{\phi(\mathbf{r})-\langle\phi(\mathbf{r})\rangle}_{\delta\phi(\mathbf{r})}\big)\Big]^2\Big\rangle \\ &= \int_{R_\xi} d^D\mathbf{r}\int_{R_\xi} d^D\mathbf{r}'\ \langle\delta\phi(\mathbf{r})\ \delta\phi(\mathbf{r}')\rangle \\ &\sim \int_{R_\xi} d^D\mathbf{r}\int_{R_\xi} d^D\mathbf{r}'\ \xi^{\frac{3-D}{2}}\frac{e^{-|\mathbf{r}-\mathbf{r}'|/\xi}}{|\mathbf{r}-\mathbf{r}'|^{\frac{D-1}{2}}} \propto \xi^{2+D}. \end{aligned} \tag{464}$$

Here R_ξ is a box of length ξ centered around the origin and we have used our result (434) for the two point function in the Gaussian approximation. This should be compared to the square of the order parameter averaged over R_ξ

$$\langle\phi\rangle^2_{R_\xi} = \xi^{2D}\langle\phi(\mathbf{0})\rangle^2\ . \tag{465}$$

We have

$$\frac{\langle(\delta\phi)^2\rangle_{R_\xi}}{\langle\phi\rangle^2_{R_\xi}} \propto \frac{\xi^{2-D}}{\langle\phi(\mathbf{0})\rangle^2} \propto |t|^{\nu(D-2)-2\beta}, \tag{466}$$

where ν and β are the correlation length and order parameter exponents respectively. The saddle point approximation has a chance of being accurate as long as the right hand side of (466) is much less than 1, which occurs if the exponent of $|t|$ is positive. Using our previous results that $\beta = \nu = \frac{1}{2}$ we therefore require

$$D > 4. \tag{467}$$

This means that our saddle point approximation is not expected to work quantitatively in dimensions less than 4. The dimension above which a saddle point (mean field) approximation becomes valid is called *upper critical dimension*. In $D < 4$ a significantly more refined analysis based on the renormalization group is required to obtain quantitative results.

Mean field theory can fail in another way: at low temperatures it predicts an ordered phase, but we have already seen for the case of the 1D Ising model and the Landau theory for a superfluid that in low numbers of spatial dimensions fluctuations can be so strong that they preclude the formation of an ordered state. This fact goes under the name of "Mermin-Wagner-Hohenberg theorem". The dimension below which there is no ordered phase is called *lower critical dimensions*. As we have seen, the formation of order is associated with the spontaneous breaking of a discrete ($\mathbb{Z}_2$ spin reflection for the Ising model) or continuous (spin rotational symmetry $SU(2)$ for the Heisenberg model, U(1) for the superfluid) symmetry. For systems with sort range interactions as the ones we have considered here, continuous symmetries can be broken spontaneously only at zero temperature in $D = 2$, and never in $D = 1$. In contrast, discrete symmetries can be broken only at zero temperature in $D = 1$. Let us summarize our findings for the example for the Landau theory of a superfluid.

- For $D > 4$ the saddle point approximation is valid (the upper critical dimension is 4). There is a transition to an ordered phase and the critical exponents are given by mean field theory.
- In the intermediate regime $2 < D < 4$ there is a transition to an ordered state at low temperatures, but fluctuations modify the mean-field behaviour and lead to different values for the critical exponents characterizing the phase transition.

- For $D < 2$ (the lower critical dimension is 2) phase fluctuations become so strong that it is no longer possible for the system to order at low temperatures.

Part III
Random Systems and Stochastic Processes

In nature there are many phenomena, in which some quantity varies in a random way. An example is *Browninan 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*

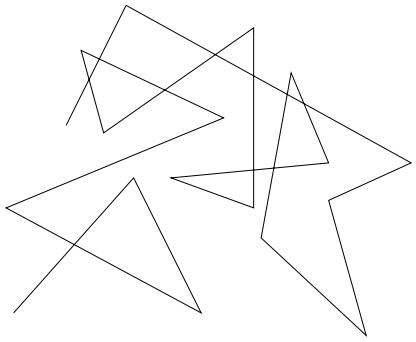


Figure 17: Cartoon of a path of a particle undergoing Brownian motion.

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

$$\text{time average} \longrightarrow \text{ensemble average.} \tag{468}$$

11 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

$$P_X(x_j) \geq 0 \ , \qquad \sum_j P_X(x_j) = 1 \text{ (this becomes an integral in the continuous case.).} \tag{469}$$

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

$$\{x_j\} = \{1,2,3,4,5,6\} \ , \qquad P_X(x_j) = \frac{1}{6}. \tag{470}$$

11.1 Some Definitions

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

$$\boxed{\langle X^n \rangle \equiv \int dx \ x^n \ P_X(x).} \tag{471}$$

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

$$\boxed{\phi_X(k) = \int dx\ P_X(x) e^{ikx} = \sum_{n=0}^{\infty} \frac{(ik)^n}{n!} \langle X^n \rangle \equiv \langle e^{ikX} \rangle.} \tag{472}$$

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

$$\langle X^n \rangle = (-i)^n \frac{d^n}{dk^n}\Big|_{k=0} \phi_X(k). \tag{473}$$

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

$$\boxed{C_n(X) = (-i)^n \frac{d^n}{dk^n}\Big|_{k=0} \ln\big(\phi_X(k)\big).} \tag{474}$$

The first few cumulants are

$$C_1(X) = \langle X \rangle\ , \quad C_2(X) = \langle X^2 \rangle - \langle X \rangle^2\ , \quad C_3(X) = \langle X^3 \rangle - 3\langle X^2 \rangle \langle X \rangle + 2\langle X \rangle^3. \tag{475}$$

The Gaussian distribution is defined by

$$P_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-x_0)^2}{2\sigma^2}}. \tag{476}$$

Its characteristic function is

$$\phi_X(k) = e^{ikx_0 - \frac{1}{2}k^2\sigma^2}. \tag{477}$$

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.

11.2 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 $(j = 1, \ldots, N)$ taking the values ± 1 with probabilities $1/2$.
- The position after r steps is $Y = \sum_{j=1}^{N} X_j$.

Clearly we have

$$\langle Y \rangle = 0\ , \qquad \langle Y^2 \rangle = N \langle X_j^2 \rangle = N\ , \tag{478}$$

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

$$\phi_{X_j}(k) = P_{X_j}(1) e^{ik1} + P_{X_j}(-1) e^{ik(-1)} = \cos(k). \tag{479}$$

The characteristic function of the random variable Y is

$$\begin{aligned} \phi_Y(k) &= \langle e^{ikY} \rangle = \langle e^{ik\sum_{j=1}^{N} X_j} \rangle = \prod_{j=1}^{N} \langle e^{ikX_j} \rangle = \big(\phi_{X_j}(k)\big)^N \\ &= \frac{1}{2^N} \left(e^{ik} + e^{-ik}\right)^N = \frac{1}{2^N} \sum_{r=0}^{N} \binom{N}{r} e^{ik(N-2r)}. \end{aligned} \tag{480}$$

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

$$\phi_Y(k) = \sum_n P_Y(n) e^{ikn} = \sum_{n=-N}^{N} p_N(n) e^{ikn}. \tag{481}$$

Equating (480) with (481) gives

$$p_N(n) = \frac{1}{2^N} \begin{pmatrix} N \\ \frac{N-n}{2} \end{pmatrix}, \tag{482}$$

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

11.3 The central limit theorem

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

$$S_N = \frac{1}{N} \sum_{j=1}^{N} X_j. \tag{483}$$

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

$$\phi_{S_N}(k) = \langle e^{ikS_N} \rangle = \langle e^{i\frac{k}{N}\sum_{j=1}^{N} X_j} \rangle = \langle e^{ik\frac{X}{N}} \rangle^N = \left(\phi_X\big(\frac{k}{N}\big) \right)^N, \tag{484}$$

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

$$C_n(S_N) = (-i)^n \frac{d^n}{dk^n}\Big|_{k=0} N \ln\Big(\phi_X\big(\frac{k}{N}\big)\Big). \tag{485}$$

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

$$\boxed{C_n(S_N) = N^{1-n} C_n(X).} \tag{486}$$

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.

$$\boxed{P_{S_N}(s) \longrightarrow \left(\frac{N}{2\pi\sigma^2} \right)^{\frac{1}{2}} \exp\left[-\frac{N}{2\sigma^2} (s - \langle X \rangle)^2 \right].} \tag{487}$$

12 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

$$\begin{aligned} P_1(y_1, t_1) &= \text{probability that } Y(t_1) = y_1; \\ &\vdots \\ P_n(y_1, t_1; \ldots; y_n, t_n) &= \text{joint probability that } Y(t_1) = y_1, \ldots, Y(t_n) = y_n. \end{aligned} \tag{488}$$

These are

- normalized

$$\int dy_1 \dots dy_n \; P_n(y_1, t_1; \dots; y_n, t_n) = 1 \; ; \tag{489}$$

- reducible

$$\int dy_n \; P_n(y_1, t_1; \dots; y_n, t_n) = P_{n-1}(y_1, t_1; \dots; y_{n-1}, t_{n-1}). \tag{490}$$

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

$$\begin{aligned} \langle Y(t_1) \rangle &= \int dy_1 \; y_1 \; P(y_1, t_1) \; , \\ \langle Y(t_1) Y(t_2) \rangle &= \int dy_1 dy_2 \; y_1 y_2 \; P_2(y_1, t_1; y_2, t_2) \; , \\ &\dots \end{aligned} \tag{491}$$

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

$$P_n(y_1, t_1; \dots; y_n, t_n) = P_n(y_1, t_1 + \tau; \dots; y_n, t_n + \tau). \tag{492}$$

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

$$P_{1|1}(y_2, t_2 | y_1, t_1). \tag{493}$$

This is normalized

$$\int dy_2 \; P_{1|1}(y_2, t_2 | y_1, t_1) = 1 \; , \tag{494}$$

and related to unconditional probabilities by

$$P_2(y_1, t_1; y_2, t_2) = P_{1|1}(y_2, t_2 | y_1, t_1) \; P_1(y_1, t_1). \tag{495}$$

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

$$P_{1|n-1}(y_n, t_n | y_1, t_1; y_2, t_2; \dots; y_{n-1}, t_{n-1}). \tag{496}$$

13 Markov Processes

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

$$\boxed{P_{1|n-1}(y_n, t_n | y_1, t_1; y_2, t_2; \dots; y_{n-1}, t_{n-1}) = P_{1|1}(y_n, t_n | y_{n-1}, t_{n-1}).} \tag{497}$$

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

$$\boxed{P_{1|1}(y_3, t_3 | y_1, t_1) = \int dy_2 \; P_{1|1}(y_3, t_3 | y_2, t_2) P_{1|1}(y_2, t_2 | y_1, t_1) \; , \quad t_3 > t_2 > t_1.} \tag{498}$$

- Evolution equation

$$\boxed{P_1(y_2,t_2) = \int dy_1 \; P_{1|1}(y_2,t_2|y_1,t_1)P_1(y_1,t_1) \;, \quad t_2 > t_1.} \tag{499}$$

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 $y_{1,2}$).

13.1 Examples of Markov Processes

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

$$\begin{aligned} P_{1|1}(y_2,t_2|y_1,t_1) &= \frac{1}{\sqrt{2\pi(t_2-t_1)}} e^{-\frac{(y_2-y_1)^2}{2(t_2-t_1)}} \;, \\ P_1(y,t) &= \frac{1}{\sqrt{2\pi t}} e^{-\frac{y^2}{2t}} \;, \end{aligned} \tag{500}$$

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

13.2 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 (494) becomes

$$\boxed{\sum_{j=1}^{N} T_{jk} = 1,} \tag{501}$$

which is often referred to as probability conservation.

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

$$\begin{aligned} P_{1|1}(1,t+1;1,t) &= 1-q \;, \\ P_{1|1}(2,t+1;1,t) &= q \;, \\ P_{1|1}(1,t+1;2,t) &= r \;, \\ P_{1|1}(2,t+1;2,t) &= 1-r \;. \end{aligned} \tag{502}$$

Introducing a two-component vector

$$\vec{p}(t) = \begin{pmatrix} P_1(1,t) \\ P_1(2,t) \end{pmatrix} \;, \tag{503}$$

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

$$\boxed{\vec{p}(t+1) = \underbrace{\begin{pmatrix} 1-q & r \\ q & 1-r \end{pmatrix}}_{T} \vec{p}(t).} \tag{504}$$

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

$$p_n(t+1) = \sum_j T_{nj} p_j(t). \tag{505}$$

Then

$$p_n(t+1) - p_n(t) = \sum_j T_{nj} p_j(t) - \underbrace{\sum_j T_{jn}}_{1}\ p_n(t) = \sum_j T_{nj} p_j(t) - T_{jn} p_n(t). \tag{506}$$

Now consider our time interval to be small instead of 1. Then (506) turns into a differential equation, called a *Master equation*

$$\boxed{\frac{dp_n(t)}{dt} = \sum_j T_{nj} p_j(t) - T_{jn} p_n(t).} \tag{507}$$

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 (504), which can be iterated to give

$$\vec{p}(t+1) = T^{t+1} \vec{p}(0). \tag{508}$$

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

$$T|R_\alpha\rangle = \lambda_\alpha |R_\alpha\rangle\ , \quad \langle L_\alpha| T = \lambda_\alpha \langle L_\alpha|\ , \quad \langle L_\alpha | R_\beta\rangle = \delta_{\alpha,\beta}\ , \tag{509}$$

and T can be represented in the form

$$T = \sum_\alpha \lambda_\alpha |R_\alpha\rangle\langle L_\alpha|. \tag{510}$$

Using the orthonormality of left and right eigenvectors we have

$$\boxed{T^{t+1} = \sum_\alpha \lambda_\alpha^{t+1} |R_\alpha\rangle\langle L_\alpha|.} \tag{511}$$

In our example

$$\lambda_1 = 1\ , \quad \langle L_1| = (1,1)\ , \quad |R_1\rangle = \frac{1}{1+r/q}\begin{pmatrix} r/q \\ 1 \end{pmatrix}. \tag{512}$$

$$\lambda_2 = 1-q-r\ , \quad \langle L_2| = (-\frac{q}{r}, 1)\ , \quad |R_2\rangle = \frac{1}{1+r/q}\begin{pmatrix} -1 \\ 1 \end{pmatrix}. \tag{513}$$

So for large t we have

$$T^{t+1} \approx |R_1\rangle\langle L_1| = \frac{1}{r+q}\begin{pmatrix} r & r \\ q & q \end{pmatrix}, \tag{514}$$

and hence

$$\vec{p}(\infty) = \frac{1}{r+q}\begin{pmatrix} r & r \\ q & q \end{pmatrix}\ \vec{p}(0). \tag{515}$$

14 Brownian Motion

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

14.1 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 $D > 1$)

$$\boxed{\frac{dv(t)}{dt} = -\gamma v(t) + \eta(t).} \tag{516}$$

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

$$\langle \eta(t)\eta(t') \rangle = \Gamma \delta(t - t'). \tag{517}$$

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

$$\frac{d}{dt'}\left(v(t')e^{\gamma t'}\right) = \left(\frac{dv}{dt'} + \gamma v\right) e^{\gamma t'} = \eta(t')e^{\gamma t'}, \tag{518}$$

where in the last step we used the Langevin equation (516). Integrating both sides of this equation between $t = 0$ and $t' = t$, we obtain

$$\boxed{v(t) = v(0)e^{-\gamma t} + \int_0^t dt'\ \eta(t')e^{-\gamma(t-t')}.} \tag{519}$$

Averaging this over the noise, we find the average velocity

$$\boxed{\langle v(t) \rangle = v(0)e^{-\gamma t} + \int_0^t dt'\ \underbrace{\langle \eta(t') \rangle}_{=0} e^{-\gamma(t-t')} = v(0)e^{-\gamma t}.} \tag{520}$$

Similarly we have

$$\begin{aligned} \langle v^2(t) \rangle &= \langle \left[v(0)e^{-\gamma t} + \int_0^t dt'\ \eta(t')e^{-\gamma(t-t')} \right] \left[v(0)e^{-\gamma t} + \int_0^t dt''\ \eta(t'')e^{-\gamma(t-t'')} \right] \rangle \\ &= v^2(0)e^{-2\gamma t} + e^{-2\gamma t} \int_0^t dt' dt''\ \underbrace{\langle \eta(t')\eta(t'') \rangle}_{\Gamma\delta(t'-t'')} e^{\gamma(t'+t'')} \end{aligned} \tag{521}$$

Carrying out the time integrals this becomes

$$\boxed{\langle v^2(t) \rangle = v^2(0)e^{-2\gamma t} + \frac{\Gamma}{2\gamma}(1 - e^{-2\gamma t}).} \tag{522}$$

The displacement of the particle is

$$x(t) - x(0) = \int_0^t dt'\ v(t') = \frac{v(0)}{\gamma}(1 - e^{-\gamma t}) + \int_0^t dt' \int_0^{t'} dt''\ \eta(t'')e^{-\gamma(t'-t'')}. \tag{523}$$

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

$$\boxed{\langle x(t) \rangle = \frac{v(0)}{\gamma}(1 - e^{-\gamma t}).} \tag{524}$$

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

$$\langle \big[x(t) - \langle x(t)\rangle\big]^2\rangle = \langle x^2(t)\rangle - \langle x(t)\rangle^2. \tag{525}$$

Substituting (523) and using again that $x(0) = 0$ this becomes

$$\int_0^t dt_1 \int_0^{t_1} dt_2 e^{-\gamma(t_1-t_2)} \int_0^t dt_1' \underbrace{\int_0^{t_1'} dt_2' e^{-\gamma(t_1'-t_2')}\langle \eta(t_2)\eta(t_2')\rangle}_{\Gamma\Theta(t_1'-t_2)e^{-\gamma(t_1'-t_2)}} = \int_0^t dt_1 \int_0^{t_1} dt_2 e^{-\gamma(t_1-2t_2)} \underbrace{\int_{t_2}^t dt_1' e^{-\gamma t_1'}}_{\gamma^{-1}(e^{-\gamma t_2}-e^{-\gamma t})}. \tag{526}$$

Carrying out the remaining two integrals we find

$$\boxed{\langle \big[x(t) - \langle x(t)\rangle\big]^2\rangle = \frac{\Gamma}{\gamma^2}t - \frac{\Gamma}{\gamma^3}(1-e^{-\gamma t}) - \frac{\Gamma}{2\gamma^3}(1-e^{-\gamma t})^2.} \tag{527}$$

So at very late times we have

$$\boxed{\langle x^2(t)\rangle = \frac{\Gamma}{\gamma^2}t + \ldots} \tag{528}$$

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

$$\langle v^2(t\to\infty)\rangle = \frac{\Gamma}{2\gamma}. \tag{529}$$

On the other hand, by equipartition we have

$$\frac{m}{2}\langle v^2\rangle \sim \frac{k_B T}{2} \tag{530}$$

Combining these two equations, we arrive at

$$\boxed{\frac{\Gamma}{2\gamma} = \frac{k_B T}{m}.} \tag{531}$$

14.2 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

$$v(t) = v(0)e^{-\gamma t} + \int_0^t dt'\ e^{-\gamma(t-t')}\eta(t'). \tag{532}$$

Our starting point is the general evolution equation

$$P_1(v,t+\Delta t) = \int du P_{1|1}(v,t+\Delta t|u,t)\ P_1(u,t). \tag{533}$$

It is convenient to consider the integral

$$\Omega = \int dv\,[P_1(v,t+\Delta t) - P_1(v,t)]\,h(v), \tag{534}$$

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

$$\int dv\,[P_1(v,t+\Delta t) - P_1(v,t)]\,h(v) = \int dv\ \frac{\partial P_1(v,t)}{\partial t}\Delta t\ h(v). \tag{535}$$

On the other hand, using (533) we have

$$\Omega = \int du\ dv\ h(v)\ P_{1|1}(v, t+\Delta t|u,t)\ P_1(u,t) - \int dv\ h(v)\ P_1(v,t). \tag{536}$$

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

$$\Omega = \int du\ P_1(u,t) \int dv\ P_{1|1}(v, t+\Delta t|u,t)\,[h(v) - h(u)]\,. \tag{537}$$

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

$$\Omega = \int du\ P_1(u,t) \sum_{n=1}^{\infty} h^{(n)}(u) \underbrace{\int dv\ P_{1|1}(v, t+\Delta t|u,t) \frac{(v-u)^n}{n!}}_{D^{(n)}(u)}\,. \tag{538}$$

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

$$\Omega = \int du\ h(u) \sum_{n=1}^{\infty} \left(-\frac{\partial}{\partial u}\right)^n P_1(u,t) D^{(n)}(u). \tag{539}$$

Using that (535) and (539) have to be equal for *any* test function $h(u)$, we conclude that

$$\boxed{\frac{\partial P_1(v,t)}{\partial t}\Delta t = \sum_{n=1}^{\infty} \left(-\frac{\partial}{\partial v}\right)^n P_1(v,t) D^{(n)}(v).} \tag{540}$$

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

$$\begin{aligned} D^{(n)}(v) &= \int dw\ P_{1|1}(w, t+\Delta t|v,t) \frac{(w-v)^n}{n!} = \int dz\ P_{1|1}(v+z, t+\Delta t|v,t) \frac{z^n}{n!} \\ &= \frac{1}{n!} \langle [v(t+\Delta t) - v(t)]^n \rangle. \end{aligned} \tag{541}$$

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

$$\begin{aligned} \langle v(t+\Delta t) - v(t) \rangle &= -\gamma v(0) e^{-\gamma t} \Delta t\ , \\ \langle \big(v(t+\Delta t) - v(t)\big)^2 \rangle &= \Gamma \Delta t + \mathcal{O}\big((\Delta t)^2\big)\ , \\ \langle \big(v(t+\Delta t) - v(t)\big)^n \rangle &= \mathcal{O}\big((\Delta t)^2\big)\ , \quad n \geq 3. \end{aligned} \tag{542}$$

Substituting these into (540) and then taking the limit $\Delta t \to 0$, we arrive at the *Fokker-Planck equation*

$$\boxed{\frac{\partial}{\partial t} P_1(v,t) = \gamma \frac{\partial}{\partial v} v P_1(v,t) + \frac{\Gamma}{2} \frac{\partial^2}{\partial v^2} P_1(v,t).} \tag{543}$$

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

$$\boxed{P_1(v,t) = \frac{1}{\sqrt{2\pi\sigma^2(t)}} \exp\left(-\frac{\big(v - \bar{v}(t)\big)^2}{2\sigma^2(t)}\right),} \tag{544}$$

where

$$\sigma^2(t) = \frac{\Gamma}{2\gamma}(1 - e^{-2\gamma t}) \ , \quad \bar{v}(t) = v_0 e^{-\gamma t}. \tag{545}$$

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

$$P_1(v,t) = \sqrt{\frac{\gamma}{\pi\Gamma}} e^{-\frac{\gamma v^2}{\Gamma}} = \sqrt{\frac{m}{2\pi k_B T}} e^{-\frac{mv^2}{2k_B T}}. \tag{546}$$

14.2.1 Moments of the Velocity Difference Probability Distribution

Let us see how to derive (542), starting from the Langevin equation

$$v(t) = v(0)e^{-\gamma t} + \int_0^t dt' \ e^{-\gamma(t-t')}\eta(t'). \tag{547}$$

For a very small time interval Δt we have

$$\begin{aligned} v(t+\Delta t) &= v(0)e^{-\gamma(t+\Delta t)} + \int_0^{t+\Delta t} dt' \ e^{-\gamma(t+\Delta t-t')}\eta(t') \\ &= v(0)e^{-\gamma t}(1-\gamma\Delta t) + \int_0^t dt' \ e^{-\gamma(t-t')}\eta(t')(1-\gamma\Delta t) + \int_t^{t+\Delta t} dt' \ e^{-\gamma(t-t')}\eta(t') + \mathcal{O}\big((\Delta t^2)\big) \\ &= (1-\gamma\Delta t) \ v(t) + \int_t^{t+\Delta t} dt' \ e^{-\gamma(t-t')}\eta(t') + \mathcal{O}\big((\Delta t^2)\big). \end{aligned} \tag{548}$$

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

$$\Delta v = -\gamma v \Delta t + \int_t^{t+\Delta t} dt' \ e^{-\gamma(t-t')}\eta(t') + \mathcal{O}\big((\Delta t^2)\big). \tag{549}$$

To derive (542) we rewrite (549) in the form

$$\Delta v(t) = v_0(t)\Delta t + \int dt' K(t,t') \ \eta(t'), \tag{550}$$

where

$$\begin{aligned} v_0(t) &= -\gamma v(0) e^{-\gamma t} \ , \\ K(t,t') &= -\gamma\Theta(t-t')\Theta(t')e^{-\gamma(t-t')}\Delta t + \Theta(t'-t)\Theta(t+\Delta t-t')e^{-\gamma(t-t')}. \end{aligned} \tag{551}$$

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

$$P[\eta(t)] = e^{-\frac{1}{2\Gamma}\int dt \ \eta^2(t)}. \tag{552}$$

Averages are then given by the path integral

$$\langle \eta(t_1)\ldots\eta(t_n)\rangle = \int \mathcal{D}\eta(t) \ \eta(t_1)\ldots\eta(t_n) \ e^{-\frac{1}{2\Gamma}\int dt' \ \eta^2(t')}. \tag{553}$$

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

$$\begin{aligned} \langle\eta(t)\rangle &= 0 \ , \\ \langle\eta(t)\eta(t')\rangle &= \Gamma\delta(t-t') \ , \\ \langle\eta(t_1)\eta(t_2)\eta(t_3)\rangle &= \langle\eta(t_1)\eta(t_2)\rangle\langle\eta(t_3)\rangle + \langle\eta(t_3)\eta(t_1)\rangle\langle\eta(t_2)\rangle + \langle\eta(t_2)\eta(t_3)\rangle\langle\eta(t_1)\rangle = 0 \ , \\ \langle\eta(t_1)\eta(t_2)\eta(t_3)\eta(t_4)\rangle &= \langle\eta(t_1)\eta(t_2)\rangle\langle\eta(t_3)\eta(t_4)\rangle + \ldots \end{aligned} \tag{554}$$

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

$$Z(\lambda) = \langle e^{\lambda \Delta v(t)} \rangle = e^{\lambda v_0(t)\Delta t} \int \mathcal{D}\eta(t)\ e^{-\frac{1}{2\Gamma}\int dt' \left[\eta^2(t') - 2\lambda\Gamma K(t,t')\eta(t')\right]}. \tag{555}$$

Changing variables to

$$\tilde{\eta}(t') = \eta(t') - \lambda\Gamma K(t,t'), \tag{556}$$

the generating function becomes

$$Z(\lambda) = e^{\lambda v_0(t)\Delta t + \frac{\lambda^2\Gamma}{2}\int dt'\ K^2(t,t')}. \tag{557}$$

It is then a straightforward matter to calculate the moments

$$\begin{aligned}
\langle \Delta v \rangle &= \frac{dZ}{d\lambda}\Big|_{\lambda=0} = v_0(t)\Delta t\ , \\
\langle (\Delta v)^2 \rangle &= \frac{d^2 Z}{d\lambda^2}\Big|_{\lambda=0} = \Gamma \int dt'\ K^2(t,t') = \Gamma \int_t^{t+\Delta t} dt'\ e^{-\gamma(t-t')} + \mathcal{O}\big((\Delta t)^2\big) \\
&= \Gamma\Delta t + \mathcal{O}\big((\Delta t)^2\big)\ , \\
\langle (\Delta v)^n \rangle &= \frac{d^n Z}{d\lambda^n}\Big|_{\lambda=0} = \mathcal{O}\big((\Delta t)^2\big)\ , \quad n \geq 3.
\end{aligned} \tag{558}$$

14.3 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 $v(t)$. On can treat the combined evolution of $x(t)$ and $v(t)$, but we will follow a simpler route. Recalling that

$$v(t) = v(0)e^{-\gamma t} + \int_0^t dt'\ e^{-\gamma(t-t')}\eta(t'), \tag{559}$$

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

$$\langle v(t) \rangle = 0\ , \qquad \langle v(t)v(t') \rangle = \frac{\Gamma}{2\gamma} e^{-\gamma|t-t'|}. \tag{560}$$

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

$$\langle v(t)v(t') \rangle = \frac{\Gamma}{2\gamma} e^{-\gamma|t-t'|} \to \frac{\Gamma}{\gamma^2}\ \delta(t-t'). \tag{561}$$

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

$$\int_{-\infty}^{\infty} dt \frac{\Gamma}{2\gamma} e^{-\gamma|t-t'|} = \frac{\Gamma}{\gamma^2}. \tag{562}$$

The differential equation for the position

$$\frac{dx(t)}{dt} = v(t), \tag{563}$$

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

$$\boxed{\frac{\partial}{\partial t} P_1(x,t) = \frac{\Gamma}{2\gamma^2}\frac{\partial^2}{\partial x^2} P_1(x,t).} \tag{564}$$

This is the *diffusion equation* with diffusion coefficient

$$D = \frac{\Gamma}{2\gamma^2}. \tag{565}$$

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

$$\boxed{P_1(x,t) = \frac{1}{\sqrt{2\pi D|t-t_0|}} \exp\left(-\frac{(x-x_0)^2}{4D|t-t_0|}\right).} \tag{566}$$

14.4 Homework Questions 12-16

Question 12. This question is concerned with the central limit theorem.
(i) Show explicitly that for $N \gg 1, pN \gg 1$ the binomial distribution

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

becomes

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

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 $x \to x + dx$ 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 13. Let $y = \pm 1$. Show that

$$P_{1|1}(y,t \mid y',t') = \frac{1}{2}\left\{1 + e^{-2\gamma(t-t')}\right\}\delta_{y,y'} + \frac{1}{2}\left\{1 - e^{-2\gamma(t-t')}\right\}\delta_{y,-y'} \tag{567}$$

obeys the Chapman-Kolmogorov equation.

Show that

$$P_1(y,t) = \frac{1}{2}(\delta_{y,1} + \delta_{y,-1}) \tag{568}$$

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

Question 14. 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-Kolmogarov equation is satisfied for

$$P_{1|1}(y_2,t_2 \mid y_1,t_1) = \frac{1}{\sqrt{2\pi(t_2-t_1)}} \exp - \left\{\frac{(y_2-y_1)^2}{2(t_2-t_1)}\right\}. \tag{569}$$

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

$$P_1(y,t) = \frac{1}{\sqrt{2\pi t}} \exp\left\{\frac{-y^2}{2t}\right\}. \tag{570}$$

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

$$\frac{\partial P}{\partial t} = D\frac{\partial^2 P}{\partial y^2} \tag{571}$$

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

Question 15. 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

$$\frac{\mathrm{d}v}{\mathrm{d}t} = -\gamma v + \eta(t),$$

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 $t > 0$

$$x(t) = \frac{v(0)}{\gamma}\left(1 - \mathrm{e}^{-\gamma t}\right) + \int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2 \; \mathrm{e}^{-\gamma(t_1 - t_2)}\, \eta(t_2)$$

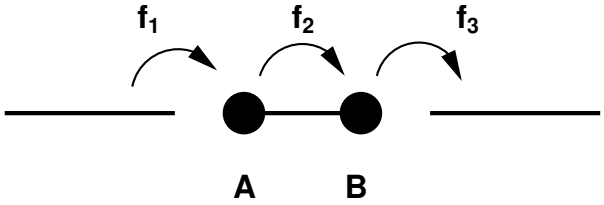
is a solution of the Langevin equation with initial condition $x(0) = 0$. Calculate the average $\langle x(t)\, v(t) \rangle$ and discuss its limiting behaviour at short and long times.

Question 16. The time evolution of a stochastic system is represented by a master equation of the form

$$\frac{\mathrm{d}p_n(t)}{\mathrm{d}t} = \sum_m W_{nm} p_m(t)\,.$$

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

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