Last updated October 12, 2022

UNIVERSITY COLLEGE LONDON

DEPARTMENT OF PHYSICS AND ASTRONOMY

PHAS0059 Theoretical Condensed Matter

Dr. Frank Krüger

Over the past few years, this lecture course has been developed in collaboration with Prof. Andrew Green. Parts of this course are based on the notes of the graduate course 'Quantum Theory of Condensed Matter' by John Chalker, Oxford University, 2013/14 (http://www-thphys.physics.ox.ac.uk/people/JohnChalker/teaching.html).

Please send any comments and corrections to f.kruger@ucl.ac.uk

These notes are intended as an aid to revision, but they do not contain all the examples and diagrams from the course and are a supplement, not a substitute for your own notes taken in lectures.
They are also still being revised, added to and corrected.

Contents

Chapter 1

From the Harmonic Oscillator to Phonons

In this chapter, we start with a review of the harmonic oscillator and its description using raising and lowering operators. We note in particular how all of its properties - expectations of operators in arbitrary states or even thermal distributions - can be calculated in terms of these raising and lowering operators without ever having to write down the wavefunction explicitly.

Next, we will proceed to couple two such oscillators and then a whole chain. The latter provides a simple model for quantum oscillations of a lattice - known as phonons. In working through this problem, we will have introduced the central ideas of quantum field theory. These will be formulated as a book-keeping tool to describe many-particle quantum systems in chapter 2 and we will use them throughout the course.

1.1 The Harmonic Oscillator

Consider a quantum particle whose motion is described by the Hamiltonian

$$\hat{\mathcal{H}} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2, \tag{1.1}$$

where $\hat{x}$ and $\hat{p}$ are the position and momentum operators, respectively. These operators are hermitian (real eigenvalues), $\hat{x} = \hat{x}^\dagger$ and $\hat{p} = \hat{p}^\dagger$, and are subject to the commuter relation $[\hat{x}, \hat{p}] = \hat{x}\hat{p} - \hat{p}\hat{x} = i\hbar$. The non-zero commutator implies that it is not possible to have perfect knowledge about both the position and momentum of the quantum particle, which is known as the Heisenberg uncertainty principle.

Rather than computing the wavefunctions of the ground state and excited states by solving the Schrödinger equation, we can solve the problem by introducing raising and lowering operators,

$$\hat{a}^\dagger = \sqrt{\frac{m\omega}{2\hbar}}\left(\hat{x} - i\frac{\hat{p}}{m\omega}\right) \quad \text{and} \quad \hat{a} = \sqrt{\frac{m\omega}{2\hbar}}\left(\hat{x} + i\frac{\hat{p}}{m\omega}\right). \tag{1.2}$$

It is easy to verify that the raising and lowering operators (ladder operators) satisfy the bosonic commutator relation $[\hat{a}, \hat{a}^\dagger] = 1$,

$$[\hat{a},\hat{a}^\dagger] = \frac{m\omega}{2\hbar}[\hat{x}+i\frac{\hat{p}}{m\omega},\hat{x}-i\frac{\hat{p}}{m\omega}] = \frac{1}{2\hbar}(-i\underbrace{[\hat{x},\hat{p}]}_{=i\hbar}+i\underbrace{[\hat{p},\hat{x}]}_{=-i\hbar}) = 1. \tag{1.3}$$

In terms of the ladder operators the Hamilton takes the form (homework problem)

$$\hat{\mathcal{H}} = \hbar\omega(\hat{a}^\dagger\hat{a}+1/2), \tag{1.4}$$

where $\hat{n} = \hat{a}^\dagger\hat{a}$ is the number operator with eigenstates $|n\rangle$, $\hat{n}|n\rangle = n|n\rangle$. The occupation number states form and orthonormal basis, $\langle n|m\rangle = \delta_{m,n}$. Obviously, these states are eigenstates of the Hamiltonian, $\hat{H}|n\rangle = E_n|n\rangle$ with energies $E_n = \hbar\omega(n+1/2)$.

The quantum mechanical groundstate $|0\rangle$ has the zero-point energy $E_0 = \hbar\omega/2$. It follows directly from the Heisenberg uncertainty principle, since the classical groundstate is forbidden as it would give perfect knowledge about both the position and momentum of the operator.

Excitations are created by the action of the raising operator on the groundstate. Using that

$$\begin{aligned}\hat{a}|n\rangle &= \sqrt{n}|n-1\rangle,\\ \hat{a}^\dagger|n\rangle &= \sqrt{n+1}|n+1\rangle,\end{aligned}$$

we can write the state with n excitations as

$$|n\rangle = \frac{(\hat{a}^\dagger)^n}{\sqrt{n!}}|0\rangle. \tag{1.5}$$

Calculations with Raising and Lowering Operators:
We can calculate all of the observable properties of the harmonic oscillator using the above relations and without ever having to obtain explicit expressions for the wavefunctions $\psi_n(x)$. The following are a few instructive examples:

a) Average in the n^{th} excited state:

$$\begin{aligned}\langle\hat{x}^2\rangle_n &= \langle n|\hat{x}^2|n\rangle\\ &= \frac{\hbar}{2m\omega}\langle n|(\hat{a}+\hat{a}^\dagger)^2|n\rangle\\ &= \frac{\hbar}{2m\omega}\langle n|(\hat{a}^2+\hat{a}^\dagger\hat{a}+\hat{a}\hat{a}^\dagger+\hat{a}^\dagger\hat{a})|n\rangle\\ &\quad \text{Using } \langle n|\hat{a}^2|n\rangle \propto \langle n|n-2\rangle = 0 \text{ and } \langle n|\hat{a}^\dagger\hat{a}^\dagger|n\rangle \propto \langle n|n+2\rangle = 0\\ &= \frac{\hbar}{2m\omega}\langle n|(\hat{a}\hat{a}^\dagger+\hat{a}^\dagger\hat{a})|n\rangle\\ &\quad \text{Using } \hat{a}\hat{a}^\dagger = 1+\hat{a}^\dagger\hat{a}\\ &= \frac{\hbar}{2m\omega}\langle n|(2\hat{a}^\dagger\hat{a}+1)|n\rangle = \frac{\hbar}{m\omega}(n+1/2)\end{aligned}$$

Note that a similar calculation (homework) gives $\langle\hat{p}^2\rangle_n = \hbar m\omega(n+1/2)$ and hence $\langle\hat{\mathcal{H}}_{\rm kin}\rangle_n = \langle\hat{\mathcal{H}}_{\rm pot}\rangle_n = \frac{1}{2}E_n$.

b) The partition function :

$$\begin{aligned}\mathcal{Z} &= \text{Tr } e^{-\beta\hat{\mathcal{H}}} = \sum_{n=0}^{\infty}\langle n|e^{-\beta\hbar\omega(\hat{n}+1/2)}|n\rangle \\ &= \sum_{n=0}^{\infty} e^{-\beta\hbar\omega(n+1/2)} = e^{-\beta\hbar\omega/2}\sum_{n=0}^{\infty}\left(e^{-\beta\hbar\omega}\right)^n = \frac{e^{-\beta\hbar\omega/2}}{1-e^{-\beta\hbar\omega}}.\end{aligned}$$

Here the factor $e^{-\beta\hbar\omega/2}$ arrises from the zero-point energy $\hbar\omega/2$. Note that this factor drops out when calculating thermal averages of observables,

$$\langle\langle\hat{A}\rangle\rangle = \frac{1}{\mathcal{Z}}\text{Tr } \hat{A}e^{-\beta\hat{\mathcal{H}}} = \frac{\text{Tr } \hat{A}e^{-\beta\hbar\omega(\hat{n}+1/2)}}{\text{Tr } e^{-\beta\hbar\omega(\hat{n}+1/2)}} = \frac{\cancel{e^{-\beta\hbar\omega/2}}}{\cancel{e^{-\beta\hbar\omega/2}}}\frac{\text{Tr } \hat{A}e^{-\beta\hbar\omega\hat{n}}}{\text{Tr } e^{-\beta\hbar\omega\hat{n}}}.$$

It is therefore convenient to drop the zero-point energy when calculating thermal averages, as the following example shows:

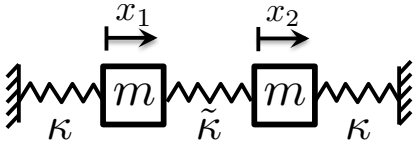
c) Finite temperature average of the number operator:

$$\begin{aligned}\langle\langle\hat{n}\rangle\rangle &= \frac{1}{\mathcal{Z}}\text{Tr } \hat{n}e^{-\beta\hat{\mathcal{H}}} = \frac{1}{\mathcal{Z}}\text{Tr } \hat{n}e^{-\beta\hbar\omega\hat{n}} \quad \left(\text{where} \quad \mathcal{Z} = \text{Tr}e^{-\beta\hbar\omega\hat{n}} = 1/(1-e^{-\beta\hbar\omega})\right) \\ &= -\frac{1}{\hbar\omega}\frac{1}{\mathcal{Z}}\text{Tr } \frac{\partial}{\partial\beta}e^{-\beta\hbar\omega\hat{n}} = -\frac{1}{\hbar\omega}\frac{1}{\mathcal{Z}}\frac{\partial\mathcal{Z}}{\partial\beta} = -\frac{1}{\hbar\omega}\frac{\partial\ln\mathcal{Z}}{\partial\beta} \\ &= \frac{1}{\hbar\omega}\frac{\partial}{\partial\beta}\ln\left(1-e^{-\beta\hbar\omega}\right) = \frac{e^{-\beta\hbar\omega}}{1-e^{-\beta\hbar\omega}} = \frac{1}{e^{\beta\hbar\omega}-1} = n_B(\hbar\omega),\end{aligned}$$

where in the last step we have defined the bose function $n_B(\epsilon) = 1/(e^{\beta\epsilon}-1)$.

1.2 Two Coupled Oscillators

Up to now this has hopefully been revision. Now let's turn to something new. We will consider a simple problem of two coupled quantum harmonic oscillators. In doing so, we will see many of the techniques that underpin quantum field theory. Consider two coupled quantum mechanical oscillators as illustrated in the following figure:

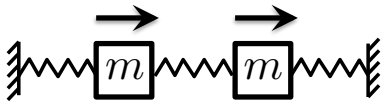


The Hamiltonian of this system is given by

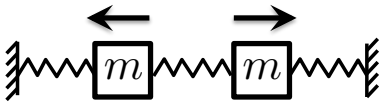
$$\hat{\mathcal{H}} = \frac{\hat{p}_1^2}{2m} + \frac{\kappa}{2}\hat{x}_1^2 + \frac{\hat{p}_2^2}{2m} + \frac{\kappa}{2}\hat{x}_2^2 + \frac{\tilde{\kappa}}{2}(\hat{x}_1-\hat{x}_2)^2, \tag{1.6}$$

where the position and momentum operators of the two particles satisfy $[\hat{x}_i, \hat{p}_j] = i\hbar\delta_{ij}$. All other commutators vanish, e.g. $[\hat{p}_i, \hat{p}_j] = 0$.

Sum and Difference Coordinates:
Our physical intuition about this situation suggests two types of periodic motion or *normal modes*:



symmetric normal mode



anti-symmetric normal mode

Accordingly, we make a change of variables to the mean displacement and difference in displacement;

$$\hat{X} = \frac{\hat{x}_1 + \hat{x}_2}{\sqrt{2}} \qquad \text{and} \qquad \hat{x} = \frac{\hat{x}_1 - \hat{x}_2}{\sqrt{2}},$$

with conjugate momenta

$$\hat{P} = \frac{\hat{p}_1 + \hat{p}_2}{\sqrt{2}} \qquad \text{and} \qquad \hat{p} = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{2}}.$$

In the position representation, $\hat{p}_1 = -i\hbar\partial_{x_1}$ and $\hat{p}_2 = -i\hbar\partial_{x_2}$. The expressions for the momenta $\hat{P}$ and $\hat{p}$ — in particular the factors of $1/\sqrt{2}$ — follow from these.

Preservation of Commutation Relations:
These definitions imply the commutation relations

$$[\hat{x}_1, \hat{p}_1] = [\hat{x}_2, \hat{p}_2] = i\hbar \quad \Leftrightarrow \quad [\hat{X}, \hat{P}] = [\hat{x}, \hat{p}] = i\hbar.$$

The requirement that the commutation relations take this form — *i.e.* that their form is preserved by our change of variables — can be used to fix the expressions (including the factors of $1/\sqrt{2}$) for $\hat{P}$ and $\hat{p}$.

$$[\hat{X}, \hat{P}] = [\frac{\hat{x}_1 + \hat{x}_2}{\sqrt{2}}, \frac{\hat{p}_1 + \hat{p}_2}{\sqrt{2}}] = [\hat{x}_1, \hat{p}_1]/2 + [\hat{x}_2, \hat{p}_2]/2 + \underbrace{[\hat{x}_1, \hat{p}_2]/2}_{=0} + \underbrace{[\hat{x}_2, \hat{p}_1]/2}_{=0} = i\hbar.$$

$$[\hat{x}, \hat{p}] = [\frac{\hat{x}_1 - \hat{x}_2}{\sqrt{2}}, \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{2}}] = [\hat{x}_1, \hat{p}_1]/2 + [\hat{x}_2, \hat{p}_2]/2 - \underbrace{[\hat{x}_1, \hat{p}_2]/2}_{=0} - \underbrace{[\hat{x}_2, \hat{p}_1]/2}_{=0} = i\hbar.$$

Substituting into the Hamiltonian(1.6), we find

$$\begin{aligned} \hat{\mathcal{H}} &= \frac{1}{2m}\left[\frac{(\hat{P}+\hat{p})^2}{2} + \frac{(\hat{P}-\hat{p})^2}{2}\right] + \frac{\kappa}{2}\left[\frac{(\hat{X}+\hat{x})^2}{2} + \frac{(\hat{X}-\hat{x})^2}{2}\right] + \tilde{\kappa}\hat{x}^2 \\ &= \frac{\hat{P}^2}{2m} + \underbrace{\kappa/2}_{=\frac{1}{2}m\omega_s^2}\hat{X}^2 + \frac{\hat{p}^2}{2m} + \underbrace{(\kappa/2+\tilde{\kappa})}_{=\frac{1}{2}m\omega_a^2}\hat{x}^2. \end{aligned}$$

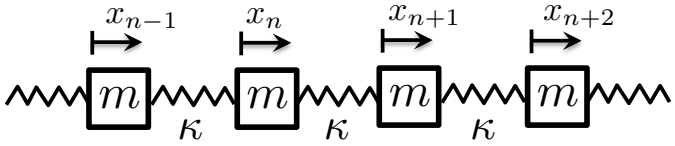
We have reduced the motion to two *independent* harmonic motions with mass m, and frequencies

$$\omega_s = \sqrt{\frac{\kappa}{m}} \quad \text{and} \quad \omega_a = \sqrt{\frac{\kappa + 2\tilde{\kappa}}{m}}.$$

This method of separating — or *diagonalising* — a Hamiltonian into independent harmonic oscillators is more general. As we shall see in the next section, it is related to a Fourier transform for a translationally invariant system; for a 2-site chain, the allowed Fourier components have wavevectors $k = 0$ and $k = 2\pi/2a$, corresponding to in-phase and anti-phase oscillation.

1.3 The Harmonic Chain

We now consider a chain of N quantum particles of mass m with harmonic spring with spring constant κ between them. The equilibrium spacing (lattice constant) between the atoms is a.



In terms of momentum and position operators the Hamiltonian of the harmonic chain is given by

$$\mathcal{H} = \sum_{n=1}^{N} \left[\frac{\hat{p}_n^2}{2m} + \frac{1}{2}\kappa(\hat{x}_n - \hat{x}_{n-1})^2 \right], \tag{1.7}$$

where the operators are subject to the canonical commutator relation $[\hat{x}_i, \hat{p}_j] = i\hbar\delta_{ij}$. We further impose periodic boundary conditions corresponding to taking $\hat{x}_0 = \hat{x}_N$. In the end, we will take the thermodynamic limit $N \to \infty$. Our solution is going to follow essentially the same steps as we used for two coupled oscillators.

1.3.1 Diagonalising by Fourier Transform

The first step is to reduce the problem to a set of independent harmonic oscillators. We achieve this by using a discrete Fourier transform. Many problems in physics are translationally invariant (or can be treated as such) and in these cases the Fourier transform takes us much of the way towards diagonalising our system.

The Fourier components of the particle displacements are given by

$$\hat{x}_k = \frac{1}{\sqrt{N}} \sum_n e^{ikna} \hat{x}_n \quad \text{with inverse} \quad \hat{x}_n = \frac{1}{\sqrt{N}} \sum_k e^{-ikna} \hat{x}_k, \tag{1.8}$$

for the position operators and

$$\hat{p}_k = \frac{1}{\sqrt{N}} \sum_n e^{-ikna} \hat{p}_n \quad \text{with inverse} \quad \hat{p}_n = \frac{1}{\sqrt{N}} \sum_k e^{ikna} \hat{p}_k, \tag{1.9}$$

for the momentum operators. Details of the Fourier transform can be found below (box A). As a result of the periodic boundary conditions, $\hat{x}_n = \hat{x}_{n+N}$, the momentum k can only take N discrete values which are integer multiples of $\Delta k = \frac{2\pi}{Na}$,

$$\frac{1}{\sqrt{N}} \sum_k e^{-ikna} \hat{x}_k = \frac{1}{\sqrt{N}} \sum_k e^{-ik(n+N)a} \hat{x}_k \Rightarrow e^{-ikNa} = 1 \Rightarrow k = \frac{2\pi}{Na} j,$$

where we could restrict j to the integer values $j = 0, \ldots, N-1$, or a range shifted by any integer. This is because a wavevector $k' = k + 2\pi/a$ leads to the same displacements of the particles; $e^{ik'na} = e^{ikna + i2\pi n} = e^{ikna}$. The standard convention is to restrict momenta to the range $-\pi/a < k \leq \pi/a$, which is known as the *first Brillouin zone*.

Note that could have written the Fourier transform in terms of cos and sin rather than complex exponentials. Since $\hat{x}_n$ and $\hat{p}_n$ are hermitian, in momentum space we obtain $\hat{x}_k^\dagger = \hat{x}_{-k}$ and $\hat{p}_k^\dagger = \hat{p}_{-k}$. This fixes the number of degrees of freedom to be correct — otherwise there would appear to be twice as many components of $\hat{x}_k$ due to the real and imaginary parts. Also note the opposite sign of wavevector k in the transformations of $\hat{x}$ and $\hat{p}$. The reasons for this will become clear shortly.

Commutation relations:
Just as in the case of the two coupled oscillators, the commutation relations are preserved by our Fourier transformation (see box B):

$$[\hat{x}_n, \hat{p}_m] = i\hbar\delta_{n,m} \quad \Leftrightarrow [\hat{x}_p, \hat{p}_q] = i\hbar\delta_{p,q}$$

The Hamiltonian expressed in terms of these transformed operators is given by

$$\hat{\mathcal{H}} = \sum_k \left[\frac{\hat{p}_k \hat{p}_{-k}}{2m} + 2\kappa \sin^2(ka/2) \hat{x}_k \hat{x}_{-k} \right]. \tag{1.10}$$

A detailed derivation can be found below (box C). Except for the presence of both $\hat{p}_k$ and $\hat{p}_{-k}$ (and $\hat{x}_k$ and $\hat{x}_{-k}$) this a Hamiltonian for one independent oscillator for each value of k. We can make it look exactly like independent harmonic oscillators by separating $\hat{x}_k$ and $\hat{p}_k$ into parts that are symmetric and anti-symmetric in k;

$$\hat{x}_k^c = \frac{\hat{x}_k + \hat{x}_{-k}}{\sqrt{2}} \quad \text{and} \quad \hat{x}_k^s = \frac{\hat{x}_k - \hat{x}_{-k}}{i\sqrt{2}}, \quad \text{and} \quad \hat{p}_k^c = \frac{\hat{p}_k + \hat{p}_{-k}}{\sqrt{2}} \quad \text{and} \quad \hat{p}_k^s = \frac{\hat{p}_k - \hat{p}_{-k}}{i\sqrt{2}}.$$

As implied by the notation, these are the coefficients of $\cos(ka)$ and $\sin(ka)$ in a real Fourier expansion. In terms of these

$$\mathcal{H} = \sum_{k,\sigma=c,s} \left[\frac{\hat{p}_k^\sigma \hat{p}_k^\sigma}{2m} + 2\kappa \sin^2(ka/2) \hat{x}_k^\sigma \hat{x}_k^\sigma \right] \tag{1.11}$$

The system has been reduced to a set of independent harmonic oscillators with frequencies given by $m\omega_k^2/2 = 2\kappa \sin^2(ka/2)$ or $\omega_k = 2\sqrt{\kappa/m}|\sin(ka/2)|$, which in the limit of small frequencies reduces to $\omega_k \approx ka\sqrt{\kappa/m}$. This corresponds to an acoustic phonon with sound velocity $c = a\sqrt{\kappa/m}$.

Details of Fourier Transform

A. Forwards and Backwards Transforms

This can be verified using the identity

$$\delta_{n,m} = \frac{1}{N}\sum_k e^{ik(n-m)a}.$$

Substituting into the Fourier transformed operators, we find

$$\begin{aligned}
\hat{x}_n &= \frac{1}{\sqrt{N}}\sum_k e^{-ikna}\hat{x}_k \\
&= \frac{1}{N}\sum_k e^{-ikna}\left(\sum_m e^{ikma}\hat{x}_m\right) \\
&= \frac{1}{N}\sum_{k,m} e^{-ik(n-m)a}\hat{x}_m \\
&= \sum_m \delta_{n,m}\hat{x}_m \\
&= \hat{x}_n
\end{aligned}$$

We can also verify the relationship between the forward and backward transforms using the wavevector version of the delta function summation formula

$$\delta_{k,q+G} = \frac{1}{N}\sum_n e^{-i(k-q)na},$$

where $G = 2\pi \times \text{integer}/a$ is a reciprocal lattice vector. This extra contribution of G effectively allows us to fold back contributions of wavevectors outside of the Brillouin zone into the Brillouin zone, since they give the same displacements. The calculation takes the form

$$\begin{aligned}
\hat{x}_k &= \frac{1}{\sqrt{N}}\sum_n e^{ikna}\hat{x}_n \\
&= \frac{1}{N}\sum_n e^{ikna}\left(\sum_q e^{-iqna}\hat{x}_q\right) \\
&= \frac{1}{N}\sum_{k,n} e^{-in(q-k)a}\hat{x}_q \\
&= \sum_k \delta_{q,k}\hat{x}_q \\
&= \hat{x}_k
\end{aligned}$$

B. Fourier transforming the commutation relations

The Fourier transform of the commutation relations proceeds in a very similar manner to the Fourier transform of the Hamiltonian

$$\begin{aligned}
[\hat{x}_p, \hat{p}_q] &= \frac{1}{N}\left[\sum_n e^{ipna}\hat{x}_n \ , \ \sum_m e^{-iqma}\hat{p}_m\right] \\
&= \frac{1}{N}\sum_{n,m} e^{i(pn-qm)a}\underbrace{[\hat{x}_n, \hat{p}_m]}_{=i\hbar\delta_{n,m}} \\
&= i\hbar\frac{1}{N}\sum_n e^{i(p-q)na} \ = \ i\hbar\delta_{p,q}
\end{aligned}$$

C. Fourier transforming the Hamiltonian

The Fourier transformation of the Hamiltonian is most easily achieved by splitting it into the kinetic and potential energy parts.
i. Kinetic Energy

$$\begin{aligned}
\hat{\mathcal{H}}_{\text{kin}} &= \sum_n \frac{\hat{p}_n^2}{2m} \\
&= \frac{1}{2m}\sum_n \left(\frac{1}{\sqrt{N}}\sum_k e^{ikna}\hat{p}_k\right)\left(\frac{1}{\sqrt{N}}\sum_q e^{iqna}\hat{p}_q\right) \\
&= \frac{1}{2m}\sum_{k,q}\underbrace{\frac{1}{N}\sum_n e^{i(k+q)na}}_{=\delta_{k,-q}}\hat{p}_k\hat{p}_q \ = \ \frac{1}{2m}\sum_k \hat{p}_k\hat{p}_{-k}
\end{aligned}$$

ii. Potential Energy

$$\begin{aligned}
\hat{\mathcal{H}}_{\text{pot}} &= \frac{\kappa}{2}\sum_n(\hat{x}_n - \hat{x}_{n-1})^2 \\
&= \frac{\kappa}{2}\sum_n \left(\frac{1}{\sqrt{N}}\sum_k e^{-ikna}(1-e^{ika})\hat{x}_k\right)\left(\frac{1}{\sqrt{N}}\sum_q e^{-iqna}(1-e^{iqa})\hat{x}_q\right) \\
&= \frac{\kappa}{2}\sum_{k,q}\underbrace{\frac{1}{N}\sum_n e^{-i(k+q)na}}_{=\delta_{k,-q}}(1-e^{ika})(1-e^{iqa})\hat{x}_k\hat{x}_q \\
&= \frac{\kappa}{2}\sum_k(1-e^{ika})(1-e^{-ika})\hat{x}_k\hat{x}_{-k} \ = \ 2\kappa\sum_k \sin^2(ka/2)\hat{x}_k\hat{x}_{-k}
\end{aligned}$$

1.3.2 Raising and Lowering Operators

Just as in the case of a single harmonic oscillator, we can now calculate properties of the harmonic chain by introducing ladder operators for each k-mode. These create or destroy quanta of energy in each of the normal modes of oscillation of the harmonic chain,

$$\hat{a}_k = \sqrt{\frac{m\omega_k}{2\hbar}}\left(\hat{x}_k + i\frac{\hat{p}_{-k}}{m\omega_k}\right), \quad \text{and} \quad \hat{a}_k^\dagger = \sqrt{\frac{m\omega_k}{2\hbar}}\left(\hat{x}_{-k} - i\frac{\hat{p}_k}{m\omega_k}\right). \tag{1.12}$$

Notice the opposite momentum labels in $\hat{x}_k$ and $\hat{p}_{-k}$, and the relative signs of k between $\hat{a}$ and $\hat{a}^\dagger$. This is required to obtain the usual commutation relations for the raising and lowering operators;

$$[\hat{a}_k, \hat{a}_q^\dagger] \;=\; \delta_{k,q} \tag{1.13}$$

It is straight forward to check the commutator relations:

D. Checking the commutation relations

$$\begin{aligned}
[\hat{a}_k, \hat{a}_q^\dagger] &= \frac{m\sqrt{\omega_k\omega_q}}{2\hbar}\left[\left(\hat{x}_k + i\frac{\hat{p}_{-k}}{m\omega_k}\right), \left(\hat{x}_{-q} - i\frac{\hat{p}_q}{m\omega_q}\right)\right] \\
&= \frac{m\sqrt{\omega_k\omega_q}}{2\hbar}\left(\frac{i}{m\omega_k}\underbrace{[\hat{p}_{-k}, \hat{x}_{-q}]}_{=-i\hbar\delta_{k,q}} - \frac{i}{m\omega_q}\underbrace{[\hat{x}_k, \hat{p}_q]}_{=i\hbar\delta_{k,q}}\right) \\
&= \delta_{k,q}
\end{aligned}$$

As you will show in a homework problem, in terms of the creation and annihilation operators the Hamiltonian takes the form

$$\begin{aligned}
\hat{\mathcal{H}} &= \sum_k \hbar\omega_k(\hat{a}_k^\dagger\hat{a}_k + 1/2) \\
\omega_k &= 2\sqrt{\kappa/m}|\sin(ka/2)|
\end{aligned} \tag{1.14}$$

Particle Interpretation/Fock Space:
Just as for the harmonic oscillator, armed with these tools we can describe the state of the harmonic chain in terms of the number of quanta — or occupation number n_k — of each mode k. These quanta of lattice vibration are quasiparticles, known as phonons. A general state may be written

$$|\psi\rangle = |n_{q_1}, n_{q_2}, n_{q_3}...\rangle = \prod_q \frac{\left(\hat{a}_q^\dagger\right)^{n_q}}{\sqrt{n_q!}}|0\rangle,$$

where $|0\rangle$ is the vacuum state with no quanta in any mode. The collection of all such states with arbitrary occupations forms a space known as Fock space.

Observables:
One can evaluate the observable properties of the harmonic chain by first translating into

$\hat{a}_k$ and $\hat{a}_k^\dagger$. For example, the mean squared displacement of a particle in the chain at finite temperature is given by

$$\begin{aligned}
\bar{x}^2 &= \frac{1}{N}\sum_n \langle\langle \hat{x}_n^2 \rangle\rangle \\
&\quad \text{Fourier transform - allows us to use the properties of independent k-modes} \\
&= \frac{1}{N}\sum_k \langle\langle \hat{x}_k \hat{x}_{-k} \rangle\rangle \\
&\quad \text{Express in terms of ladder operators} \\
&= \frac{1}{N}\sum_k \frac{\hbar}{2m\omega_k} \langle\langle (\hat{a}_k + \hat{a}_{-k}^\dagger)(\hat{a}_{-k} + \hat{a}_k^\dagger) \rangle\rangle \\
&= \frac{1}{N}\sum_k \frac{\hbar}{2m\omega_k} \langle\langle (\underbrace{\hat{a}_k \hat{a}_{-k}}_{\to 0} + \underbrace{\hat{a}_k \hat{a}_k^\dagger}_{=\hat{a}^\dagger\hat{a}+1=\hat{n}_k+1} + \underbrace{\hat{a}_{-k}^\dagger \hat{a}_{-k}}_{=\hat{n}_{-k}} + \underbrace{\hat{a}_{-k}^\dagger \hat{a}_k^\dagger}_{\to 0}) \rangle\rangle \\
&= \frac{1}{N}\sum_k \frac{\hbar}{2m\omega_k} \left(\langle\langle \hat{n}_k \rangle\rangle + \langle\langle \hat{n}_k \rangle\rangle + 1 \right) \\
&= \frac{1}{N}\sum_k \frac{\hbar}{m\omega_k} \left(\frac{1}{e^{\hbar\omega_k\beta} - 1} + 1/2 \right).
\end{aligned}$$

Note that in the thermodynamic limit ($N \to \infty$) the sums over N discrete momenta with spacing $\Delta k = \frac{2\pi}{Na}$ can be converted into integrals over the Brillouin zone,

$$\frac{1}{N}\sum_k f(k) = \frac{a}{2\pi}\sum_k \Delta k\, f(k) \simeq \frac{a}{2\pi}\int_{-\pi/a}^{\pi/a} dk\, f(k).$$

This generalises straightforwardly to d dimensions with $N = N_1 \cdot N_2 \cdot \ldots \cdot N_d$ and a hyper-cubic Brillouin zone $\mathcal{BZ} = [-\pi/a, \pi/a]^d$,

$$\frac{1}{N}\sum_{\mathbf{k}} f(\mathbf{k}) = \left(\frac{a}{2\pi}\right)^d \sum_{\mathbf{k}} (\Delta k)^d\, f(\mathbf{k}) \simeq \left(\frac{a}{2\pi}\right)^d \int_{\mathbf{k}\in\mathcal{BZ}} d^d k\, f(\mathbf{k}).$$

Note that in general, e.g. for other lattices. the prefactor is equal to $1/V_{\mathcal{BZ}}$, where $V_{\mathcal{BZ}}$ is the volume of the Brillouin zone.

Zero-point Energy and Normal ordering:
Just as for the harmonic oscillator, the Hamiltonian for the harmonic chain has a constant term $\hbar\omega_k/2$ for each wavevector. This implies a minimum energy proportional to N,

$$\hat{\mathcal{H}}_{\text{Zero-point}} = \sum_k \hbar\omega_k/2.$$

This energy cannot usually be extracted (except by changing boundary conditions *c.f.* the Casimir effect) and so is often ignored in field theory. For an infinite system the zero-point energy $\to \infty$. In the continuum limit, the situation is even more severe and the *energy density* $\to \infty$. This is the first infinity that we have encountered. Quantum field theory

is plagued by them. To deal properly with all of the infinities that occur, one needs the renormalisation group.

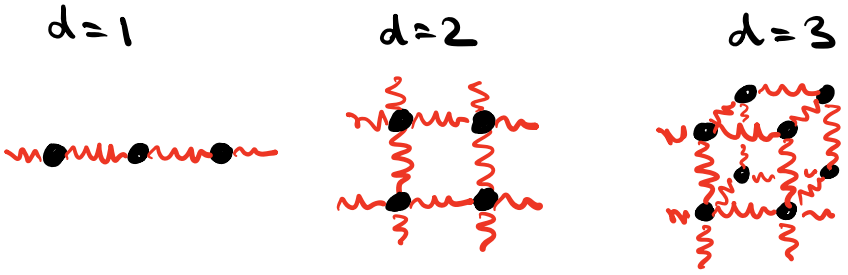
The infinite zero-point energy can be dealt with rather straightforwardly by *normal ordering*. The infinite zero-point energy arises from the form of the Hamiltonain $\mathcal{H} = \sum_k \hbar\omega_k(\hat{a}_k^\dagger \hat{a}_k + \hat{a}_k \hat{a}_k^\dagger)/2$, when we use the commutation relations to reorder the second term so that it can be written in terms of number operators. Similar infinities can arise from interaction terms — terms higher than quadratic order — in Hamiltonians. The trick to dealing with these is to find a consistent way to ignore them! Basically, we define the operation of normal ordering to consist of ordering the creation and annihilation operators so that the creation operators are on the left and the annihilation operators on the right:

$$: \hat{a}_k^\dagger \hat{a}_k + \hat{a}_k \hat{a}_k^\dagger := 2\hat{a}_k^\dagger \hat{a}_k \tag{1.15}$$

Applied to quadratic terms, normal ordering essentially ignores zero-point energy. Applied to interaction terms, it ignores self-interaction of particles that can lead to other unphysical divergences.

1.4 Specific Heat of Phonons in the Harmonic Crystal

We now consider a harmonic crystal in d dimensions, e.g. a hyper-cubic lattice of atoms with elastic springs between them.



Neglecting the zero-point energy, the Hamiltonian in terms of phonon creation and annihilation operators in momentum space is given by

$$\hat{\mathcal{H}} = \sum_{\mathbf{k}} \hbar\omega_{\mathbf{k}} \hat{a}_{\mathbf{k}}^\dagger \hat{a}_{\mathbf{k}}, \tag{1.16}$$

with momenta $\mathbf{k} = (k_1, \ldots, k_d)$ on a d-dimensional momentum grid within the Brillouin zone $\mathcal{BZ} = [-\pi/a, \pi/a]^d$. For simplicity, we will use a linearised phonon dispersion $\omega_{\mathbf{k}} = c|\mathbf{k}|$, where c is the speed of sound of the acoustic phonon. In order to compute the specific heat per atom,

$$c_V = \frac{C_V}{N} = \frac{1}{N}\left(\frac{\partial E}{\partial T}\right)_V, \tag{1.17}$$

we first need to compute the energy $E = \langle\langle \hat{\mathcal{H}} \rangle\rangle$,

$$E = \frac{1}{\mathcal{Z}} \mathrm{Tr}\, \hat{\mathcal{H}} e^{-\beta\hat{\mathcal{H}}} = -\frac{1}{\mathcal{Z}} \frac{\partial}{\partial\beta} \mathrm{Tr}\, e^{-\beta\hat{\mathcal{H}}} = -\frac{1}{\mathcal{Z}} \frac{\partial \mathcal{Z}}{\partial\beta} = -\frac{\partial}{\partial\beta} \ln \mathcal{Z} \tag{1.18}$$

The partition function is equal to

$$\begin{aligned}\mathcal{Z} &= \mathrm{Tr} e^{-\beta\hat{\mathcal{H}}} = \sum_{\{n_{\mathbf{k}}\}} e^{-\beta\sum_{\mathbf{k}}\hbar c|\mathbf{k}|n_{\mathbf{k}}} = \sum_{\{n_{\mathbf{k}}\}}\prod_{\mathbf{k}} e^{-\beta\hbar c|\mathbf{k}|n_{\mathbf{k}}} \\ &= \prod_{\mathbf{k}}\left(\sum_{n=0}^{\infty} e^{-\beta\hbar c|\mathbf{k}|n}\right) = \prod_{\mathbf{k}}\frac{1}{1-e^{-\beta\hbar c|\mathbf{k}|}}. \end{aligned} \tag{1.19}$$

Inserting $\mathcal{Z}$ into Eq. (1.18) and dividing by the number of lattice sites, we obtain

$$\frac{E}{N} = \frac{1}{N}\frac{\partial}{\partial\beta}\sum_{\mathbf{k}}\ln(1-e^{-\beta\hbar c|\mathbf{k}|}) = \frac{1}{N}\sum_{\mathbf{k}}\underbrace{\frac{\hbar c|\mathbf{k}|}{e^{\beta\hbar c|\mathbf{k}|}-1}}_{=\hbar\omega_{\mathbf{k}}\, n_B(\hbar\omega_{\mathbf{k}})}. \tag{1.20}$$

Let us first evaluate the sum for a one-dimensional harmonic chain by converting the momentum sum to an integral for $N\to\infty$,

$$\frac{E}{N} = \frac{a}{2\pi}\int_{-\pi/a}^{\pi/a} dk\frac{\hbar c|k|}{e^{\beta\hbar c|k|}-1} = \frac{a}{\pi}\int_0^{\pi/a} dk\frac{\hbar ck}{e^{\beta\hbar ck}-1} = \frac{T^2}{\hbar\omega_D}\int_0^{\hbar\omega_D/T} dx\frac{x}{e^x-1},$$

where in the last step we substituted $x=\beta\hbar ck$ and defined the Debye frequency $\omega_D = c\pi/a$ as the cut-off frequency. We can evaluate the integral in the high- and low-temperature regimes. At high temperatures, $\hbar\omega_D \ll T$ and therefore $x\ll 1$. In this case we can approximate $x/(e^x-1)\approx 1$, resulting in $E/N\approx T$ and therefor a constant specific heat, $c_V = 1$.

In the low-temperature regime, $T\ll\hbar\omega_D$, we can replace the upper limit of the integral by infinity, resulting in

$$\frac{E}{N} \approx \frac{T^2}{\hbar\omega_D}\underbrace{\int_0^{\infty} dx\frac{x}{e^x-1}}_{=\pi^2/6} \Rightarrow c_V = \frac{\pi^2}{3}\frac{T}{\hbar\omega_D}. \tag{1.21}$$

It is possible to perform the same calculation in general dimension (homework problem). At low temperatures, $T\ll\hbar\omega_D$, the phonon specific heat scales as $c_V\sim T^d$.

Chapter 2

Second Quantisation

In the first chapter, we saw how creation and annihilation operators are a very convenient way of dealing with systems of many harmonic oscillators — armed with their commutation relations, we can calculate observable properties without ever having to write down a wavefunction. Phonons are *quasiparticles* (quanta of energy of lattice vibrations created by $\hat{a}^\dagger_{\mathbf{k}}$). The idea of second quantisation — and field theory in general — is to use this machinery to study many-body quantum systems, e.g. systems of electrons in a solid. In first quantisation, the wavefunction of N bosonic or fermionic particles is obtained by solving the N-particle Schrödinger equation. In 2nd quantisation we introduce creation and annihilation operators operators:

- $\hat{a}^\dagger_{\mathbf{k}}$ – Creates a particle with momentum $\mathbf{k}$ (in a plane wave state)
- $\hat{a}^\dagger_{\mathbf{r}}$ – Creates a particle at the point $\mathbf{r}$ (in a delta-function wavefunction state at $\mathbf{r}$)

2.1 Identical Particles/Many-Particle States

Many-body quantum mechanics in general — and theoretical condensed matter in particular — is concerned with the collective quantum behaviour of many identical quantum particles (for example electrons or phonons in a solid). In this chapter, we will review the formalism required to describe such systems. Consider N identical quantum particles with a generic Hamiltonian

$$\hat{\mathcal{H}} = \underbrace{\sum_{i=1}^{N} \frac{\hat{\mathbf{p}}_i^2}{2m}}_{\text{kinetic energy}} + \underbrace{\sum_{i=1}^{N} U(\hat{\mathbf{r}}_i)}_{\text{potential energy}} + \underbrace{\frac{1}{2}\sum_{i,j=1}^{N} V(\hat{\mathbf{r}}_i - \hat{\mathbf{r}}_j)}_{\text{interaction energy}},$$

with particle labels i and j and canonical commutator relations

$$[\hat{\mathbf{r}}_i, \hat{\mathbf{p}}_j] = i\hbar\delta_{i,j}.$$

Solution for Non-interacting, Distinguishable Particles
In the absence of interactions, $V(\mathbf{r}) \equiv 0$, the time-independent N-particle Schrödinger equation in real space ($\hat{\mathbf{p}}_i = -i\hbar\nabla_i$) is given by

$$\sum_i \left(-\frac{\hbar^2}{2m}\nabla_i^2 + U(\mathbf{r}_i) \right) \psi(\mathbf{r}_1, \ldots, \mathbf{r}_N) = E\psi(\mathbf{r}_1, \ldots, \mathbf{r}_N). \tag{2.1}$$

Since the Hamiltonian is a sum of single-particle Hamiltonians, $\hat{\mathcal{H}} = \sum_i \hat{h}_i$, a mathematical solution of the differential equation is given by a product of single-particle wavefunctions,

$$\psi(\mathbf{r}_1, \ldots, \mathbf{r}_N) = \phi_{\alpha_1}(\mathbf{r}_1)\phi_{\alpha_2}(\mathbf{r}_2) \cdot \ldots \cdot \phi_{\alpha_N}(\mathbf{r}_N), \tag{2.2}$$

where $\phi_\alpha(\mathbf{r}) = \langle \mathbf{r}|\alpha\rangle$ are solutions of the single-particle Schrödinger equation

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + U(\mathbf{r}) \right) \phi_\alpha(\mathbf{r}) = E_\alpha \phi_\alpha(\mathbf{r}) \tag{2.3}$$

with energy E_α.[1] The total energy of the N-particle system is

$$E = E_{\alpha_1} + E_{\alpha_2} + \ldots + E_{\alpha_N}. \tag{2.4}$$

Identical Particles and Exchange
In the above considerations we have assumed that the particles are distinguishable, e.g. we constructed a solution where particle 1 is in state α_1, particle 2 in state α_2 and so on. In order to construct an N-particle wavefunction of *indistinguishable* quantum particles we need to sum over all possible particle permutations. We define the pair permutation operator as

$$\hat{P}_{ij}\psi(\ldots, \mathbf{r}_i, \ldots \mathbf{r}_j, \ldots) = \psi(\ldots, \mathbf{r}_j, \ldots \mathbf{r}_i, \ldots). \tag{2.5}$$

A general permutation $\hat{P}$ can be written as a product of pair permutations. We define $\mathrm{par}(\hat{P})$ as the number of pair permutations that is required to represent $\hat{P}$. The most fundamental way to define bosons and fermions is through the braiding of two particles. While for bosons the wavefunction remains invariant under the exchange of two particles, the fermionic wavefunction changes sign,

$$\begin{aligned} \hat{P}_{ij}\psi(\mathbf{r}_1, \ldots, \mathbf{r}_N) &= \pm\, \psi(\mathbf{r}_1, \ldots, \mathbf{r}_N), \\ &+ \quad \text{Bosons} \\ &- \quad \text{Fermions} \end{aligned}$$

This implies that for a general permutation $\hat{P}$ we obtain

$$\hat{P}\psi(\mathbf{r}_1, \ldots, \mathbf{r}_N) = (\pm 1)^{\mathrm{par}(\hat{P})}\, \psi(\mathbf{r}_1, \ldots, \mathbf{r}_N), \tag{2.6}$$

where we have used the same sign convention as above. Note that expectation values $\langle\psi|\hat{A}|\psi\rangle$ remain invariant under particle permutation $|\psi\rangle \to \hat{P}|\psi\rangle$. This of course implies that $|\psi\rangle$ and $\hat{P}|\psi\rangle$ have the same energy.

[1] We assume that the wavefunctions form an orthonormal set, $\langle\alpha|\beta\rangle = \frac{1}{V}\int d^d\mathbf{r}\phi_\alpha^*(\mathbf{r})\phi_\beta(\mathbf{r}) = \delta_{\alpha\beta}$.

Starting from the product solution (2.2) for distinguishable particles, it is straightforward to construct N-particle wavefunctions that are solutions of the N-particle Schrödinger equation and symmetric or anti-symmetric under particle exchange. Let us first illustrate this for $N = 2$ with particles in different single-particle eigenstates α_1, α_2. If $\phi_{\alpha_1}(\mathbf{r}_1)\phi_{\alpha_2}(\mathbf{r}_2)$ is a solution then of course also $\hat{P}_{1,2}\phi_{\alpha_1}(\mathbf{r}_1)\phi_{\alpha_2}(\mathbf{r}_2) = \phi_{\alpha_2}(\mathbf{r}_1)\phi_{\alpha_1}(\mathbf{r}_2)$ and both solutions have the same energy. Because of the linearity of the Schrödinger equation, any linear combination of the two solutions is a solution of the same energy. Defining

$$\psi_{B/F}(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{\sqrt{2}} \left(\phi_{\alpha_1}(\mathbf{r}_1)\phi_{\alpha_2}(\mathbf{r}_2) \pm \phi_{\alpha_2}(\mathbf{r}_1)\phi_{\alpha_1}(\mathbf{r}_2)\right) \tag{2.7}$$

we constructed a symmetric (ψ_B) and anti-symmetric (ψ_F) two-particle wavefunction, as required for indistinguishable bosons and fermions, respectively. We can generalise this construction easily to the N-particle case,

$$\psi_{B/F}(\mathbf{r}_1, \ldots, \mathbf{r}_N) = \mathcal{N} \sum_{\hat{P} \in S_N} (\pm 1)^{\mathrm{par}(\hat{P})}\ \phi_{\alpha_{P(1)}}(\mathbf{r}_1) \cdot \ldots \cdot \phi_{\alpha_{P(N)}}(\mathbf{r}_N), \tag{2.8}$$

where the sum is over the $N!$ elements of the permutation group S_N and $\mathcal{N}$ is a normalisation constant,[2] which is not very important.

Slater Determinants and Permanents
The wavefunction for fermions may be written in the following form:

$$\psi_F(\mathbf{r}_1, \ldots, \mathbf{r}_N) = \mathcal{N} \begin{vmatrix} \phi_{\alpha_1}(\mathbf{r}_1) & \ldots & \ldots & \phi_{\alpha_1}(\mathbf{r}_N) \\ \phi_{\alpha_2}(\mathbf{r}_1) & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ \phi_{\alpha_N}(\mathbf{r}_1) & \ldots & \ldots & \phi_{\alpha_N}(\mathbf{r}_N) \end{vmatrix}. \tag{2.9}$$

The alternating signs $(-1)^{\mathrm{par}(\hat{P})}$ required for fermions are taken care of by the minus signs coming from the determinant. Note that the determinant is zero if any two or more of the single-particle states are the same. The determinant — known as a *Slater determinant* — therefore properly encodes the Pauli principle. For bosons, we may construct a similar representation in terms of a permanent. This is essentially the same object as a determinant but without the minus signs.

2.2 Occupation Numbers and Fock Space

The states that we identified as many-particle basis states can be specified completely by the number of particles in each single-particle basis state. These are usually denoted

$$|n_1, n_2, \ ...\rangle,$$

where n_i is the number of quanta/particles in the state i. n_i takes values $0, 1$ for fermions and values $0, 1, 2, \ ..., \infty$ for bosons. This notation assumes: i. a particular identification

[2] For a state with n_1 particles in the $\alpha = 1$ state, n_2 particles in $\alpha = 2$ state and so on, we obtain $\mathcal{N} = 1/\sqrt{N!n_1!n_2! \cdot \ldots}$. Note that for fermions the occupation numbers can only take values $n_\alpha = 0, 1$, resulting in $\mathcal{N} = 1/\sqrt{N!}$.

of single-particle states. In condensed matter, we typically use either momentum or plane wave states, or an orthonormal set of states constructed from orbitals on particular atoms, *i.e* position states (known as Wannier states); of course, we could also use the single particle energy eigenstates $\{|\alpha\rangle\}$ but this is not a requirement. ii. that the states are appropriately symmetrized or anti-symmetrized for bosons or fermions, respectively.

Fock Space: is the set of states with all possible combinations of the occupation numbers.

The Vacuum State: is the state in which none of the particle states are occupied. It is written as $|0\rangle$ and normalized $\langle 0|0\rangle = 1$.

Comparison with the Harmonic Oscillator:
This structure is very similar to that of the eigenstates of the harmonic oscillator and harmonic chain considered previously. In the case of the harmonic chain, for example, a generic eigenstate could be written $|n_{k_1}, n_{k_2}, \ ...\rangle$ describing a state with n_{k_i} quanta in the k_i^{th} wavevector mode. These modes are bosons and are given the name phonons when applied to vibrational modes of a crystal lattice. We have been using Fock space all along to describe the harmonic oscillator and its derivatives.

2.3 Creation and Annihilation Operators

Since the structure of Fock space discussed above is so reminiscent of the state space of the harmonic oscillator, it is natural to expect that we can also find operators that are the analogue of the ladder/creation and annihilation operators. These operators will allow us to navigate Fock space and, moreover, to calculate properties of our many-body system without ever having to write down its wavefunction.

The basic notion is to identify creation operators $\hat{c}_\ell^\dagger$ that create a particle in the single-particle state $|\ell\rangle$. The type of single-particle states is not important for what follows. We only postulate that the states $\{|\ell\rangle|\ell = 1, 2, \ldots\}$ form an orthonormal set, $\langle m|\ell\rangle = \delta_{m,\ell}$. We will then manipulate the resulting states to determine the properties of the operators $\hat{c}_\ell^\dagger$. A many-particle state in Fock space can be constructed from the action of many such operators,

$$|n_1, n_2, \ldots\rangle = \frac{1}{n_1! n_2! \ldots}(\hat{c}_1^\dagger)^{n_1}(\hat{c}_2^\dagger)^{n_2} \ldots |0\rangle, \tag{2.10}$$

where $|0\rangle = |0, 0, \ldots\rangle$ denotes the vacuum state with no particles present. The single particle states are included in Fock space,

$$|\ell\rangle = \hat{c}_\ell^\dagger|0\rangle = |0, \ldots, 0, \underbrace{1}_{\ell}, 0, \ldots\rangle. \tag{2.11}$$

Hermitian Conjugation and Annihilation of the Vacuum
By definition, annihilation operators are given by the hermitian conjugate of the creation operators. Their action on the vacuum is as follows:

$$\hat{c}_\ell|0\rangle = 0 \quad \text{and} \quad \langle 0|\hat{c}_\ell^\dagger = 0.$$

Commutation and Anti-commutation
Considering the action of the permutation operator on a two-particle state, we have

$$\hat{c}_m^\dagger \hat{c}_\ell^\dagger |0\rangle = \pm \hat{c}_\ell^\dagger \hat{c}_m^\dagger |0\rangle$$

for bosons and fermions, respectively. This implies that

$$\begin{aligned} \left[\hat{c}_m^\dagger, \hat{c}_\ell^\dagger\right] &= 0 = \hat{c}_m^\dagger \hat{c}_\ell^\dagger - \hat{c}_\ell^\dagger \hat{c}_m^\dagger \quad \text{Bosons} \\ \left\{\hat{c}_m^\dagger, \hat{c}_\ell^\dagger\right\} &= 0 = \hat{c}_m^\dagger \hat{c}_\ell^\dagger + \hat{c}_\ell^\dagger \hat{c}_m^\dagger \quad \text{Fermions} \end{aligned} \tag{2.12}$$

and (after hermitian conjugation)

$$\begin{aligned} [\hat{c}_m, \hat{c}_\ell] &= 0 \quad \text{Bosons} \\ \{\hat{c}_m, \hat{c}_\ell\} &= 0 \quad \text{Fermions} \end{aligned} \tag{2.13}$$

As we will prove below (box), the creation and annihilation operators satisfy the (anti-) commutator relations

$$\begin{aligned} \left[\hat{c}_m, \hat{c}_\ell^\dagger\right] &= \delta_{m,\ell} \quad \text{Bosons} \\ \left\{\hat{c}_m, \hat{c}_\ell^\dagger\right\} &= \delta_{m,\ell} \quad \text{Fermions} \end{aligned} \tag{2.14}$$

Before proving Eq. (2.14), we need to understand the action of the creation and annihilation operators on a general state $|n_1, n_2, \ldots\rangle$ in Fock space. Using Eq. (2.12), we obtain

$$\hat{c}_\ell^\dagger |n_1, n_2, \ldots, n_\ell, \ldots\rangle = (\pm 1)^{n_1+n_2+\ldots+n_{\ell-1}} \sqrt{n_\ell + 1} |n_1, n_2, \ldots, n_\ell + 1, \ldots\rangle, \tag{2.15}$$

where the negative signs in the fermionic case come from commuting $\hat{c}_\ell^\dagger$ through all of the creation operators with labels $m = 1, \ldots, \ell - 1$. Note also that for fermions $(\hat{c}_\ell^\dagger)^2 = 0$ which implies that the right hand side of the above expression is zero when $n_\ell = 1$.
In order to compute the action of the annihilation operator, we evaluate

$$\begin{aligned} \langle \ldots, n_\ell - 1, \ldots | \hat{c}_\ell | \ldots, n_\ell, \ldots\rangle &= \left(\langle \ldots, n_\ell, \ldots | \hat{c}_\ell^\dagger | \ldots, n_\ell - 1, \ldots\rangle\right)^* \\ &\overset{(2.15)}{=} (\pm 1)^{n_1+n_2+\ldots+n_{\ell-1}} \sqrt{n_\ell - 1 + 1}, \end{aligned}$$

from which it immediately follows that

$$\hat{c}_\ell |n_1, n_2, \ldots, n_\ell, \ldots\rangle = (\pm 1)^{n_1+n_2+\ldots+n_{\ell-1}} \sqrt{n_\ell} |n_1, n_2, \ldots, n_\ell - 1, \ldots\rangle, \tag{2.16}$$

This is zero for both bosons and fermions if $n_\ell = 0$.

From Eqs. (2.15) and (2.16) we deduce that

$$\hat{c}_\ell^\dagger \hat{c}_\ell |n_1, n_2, \ldots, n_\ell, \ldots\rangle = n_\ell |n_1, n_2, \ldots, n_\ell - 1, \ldots\rangle, \tag{2.17}$$

as expected for the number operator $\hat{n}_\ell = \hat{c}_\ell^\dagger \hat{c}_\ell$.

Proof of commutator relation $\left[\hat{c}_m, \hat{c}_\ell^\dagger\right] = \delta_{m,\ell}$ for bosons and anti-commutator relation $\left\{\hat{c}_m, \hat{c}_\ell^\dagger\right\} = \delta_{m,\ell}$ for fermions

Let us first assume $\ell < m$ and act with $\hat{c}_\ell^\dagger \hat{c}_m$ and $\hat{c}_m \hat{c}_\ell^\dagger$ on a general Fock-space state $|n_1, n_2, \ldots\rangle$:

$$\begin{aligned}
\hat{c}_\ell^\dagger \hat{c}_m |\ldots, n_\ell, \ldots, n_m, \ldots\rangle &= \sqrt{n_m}(\pm 1)^{n_1+\ldots+n_{m-1}} \hat{c}_\ell^\dagger |\ldots, n_\ell, \ldots, n_m - 1, \ldots\rangle \\
&= \sqrt{n_m}\sqrt{n_\ell + 1}(\pm 1)^{n_\ell+\ldots+n_{m-1}} |\ldots, n_\ell + 1, \ldots, n_m - 1, \ldots\rangle \\
\hat{c}_m \hat{c}_\ell^\dagger |\ldots, n_\ell, \ldots, n_m, \ldots\rangle &= \sqrt{n_\ell + 1}(\pm 1)^{n_1+\ldots+n_{\ell-1}} \hat{c}_m |\ldots, n_\ell + 1, \ldots, n_m, \ldots\rangle \\
&= \sqrt{n_m}\sqrt{n_\ell + 1}(\pm 1)^{1+n_\ell+\ldots+n_{m-1}} |\ldots, n_\ell + 1, \ldots, n_m - 1, \ldots\rangle \\
&= \pm \hat{c}_\ell^\dagger \hat{c}_m |\ldots, n_\ell, \ldots, n_m, \ldots\rangle,
\end{aligned}$$

from which it follows that $\hat{c}_m \hat{c}_\ell^\dagger \mp \hat{c}_\ell^\dagger \hat{c}_m = 0$, where the minus sign is for bosons, the plus sign for fermions. This oncludes the proof for $\ell < m$. The case $\ell > m$ simply follows from hermitian conjugation.
Let us now check the case $m = \ell$. Again, we act with $\hat{c}_\ell^\dagger \hat{c}_\ell$ and $\hat{c}_\ell \hat{c}_\ell^\dagger$ on a general state. In the case of bosons we obtain

$$\begin{aligned}
\hat{c}_\ell^\dagger \hat{c}_\ell |\ldots, n_\ell, \ldots\rangle &= \sqrt{n_\ell} \hat{c}_\ell^\dagger |\ldots, n_\ell - 1, \ldots\rangle = n_\ell |\ldots, n_\ell, \ldots\rangle \\
\hat{c}_\ell \hat{c}_\ell^\dagger |\ldots, n_\ell, \ldots\rangle &= \sqrt{n_\ell + 1} \hat{c}_\ell |\ldots, n_\ell + 1, \ldots\rangle = (n_\ell + 1) |\ldots, n_\ell, \ldots\rangle.
\end{aligned}$$

From this it follows that $(\hat{c}_\ell \hat{c}_\ell^\dagger - \hat{c}_\ell^\dagger \hat{c}_\ell)|\ldots, n_\ell, \ldots\rangle = |\ldots, n_\ell, \ldots\rangle$ and hence $[\hat{c}_\ell, \hat{c}_\ell^\dagger] = 1$. For fermions we need to be careful with the minus signs and make sure we satisfy Pauli's exclusion principle:

$$\begin{aligned}
\hat{c}_\ell^\dagger \hat{c}_\ell |\ldots, n_\ell, \ldots\rangle &= n_\ell |\ldots, n_\ell, \ldots\rangle. \\
\hat{c}_\ell \hat{c}_\ell^\dagger |\ldots, n_\ell, \ldots\rangle &= \begin{cases} (-1)^{n_1+\ldots+n_{\ell-1}} \hat{c}_\ell |\ldots, 1, \ldots\rangle & \text{for } n_\ell = 0 \\ 0 & \text{for } n_\ell = 1 \end{cases} \\
&= \begin{cases} |\ldots, 0, \ldots\rangle & \text{for } n_\ell = 0 \\ 0 & \text{for } n_\ell = 1 \end{cases}
\end{aligned}$$

From this we obtain

$$\begin{aligned}
(\hat{c}_\ell \hat{c}_\ell^\dagger + \hat{c}_\ell^\dagger \hat{c}_\ell)|\ldots, n_\ell, \ldots\rangle &= \begin{cases} |\ldots, 0, \ldots\rangle & \text{for } n_\ell = 0 \\ |\ldots, 1, \ldots\rangle & \text{for } n_\ell = 1 \end{cases} \\
&= |\ldots, n_\ell, \ldots\rangle,
\end{aligned}$$

from what follows that $\{\hat{c}_\ell, \hat{c}_\ell^\dagger\} = 1$.

2.4 Transformation Between Bases

It is often useful to address different physical questions with reference to different single particle bases (such as the position and momentum bases). We need then to understand how to transform our creation and annihilation operators between such bases. We have already seen an example of this for the harmonic chain. Let us think about this more generally for a moment.

2.4.1 The Transformation

Consider two sets of single-particle states, $\{|\ell\rangle\}$ and $\{|\alpha\rangle\}$, which we assume to be orthonormal,

$$\langle \ell|\ell'\rangle = \delta_{\ell\ell'} \quad \text{and} \quad \langle \alpha|\alpha'\rangle = \delta_{\alpha\alpha'}. \tag{2.18}$$

In terms of creation operators the single-particle states can be written as

$$|\ell\rangle = \hat{c}_\ell^\dagger|0\rangle \quad \text{and} \quad |\alpha\rangle = \hat{d}_\alpha^\dagger|0\rangle. \tag{2.19}$$

The basis transformation between the two sets is achieved by a matrix $\mathbf{U}$,

$$|\alpha\rangle = \sum_\ell U_{\alpha\ell}|\ell\rangle, \tag{2.20}$$

where the matrix elements are given by the overlap between states, $U_{\alpha\ell} = \langle \ell|\alpha\rangle$. Orthonormality and completeness of the basis sets implies that the transformation matrix is unitary, $\mathbf{U}\mathbf{U}^\dagger = \mathbf{1}$,

$$\begin{aligned}(\mathbf{U}\mathbf{U}^\dagger)_{\alpha\alpha'} &= \sum_\ell (\mathbf{U})_{\alpha\ell}(\mathbf{U}^\dagger)_{\ell\alpha'} = \sum_\ell U_{\alpha\ell}U^*_{\alpha'\ell} = \sum_\ell \langle \ell|\alpha\rangle\langle \ell|\alpha'\rangle^* \\ &= \sum_\ell \langle \ell|\alpha\rangle\langle \alpha'|\ell\rangle = \sum_\ell \langle \alpha'|\ell\rangle\langle \ell|\alpha\rangle = \langle \alpha'|\underbrace{(\sum_\ell |\ell\rangle\langle \ell|)}_{=\mathbf{1}}|\alpha\rangle = \langle \alpha'|\alpha\rangle = \delta_{\alpha\alpha'}\end{aligned}$$

Next, let us determine the corresponding transformation rule for the creation and annihilation operators. From

$$\hat{d}_\alpha^\dagger|0\rangle = |\alpha\rangle = \sum_\ell U_{\alpha\ell}|\ell\rangle = \sum_\ell U_{\alpha\ell}\hat{c}_\ell^\dagger|0\rangle$$

we obtain

$$\hat{d}_\alpha^\dagger = \sum_\ell U_{\alpha\ell}\hat{c}_\ell^\dagger, \tag{2.21}$$

and by hermitian conjugation

$$\hat{d}_\alpha = \sum_\ell U^*_{\alpha\ell}\hat{c}_\ell = \sum_\ell \hat{c}_\ell(\mathbf{U}^\dagger)_{\ell\alpha}. \tag{2.22}$$

It can be shown (homework problem) that the bosonic commutation relations and the fermionic anti-commutation relations are preserved under unitary transformations of the corresponding single-particle basis sets. The invariance of the commutation relations between position and wavevector basis that we noted in the case of the harmonic chain is just a special case of the invariance under general unitary transformations noted here.

2.4.2 Transforming Between Momentum and Position Bases

An important example of two basis sets of single-particle states are position states, $|\mathbf{r}\rangle = \hat{c}^\dagger_{\mathbf{r}}|0\rangle$, and momentum states, $|\mathbf{k}\rangle = \hat{d}^\dagger_{\mathbf{k}}|0\rangle$. In this section, we will show that the unitary transformation between the position and momentum single-particle states is nothing but the conventional Fourier transform already used in Chapter.[3]

Using the general notation from the previous section and assuming a continuous d-dimensional position space, we can write the unitary transformation as

$$\hat{d}^\dagger_{\mathbf{k}} = \int d^d\mathbf{r} U_{\mathbf{k},\mathbf{r}} \hat{c}^\dagger_{\mathbf{r}},$$

where the matrix elements of the unitary transformation are given by the overlap of single-particle position and momentum eigenstates, $U_{\mathbf{k},\mathbf{r}} = \langle \mathbf{r}|\mathbf{k}\rangle$. This overlap is simply the normalised wavefunction $\psi_{\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{V}} e^{i\mathbf{k}\mathbf{r}}$ of a plane-wave state. Note that in order to normalise the plane wave state we had to assume a finite d-dimensional volume, $V = L^d$. We therefore obtain

$$\hat{d}^\dagger_{\mathbf{k}} = \frac{1}{\sqrt{V}} \int_V d^d\mathbf{r} e^{i\mathbf{k}\mathbf{r}} \hat{c}^\dagger_{\mathbf{r}} \quad \text{and} \quad \hat{d}_{\mathbf{k}} = \frac{1}{\sqrt{V}} \int_V d^d\mathbf{r} e^{-i\mathbf{k}\mathbf{r}} \hat{c}_{\mathbf{r}} \tag{2.23}$$

which is simply a Fourier transform.[4] Periodic boundary conditions require that the momenta are on a grid of spacing $\Delta k = 2\pi/L$:

$$e^{i\mathbf{k}\mathbf{r}} = e^{i\mathbf{k}(\mathbf{r}+L\hat{\mathbf{e}}_j)} \Rightarrow e^{iLk_j} = 1 \Rightarrow k_j = \frac{2\pi}{L} n_j \ (n_j \in \mathbb{Z}) \Rightarrow \mathbf{k} = \frac{2\pi}{L}(n_1, \ldots, n_d)$$

The inverse transformation is given by

$$\hat{c}^\dagger_{\mathbf{r}} = \frac{1}{\sqrt{V}} \sum_{\mathbf{k}} e^{-i\mathbf{k}\mathbf{r}} \hat{d}^\dagger_{\mathbf{k}} \quad \text{and} \quad \hat{c}_{\mathbf{r}} = \frac{1}{\sqrt{V}} \sum_{\mathbf{k}} e^{i\mathbf{k}\mathbf{r}} \hat{d}_{\mathbf{k}}, \tag{2.24}$$

and the condition that the transformation be unitary simply reduces to the integral representation of the δ-function:

$$\delta_{\mathbf{k},\mathbf{q}} = \int d^d\mathbf{r} U_{\mathbf{k},\mathbf{r}} U^\dagger_{\mathbf{r},\mathbf{q}} = \frac{1}{V} \int d^d\mathbf{r} e^{i(\mathbf{k}-\mathbf{q})\mathbf{r}}. \tag{2.25}$$

Fourier transform on the lattice

For comparison, we list the corresponding equations for the case of a d dimensional lattice of N sites ($N = M^d$) with periodic boundary conditions and lattice constant a. The momenta can take values $k_i = \frac{2\pi}{Ma} n_i$ with $n_i = 1, \ldots, M$.

$$\hat{c}^\dagger_{\mathbf{k}} = \frac{1}{\sqrt{N}} \sum_{\mathbf{r}} e^{i\mathbf{k}\mathbf{r}} \hat{c}^\dagger_{\mathbf{r}} \quad \text{and} \quad \hat{c}_{\mathbf{k}} = \frac{1}{\sqrt{N}} \sum_{\mathbf{r}} e^{-i\mathbf{k}\mathbf{r}} \hat{c}_{\mathbf{r}}, \tag{2.26}$$

with inverse

$$\hat{c}^\dagger_{\mathbf{r}} = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} e^{-i\mathbf{k}\mathbf{r}} \hat{c}^\dagger_{\mathbf{k}} \quad \text{and} \quad \hat{c}_{\mathbf{r}} = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} e^{i\mathbf{k}\mathbf{r}} \hat{c}_{\mathbf{k}}, \tag{2.27}$$

and resolution of the delta function $\delta_{\mathbf{k},\mathbf{q}} = \frac{1}{N} \sum_{\mathbf{r}} e^{i(\mathbf{k}-\mathbf{q})\mathbf{r}}$.

[3] Note that we often use momentum $\mathbf{p}$ and wavevector $\mathbf{k}$ interchangeably. However, strictly speaking $\mathbf{p} = \hbar\mathbf{k}$. In theoretical physics we often work in units $\hbar = 1$.

[4] Note that often we use the same symbols for operators in real and momentum space.

2.5 Single- and Two-Particle Operators

2.5.1 Single-Particle Operators

We are going to build up gradually to a way of writing general single particle operators in a second quantised form - *i.e.* in terms of creation and annihilation operators. Single particle operators are completely characterised by their action on single-particle states. When acting on an N-particle state, single particle operators take the form

$$\hat{\Omega} = \sum_{i=1}^{N} \hat{\omega}_i, \tag{2.28}$$

where ω_i acts on particle i. Examples are the kinetic and potential energy operators,

$$\hat{T} = \sum_i \frac{\hat{\mathbf{p}}^2}{2m}, \quad \hat{U} = \sum_i U(\hat{\mathbf{r}}_i).$$

The operators $\hat{\omega}_i$ are of exactly the same form for all i and we could therefore express $\hat{\Omega}$ as a tensor product,

$$\hat{\Omega} = \sum_{i=1}^{N} \underbrace{\mathbf{1} \otimes \ldots \otimes \mathbf{1}}_{i-1} \otimes \hat{\omega} \otimes \mathbf{1} \otimes \ldots \otimes \mathbf{1}.$$

Let us first assume that the single-particle basis $\{|\ell\rangle\}$ is formed of eigenstates of $\hat{\omega}$, $\hat{\omega}|\ell\rangle = \omega_\ell|\ell\rangle$. Acting with $\hat{\Omega}$ on N particles in states $\ell_1, \ell_2, \ldots, \ell_N$ we obtain

$$\hat{\Omega}|\ell_1, \ell_2, \ldots, \ell_N\rangle = \left(\sum_{i=1}^{N} \omega_{\ell_i}\right) |\ell_1, \ell_2, \ldots, \ell_N\rangle. \tag{2.29}$$

We now consider a state in Fock space with n_1 particles in state $\ell = 1$, n_2 particles in state $\ell = 2$, and so on. This state can be identified with an $N = n_1 + n_2 + \ldots$ particle state,

$$|n_1, n_2, \ldots\rangle = |\underbrace{1, \ldots, 1}_{n_1}, \underbrace{2, \ldots, 2}_{n_2}, 3, \ldots\rangle,$$

from which it immediately follows that

$$\hat{\Omega}|n_1, n_2, \ldots\rangle = \left(\sum_\ell \omega_\ell n_\ell\right) |n_1, n_2, \ldots\rangle. \tag{2.30}$$

We therefore obtain the form of $\hat{\Omega}$ in second quantisation in terms of eigenstates of ω,

$$\hat{\Omega} = \sum_\ell \omega_\ell \hat{n}_\ell = \sum_\ell \omega_\ell \hat{c}_\ell^\dagger \hat{c}_\ell. \tag{2.31}$$

In order to obtain the form of $\hat{\Omega}$ in a general basis we perform a unitary transformation $\{|\ell\rangle\} \to \{|\alpha\rangle\}$,

$$\hat{c}_\ell^\dagger = \sum_\alpha U_{\ell\alpha} \hat{d}_\alpha^\dagger = \sum_\alpha \langle\alpha|\ell\rangle \hat{d}_\alpha^\dagger,$$

and use that $\langle \ell|\hat{\omega}|\ell'\rangle = \omega_\ell \delta_{\ell\ell'}$:

$$\begin{aligned}
\hat{\Omega} &= \sum_{\ell\ell'} \langle \ell|\hat{\omega}|\ell'\rangle \hat{c}_\ell^\dagger \hat{c}_{\ell'} \\
&= \sum_{\ell\ell'} \langle \ell|\hat{\omega}|\ell'\rangle \sum_\alpha \langle \alpha|\ell\rangle \hat{d}_\alpha^\dagger \sum_{\alpha'} \langle \alpha'|\ell'\rangle^* \hat{d}_{\alpha'} \\
&= \sum_{\ell\ell'} \sum_{\alpha\alpha'} \langle \alpha|\ell\rangle \langle \ell|\hat{\omega}|\ell'\rangle \langle \ell'|\hat{\omega}|\alpha'\rangle \hat{d}_\alpha^\dagger \hat{d}_{\alpha'} \\
&= \sum_{\alpha\alpha'} \langle \alpha|\hat{\omega}|\alpha'\rangle \hat{d}_\alpha^\dagger \hat{d}_{\alpha'}. \qquad (2.32)
\end{aligned}$$

This is the 2nd quantised form of $\hat{\Omega}$ in a general single particle basis $\{|\alpha\rangle\}$, in which $\hat{\omega}$ is not diagonal and has matrix elements $\langle \alpha|\hat{\omega}|\alpha'\rangle = \omega_{\alpha\alpha'}$.
As an important example we will express the non-interacting N-particle Hamiltonian

$$\hat{\mathcal{H}} = \hat{T} + \hat{U} = \sum_i \frac{\hat{\mathbf{p}}_i^2}{2m} + \sum_i U(\hat{\mathbf{r}}_i), \qquad (2.33)$$

in terms of creation and annihilation operators $\hat{c}_{\mathbf{k}}^\dagger$, $\hat{c}_{\mathbf{k}}$ in momentum space. As you will show in a homework problem, the 2nd quantised Hamiltonian is given by

$$\hat{\mathcal{H}} = \sum_{\mathbf{k}} \frac{\hbar^2 \mathbf{k}^2}{2m} \hat{c}_{\mathbf{k}}^\dagger \hat{c}_{\mathbf{k}} + \sum_{\mathbf{k},\mathbf{k}'} \tilde{U}(\mathbf{k}' - \mathbf{k}) \hat{c}_{\mathbf{k}}^\dagger \hat{c}_{\mathbf{k}'}, \qquad (2.34)$$

where $\tilde{U}(\mathbf{q}) = \frac{1}{V} \int d^d\mathbf{r}\, U(\mathbf{r}) e^{i\mathbf{q}\mathbf{r}}$ is the Fourier transform of the potential $U(\mathbf{r})$.

2.5.2 Two-Particle Operators

Two-body operators, such as the interaction potential

$$\hat{\mathcal{H}}_{\text{int}} = \frac{1}{2} \sum_{i,j=1}^{N} V(\hat{\mathbf{r}}_i - \hat{\mathbf{r}}_j), \qquad (2.35)$$

act on pairs of particles. In general, we can write such two-body operators as

$$\hat{\Theta} = \frac{1}{2} \sum_{i,j} \hat{\theta}_{ij}. \qquad (2.36)$$

Expressing two-body operators in 2nd quantisation follows essentially the same steps as in the case of one-body operators. However, the calculation is more tedious. Here we only give the answer, which is of the expected form:

$$\hat{\Theta} = \frac{1}{2} \sum_{\ell\ell' mm'} \langle \ell\ell'|\hat{\theta}|mm'\rangle \hat{c}_\ell^\dagger \hat{c}_{\ell'}^\dagger \hat{c}_{m'} \hat{c}_m, \qquad (2.37)$$

where $|mm'\rangle$ denotes the state of 2 particles in single particle states m and m'.

Let us return to the generic 2-body interaction $\hat{\mathcal{H}}_{\rm int}$, which becomes very simple in terms of single particle creation and annihilation operators in position space, $\hat{c}^\dagger_{\mathbf{r}}$, $\hat{c}_{\mathbf{r}}$. This is because

$$\langle \mathbf{r}, \mathbf{r}' | V(\hat{\mathbf{r}} - \hat{\mathbf{r}}') | \tilde{\mathbf{r}}, \tilde{\mathbf{r}}' \rangle = V(\mathbf{r} - \mathbf{r}')\delta(\mathbf{r} - \tilde{\mathbf{r}})\delta(\mathbf{r}' - \tilde{\mathbf{r}}'). \tag{2.38}$$

We therefore obtain

$$\hat{\mathcal{H}}_{\rm int} = \frac{1}{2}\int d^d\mathbf{r} \int d^d\mathbf{r}' V(\mathbf{r} - \mathbf{r}')\hat{c}^\dagger_{\mathbf{r}}\hat{c}^\dagger_{\mathbf{r}'}\hat{c}_{\mathbf{r}'}\hat{c}_{\mathbf{r}} = \frac{1}{2}\int d^d\mathbf{r} \int d^d\mathbf{r}' V(\mathbf{r} - \mathbf{r}')\hat{n}_{\mathbf{r}}\hat{n}_{\mathbf{r}'}, \tag{2.39}$$

where in the last step we have used that for $\mathbf{r} \neq \mathbf{r}'$

$$\hat{c}^\dagger_{\mathbf{r}}\hat{c}^\dagger_{\mathbf{r}'}\hat{c}_{\mathbf{r}'}\hat{c}_{\mathbf{r}} = \pm\hat{c}^\dagger_{\mathbf{r}}\hat{c}^\dagger_{\mathbf{r}'}\hat{c}_{\mathbf{r}}\hat{c}_{\mathbf{r}'} = (\pm 1)^2\hat{c}^\dagger_{\mathbf{r}}\hat{c}_{\mathbf{r}}\hat{c}^\dagger_{\mathbf{r}'}\hat{c}_{\mathbf{r}'}.$$

2.6 Diagonalising Quantum Hamiltonians

The aim in using second quantization techniques is to reduce the expectation values that we are interested in — as far as possible — to number operators. As we saw in the case of the harmonic oscillator, it is then a simple matter to calculate the physical properties. Single particle operators can always be written in this form. Higher order many-particle operators can be expanded in terms of single-particle operators using mean-field theory and perturbative extensions (not part of this course).

2.6.1 Unitary transformations

Non-interacting Hamiltonians that conserve particle number can always be written as

$$\hat{\mathcal{H}} = \sum_{ij} \mathcal{H}_{ij}\hat{a}^\dagger_i\hat{a}_j = (\hat{a}^\dagger_1, \ldots, \hat{a}^\dagger_n)\mathbf{H}\begin{pmatrix} \hat{a}_1 \\ \vdots \\ \hat{a}_n \end{pmatrix} = \hat{\mathbf{a}}^\dagger\mathbf{H}\hat{\mathbf{a}}, \tag{2.40}$$

where $\hat{a}^\dagger_i$, $\hat{a}_i$ could be bosonic or fermionic and i labels single-particle basis states $|i\rangle = \hat{a}^\dagger_i|0\rangle$. The matrix $\mathbf{H}$ is hermitian, $\mathbf{H}^\dagger = \mathbf{H}$. It can therefore be diagonalised by a unitary transformation $\mathbf{U}$ ($\mathbf{U}^\dagger\mathbf{U} = \mathbf{U}\mathbf{U}^\dagger = \mathbf{1}$),

$$\mathbf{U}^\dagger\mathbf{H}\mathbf{U} = \begin{pmatrix} \epsilon_1 & & \\ & \ddots & \\ & & \epsilon_n \end{pmatrix}, \tag{2.41}$$

where ϵ_i are the energy eigenvalues. The Hamiltonian transforms accordingly,

$$\hat{\mathbf{a}}^\dagger\mathbf{H}\hat{\mathbf{a}} = \underbrace{\hat{\mathbf{a}}^\dagger\mathbf{U}}_{\hat{\mathbf{d}}^\dagger}\begin{pmatrix} \epsilon_1 & & \\ & \ddots & \\ & & \epsilon_n \end{pmatrix}\underbrace{\mathbf{U}^\dagger\hat{\mathbf{a}}}_{\hat{\mathbf{d}}} = \sum_\ell \epsilon_\ell\hat{d}^\dagger_\ell\hat{d}_\ell. \tag{2.42}$$

The creation and annihilation operators transform as $\hat{\mathbf{d}}^\dagger = \hat{\mathbf{a}}^\dagger \mathbf{U}$ and $\hat{\mathbf{d}} = \mathbf{U}^\dagger \hat{\mathbf{a}}$ with inverse $\hat{\mathbf{a}}^\dagger = \hat{\mathbf{d}}^\dagger \mathbf{U}^\dagger$ and $\hat{\mathbf{a}} = \mathbf{U}\hat{\mathbf{d}}$. The bosonic commutation and fermionic anti-commutation relations are preserved under such unitary transformations.

Solving a problem defined by a particular $\hat{\mathcal{H}}$ amounts to finding the unitary rotation that diagonalises it. This reduces to a simple matrix diagonalisation and hence a simple linear algebra problem: the energy eigenvalues are obtained from solving the equation $\det(\mathbf{H} - \epsilon \mathbf{1}) = 0$, the columns of the unitary transformation matrix $\mathbf{U}$ are given by the corresponding eigenvectors. As we saw for the harmonic chain, often a Fourier transform diagonalises the Hamiltonian. We will see a number of examples in this course where there is a residual diagonalisation required - often of just a 2×2 or 4×4 matrix.

2.6.2 Bogoliubov Transformation

There are lots of physical systems for which a mean-field treatment of the interactions between particles leads to *anomalous* terms $\hat{c}_i^\dagger \hat{c}_j^\dagger$ and $\hat{c}_i \hat{c}_j$ in the Hamiltonian. We will encounter this in the case of Bose condensates (chapter 3) and superconductors (chapter 7). Anomalous terms also appear in the linear spin-wave theory of quantum antiferromagnets (chapter 4).

In the presence of anomalous terms, $\hat{\mathcal{H}}$ does not conserve the particle number and the ground state of the system will be formed by a superposition of states with different particle numbers. This is one of the main reasons why we use Fock space. In this course, we will encounter fermionic and bosonic Hamiltonians which are of the form

$$\hat{\mathcal{H}} = \epsilon_1 \hat{c}_1^\dagger \hat{c}_1 + \epsilon_2 \hat{c}_2^\dagger \hat{c}_2 + \lambda \hat{c}_1^\dagger \hat{c}_2^\dagger + \lambda^* \hat{c}_2 \hat{c}_1,$$

where the operators have additional momentum dependence, which we suppress here, for simplicity. We will also assume $\epsilon_1 = \epsilon_2$ since this will be the case in later applications, and $\lambda \in \mathbb{R}$, for reasons we will explain later.

A unitary transformation cannot be used to diagonalise these Hamiltonians. Instead, one must use a Bogoliubov transformation. This takes a slightly different form for bosons and fermions. The procedure is very similar in the two cases with some additional negative signs in the latter case, due to the fermionic anti-commutation relations.

Bosonic Bogoliubov
The bosonic form of the Bogoliubov transformation is used for example for Bose condensates and anti-ferromagnets. We can write the Hamiltonian in a 2×2 matrix structure,

$$\hat{\mathcal{H}} = \epsilon(\hat{c}_1^\dagger \hat{c}_1 + \underbrace{\hat{c}_2^\dagger \hat{c}_2}_{=\hat{c}_2 \hat{c}_2^\dagger - 1}) + \lambda(\hat{c}_1^\dagger \hat{c}_2^\dagger + \hat{c}_2 \hat{c}_1) = (\hat{c}_1^\dagger, \hat{c}_2) \begin{pmatrix} \epsilon & \lambda \\ \lambda & \epsilon \end{pmatrix} \begin{pmatrix} \hat{c}_1 \\ \hat{c}_2^\dagger \end{pmatrix} - \epsilon. \tag{2.43}$$

The operators

$$\hat{\psi} = \begin{pmatrix} \hat{c}_1 \\ \hat{c}_2^\dagger \end{pmatrix} \quad \text{and} \quad \hat{\psi}^\dagger = (\hat{c}_1^\dagger, \hat{c}_2) \tag{2.44}$$

are called Nambu spinors. The aim is to transform to a new set of boson operators $\hat{d}_i^\dagger$, $\hat{d}_i$ $(i = 1, 2)$ such that the Hamiltonian becomes diagonal in Nambu space (no longer contains anomalous terms). This can be achieved by a transformation

$$\begin{pmatrix} \hat{c}_1 \\ \hat{c}_2^\dagger \end{pmatrix} = \begin{pmatrix} u & v \\ v & u \end{pmatrix} \begin{pmatrix} \hat{d}_1 \\ \hat{d}_2^\dagger \end{pmatrix}, \tag{2.45}$$

with $u, v \in \mathbb{R}$. By hermitian conjugation this implies

$$(\hat{c}_1^\dagger, \hat{c}_2) = (\hat{d}_1^\dagger, \hat{d}_2) \begin{pmatrix} u & v \\ v & u \end{pmatrix}. \tag{2.46}$$

Importantly, the bosonic commutation relations for the two sets of operators impose the constraint

$$u^2 - v^2 = 1 \tag{2.47}$$

on the coefficients of the transformation matrix. This can be seen from the following calculation,

$$1 = [\hat{c}_1, \hat{c}_1^\dagger] = [u\hat{d}_1 + v\hat{d}_2^\dagger, u\hat{d}_1^\dagger + v\hat{d}_2] = u^2 \underbrace{[\hat{d}_1, \hat{d}_1^\dagger]}_{=1} + v^2 \underbrace{[\hat{d}_2^\dagger, \hat{d}_2]}_{=-1} = u^2 - v^2.$$

We can use the transformation to express the Hamiltonian in terms of the new set of boson operators,

$$\begin{aligned} \hat{\mathcal{H}} &= (\hat{c}_1^\dagger, \hat{c}_2) \begin{pmatrix} \epsilon & \lambda \\ \lambda & \epsilon \end{pmatrix} \begin{pmatrix} \hat{c}_1 \\ \hat{c}_2^\dagger \end{pmatrix} - \epsilon \\ &= (\hat{d}_1^\dagger, \hat{d}_2) \begin{pmatrix} u & v \\ v & u \end{pmatrix} \begin{pmatrix} \epsilon & \lambda \\ \lambda & \epsilon \end{pmatrix} \begin{pmatrix} u & v \\ v & u \end{pmatrix} \begin{pmatrix} \hat{d}_1 \\ \hat{d}_2^\dagger \end{pmatrix} - \epsilon \\ &= (\hat{d}_1^\dagger, \hat{d}_2) \begin{pmatrix} (u^2+v^2)\epsilon + 2uv\lambda & 2uv\epsilon + (u^2+v^2)\lambda \\ 2uv\epsilon + (u^2+v^2)\lambda & (u^2+v^2)\epsilon + 2uv\lambda \end{pmatrix} \begin{pmatrix} \hat{d}_1 \\ \hat{d}_2^\dagger \end{pmatrix} - \epsilon. \end{aligned} \tag{2.48}$$

We can now determine u and v from the condition that the anomalous terms vanish,

$$2uv\epsilon + (u^2 + v^2)\lambda = 0. \tag{2.49}$$

Because of the constraint, $u^2 - v^2 = 1$, u and v are not independent. Rather than using the constraint to express v in terms of u, we parametrise

$$u = \cosh(\theta) \quad \text{and} \quad v = \sinh(\theta), \tag{2.50}$$

which satisfies the constraint. Inserting into Eq. (2.49), we obtain the condition for θ that diagonalises $\hat{\mathcal{H}}$,

$$\sinh(2\theta)\epsilon + \cosh(2\theta) = 0 \Rightarrow \tanh(2\theta) = -\frac{\lambda}{\epsilon}. \tag{2.51}$$

This brings the Hamiltonian to the final diagonal form

$$\begin{aligned} \hat{\mathcal{H}} &= (\hat{d}_1^\dagger, \hat{d}_2) \begin{pmatrix} \tilde{\epsilon} & 0 \\ 0 & \tilde{\epsilon} \end{pmatrix} \begin{pmatrix} \hat{d}_1 \\ \hat{d}_2^\dagger \end{pmatrix} - \epsilon \\ &= \tilde{\epsilon}(\hat{d}_1^\dagger \hat{d}_1 + \hat{d}_2^\dagger \hat{d}_2) - \epsilon + \tilde{\epsilon}, \end{aligned} \tag{2.52}$$

where the energy of the Bogoliubov quasiparticles is given by

$$\begin{aligned}\tilde{\epsilon} &= (u^2+v^2)\epsilon + 2uv\lambda = \epsilon\cosh(2\theta) + \lambda\sinh(2\theta) = \epsilon\cosh(2\theta)[1+\frac{\lambda}{\epsilon}\tanh(2\theta)] \\ &= \frac{\epsilon}{\sqrt{1-\tanh^2(2\theta)}}[1+\frac{\lambda}{\epsilon}\tanh(2\theta)] = \frac{\epsilon}{\sqrt{1-\lambda^2/\epsilon^2}}(1-\lambda^2/\epsilon^2) \\ &= \sqrt{\epsilon^2-\lambda^2}. \end{aligned} \tag{2.53}$$

Fermionic Bogoliubov
This is used for example for superconductivity and superfluidity of fermions (*e.g.* in ^3He). The steps are very similar to those for bosons, but crucially, negative signs from the anti-commutation of fermionic operators change things somewhat. We first write the Hamiltonian in Nambu spinor representation:

$$\hat{\mathcal{H}} = \epsilon(\hat{c}_1^\dagger\hat{c}_1 + \underbrace{\hat{c}_2^\dagger\hat{c}_2}_{=-\hat{c}_2\hat{c}_2^\dagger+1}) + \lambda(\hat{c}_1^\dagger\hat{c}_2^\dagger + \hat{c}_2\hat{c}_1) = (\hat{c}_1^\dagger, \hat{c}_2)\begin{pmatrix}\epsilon & \lambda\\ \lambda & -\epsilon\end{pmatrix}\begin{pmatrix}\hat{c}_1\\ \hat{c}_2^\dagger\end{pmatrix} + \epsilon, \tag{2.54}$$

where the opposite signs obtained for fermions are highlighted in red. The fermionic Bogoliubov transformation is achieved by

$$\begin{pmatrix}\hat{c}_1\\ \hat{c}_2^\dagger\end{pmatrix} = \begin{pmatrix}u & v\\ -v & u\end{pmatrix}\begin{pmatrix}\hat{d}_1\\ \hat{d}_2^\dagger\end{pmatrix}, \tag{2.55}$$

with $u, v \in \mathbb{R}$. By hermitian conjugation this implies

$$(\hat{c}_1^\dagger, \hat{c}_2) = (\hat{d}_1^\dagger, \hat{d}_2)\begin{pmatrix}u & -v\\ v & u\end{pmatrix}. \tag{2.56}$$

The anti-commutation relations must be preserved by the fermionic Bogoliubov transformation. This implies that

$$1 = \{\hat{c}_1, \hat{c}_1^\dagger\} = \{u\hat{d}_1 + v\hat{d}_2^\dagger, u\hat{d}_1^\dagger + v\hat{d}_2\} = u^2\{\hat{d}_1, \hat{d}_1^\dagger\} + v^2\{\hat{d}_2^\dagger, \hat{d}_2\} = u^2+v^2, \tag{2.57}$$

which suggests taking $u = \cos\theta$ and $v = \sin\theta$.

We can use the transformation to express the Hamiltonian in terms of the new set of fermion operators,

$$\begin{aligned}\hat{\mathcal{H}} &= (\hat{c}_1^\dagger, \hat{c}_2)\begin{pmatrix}\epsilon & \lambda\\ \lambda & -\epsilon\end{pmatrix}\begin{pmatrix}\hat{c}_1\\ \hat{c}_2^\dagger\end{pmatrix} + \epsilon \\ &= (\hat{d}_1^\dagger, \hat{d}_2)\begin{pmatrix}u & -v\\ v & u\end{pmatrix}\begin{pmatrix}\epsilon & \lambda\\ \lambda & -\epsilon\end{pmatrix}\begin{pmatrix}u & v\\ -v & u\end{pmatrix}\begin{pmatrix}\hat{d}_1\\ \hat{d}_2^\dagger\end{pmatrix} + \epsilon \\ &= (\hat{d}_1^\dagger, \hat{d}_2)\begin{pmatrix}(u^2-v^2)\epsilon - 2uv\lambda & 2uv\epsilon + (u^2-v^2)\lambda\\ 2uv\epsilon + (u^2-v^2)\lambda & -(u^2-v^2)\epsilon + 2uv\lambda\end{pmatrix}\begin{pmatrix}\hat{d}_1\\ \hat{d}_2^\dagger\end{pmatrix} + \epsilon. \end{aligned} \tag{2.58}$$

The Hamiltonian is diagonalised for

$$0 = 2uv\epsilon + (u^2-v^2)\lambda = \epsilon\sin(2\theta) + \lambda\cos(2\theta) \Rightarrow \tan(2\theta) = -\frac{\lambda}{\epsilon}, \tag{2.59}$$

and takes the form

$$\begin{aligned}\hat{\mathcal{H}} &= (\hat{d}_1^\dagger, \hat{d}_2) \begin{pmatrix} \tilde{\epsilon} & 0 \\ 0 & -\tilde{\epsilon} \end{pmatrix} \begin{pmatrix} \hat{d}_1 \\ \hat{d}_2^\dagger \end{pmatrix} + \epsilon \\ &= \tilde{\epsilon}(\hat{d}_1^\dagger \hat{d}_1 + \hat{d}_2^\dagger \hat{d}_2) + \epsilon - \tilde{\epsilon}, \end{aligned} \tag{2.60}$$

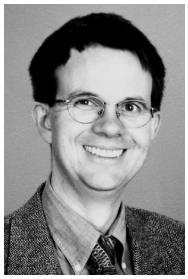
where the energy of the Bogoliubov quasiparticles is given by

$$\begin{aligned}\tilde{\epsilon} &= (u^2 - v^2)\epsilon - 2uv\lambda = \epsilon\cos(2\theta) - \lambda\sin(2\theta) = \epsilon\cos(2\theta)[1 - \frac{\lambda}{\epsilon}\tan(2\theta)] \\ &= \frac{\epsilon}{\sqrt{1+\tan^2(2\theta)}}[1 - \frac{\lambda}{\epsilon}\tan(2\theta)] = \frac{\epsilon}{\sqrt{1+\lambda^2/\epsilon^2}}(1 + \lambda^2/\epsilon^2) \\ &= \sqrt{\epsilon^2 + \lambda^2}. \end{aligned} \tag{2.61}$$

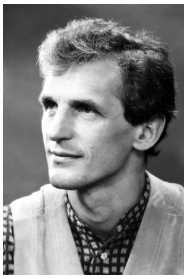
Chapter 3

The Weakly-Interacting Bose Gas

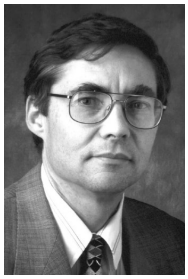
Superfluidity (flow without loss of kinetic energy and hence with zero viscosity) was first observed in a liquid of bosonic ^{4}He atoms below about 2.1K. Similar phenomena - at least the phenomenon of Bose-Einstein condensation - have since been observed in dilute ultra-cold gases of bosonic alkali atoms, resulting in a Nobel Prize in Physics in 2001 for Cornell, Ketterle, and Wieman. After the rather formal interlude of the previous chapter, we now poses the analytical tools to describe Bose-Einstein condensation in detail.



Eric A. Cornell



Wolfgang Ketterle



Carl E. Wieman

Nobel Prize in Physics
for the achievement of Bose-Einstein condensation in dilute gases of alkali atoms, and for early fundamental studies of the properties of the condensate
2001

3.1 Bose-Einstein Condensation

In order to understand the phenomenon of Bose condensation, let us first consider an ideal Bose gas, a gas of bosons without interactions between the particles. There exist different types of bosons in nature, e.g. the phonons discussed in chapter 1. The number of phonons is not conserved and hence their chemical potential is zero. On the other hand, if we consider bosonic atoms such as ^{4}He the particle number is conserved and we require a non-zero chemical potential $\mu = \mu(T)$ to satisfy the fixed particle number constraint.

Let us consider a system of N bosonic particles of mass m in a cubic box of volume $V = L^3$. In this case the density $\rho = N/V$ is fixed. The non-interacting Hamiltonian is conveniently written in terms of bosonic creation and annihilation operators $\hat{c}^\dagger_{\mathbf{k}}$, $\hat{c}_{\mathbf{k}}$ in momentum space,

$$\hat{\mathcal{H}} = \sum_{i=1}^{N} \frac{\hat{\mathbf{p}}_i^2}{2m} = \sum_{\mathbf{k}} \underbrace{\frac{\hbar^2 \mathbf{k}^2}{2m}}_{=\epsilon_{\mathbf{k}}} \hat{c}^\dagger_{\mathbf{k}} \hat{c}_{\mathbf{k}}. \tag{3.1}$$

Assuming periodic boundary conditions, the momenta are restricted to a grid of spacing $\Delta k = \frac{2\pi}{L}$, $\mathbf{k} = \Delta k(n_x, n_y, n_z)$ with $n_i \in \mathbb{Z}$. The N=particle ground state is given by the state in which all bosons occupy the $\mathbf{k} = 0$ state, which is the single-particle state with the lowest energy,

$$|\mathrm{GS}\rangle = \frac{1}{\sqrt{N!}}(\hat{c}^\dagger_{\mathbf{k}=0})^N|0\rangle. \tag{3.2}$$

In order to understand the finite-temperature behaviour and to investigate the relation between particle number and chemical potential, we consider the grand canonical potential with partition sum $\mathcal{Z}_G$ and grand potential Ω,

$$\mathcal{Z}_G = \mathrm{Tr}\; e^{-\beta(\hat{\mathcal{H}}-\mu\hat{N})} = e^{-\beta\Omega}, \tag{3.3}$$

which allows us to compute the particle number

$$N = \langle\hat{N}\rangle = \frac{1}{\mathcal{Z}_G}\mathrm{Tr}\;\hat{N}e^{-\beta(\hat{\mathcal{H}}-\mu\hat{N})} = \frac{1}{\beta\mathcal{Z}_G}\frac{\partial\mathcal{Z}_G}{\partial\mu} = \frac{1}{\beta}\frac{\partial}{\partial\mu}\ln\mathcal{Z}_G = -\frac{\partial\Omega}{\partial\mu}. \tag{3.4}$$

It is trivial to evaluate $\mathcal{Z}_G$ for a system of non-interacting bosons,

$$\mathcal{Z}_G = \sum_{\{n_\mathbf{k}=0,1,\ldots\}}\prod_\mathbf{k} e^{-\beta(\epsilon_\mathbf{k}-\mu)n_\mathbf{k}} = \prod_\mathbf{k}\left(\sum_{n=0}^\infty e^{-\beta(\epsilon_\mathbf{k}-\mu)n}\right) = \prod_\mathbf{k}\frac{1}{1-e^{-\beta(\epsilon_\mathbf{k}-\mu)}}. \tag{3.5}$$

Note that convergence of the geometric series requires that $\mu < 0$. This results in

$$\Omega = -\frac{1}{\beta}\ln\mathcal{Z}_G = \frac{1}{\beta}\sum_\mathbf{k}\ln(1-e^{-\beta(\epsilon_\mathbf{k}-\mu)}) \Rightarrow N = -\frac{\partial\Omega}{\partial\mu} = \sum_\mathbf{k}\underbrace{\frac{1}{e^{\beta(\epsilon_\mathbf{k}-\mu)}-1}}_{=\langle\hat{n}_\mathbf{k}\rangle = n_B(\epsilon_\mathbf{k}-\mu)}. \tag{3.6}$$

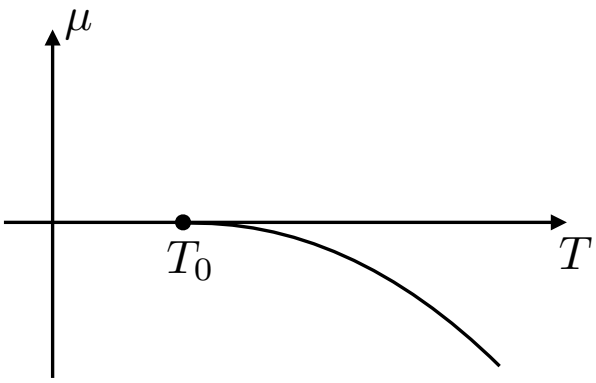
Of course we could have written the last expression without derivation. In the thermodynamic limit $L \to \infty$ ($\rho = N/L^3$ fixed) we can convert the sum into an integral and make use of the rotational symmetry,

$$\begin{aligned}\rho &= \frac{N}{V} = \frac{1}{V}\sum_\mathbf{k} n_B(\epsilon_\mathbf{k}-\mu) = \frac{1}{(2\pi)^3}\sum_\mathbf{k}(\Delta k)^3 n_B(\epsilon_\mathbf{k}-\mu)\\ &\approx \frac{1}{(2\pi)^3}\int d^3\mathbf{k}\; n_B(\epsilon_\mathbf{k}-\mu) = \frac{1}{2\pi^2}\int_0^\infty dk\; k^2 n_B(\epsilon_k-\mu).\end{aligned}$$

This can be converted into an integral over energy by substitution $\epsilon = \epsilon_k = \hbar^2k^2/(2m)$,

$$\rho = \frac{N}{V} = \frac{1}{4\pi^2}\sqrt{\frac{2m}{\hbar^2}}^3\int_0^\infty d\epsilon\frac{\sqrt{\epsilon}}{e^{\beta(\epsilon-\mu)}-1} \tag{3.7}$$

This integral equation serves as an implicit equation for μ and can be solved numerically. We find that the chemical potential is negative and increases as temperature is lowered until it becomes zero at a temperature T_0 which depends on the choice of ρ and m. This generic behaviour is shown in the following figure:



Disturbingly, the integral equation for the chemical potential has no solution for $T < T_0$, suggesting that we can no longer satisfy the fixed density constraint. There is a simple resolution of this apparent paradox. We have converted the momentum sum into an integral, ignoring a macroscopic occupation of the $\mathbf{k} = 0$ state, which we already know will happen in the groundstate. The integral only gives the density $\rho_{\epsilon>0}$ of particles that are not condensed into the $\mathbf{k} = 0$ state. Bose condensation occurs for $T < T_0$ with a certain fraction of particles in the condensate.

We can determine T_0 by solving the integral equation for $\mu = 0$. In this case we can scale out temperature from the integral by the substitution $z = \beta\epsilon = \epsilon/T$, resulting in a purely numerical integral,

$$\rho = \frac{1}{4\pi^2}\sqrt{\frac{2m}{\hbar^2}}^3 T_0^{3/2} \underbrace{\int_0^\infty dz \frac{\sqrt{z}}{e^z - 1}}_{=\frac{\sqrt{\pi}}{2}\zeta(3/2)\approx 2.61}, \tag{3.8}$$

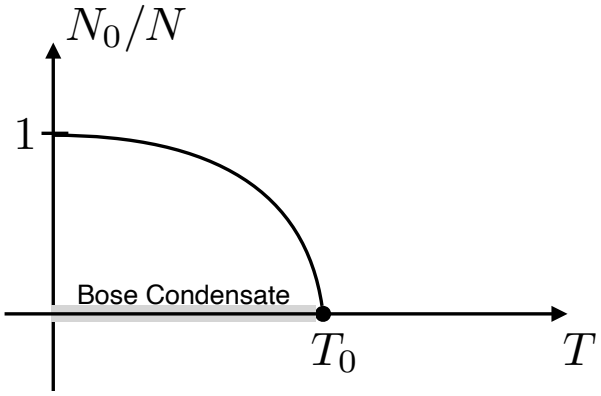
where $\zeta(x)$ denotes the zeta-function. Hence we obtain $T_0 \sim \frac{\hbar^2}{2m}\rho^{2/3}$. For $T < T_0$ and $\mu = 0$ the integral (3.7) gives the density of particles that are not in the condensate,

$$\rho_{\epsilon>0} = \frac{1}{4\pi^2}\sqrt{\frac{2m}{\hbar^2}}^3 T^{3/2}\frac{\sqrt{\pi}}{2}\zeta(3/2). \tag{3.9}$$

The fraction of particles in the condensate is therefore given by

$$\frac{N_0}{N} = \frac{\rho_0}{\rho} = \frac{\rho - \rho_{\epsilon>0}}{\rho} = 1 - \left(\frac{T}{T_0}\right)^{3/2} \quad \text{for } T < T_0. \tag{3.10}$$

This is illustrated in the figure below. We obtain a Bose condensate for $T < T_0$ with the condensate fraction approaching $N_0/N = 1$ at $T = 0$. This is expected since in the groundstate all particles are in the $\mathbf{k} = 0$ state.



3.2 The Hamiltonian of the Interacting Bose Gas

While the essence of Bose-Einstein condensation can be understood in terms of a non-interacting particle picture, we require interactions to understand the phenomenon of *superfluidity*. As long as the interactions are not too strong we expect that the Bose-Einstein condensate will persist in the interacting system. We will consider a generic density-density interaction between the bosonic atoms. In second quantisation the contribution to the Hamiltonian can be written as

$$\hat{\mathcal{H}}_{\rm int} = \frac{1}{2}\int d^3\mathbf{r}\int d^3\mathbf{r}'\, U(\mathbf{r}-\mathbf{r}')\hat{n}_{\mathbf{r}}\hat{n}_{\mathbf{r}'}, \tag{3.11}$$

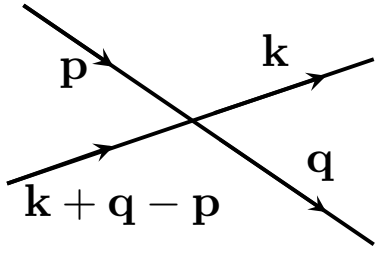
where $\hat{n}_{\mathbf{r}} = \hat{c}^\dagger_{\mathbf{r}}\hat{c}_{\mathbf{r}}$ denote occupation number operators in real space. Here the spatial integrals run over the cubic box of volume $V = L^3$. Considering the atoms as charge-neutral hard balls we would expect that the atoms are subject to a short-range local repulsion. For simplicity, we will assume a repulsive delta-function interaction,

$$U(\mathbf{r}-\mathbf{r}') = U\delta(\mathbf{r}-\mathbf{r}'), \tag{3.12}$$

with $U > 0$. As you will show in a homework problem, in momentum space the interaction Hamiltonian takes the form

$$\hat{\mathcal{H}}_{\rm int} = \frac{U}{2V}\sum_{\mathbf{k},\mathbf{p},\mathbf{q}} \hat{c}^\dagger_{\mathbf{k}}\hat{c}^\dagger_{\mathbf{q}}\hat{c}_{\mathbf{p}}\hat{c}_{\mathbf{k}-\mathbf{p}+\mathbf{q}}. \tag{3.13}$$

Notice that the momenta of the ingoing bosons that are annihilated matches that of the outgoing, created bosons, as required by momentum conservation. This can be represented by a Feynman diagram:



3.3 Mean Field Theory

At the end of Chapter 2, we learnt how to diagonalise quadratic, second quantised Hamiltonian. In the present case, however, the interaction is quartic in creation and annihilation operators. Based upon our knowledge of the non-interacting Bose gas, we can develop a leading-order approximation that reduces the Hamiltonian to quadratic form.

We know that in the groundstate of the non-interacting system, all N particles condense into the $\mathbf{k} = 0$ state. The main assumption is that in the presence of weak repulsion U the occupation of the $\mathbf{k} = 0$ state remains macroscopic,

$$\langle \hat{c}_{\mathbf{0}}^\dagger c_{\mathbf{0}} \rangle = N_0 \gg 1. \tag{3.14}$$

On the other hand, the "quantum-ness" of the operators is manifested in the bosonic commutator,

$$[\hat{c}_{\mathbf{0}}, \hat{c}_{\mathbf{0}}^\dagger] = 1. \tag{3.15}$$

Since $1 \ll N_0$, this suggests that for the $\mathbf{k} = 0$ state the "quantum-ness" of the operators can be ignored and we can treat $\hat{c}_{\mathbf{0}}^\dagger$ and $\hat{c}_{\mathbf{0}}$ as numbers

$$\hat{c}_{\mathbf{0}}^\dagger \approx \sqrt{N_0} \quad \text{and} \quad \hat{c}_{\mathbf{0}} \approx \sqrt{N_0}. \tag{3.16}$$

By splitting off the $\mathbf{k} = 0$ terms from the sums we can expand the interaction Hamiltonian,

$$\begin{aligned} \hat{\mathcal{H}}_{\text{int}} &= \frac{U}{2V} \sum_{\mathbf{k},\mathbf{p},\mathbf{q},\mathbf{l}} \delta(\mathbf{k}+\mathbf{q}-\mathbf{p}-\mathbf{l}) \hat{c}_{\mathbf{k}}^\dagger \hat{c}_{\mathbf{q}}^\dagger \hat{c}_{\mathbf{p}} \hat{c}_{\mathbf{l}} \\ &= \frac{U}{2V} N_0^2 + \frac{U}{2V} N_0 \sum_{\mathbf{k}\neq 0} (4 \hat{c}_{\mathbf{k}}^\dagger \hat{c}_{\mathbf{k}} + \hat{c}_{\mathbf{k}}^\dagger \hat{c}_{-\mathbf{k}}^\dagger + \hat{c}_{-\mathbf{k}} \hat{c}_{\mathbf{k}}) + \dots \end{aligned}$$

Note that there are no linear terms $\sim c_{\mathbf{k}}^\dagger$ or $\sim \hat{c}_{\mathbf{k}}$ with $\mathbf{k} \neq 0$ since after setting three momenta to zero, the delta function renders the forth momentum zero as well. Importantly, the quadratic terms dominate over higher-order interaction terms, which are $\mathcal{O}(1)$ in N_0. The above expansion is therefore controlled by the small parameter $1/N_0$. In the mean-field approximation we only retain the constant and quadratic terms, resulting in an effective non-interacting Hamiltonian that can be diagonalised. We further substitute

$$N_0 = N - \sum_{\mathbf{k}\neq 0} \hat{c}_{\mathbf{k}}^\dagger \hat{c}_{\mathbf{k}} \tag{3.17}$$

and obtain

$$\hat{\mathcal{H}}_{\text{int}} = \underbrace{\frac{U}{2V} N^2}_{=\frac{1}{2}UV\rho^2} + \underbrace{\frac{U}{2V} N}_{=\frac{1}{2}U\rho} \sum_{\mathbf{k}\neq 0} (2 \hat{c}_{\mathbf{k}}^\dagger \hat{c}_{\mathbf{k}} + \hat{c}_{\mathbf{k}}^\dagger \hat{c}_{-\mathbf{k}}^\dagger + \hat{c}_{-\mathbf{k}} \hat{c}_{\mathbf{k}}), \tag{3.18}$$

up to quadratic order. Combining with the kinetic energy contribution, the mean-field Hamiltonian becomes

$$\hat{\mathcal{H}} = \sum_{\mathbf{k}\neq 0} \Big\{ \underbrace{\left(\frac{\hbar^2 \mathbf{k}^2}{2m} + U\rho \right)}_{=\epsilon(\mathbf{k})} \hat{c}_{\mathbf{k}}^\dagger \hat{c}_{\mathbf{k}} + \frac{U\rho}{2} \left(\hat{c}_{\mathbf{k}}^\dagger \hat{c}_{-\mathbf{k}}^\dagger + \hat{c}_{-\mathbf{k}} \hat{c}_{\mathbf{k}} \right) \Big\} + \frac{1}{2} UV\rho^2. \tag{3.19}$$

In order to write the Hamiltonian in a 2x2 matrix form in terms of Nambu spinors,

$$\hat{\psi}^\dagger = (\hat{c}^\dagger_{\mathbf{k}}, \hat{c}_{-\mathbf{k}}) \quad \text{and} \quad \hat{\psi} = \begin{pmatrix} \hat{c}_k \\ \hat{c}^\dagger_{-\mathbf{k}} \end{pmatrix}, \tag{3.20}$$

we write the regular terms in symmetric form with respect to $\mathbf{k}$ and $-\mathbf{k}$, using that $\sum_{\mathbf{k}\neq 0} f(\mathbf{k}) = \frac{1}{2}\sum_{\mathbf{k}\neq 0}[f(\mathbf{k}) + f(-\mathbf{k})]$, resulting in

$$\hat{\mathcal{H}} = \frac{1}{2}\sum_{\mathbf{k}\neq 0}(\hat{c}^\dagger_{\mathbf{k}}, \hat{c}_{-\mathbf{k}}) \begin{pmatrix} \epsilon(\mathbf{k}) & U\rho \\ U\rho & \epsilon(\mathbf{k}) \end{pmatrix} \begin{pmatrix} \hat{c}_k \\ \hat{c}^\dagger_{-\mathbf{k}} \end{pmatrix} - \frac{1}{2}\sum_{\mathbf{k}\neq 0}\epsilon(\mathbf{k}) + \frac{1}{2}UV\rho^2, \tag{3.21}$$

where we have used that $\hat{c}^\dagger_{-\mathbf{k}}\hat{c}_{-\mathbf{k}} = \hat{c}_{-\mathbf{k}}\hat{c}^\dagger_{-\mathbf{k}} - 1$. We can diagonalise this Hamiltonian by a Bogoliubov transformation to new boson operators,

$$\begin{pmatrix} \hat{c}_k \\ \hat{c}^\dagger_{-\mathbf{k}} \end{pmatrix} = \begin{pmatrix} \cosh(\theta_{\mathbf{k}}) & \sinh(\theta_{\mathbf{k}}) \\ \sinh(\theta_{\mathbf{k}}) & \cosh(\theta_{\mathbf{k}}) \end{pmatrix} \begin{pmatrix} \hat{d}_{\mathbf{k}} \\ \hat{d}^\dagger_{-\mathbf{k}} \end{pmatrix}. \tag{3.22}$$

As you will show in a homework problem, this is achieved by

$$\tanh(2\theta_{\mathbf{k}}) = -\frac{U\rho}{\epsilon(\mathbf{k})}, \tag{3.23}$$

resulting in the diagonal Hamiltonian

$$\hat{\mathcal{H}} = \sum_{\mathbf{k}\neq 0} E(\mathbf{k})\hat{d}^\dagger_{\mathbf{k}}\hat{d}_{\mathbf{k}} + \frac{1}{2}\sum_{\mathbf{k}\neq 0}[E(\mathbf{k}) - \epsilon(\mathbf{k})] + \frac{1}{2}UV\rho^2, \tag{3.24}$$

with Bogoliubov quasiparticle dispersion

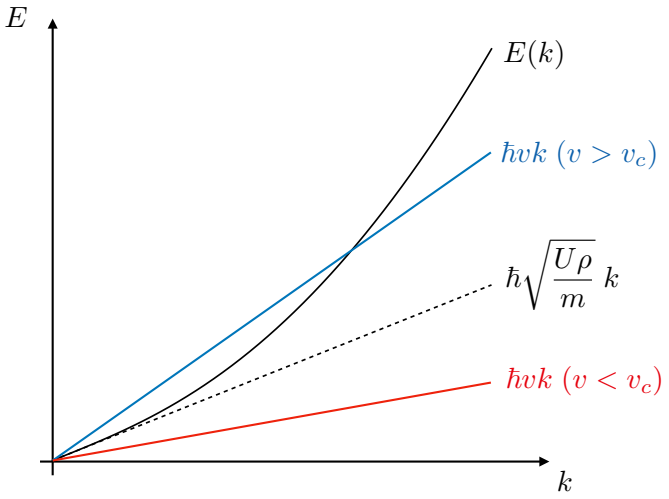
$$E(\mathbf{k}) = \sqrt{\epsilon^2(\mathbf{k}) - (U\rho)^2} = \sqrt{\left(\frac{\hbar^2\mathbf{k}^2}{2m} + U\rho\right)^2 - (U\rho)^2}. \tag{3.25}$$

3.4 Landau's Critical Superfluid Velocity

While the dispersion in the non-interacting Bose gas is quadratic, the dispersion $E(\mathbf{k})$ of excitations in the interacting system has different asymptotic behaviour at large and small momenta,

$$E(\mathbf{k}) \simeq \begin{cases} \frac{\hbar^2\mathbf{k}^2}{2m} & \text{for} \quad \frac{\hbar^2\mathbf{k}^2}{2m} \gg U\rho \\ \hbar\sqrt{\frac{U\rho}{m}}|\mathbf{k}| & \text{for} \quad \frac{\hbar^2\mathbf{k}^2}{2m} \ll U\rho \end{cases}, \tag{3.26}$$

hence it crosses over from linear at small $\mathbf{k}$ to quadratic at large $\mathbf{k}$, as shown in the figure below. The linear dispersion $E(\mathbf{k}) \sim |\mathbf{k}|$ at small momenta is responsible for the superfluid properties as can be seen by the following argument due to Landau:



Suppose that the liquid of total mass M moves with velocity v along the direction of a capillary (thin tube). The liquid has kinetic energy $E = \frac{1}{2}Mv^2$ and momentum $P = Mv$. The only way that friction can arise is by creating excitations in the liquid. Let us consider a single quasi-particle excitation with momentum $\hbar k$ in the direction of v. This excitation has energy $E(k)$ as calculated in the previous section. Such an excitation requires energy and momentum transfer and hence a reduction of the velocity by a small amount $v \to v - \Delta v$.

Conservation of Momentum:

$$Mv = M(v - \Delta v) + \hbar k$$

Conservation of Energy:

$$\frac{1}{2}Mv^2 = \frac{1}{2}M(v - \Delta v)^2 + E(k).$$

Linearising the second equation in Δv and combining with the other equation, we obtain

$$E(k) = \hbar v k. \tag{3.27}$$

Hence, the momentum k of the excitation is determined from the intersection of the quasi-particle dispersion $E(k)$ with the straight line $\hbar v k$. Of course, a solution is only possible if $\hbar v$ is bigger than the slope $\hbar\sqrt{\frac{U\rho}{2m}}$ of $E(k)$ at mall momenta, as illustrated in the figure. This determines a critical velocity

$$v_c = \sqrt{\frac{U\rho}{2m}}. \tag{3.28}$$

For velocities $v < v_c$, it is not possible to transfer energy and momentum by creating an excitation in the fluid. As a result, the flow remains frictionless, resulting in superfluidity. In the absence of interactions ($U = 0$) the critical velocity is zero and superfluidity is not possible.

3.5 Superfluid Fraction

In approximating the interacting boson Hamiltonian by an effective free-particle Hamiltonian with anomalous terms, we have assumed that the number of particles

$$N_0 = N - \sum_{\mathbf{k}\neq 0} \langle \hat{c}^\dagger_{\mathbf{k}} \hat{c}_{\mathbf{k}} \rangle \tag{3.29}$$

in the $\mathbf{k} = 0$ state, forming a condensate, is large. In the non-interacting grounds state, all particles occupy the $\mathbf{k} = 0$ state, corresponding to a condensate fraction $N_0/N = 1$. Interactions are expected to deplete the condensate and we are now going to investigate how N_0 depends on the interaction U. This serves as a consistency check if our assumption $N_0 \gg 1$ remains valid.

In order to calculate the expectation value $\langle \hat{c}^\dagger_{\mathbf{k}} \hat{c}_{\mathbf{k}} \rangle$ we can simply insert the Bogoliubov transformation to the new operators $\hat{d}^\dagger_{\mathbf{k}}$ and $\hat{d}_{\mathbf{k}}$ in which the Hamiltonian is diagonal and use that

$$\langle \hat{d}^\dagger_{\mathbf{k}} \hat{d}_{\mathbf{k}} \rangle = \frac{1}{e^{\beta E(\mathbf{k})} - 1}, \tag{3.30}$$

where $E(\mathbf{k})$ is the dispersion of the Bogoliubov quasiparticles, given in Eq. (3.25). Since the Hamiltonian is diagonal in terms of the new operators, it also follows that $\langle \hat{d}^\dagger_{\mathbf{k}} \hat{d}^\dagger_{-\mathbf{k}} \rangle = \langle \hat{d}_{-\mathbf{k}} \hat{d}_{\mathbf{k}} \rangle = 0$. As you will show in a homework problem. the density $\rho_0 = N_0/V$ of particles in the condensate of the interacting gas at temperature T is given by

$$\rho_0 = \rho - \frac{1}{V} \sum_{\mathbf{k}\neq 0} \left[\cosh(2\theta_{\mathbf{k}}) \frac{1}{e^{\beta E(\mathbf{k})} - 1} + \sinh^2(\theta_{\mathbf{k}}) \right], \tag{3.31}$$

where $\theta_{\mathbf{k}}$ is defined by $\tanh(2\theta_{\mathbf{k}}) = -U\rho/\epsilon(\mathbf{k}) = -U\rho/(\frac{\hbar^2\mathbf{k}^2}{2m} + U\rho)$. The first term in the sum describes the depletion of condensate due to thermal population of $\mathbf{k} \neq 0$ states. This contribution vanishes at $T = 0$ where all Bogoliubov quasiparticles occupy the $\mathbf{k} = 0$ state. The second term in the sum is present at $T = 0$ and corresponds to the depletion of the condensate by interactions.

In the homework problem you will evaluate the superfluid density ρ_0 at $T = 0$ by converting the momentum sum into an integral over energy, as we have done before. The final result is

$$\rho_0 = \rho - \frac{1}{6\sqrt{2}\pi^2} \left(\frac{2m}{\hbar^2} U\rho \right)^{3/2}. \tag{3.32}$$

As to be expected, this result reflects that in the absence of interactions, all particles are in the condensate ($\rho_0 = \rho$). The superfluid fraction $\rho_0/\rho = N_0/N$ decreases with the interaction U but remains close to 1 for sufficiently weak interactions. In this regime, the mean-field approximation is justified.

Chapter 4

Quantum Magnets

In this chapter, we will use the language of second quantisation to discuss properties of quantum magnets — magnets in which atoms in a lattice have unpaired electron spins that may interact with the spins on other sites. We will focus upon insulating magnets in which the electrons that carry the spins are localised on atomic sites, as opposed to itinerant systems, where they would be free to hop around. Note that strong Hund's coupling could potentially give rise to the formation of local moment spins that are much larger than the electron spin. Despite the apparent simplicity of this set up, the collective quantum behaviour of such systems is incredibly rich. Just about every phenomenon of modern quantum physics is revealed by these systems.

The spins could be arranged in one-dimensional chains, two-dimensional lattices, such as square-, triangular- or Kagome lattices, or three dimensional lattices such as cubic or Pyrochlore lattices. The competition between different interactions can lead rich magnetic phase diagrams with magnetically ordered phases that show complicated spin textures. Geometric frustration in low-dimensional systems, e.g. on the two-dimensional Kagome lattice, can prevent the system from ordering altogether. In such situations the spins remain strongly fluctuating down to zero temperature, giving rise to a quantum spin liquid which is characterised by emergent fractionalised excitations. Quantum spin systems also provide the first example of topology in quantum physics (through the Haldane conjecture that we will discuss later) and a magnetic model (due to Kitaev) is the basis of some of the most widely studied quantum error correcting codes.

4.1 The Heisenberg Model

The Hamiltonian of the Heisenberg model is given by

$$\hat{\mathcal{H}} = \frac{1}{2}\sum_{i,j} J_{ij}\hat{\mathbf{S}}_i \cdot \hat{\mathbf{S}}_j = \frac{1}{2}\sum_{i,j} J_{ij}\left(\hat{S}_i^x\hat{S}_j^x + \hat{S}_i^y\hat{S}_j^y + \hat{S}_i^z\hat{S}_j^z\right), \tag{4.1}$$

where i, j label the sites of a given lattice. The exchange couplings J_{ij} between spins on sites i and j could be positive or negative, favouring antiferromagnet (anti-parallel) or ferromagnetic (parallel) alignment, respectively. However, often there exists a competition between different exchanges and not all bonds can be simultaneously satisfied. Different

microsopic mechanisms contribute to the interaction between spins, including direct dipole-dipole coupling and antiferromagnetic superexchange, which we will discuss in Chapter 8. Orbital degrees of freedom can also play a crucial role in determining the strength and sign of the exchange.

The Heisenberg model is isotropic in spin space. While the relative orientation between spins depends on the exchange couplings, the Hamiltonian is invariant under a global rotation of all the spins. This symmetry can be broken or lowered by anisotropy, e.g exchange anisotropy, $J_x\hat{S}_i^x\hat{S}_j^x + J_y\hat{S}_i^y\hat{S}_j^y + J_z\hat{S}_i^z\hat{S}_j^z$, or single-ion anisotropy, $(\hat{S}_i^z)^2$.

Let us briefly review the properties of spin operators. While spin operators on different sites commute, the components of the vector spin operator on a given site satisfy the *commutation relations*

$$\left[\hat{S}^\alpha, \hat{S}^\beta\right] = i\hbar\epsilon_{\alpha\beta\gamma}\hat{S}^\gamma. \tag{4.2}$$

From this it follows that $\left[\hat{\mathbf{S}}^2, \hat{S}^z\right] = 0$, which implies that $\hat{\mathbf{S}}^2$ and $\hat{S}^z$ have a simultaneous eigenbasis,

$$\begin{aligned} \hat{\mathbf{S}}^2|S,m\rangle &= \hbar^2 S(S+1)|S,m\rangle \\ \hat{S}^z|S,m\rangle &= \hbar m|S,m\rangle \end{aligned} \tag{4.3}$$

with quantum numbers S (determining the size of the spin, $S = \frac{1}{2}, 1, \frac{3}{2}, 2, \ldots$) and $m = -S, -S+1, \ldots, S-1, S$. Hence, for a given S there exists a tower of $2S+1$ states. Rather than working with $\hat{S}^x$ and $\hat{S}^y$ it is often convenient to use the ladder operators

$$\hat{S}^\pm = \hat{S}_x \pm i\hat{S}_y, \tag{4.4}$$

which satisfy the commutation relations

$$[\hat{S}^+, \hat{S}^-] = 2\hbar\hat{S}^z \quad \text{and} \quad [\hat{S}^\pm, \hat{S}^z] = \mp\hbar\hat{S}^\pm, \tag{4.5}$$

as can be easily derived from Eq. (4.2). The ladder operators raise and lower the m quantum number by one,

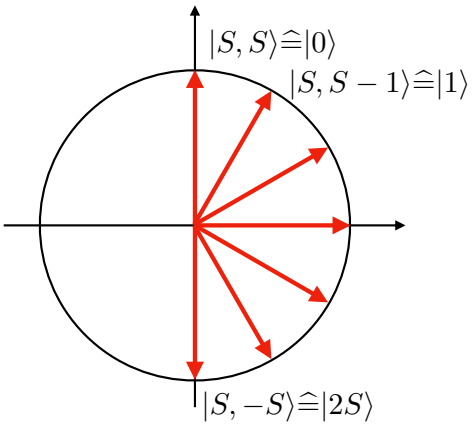
$$\hat{S}^\pm|S,m\rangle = \hbar\sqrt{S(S+1) - m(m\pm 1)}|S, m\pm 1\rangle, \tag{4.6}$$

and $\hat{S}^+|S,S\rangle = \hat{S}^-|S,-S\rangle = 0$.

4.2 Holstein-Primakoff Transformation

The non-trivial commutation relations of the spin operators make them tricky to deal with when facing a system of many interacting quantum spins. Bosons are much easier to deal with, as we have seen in previous Chapters. The Holstein-Primakoff transformation allows to map quantum spin-operators to bosonic creation and annihilation operators. Similar to boson operators, spin operators on different sites commute. We can therefore focus on a single site in the following.

The idea is to identify the maximally polarised spin state $|S,S\rangle$ with the vacuum state $|0\rangle$ of having no boson, the state $|S,S-1\rangle$ with the one-boson state $|1\rangle$ and so on. This suggests that we should identify $\hat{S}^z = \hbar(S - \hat{b}^\dagger\hat{b})$.



In order to match the dimension of the spin and boson Hilbert spaces we need to truncate the bosonic Hilbert space and allow for a maximum occupation number of $n_{\max} = 2S$, corresponding to the state $|S,-S\rangle$. Increasing the boson number by one is equivalent to lowering the m quantum number of the spin state by one and vice verse, which suggests that $\hat{S}^- \sim \hat{b}^\dagger$ and $\hat{S}^+ \sim \hat{b}$. We need to be careful, however, since the identification between spin and boson operators needs to respect boson and spin commutation relations. The correct mapping is known as the Holstein-Primakoff transformation,

$$\begin{aligned} \hat{S}^z &= \hbar(S - \hat{b}^\dagger \hat{b}), \\ \hat{S}^+ &= \hbar \left(2S - \hat{b}^\dagger \hat{b}\right)^{1/2} \hat{b}, \\ \hat{S}^- &= \hbar b^\dagger \left(2S - \hat{b}^\dagger \hat{b}\right)^{1/2}. \end{aligned} \tag{4.7}$$

One can easily check[1] that the bosonic commutation relations $[\hat{b}, \hat{b}^\dagger] = 1$ reproduce the spin commutation relations in the form $[\hat{S}^+, \hat{S}^-] = 2\hbar \hat{S}^z$. Identifying our boson operator per site of the lattice, the fact that spin operators on different sites commute is readily accommodated by the similar commutation of bosonic operators on different sites.

4.3 Linear Spin-Wave Theory

The idea is to look at small fluctuations around the classical groundstate. To obtain the classical groundstate we treat all spins as classical vectors and minimise the energy. This could potentially be complicated and result in non-collinear groundstates. An example of such a non-collinear classical grounstate is the Heisenberg antiferromagnet on the triangular lattice where neighbouring spins in the classical groundstate enclose an angle of 120°.

The classical spin direction on a given site i determines the local quantisation axes (the z axis in the local frame), along which we define the state $|S,S\rangle_i$. We then use the Holstein-Primakoff transformation in this frame to express the spin operators $\hat{S}_i^\alpha$ in terms of boson operator $\hat{b}_i^\dagger$ and $\hat{b}_i$. The Holtstein-Primakoff bosons represent the quantised spin-wave excitations, called magnons.

[1] And you will do so in a homework problem.

Rewriting a given spin Hamiltonian in terms of Holstein-Primakoff bosons, we will obtain a constant part, corresponding to the classical ground-state energy, a part quadratic in the boson creation and annihilation operators, as well as higher order terms. In *linear spin-wave theory* we neglect all terms beyond quadratic order. The resulting Hamiltonian is one of non-interacting bosons which can be diagonalised by Fourier transform, unitary transformation or, if the Hamiltonian contains anomalous terms, by Bogoliubov transformation.

The quartic interaction terms are by a factor $1/S$ smaller than the quadratic terms. This implies that linear-spin wave theory is a good approximation for large S and that it takes the leading quantum fluctuations into account. For large S, we can also ignore the truncation of the bosonic Hilbert space.

Expanding the occupation number operators out of the square roots in the Holstein-Primakoff transformation results in terms beyond quadratic order. If we are interested in fluctuations on the level of linear-spin wave theory we can therefore work with the approximate transformation,

$$\begin{aligned} \hat{S}_i^z &= \hbar(S - \hat{b}_i^\dagger \hat{b}_i), \\ \hat{S}_i^+ &\approx \hbar\sqrt{2S}\hat{b}_i, \text{ and } \hat{S}_i^- \approx \hbar\sqrt{2S}\hat{b}_i^\dagger. \end{aligned} \tag{4.8}$$

4.4 The Heisenberg Ferromagnet

Let us now consider a Heisenberg ferromagnet on a d-dimensional hypercubic lattice with N sites and periodic boundary conditions. For simplicity, we only consider a nearest-neighbour interaction between the spins. The Hamiltonian is given by

$$\hat{\mathcal{H}} = -J \sum_{\langle i,j \rangle} \hat{\mathbf{S}}_i \cdot \hat{\mathbf{S}}_j, \tag{4.9}$$

where $J > 0$. Instead of summing over nearest-neighbour bonds we could sum over discrete lattice-site positions $\mathbf{r}$ and lattice vectors $\mathbf{a}_i = a\hat{\mathbf{e}}_i$ ($i = 1, \ldots, d$ and a the lattice constant),

$$\hat{\mathcal{H}} = -J \sum_{\mathbf{r}} \sum_{i=1}^{d} \hat{\mathbf{S}}_{\mathbf{r}} \cdot \hat{\mathbf{S}}_{\mathbf{r}+\mathbf{a}_i}, \tag{4.10}$$

which covers every bond exactly once. The classical groundstate is trivial: all spins align parallel, where the direction is spontaneously picked. We will call this direction the positive z-axis.

4.4.1 The Groundstate

The groundstate is such that all of the spins are aligned — say in the z-directions — so that $\hat{S}_i^z |S, S\rangle_i = \hbar S |S, S\rangle_i$ on each site. The product state

$$|\psi\rangle = \otimes_i |S, S\rangle_i \tag{4.11}$$

is indeed an eigenstate of $\hat{\mathcal{H}}$ and hence the quantum-mechanical ground-state of the many-body system,

$$\hat{\mathcal{H}}|\psi\rangle = -J\sum_{\langle i,j\rangle}\left\{\hat{S}_i^z\hat{S}_j^z + \frac{1}{2}\left(\hat{S}_i^+\hat{S}_j^- + \hat{S}_i^-\hat{S}_j^+\right)\right\}|\psi\rangle = -J\sum_{\langle i,j\rangle}\hat{S}_i^z\hat{S}_j^z|\psi\rangle = E_0|\psi\rangle, \quad (4.12)$$

where $E_0 = -J\hbar^2S^2Nd$ is the classical groundstate energy.[2]

4.4.2 Linear Spin-Wave Exciations

We now use the approximate form of the Holstein-Primakoff transformation, Eq. (4.7), to express hamiltonian of the Heisenberg ferromagnet in terms of bosonic creation and annihilation operators, keeping terms up to quadratic order,

$$\begin{aligned}
\hat{\mathcal{H}} &= -J\sum_{\mathbf{r}}\sum_{i=1}^{d}\hat{\mathbf{S}}_{\mathbf{r}}\cdot\hat{\mathbf{S}}_{\mathbf{r}+\mathbf{a}_i} = -J\sum_{\mathbf{r}}\sum_{i=1}^{d}\left\{\hat{S}_{\mathbf{r}}^z\hat{S}_{\mathbf{r}+\mathbf{a}_i}^z + \frac{1}{2}\left(\hat{S}_{\mathbf{r}}^+\hat{S}_{\mathbf{r}+\mathbf{a}_i}^- + \hat{S}_{\mathbf{r}}^-\hat{S}_{\mathbf{r}+\mathbf{a}_i}^+\right)\right\} \\
&= -J\hbar^2S^2Nd + J\hbar^2S\sum_{\mathbf{r}}\sum_{i=1}^{d}(\hat{b}_{\mathbf{r}}^\dagger\hat{b}_{\mathbf{r}} + \hat{b}_{\mathbf{r}+\mathbf{a}_i}^\dagger\hat{b}_{\mathbf{r}+\mathbf{a}_i}) \\
&\quad -J\hbar^2S\sum_{\mathbf{r}}\sum_{i=1}^{d}(\hat{b}_{\mathbf{r}}\hat{b}_{\mathbf{r}+\mathbf{a}_i}^\dagger + \hat{b}_{\mathbf{r}}^\dagger\hat{b}_{\mathbf{r}+\mathbf{a}_i}) \\
&= -J\hbar^2S^2Nd + 2J\hbar^2Sd\sum_{\mathbf{r}}\hat{b}_{\mathbf{r}}^\dagger\hat{b}_{\mathbf{r}} - J\hbar^2S\sum_{\mathbf{r}}\sum_{i=1}^{d}(\hat{b}_{\mathbf{r}}\hat{b}_{\mathbf{r}+\mathbf{a}_i}^\dagger + \hat{b}_{\mathbf{r}}^\dagger\hat{b}_{\mathbf{r}+\mathbf{a}_i}). \quad (4.13)
\end{aligned}$$

Our task now is to diagonalise this Hamiltonian. Because the system is translationally invariant, the first step towards this (actually the only step in this case) is to Fourier transform,

$$\hat{b}_{\mathbf{r}}^\dagger = \frac{1}{\sqrt{N}}\sum_{\mathbf{k}}e^{-i\mathbf{k}\mathbf{r}}\hat{b}_{\mathbf{k}}^\dagger, \quad \hat{b}_{\mathbf{r}} = \frac{1}{\sqrt{N}}\sum_{\mathbf{k}}e^{i\mathbf{k}\mathbf{r}}\hat{b}_{\mathbf{k}}. \quad (4.14)$$

resulting in

$$\begin{aligned}
\hat{\mathcal{H}} &= -J\hbar^2S^2Nd + 2J\hbar^2Sd\sum_{\mathbf{k}}\hat{b}_{\mathbf{k}}^\dagger\hat{b}_{\mathbf{k}} \\
&\quad -J\hbar^2S\frac{1}{N}\sum_{\mathbf{r}}\sum_{\mathbf{k},\mathbf{k}'}\sum_{i=1}^{d}\left[e^{i\mathbf{k}\mathbf{r}}e^{-i\mathbf{k}'(\mathbf{r}+\mathbf{a}_i)}\hat{b}_{\mathbf{k}}\hat{b}_{\mathbf{k}'}^\dagger + e^{-i\mathbf{k}\mathbf{r}}e^{i\mathbf{k}'(\mathbf{r}+\mathbf{a}_i)}\hat{b}_{\mathbf{k}}^\dagger\hat{b}_{\mathbf{k}'}\right] \\
&= -J\hbar^2S^2Nd + 2J\hbar^2Sd\sum_{\mathbf{k}}\hat{b}_{\mathbf{k}}^\dagger\hat{b}_{\mathbf{k}} - J\hbar^2S\sum_{\mathbf{k}}\sum_{i=1}^{d}(e^{-i\mathbf{k}\mathbf{a}_i} + e^{i\mathbf{k}\mathbf{a}_i})\hat{b}_{\mathbf{k}}^\dagger\hat{b}_{\mathbf{k}} \\
&= -J\hbar^2S^2Nd + 2J\hbar^2S\sum_{\mathbf{k}}\left[d - \sum_{i=1}^{d}\cos(ak_i)\right]\hat{b}_{\mathbf{k}}^\dagger\hat{b}_{\mathbf{k}}. \quad (4.15)
\end{aligned}$$

[2]In the above calculation we have used that $\hat{S}^+|S,S\rangle = 0$.

We have diagonalised the linear spin-wave Hamiltonian for the ferromagnet by Fourier transform. Notice that since the classical groundstate is an eigenstate of the Hamiltonian, there are no zero-point fluctuations The spin-wave excitation have the dispersion

$$\epsilon(\mathbf{k}) = 2J\hbar^2 S\left[d - \sum_{i=1}^{d}\cos(ak_i)\right], \tag{4.16}$$

which for small momenta can be approximated as

$$\epsilon(\mathbf{k}) \simeq J\hbar^2 S a^2 \mathbf{k}^2. \tag{4.17}$$

These modes are gapless — their energy goes to zero as $|\mathbf{k}|$ does. This is because they are Goldstone modes; they arise because the spin rotational symmetry of the Hamiltonian is broken by the groundstate — it costs zero energy to rotate between these groundstates and this corresponds to precisely the $\mathbf{k} = 0$ limit of the excitations described here.

4.4.3 Thermal Fluctuations

The breaking of rotational symmetry by the ferromagnetic state can be quantified by an order parameter - the magnetisation $\mathbf{M}$. The free energy of the system may in fact be expanded as a function of this order parameter. This powerful technique — known as a Ginzburg-Landau expansion — can reveal key generic properties of thermodynamic phases and transitions between them. Here we shall satisfy ourselves by considering the effect of thermal fluctuations upon the magnetisation of our system. Thermal fluctuation will reduce the magnetisation of the system up to a point where it vanishes at a critical temperature T_c, corresponding to a continuous phase transition.

The magnetisation per site (assuming polarisation in the z-direction, and in units of $\hbar$) is given by

$$M = \frac{1}{N}\sum_{\mathbf{r}}\langle\hat{S}^z_{\mathbf{r}}\rangle = S - \frac{1}{N}\sum_{\mathbf{r}}\langle\hat{b}^\dagger_{\mathbf{r}}\hat{b}_{\mathbf{r}}\rangle. \tag{4.18}$$

Since the Hamiltonian is diagonal in momentum space, we perform a Fourier transform,

$$M = S - \frac{1}{N^2}\sum_{\mathbf{r}}\sum_{\mathbf{k}\mathbf{k}'} e^{-i\mathbf{k}\mathbf{r}}e^{i\mathbf{k}'\mathbf{r}}\langle\hat{b}^\dagger_{\mathbf{k}}\hat{b}_{\mathbf{k}'}\rangle = S - \frac{1}{N}\sum_{\mathbf{k}}\langle\hat{b}^\dagger_{\mathbf{k}}\hat{b}_{\mathbf{k}}\rangle = S - \frac{1}{N}\sum_{\mathbf{k}}\frac{1}{e^{\beta\epsilon(\mathbf{k})}-1}, \tag{4.19}$$

where in the last step we have used that $\langle\hat{b}^\dagger_{\mathbf{k}}\hat{b}_{\mathbf{k}}\rangle$ is equal to the Bose function $n_B(\epsilon_{\mathbf{k}})$. In the thermodynamic limit $N \to \infty$ the sum over momenta becomes an integral over the Brillouin zone $\mathcal{BZ} = [-\pi/a, \pi/a]^d$ of volume $V_{\mathcal{BZ}} = (2\pi/a)^d$, $\frac{1}{N}\sum_{\mathbf{k}} \to \frac{1}{V_{\mathcal{BZ}}}\int_{\mathcal{BZ}} d^d\mathbf{k}$. Unfortunately, this integral can't be performed analytically.

However, we can draw important conclusions from the contribution to the integral from small momenta (IR limit) where $\epsilon(\mathbf{k}) \sim \mathbf{k}^2$. At any given temperature, we can expand the exponential function, as long as $|\mathbf{k}|$ is sufficiently small, resulting in the following IR contribution to the integral (assume from small d-dimensional ball of radius Λ around $\mathbf{k} = 0$),

$$\int d^d\mathbf{k}\frac{1}{e^{\beta\epsilon(\mathbf{k})}-1} \sim \int d^d\mathbf{k}\frac{1}{\mathbf{k}^2} \sim \int_0^\Lambda dk\frac{k^{d-1}}{k^2} = \int_0^\Lambda dk k^{d-3}. \tag{4.20}$$

This integral is IR divergent (divergence coming for $k \to 0$ limit) in $d \leq 2$. To be more precise, we encounter a power-law divergence in $d = 1$ and a logarithmic divergence in $d = 2$ This implies that in one and two dimensions thermal fluctuations are so strong that they prevent ferromagnetic order. This is an example of the Mermin-Wagner theorem which states that a continuous symmetry cannot be broken at finite temperature in $d \leq 2$.

In $d = 3$ the integral is convergent and the magnetisation will be finite at sufficiently low temperatures. We could determine the critical temperature T_c of the phase transition by solving the integral equation for $M = 0$,

$$S = \int_{[-\pi,\pi]^3} \frac{d^3\mathbf{q}}{(2\pi)^3} \frac{1}{e^{[3-\cos q_x - \cos q_y - \cos q_z]/t_c} - 1}, \tag{4.21}$$

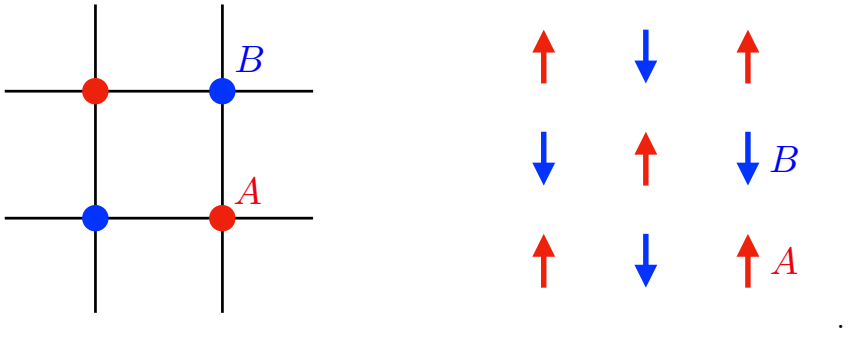
where we have substituted $q_i = ak_i$ and introduced the dimensionless critical temperature $t_c = T_c/(2J\hbar^2 S)$. This equation could be solved numerically for t_c, using the bisection method. As a next step one might numerically compute the magnetisation at temperatures t slightly below t_c and fit a power law $M \sim (t_c - t)^\beta$ to extract the order parameter critical exponent β.

4.5 The Heisenberg Antiferromagnet

Let us now consider the Heisenberg model with an antiferromagnetic nearest-neighbour coupling,

$$\hat{\mathcal{H}} = J \sum_{\langle i,j \rangle} \hat{\mathbf{S}}_i \cdot \hat{\mathbf{S}}_j, \tag{4.22}$$

with $J > 0$. As before, we consider a d-dimensional hypercubic lattice, which is a bibartite lattice. Such lattices can be divided into two sets A and B such that the sites from one set only have nearest neighbours from the other set. This is illustrated in the case of the square lattice in the figure below.



4.5.1 The Néel State

At the classical level, the energy of the Heisenberg antiferromagnet on a bipartite lattice is minimised when the spins on the A and B sublattices are anti-parallel, as shown for the square lattice in the figure. Such a checkerboard arrangement is called a Néel state. [3]

[3] Note that on the non-bipartite triangular lattice the groundstate is not a Néel state but a more complex non-collinear state with neighbouring spins enclosing an angle of 120°.

The total magnetisation of a Néel state is zero since the magnetisations on the A and B sublattices exactly cancel. Nevertheless, the spin rotation symmetry is clearly broken in this antiferromagnetically ordered state. We can define an order parameter, which is called the *staggered magnetisation*,

$$\mathbf{M} = \frac{1}{N}\sum_i \sigma_i \mathbf{m}_i, \quad \text{with } \sigma_i = \begin{cases} +1 & i \in A \\ -1 & i \in B \end{cases} \tag{4.23}$$

Unlike in the case of the ferromagnet the classical groundstate (Néel state)

$$|\psi\rangle = \otimes_i |S, \sigma_i S\rangle_i \tag{4.24}$$

is not an eigenstate of the Heisenberg Hamiltonian. To see this let us consider two neighbouring sites in the state

$$|\psi\rangle = |S, S\rangle_i \otimes |S, -S\rangle_j,$$

and act with the Hamiltonian on this state,

$$\begin{aligned} \hat{\mathcal{H}}_{ij}|\psi\rangle &= J\left[\hat{S}_i^z\hat{S}_j^z + \frac{1}{2}\left(\hat{S}_i^+\hat{S}_j^- + \hat{S}_i^-\hat{S}_j^+\right)\right]|S,S\rangle_i \otimes |S,-S\rangle_j \\ &= \underbrace{-J\hbar^2S^2}_{=E_0}|\psi\rangle + J\hbar^2 S|S,S-1\rangle_i \otimes |S,-S+1\rangle_j. \end{aligned}$$

The Néel state is an eigenstate of $\hat{S}_i^z\hat{S}_j^z$ with E_0 the classical ground-state energy. However, the action of $\hat{S}_i^+\hat{S}_j^- + \hat{S}_i^-\hat{S}_j^+$ results in an excited state. This is an important difference from the ferromagnet. In the antiferromagnet, there are quantum fluctuations around the classical groundstate.

4.5.2 Holstein-Primakoff Transformation

To study these quantum fluctuations as well as thermal fluctuations away from the Néel state we use the Holstein-Primakoff transformation. The transformation is applied with respect to the classical Néel groundstate which implies that we need a slightly different form of the transformation on the two sublattices.

Alternatively, we could rotate the spin operators on the B sublattice by 180°, such that in terms of the new spin operators the classical groundstate is ferromagnetic. In our convention, the classical groundstate is $\mathbf{S}_i = \sigma_i S\hat{\mathbf{e}}_z$. We therefore require a rotation on sublattice B that flips the sign of the z component of the spins, e.g. a rotation by 180° around the x axis,

$$\begin{aligned} \hat{T}_i^x &= \hat{S}_i^x \\ \hat{T}_i^y &= \sigma_i \hat{S}_i^y \\ \hat{T}_i^z &= \sigma_i \hat{S}_i^z, \end{aligned} \tag{4.25}$$

where it is straightforward to check that the new spin operators satisfy the commutation relations

$$[\hat{T}_i^\alpha, \hat{T}_i^\beta] = i\hbar\epsilon_{\alpha\beta\gamma}\hat{T}_i^\gamma, \tag{4.26}$$

as required. Note that on sublattice A we have $\hat{T}_i^\alpha = \hat{S}_i^\alpha$ since $\sigma_i = 1$. The classical groundstate in the new basis is ferromagnetic, $\mathbf{T}_i = S\hat{\mathbf{e}}_z$, and we can therefore use the same Holstein-Primakoff transformation on both sublattices. However, as a first step we need to transform the Hamiltonian,

$$\begin{aligned}\hat{\mathcal{H}} &= J\sum_{\langle i,j\rangle}\hat{\mathbf{S}}_i\cdot\hat{\mathbf{S}}_j = J\sum_{\langle i,j\rangle}\left(\hat{S}_i^x\hat{S}_j^x + S_i^yS_j^y + \hat{S}_i^z\hat{S}_j^z\right)\\ &= J\sum_{\langle i,j\rangle}\left(\hat{T}_i^x\hat{T}_j^x - T_i^yT_j^y - \hat{T}_i^z\hat{T}_j^z\right)\\ &= -J\sum_{\langle i,j\rangle}\left\{\hat{T}_i^z\hat{T}_j^z - \frac{1}{2}\left(\hat{T}_i^+\hat{T}_j^+ + \hat{T}_i^-\hat{T}_j^-\right)\right\}, \end{aligned} \tag{4.27}$$

where in the last step we have used $\hat{T}^\pm = \hat{T}^x \pm i\hat{T}^y$. We can now use the approximate Holstein-Primakoff transformation

$$\begin{aligned}\hat{T}_i^z &= \hbar(S - \hat{b}_i^\dagger\hat{b}_i),\\ \hat{T}_i^+ &\approx \hbar\sqrt{2S}b_i, \quad \hat{T}_i^- \approx \hbar\sqrt{2S}b_i^\dagger,\end{aligned} \tag{4.28}$$

to obtain the linear spin-wave Hamiltonian

$$\begin{aligned}\hat{\mathcal{H}} &= -J\hbar^2\sum_{\mathbf{r}}\sum_{i=1}^{d}\left\{(S-\hat{b}_{\mathbf{r}}^\dagger\hat{b}_{\mathbf{r}})(S-\hat{b}_{\mathbf{r}+\mathbf{a}_i}^\dagger\hat{b}_{\mathbf{r}+\mathbf{a}_i}) - S(\hat{b}_{\mathbf{r}}^\dagger\hat{b}_{\mathbf{r}+\mathbf{a}_i}^\dagger + \hat{b}_{\mathbf{r}}\hat{b}_{\mathbf{r}+\mathbf{a}_i})\right\}\\ &\approx -J\hbar^2S^2Nd + J\hbar^2S\sum_{\mathbf{r}}\sum_{i=1}^{d}\left(\hat{b}_{\mathbf{r}}^\dagger\hat{b}_{\mathbf{r}} + \hat{b}_{\mathbf{r}+\mathbf{a}_i}^\dagger\hat{b}_{\mathbf{r}+\mathbf{a}_i} + \hat{b}_{\mathbf{r}}^\dagger\hat{b}_{\mathbf{r}+\mathbf{a}_i}^\dagger + \hat{b}_{\mathbf{r}}\hat{b}_{\mathbf{r}+\mathbf{a}_i}\right).\end{aligned} \tag{4.29}$$

An important first step towards diagonalisation of the Hamiltonian is to perform a Fourier transform, $\hat{b}_{\mathbf{r}}^\dagger = \frac{1}{\sqrt{N}}\sum_{\mathbf{k}}e^{-i\mathbf{kr}}\hat{b}_{\mathbf{k}}^\dagger$ and $\hat{b}_{\mathbf{r}} = \frac{1}{\sqrt{N}}\sum_{\mathbf{k}}e^{i\mathbf{kr}}\hat{b}_{\mathbf{k}}$, resulting in

$$\hat{\mathcal{H}} = -J\hbar^2S^2Nd + J\hbar^2S\sum_{\mathbf{k}}\sum_{i=1}^{d}\left(2\hat{b}_{\mathbf{k}}^\dagger\hat{b}_{\mathbf{k}} + e^{i\mathbf{ka}_i}\hat{b}_{\mathbf{k}}^\dagger\hat{b}_{-\mathbf{k}}^\dagger + e^{-i\mathbf{ka}_i}\hat{b}_{\mathbf{k}}\hat{b}_{-\mathbf{k}}\right). \tag{4.30}$$

In order to write the Hamiltonian in 2x2 Nambu spinor space we use that $\sum_{\mathbf{k}}f(\mathbf{k}) = \frac{1}{2}\sum_{\mathbf{k}}[f(\mathbf{k}) + f(-\mathbf{k})]$,

$$\begin{aligned}\hat{\mathcal{H}} &= -J\hbar^2S^2Nd + J\hbar^2S\sum_{\mathbf{k}}\sum_{i=1}^{d}\left\{b_{\mathbf{k}}^\dagger\hat{b}_{\mathbf{k}} + \underbrace{b_{-\mathbf{k}}^\dagger\hat{b}_{-\mathbf{k}}}_{=\hat{b}_{-\mathbf{k}}b_{-\mathbf{k}}^\dagger - 1} + \cos(\mathbf{ka}_i)\left(\hat{b}_{\mathbf{k}}^\dagger\hat{b}_{-\mathbf{k}}^\dagger + \hat{b}_{\mathbf{k}}\hat{b}_{-\mathbf{k}}\right)\right\}\\ &= -J\hbar^2S(S+1)Nd + J\hbar^2Sd\sum_{\mathbf{k}}(\hat{b}_{\mathbf{k}}^\dagger, \hat{b}_{-\mathbf{k}})\begin{pmatrix}1 & \gamma(\mathbf{k})\\ \gamma(\mathbf{k}) & 1\end{pmatrix}\begin{pmatrix}\hat{b}_{\mathbf{k}}\\ \hat{b}_{-\mathbf{k}}^\dagger\end{pmatrix}\end{aligned} \tag{4.31}$$

where we have defined

$$\gamma(\mathbf{k}) = \frac{1}{d}\sum_{i=1}^{d}\cos(\mathbf{ka}_i). \tag{4.32}$$

4.5.3 Bogoliubov Transformation

Evidently the Fourier transform alone is not enough to diagonalise the Hamiltonian. The anomalous terms such as $\hat{b}_{\mathbf{k}}\hat{b}_{-\mathbf{k}}$ require us to use a bosonic Bogoliubov transformation as developed in Sec. 2.6.2 and used in the context of the weakly interacting Bose gas, Sec. 3.3. Rather than going through this step-by-step again, we simply use the established results. The Hamiltonian becomes diagonal in terms of the new boson operators $\hat{d}^\dagger_{\mathbf{k}}$, $\hat{d}_{\mathbf{k}}$,

$$\begin{pmatrix} \hat{b}_k \\ \hat{b}^\dagger_{-\mathbf{k}} \end{pmatrix} = \begin{pmatrix} \cosh(\theta_{\mathbf{k}}) & \sinh(\theta_{\mathbf{k}}) \\ \sinh(\theta_{\mathbf{k}}) & \cosh(\theta_{\mathbf{k}}) \end{pmatrix} \begin{pmatrix} \hat{d}_{\mathbf{k}} \\ \hat{d}^\dagger_{-\mathbf{k}} \end{pmatrix}, \tag{4.33}$$

with $\tanh(2\theta_{\mathbf{k}}) = -\frac{\gamma(\mathbf{k})}{1} = -\gamma(\mathbf{k})$ and takes the diagonal form

$$\begin{aligned} \hat{\mathcal{H}} &= -J\hbar^2 S(S+1)Nd + J\hbar^2 Sd \sum_{\mathbf{k}} (\hat{d}^\dagger_{\mathbf{k}}, \hat{d}_{-\mathbf{k}}) \begin{pmatrix} \sqrt{1-\gamma^2(\mathbf{k})} & 0 \\ 0 & \sqrt{1-\gamma^2(\mathbf{k})} \end{pmatrix} \begin{pmatrix} \hat{d}_{\mathbf{k}} \\ \hat{d}^\dagger_{-\mathbf{k}} \end{pmatrix} \\ &= 2J\hbar^2 Sd \sum_{\mathbf{k}} \sqrt{1-\gamma^2(\mathbf{k})} \hat{d}^\dagger_{\mathbf{k}} \hat{d}_{\mathbf{k}} \\ &\quad + J\hbar^2 Sd \sum_{\mathbf{k}} \sqrt{1-\gamma^2(\mathbf{k})} - J\hbar^2 S(S+1)Nd. \end{aligned} \tag{4.34}$$

Hence the linear spin-wave dispersion of the quantum antiferromagnet on the d-dimensional hypercubic lattice is given by

$$\epsilon(\mathbf{k}) = 2J\hbar^2 Sd\sqrt{1-\gamma^2(\mathbf{k})}. \tag{4.35}$$

Using that for small $|\mathbf{k}|$, $\gamma(\mathbf{k}) = \frac{1}{d}\sum_{i=1}^d \cos(ak_i) \approx \frac{1}{d}\sum_{i=1}^d (1 - \frac{1}{2}a^2 k_i^2) = 1 - \frac{a^2}{2d}\mathbf{k}^2$, we obtain

$$\epsilon(\mathbf{k}) \approx 2J\hbar^2 Sd\sqrt{1-\left(1-\frac{a^2}{2d}\mathbf{k}^2\right)^2} \approx 2J\hbar^2 Sd\sqrt{\frac{a^2}{d}\mathbf{k}^2} = 2J\hbar^2 Sa\sqrt{d}|\mathbf{k}|. \tag{4.36}$$

Note that unlike for the ferromagnet, the spin-wave dispersion for the antiferromagnet is linear at small momenta. Identifying $\epsilon(\mathbf{k}) = \hbar c|\mathbf{k}|$, we obtain the spin-wave velocity

$$c = 2J\hbar Sa\sqrt{d}. \tag{4.37}$$

4.5.4 Fluctuations of the Antiferromagnet

As we have already noted, the classical Néel state is not an eigenstate of the anti-ferromagnetic Heisenberg model. This has several consequences. Firstly, it leads to a zero-point contribution to the energy. These zero-point fluctuations also reduce the sub-lattice magnetisation even at zero temperature. In addition, we will have a reduction by thermal fluctuations.

The average staggered magnetisation per site is given by ($\hbar = 1$)

$$M = \frac{1}{N}\sum_{\mathbf{r}} \sigma_{\mathbf{r}} \langle \hat{S}^z_{\mathbf{r}} \rangle = \frac{1}{N}\sum_{\mathbf{r}} \langle \hat{T}^z_{\mathbf{r}} \rangle = S - \frac{1}{N}\sum_{\mathbf{r}} \langle \hat{b}^\dagger_{\mathbf{r}} \hat{b}_{\mathbf{r}} \rangle = S - \frac{1}{N}\sum_{\mathbf{k}} \langle \hat{b}^\dagger_{\mathbf{k}} \hat{b}_{\mathbf{k}} \rangle. \tag{4.38}$$

We can evaluate the expectation value by inserting the Bogoliubov transformation to the diagonal basis,

$$\begin{aligned}\langle \hat{b}^\dagger_{\mathbf{k}} \hat{b}_{\mathbf{k}} \rangle &= \left\langle \left(\cosh(\theta_{\mathbf{k}}) \hat{d}^\dagger_{\mathbf{k}} + \sinh(\theta_{\mathbf{k}}) \hat{d}_{-\mathbf{k}} \right) \left(\cosh(\theta_{\mathbf{k}}) \hat{d}_{\mathbf{k}} + \sinh(\theta_{\mathbf{k}}) \hat{d}^\dagger_{-\mathbf{k}} \right) \right\rangle \\ &= \cosh^2(\theta_{\mathbf{k}}) \langle \hat{d}^\dagger_{\mathbf{k}} \hat{d}_{\mathbf{k}} \rangle + \sinh^2(\theta_{\mathbf{k}}) \langle \hat{d}_{-\mathbf{k}} \hat{d}^\dagger_{-\mathbf{k}} \rangle \\ &= \cosh(2\theta_{\mathbf{k}}) \langle \hat{d}^\dagger_{\mathbf{k}} \hat{d}_{\mathbf{k}} \rangle + \sinh^2(\theta_{\mathbf{k}}). \end{aligned} \tag{4.39}$$

Using that $\langle \hat{d}^\dagger_{\mathbf{k}} \hat{d}_{\mathbf{k}} \rangle = n_B(\epsilon_{\mathbf{k}})$ is simply the Bose function, we obtain the final expression for the staggered magnetisation,

$$M = S - \frac{1}{N} \sum_{\mathbf{k}} \left(\cosh(2\theta_{\mathbf{k}}) \frac{1}{e^{\beta\epsilon(\mathbf{k})} - 1} + \sinh^2(\theta_{\mathbf{k}}) \right). \tag{4.40}$$

The first term in the sum corresponds to the reduction of the order parameter by thermal fluctuations. At $T = 0$, this contribution vanishes and we are left with the second contribution which reflects the reduction of the staggered magnetisation by zero-point quantum fluctuations. Using that

$$\sinh^2(\theta_{\mathbf{k}}) = \frac{1}{2} \left(\frac{1}{\sqrt{1 - \tanh^2(2\theta_{\mathbf{k}})}} - 1 \right) = \frac{1}{2} \left(\frac{1}{\sqrt{1 - \gamma^2(\mathbf{k})}} - 1 \right), \tag{4.41}$$

the expression for the staggered magnetisation at $T = 0$ becomes

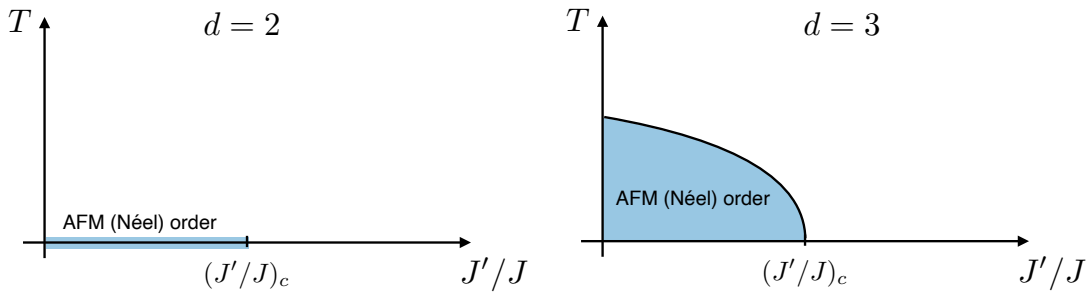
$$M = S - \frac{1}{2V_{\mathcal{BZ}}} \int_{\mathcal{BZ}} d^d\mathbf{k} \left(\frac{1}{\sqrt{1 - \gamma^2(\mathbf{k})}} - 1 \right), \tag{4.42}$$

where we have taken the thermodynamic limit and replaced the momentum sum by an integral over the Brillouin zone $\mathcal{BZ} = [-\pi/a, \pi/a]^d$. Let us investigate the small $\mathbf{k}$ IR contribution to the integral:

$$\int d^d\mathbf{k} \left(\frac{1}{\sqrt{1 - \gamma^2(\mathbf{k})}} - 1 \right) \sim \int_0^\Lambda dk k^{d-1} \frac{1}{k} = \int_0^\Lambda dk k^{d-2} \tag{4.43}$$

This result shows that the integral is infra-red convergent for $d \geq 2$, but logarithmically divergent in $1d$ (it always converges in the ultra-violet because of the cut-off at the Brillouin zone edge). Hence, for $d \geq 2$ antiferromagnetic order is possible at $T = 0$, but in $d = 1$ quantum fluctuations are too strong for the system to order.

Let us further discuss the difference between two- and three-dimensional antiferromagnets. An analysis of the thermal fluctuations reveals that the corresponding momentum integral is IR divergent in $d = 2$ and convergent in $d = 3$. While in three dimensions we can have antiferromagnetic order for $T > 0$, in two-dimensional systems this is not possible and in fact ruled out by the Mermin Wagner theorem. Typical phase diagrams in $d = 2$ and $d = 3$ are shown below, where we have introduced an additional control parameter that destabilises the magnetic order. This could be an antiferromagnetic second-neighbour exchange J', which causes magnetic frustration.



4.6 One-Dimensional Spin Chains

As we have seen in the previous Section, in one spatial dimension fluctuations destroy any potential ferromagnetic or antiferromagnet order, even at $T = 0$. The same is true for any order associated with spontaneous continuous symmetry breaking. Nevertheless, despite the lack of ordering the physics of one-dimensional spin chains is very rich, including phenomena such as fractionalisation and topological order.

4.6.1 Haldane's Conjecture

Let us consider a one-dimensional quantum antiferromagnet, described by the Hamiltonian

$$\hat{\mathcal{H}} = J \sum_i \hat{\mathbf{S}}_i \cdot \hat{\mathbf{S}}_{i+1}, \tag{4.44}$$

where i labels the sites in the chain. It was first realised by Duncan Haldane [Phys. Rev. Lett **50**, 1153 (1983)] that the physics crucially depends on the value S of the spin operators. He considered topological effects of quantum spins, known as instantons, which can be understood in terms of a spin-coherent-state path integral. Haldane concluded that for half-integer spins, $S = 1/2, 3/2, 5/2, \ldots$, the spectrum of excitations remains gapless, while for integer spins, $S = 1, 2, 3, \ldots$, there is an energy gap between the ground state and the first excited state.

It is now well understood and rigorously proven that the antiferromagnetic $S = 1$ chain (often referred to as the "Haldane spin chain") has a gapped excitation spectrum and hidden topological order, Such topological order is not characterised by a local order parameter, such as the magnetisation, but instead by a non-local string order parameter. Another feature of topological order is the occurrence of edge states at the ends of the chain. We will briefly discuss this at the end of this Chapter.

4.6.2 Jordan-Wigner Transformation for $S = 1/2$ Spin Chains

We will now focus on $S = 1/2$ spin chains and investigate the excitation spectrum of the one-dimensional quantum antiferromagnet, which according to Haldane should be gapless. The one-dimensional spin chain has no long-range magnetic order and we therefore can't use linear spin-wave theory (LSWT) to expand in terms of small transverse fluctuations around a classical ground state. Moreover, LSWT is a controlled approximation for large S and not applicable in the extreme quantum limit $S = 1/2$.

A spin $1/2$ operator has two possible eigenstates, $|\uparrow\rangle = |1/2, 1/2\rangle$ and $|\downarrow\rangle = |1/2, -1/2\rangle$. Rather than representing the spin operators by bosonic creation and annihilation operators and imposing a constraint on the bosonic Hilbert space, we will use *spinless fermion operators*. Because of the Pauli principle, we can only have a maximum of one fermion on a given site, automatically taking care of the constraint. Identifying

$$\begin{aligned} |\uparrow\rangle &= |0\rangle \quad \text{no fermion,} \\ |\downarrow\rangle &= |1\rangle \quad \text{1 fermion,} \end{aligned}$$

we obtain a simple relation between $\hat{S}^z$ and the fermionic number operator $\hat{n} = \hat{c}^\dagger \hat{c}$,

$$\hat{S}^z = \frac{\hbar}{2}(1 - 2\hat{c}^\dagger \hat{c}). \tag{4.45}$$

The spin raising and lowering operators $\hat{S}^+$ and $\hat{S}^-$ act in the same way as the fermionic annihilation and creation operators $\hat{c}$ and $\hat{c}^\dagger$:

$$\begin{array}{ll} \hat{S}^+|\uparrow\rangle = 0 & \hat{c}|0\rangle = 0 \\ \hat{S}^-|\uparrow\rangle = \hbar|\downarrow\rangle & \hat{c}^\dagger|0\rangle = |1\rangle \\ \hat{S}^+|\downarrow\rangle = \hbar|\uparrow\rangle & \hat{c}|1\rangle = |0\rangle \\ \hat{S}^-|\downarrow\rangle = 0 & \hat{c}^\dagger|1\rangle = 0 \end{array} \tag{4.46}$$

This suggests that we should identify $\hat{S}^+ = \hbar\hat{c}$ and $\hat{S}^- = \hbar\hat{c}^\dagger$. This is indeed consistent with the fermionic anti-commutation and spin commutation relations. Let us for example assume that $\{\hat{c}, \hat{c}^\dagger\} = 1$ and check if the spin commutation relation is satisfied:

$$[\hat{S}^+, \hat{S}^-] = \hbar^2[\hat{c}, \hat{c}^\dagger] = \hbar^2(\underbrace{\hat{c}\hat{c}^\dagger}_{=1-\hat{c}^\dagger\hat{c}} - \hat{c}^\dagger\hat{c}) = \hbar^2(1 - 2\hat{c}^\dagger\hat{c}) = 2\hbar\hat{S}^z,$$

as required. We have therefore established a mapping between spin-$1/2$ operators and spinless fermion operators that works perfectly for a single site. However, spin operators on different sites commute while fermion operators anti-commute. As first realised by Jordan and Wigner, this mismatch can be repaired by introducing a semi-infinite string operator,

$$\hat{D}_i = \prod_{\ell<i}(1 - 2\hat{c}_\ell^\dagger \hat{c}_\ell), \tag{4.47}$$

and defining the mapping for a given site i on the chain as

$$\begin{aligned} \hat{S}_i^z &= \frac{\hbar}{2}(1 - 2\hat{c}_i^\dagger \hat{c}_i), \\ \hat{S}_i^+ &= \hbar\hat{D}_i\hat{c}_i, \quad \hat{S}_i^- = \hbar\hat{c}_i^\dagger\hat{D}_i, \end{aligned} \tag{4.48}$$

which is known as the Jordan-Wigner transformation (JWT). Note that in the last equation we used that $\hat{D}_i^\dagger = \hat{D}_i$. Since the string operator $\hat{D}_i$ only involves sites to the left of i, not the site i itself, and since it only contains pairs of fermion operators, we obtain $[\hat{c}_i, \hat{D}_i] = [\hat{c}_i^\dagger, \hat{D}_i] = 0$. Another useful property of the string operator is that $\hat{D}_i^2 = 1$, which is easy to prove:

$$\hat{D}_i^2 = \prod_{\ell<i}(1 - 2\hat{c}_\ell^\dagger\hat{c}_\ell)^2 = \prod_{\ell<i}(1 - 4\hat{c}_\ell^\dagger\hat{c}_\ell + 4\hat{c}_\ell^\dagger \underbrace{\hat{c}_\ell\hat{c}_\ell^\dagger}_{1-\hat{c}_\ell^\dagger\hat{c}_\ell} \hat{c}_\ell) = \prod_{\ell<i}(1 - 4\underbrace{\hat{c}_\ell^\dagger\hat{c}_\ell^\dagger\hat{c}_\ell\hat{c}_\ell}_{=0}) = 1.$$

A proof that the JWT is consistent with spin-commutation and fermion anti-commutation relations is given in the box below.

Let us assume that the creation and annihilation operators satisfy the fermionic anti-commutation relations, $\{\hat{c}_i, \hat{c}_j\} = \{\hat{c}_i^\dagger, \hat{c}_j^\dagger\} = 0$, $\{\hat{c}_i, \hat{c}_j^\dagger\} = \delta_{ij}$. We will prove that the spin operators defined through the JWT indeed satisfy $[\hat{S}_i^+, \hat{S}_j^-] = 2\delta_{ij}\hbar\hat{S}_i^z$.

(i) case $i = j$:

$$\begin{aligned}[\hat{S}_i^+, \hat{S}_i^-] &= \hbar^2[\hat{D}_i\hat{c}_i, \hat{c}_i^\dagger\hat{D}_i] = \hbar^2(\hat{D}_i\hat{c}_i\hat{c}_i^\dagger\hat{D}_i - \hat{c}_i^\dagger\hat{D}_i\hat{D}_i\hat{c}_i) = \hbar^2(\hat{c}_i\hat{D}_i^2\hat{c}_i^\dagger - \hat{c}_i^\dagger\hat{D}_i^2\hat{c}_i) \\ &= \hbar^2(\hat{c}_i\hat{c}_i^\dagger - \hat{c}_i^\dagger\hat{c}_i) = \hbar^2(1 - 2\hat{c}_i^\dagger\hat{c}_i) = 2\hbar\hat{S}_i^z.\end{aligned}$$

(ii) case $i < j$ ($j = i + n$ with $n \geq 1$)

$$\begin{aligned}[\hat{S}_i^+, \hat{S}_{i+n}^-] &= \hbar^2[\hat{D}_i\hat{c}_i, \hat{c}_{i+n}^\dagger\hat{D}_{i+n}] = \hbar^2(\hat{D}_i\hat{c}_i\hat{c}_{i+n}^\dagger\hat{D}_{i+n} - \hat{c}_{i+n}^\dagger\hat{D}_{i+n}\hat{D}_i\hat{c}_i) \\ &= \hbar^2(\hat{c}_i\hat{D}_i\hat{D}_{i+n}\hat{c}_{i+n}^\dagger - \hat{c}_{i+n}^\dagger\hat{D}_{i+n}\hat{D}_i\hat{c}_i) \\ &= \hbar^2\left[\hat{c}_i\prod_{\ell=i}^{i+n-1}(1 - 2\hat{c}_\ell^\dagger\hat{c}_\ell)\hat{c}_{i+n}^\dagger - \hat{c}_{i+n}^\dagger\prod_{\ell=i}^{i+n-1}(1 - 2\hat{c}_\ell^\dagger\hat{c}_\ell)\hat{c}_i\right] \\ &= \hbar^2\left[\hat{c}_i(1 - 2\hat{c}_i^\dagger\hat{c}_i)\prod_{\ell=i+1}^{i+n-1}(1 - 2\hat{c}_\ell^\dagger\hat{c}_\ell)\hat{c}_{i+n}^\dagger\right. \\ &\qquad \left. -\hat{c}_{i+n}^\dagger\prod_{\ell=i+1}^{i+n-1}(1 - 2\hat{c}_\ell^\dagger\hat{c}_\ell)(1 - 2\hat{c}_i^\dagger\hat{c}_i)\hat{c}_i\right] \\ &\overset{(*)}{=} -\hbar^2\left[\hat{c}_i\prod_{\ell=i+1}^{i+n-1}(1 - 2\hat{c}_\ell^\dagger\hat{c}_\ell)\hat{c}_{i+n}^\dagger + \hat{c}_{i+n}^\dagger\prod_{\ell=i+1}^{i+n-1}(1 - 2\hat{c}_\ell^\dagger\hat{c}_\ell)\hat{c}_i\right] \\ &= -\hbar^2\prod_{\ell=i+1}^{i+n-1}(1 - 2\hat{c}_\ell^\dagger\hat{c}_\ell)\underbrace{(\hat{c}_i\hat{c}_{i+n}^\dagger + \hat{c}_{i+n}^\dagger\hat{c}_i)}_{=\{\hat{c}_i, \hat{c}_{i+n}^\dagger\}=0} = 0.\end{aligned}$$

$$\begin{aligned}(*)\quad \hat{c}_i(1 - 2\hat{c}_i^\dagger\hat{c}_i) &= \hat{c}_i - 2\underbrace{\hat{c}_i\hat{c}_i^\dagger}_{1-\hat{c}_i^\dagger\hat{c}_i}\hat{c}_i = -\hat{c}_i + 2\underbrace{\hat{c}_i^\dagger\hat{c}_i\hat{c}_i}_{=0} = -\hat{c}_i, \\ (1 - 2\hat{c}_i^\dagger\hat{c}_i)\hat{c}_i &= \hat{c}_i - 2\underbrace{\hat{c}_i^\dagger\hat{c}_i\hat{c}_i}_{=0} = \hat{c}_i.\end{aligned}$$

(ii) case $i > j$ ($i = j + n$ with $n \geq 1$): $[\hat{S}_{j+n}^+, \hat{S}_j^-] = [\hat{S}_j^+, \hat{S}_{j+n}^-]^\dagger = 0^\dagger = 0$.

4.6.3 The Antiferromagnetic Spin-1/2 Chain

We consider a spin-1/2 chain with antiferromagnetic nearest neighbour exchange. For reasons that will become clear later, we allow the coupling between the z components of the spins to be different from the coupling between x and y components, $J_x = J_y = J$ and $J_z = \alpha J$. Such an anisotropy is called *exchange anisotropy*. The Hamiltonian is given by

$$\hat{\mathcal{H}} = J \sum_i \left(\hat{S}_i^x \hat{S}_{i+1}^x + \hat{S}_i^y \hat{S}_{i+1}^y + \alpha \hat{S}_i^z \hat{S}_{i+1}^z \right). \tag{4.49}$$

For $\alpha = 1$ the Hamiltonian describes the O(3) Heisenberg model, for $\alpha = 0$ it reduces to an O(2) XY model. We will use the Jordan-Wigner transformation, Eq. (4.48) to express the Hamiltonian in terms of spinless fermion operator. Let us first apply the JWT to the terms involving x and y spin components,

$$\begin{aligned}
\hat{S}_i^x \hat{S}_{i+1}^x + \hat{S}_i^y \hat{S}_{i+1}^y &= \frac{1}{2}\left(\hat{S}_i^+ \hat{S}_{i+1}^- + \hat{S}_i^- \hat{S}_{i+1}^+ \right) \\
&= \frac{\hbar^2}{2}\left(\hat{D}_i \hat{c}_i \hat{c}_{i+1}^\dagger \hat{D}_{i+1} + \hat{c}_i^\dagger \hat{D}_i \hat{D}_{i+1} \hat{c}_{i+1} \right) \\
&= \frac{\hbar^2}{2}\big[\underbrace{\hat{c}_i(1 - 2\hat{c}_i^\dagger \hat{c}_i)}_{=-\hat{c}_i \text{ (*) in box}} \hat{c}_{i+1}^\dagger + \underbrace{\hat{c}_i^\dagger(1 - 2\hat{c}_i^\dagger \hat{c}_i)}_{=\hat{c}_i^\dagger} \hat{c}_{i+1} \big] \\
&= \frac{\hbar^2}{2}\left(-\hat{c}_i \hat{c}_{i+1}^\dagger + \hat{c}_i^\dagger \hat{c}_{i+1} \right) = \frac{\hbar^2}{2}\left(\hat{c}_i^\dagger \hat{c}_{i+1} + \hat{c}_{i+1}^\dagger \hat{c}_i \right).
\end{aligned}$$

For the z components we obtain

$$\begin{aligned}
\hat{S}_i^z \hat{S}_{i+1}^z &= \frac{\hbar^2}{4}(1 - 2\hat{c}_i^\dagger \hat{c}_i)(1 - 2\hat{c}_{i+1}^\dagger \hat{c}_{i+1}) \\
&= \frac{\hbar^2}{4} - \frac{\hbar^2}{2}(\hat{c}_i^\dagger \hat{c}_i + \hat{c}_{i+1}^\dagger \hat{c}_{i+1}) + \hbar^2 c_i^\dagger \hat{c}_i \hat{c}_{i+1}^\dagger \hat{c}_{i+1}.
\end{aligned}$$

Putting everything together we obtain the Hamiltonian

$$\hat{\mathcal{H}} = \frac{\hbar^2}{4}\alpha J N + \frac{\hbar^2}{2} J \sum_i \left[\hat{c}_i^\dagger \hat{c}_{i+1} + \hat{c}_{i+1}^\dagger \hat{c}_i - \alpha(\hat{c}_i^\dagger \hat{c}_i + \hat{c}_{i+1}^\dagger \hat{c}_{i+1}) \right] + \hbar^2 \alpha J \sum_i c_i^\dagger \hat{c}_i \hat{c}_{i+1}^\dagger \hat{c}_{i+1}. \tag{4.50}$$

The last term in the Hamiltonian describes a repulsive interaction between fermions on neighbouring lattice sites. The strength of the interaction is tuneable by the exchange anisotropy α and vanishes in the limit of the XY chain, $\alpha = 0$. In the following, we will consider this limit, in which the Hamiltonian turns into one of non-interacting spinless fermions

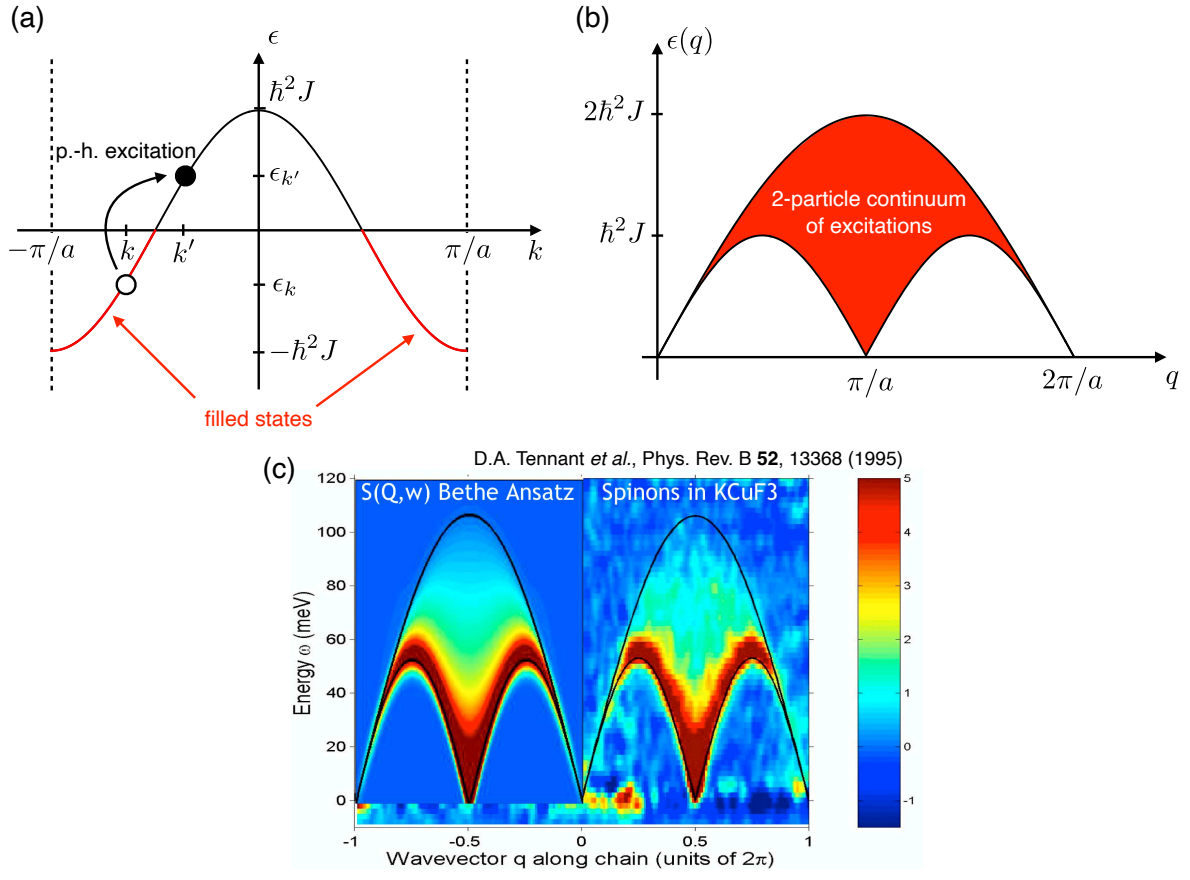
$$\hat{\mathcal{H}} = \frac{\hbar^2}{2} J \sum_i \left(\hat{c}_i^\dagger \hat{c}_{i+1} + \hat{c}_{i+1}^\dagger \hat{c}_i \right). \tag{4.51}$$

The term $\hat{c}_{i+1}^\dagger \hat{c}_i$ describes the hopping of a fermion from site i to site $i+1$, whereas $\hat{c}_i^\dagger \hat{c}_{i+1}$ is the hopping process in the opposite direction. The hopping amplitude is equal to $t = \frac{\hbar^2}{2} J$. This non-interacting Hamiltonian can be diagonalised by Fourier transform: (in

order not to confuse the site label with the i in the complex Fourier transform, we use n to label the sites in the chain)

$$\begin{aligned}
\hat{\mathcal{H}} &= t\sum_n \left(\hat{c}_n^\dagger \hat{c}_{n+1} + \hat{c}_{n+1}^\dagger \hat{c}_n\right) \\
&= t\frac{1}{N}\sum_n \sum_{k,q} \left(e^{-ikna} e^{iq(n+1)a} \hat{c}_k^\dagger \hat{c}_q + e^{-ik(n+1)a} e^{iqna} \hat{c}_k^\dagger \hat{c}_q\right) \\
&= t\sum_{k,q} \underbrace{\frac{1}{N}\sum_n e^{-i(k-q)na}}_{=\delta_{k,q}} \left(e^{iqa} + e^{-ika}\right) \hat{c}_k^\dagger \hat{c}_q \\
&= t\sum_k \left(e^{ika} + e^{-ika}\right) \hat{c}_k^\dagger \hat{c}_k = \sum_k \epsilon_k \hat{c}_k^\dagger \hat{c}_k, \qquad (4.52)
\end{aligned}$$

with $\epsilon_k = 2t\cos(ka) = \hbar^2 J\cos(ka)$. The dispersion is shown in Fig.(a) below.



Since we are dealing with fermions, the groundstate is given by the filled Fermi sea, corresponding to the state where all the single-particle momentum states with negative

energy (below the chemical potential $\mu = 0$) are occupied.[4] The excitations out of this ground state are so-called particle-hole excitations: a particle from below the Fermi level is excited into an empty state above the Fermi level [see Fig. (a)].

Since the momentum and energy of the excitation can be split in various ways between the particle and the hole the excitation spectrum will be continuous and not be given by a sharp energy-momentum relation, as for the magnon quasiparticle excitations in an ordered magnet. This is an example of fractionalisation of an excitation into a pair of particles, resulting in a two-particle continuum of excitations.

To understand this in more detail, let's assume that the particle has momentum k' and energy $\epsilon_{k'} = \hbar^2 J \cos(k'a)$, and the hole momentum k and energy $\epsilon_k = \hbar^2 J \cos(ka)$, resulting in a particle-hole excitation of momentum $q = k' - k$ and energy $\epsilon(q) = \epsilon_{k'} - \epsilon_k = \hbar^2 J[\cos(k'a) - \cos(ka)]$. Let's for example take $q = \pi/a$. From Fig. (a) it is immediately clear that the smaller possible energy in this case is for $k = -\pi/(2a)$ and $k' = \pi/(2a)$, resulting in $\epsilon_{\min}(\pi/a) = 0 - 0 = 0$. On the other, if we take $k = -\pi/a$ and $k' = 0$ we obtain the largest possible excitation energy $\epsilon_{\max}(\pi/a) = \hbar^2 J - (-\hbar^2 J) = 2\hbar^2 J$. By shifting k and k' while keeping $q = k' - k = \pi/a$ fixed we can reach all energies between 0 and $2\hbar^2 J$

Consider now a general $q < \pi/a$. Since the slope of ϵ_k decreases as we move away from the node at $-\pi/(2a)$, the largest possible excitation energy is obtained if k and k' are symmetric around the node, $k = -\pi/(2a) - q/2$ and $k' = -\pi/(2a) + q/2$, resulting in

$$\begin{aligned} \epsilon_{\max}(q) &= \hbar^2 J[\cos(qa/2 - \pi/2) - \cos(-qa/2 - \pi/2)] = \hbar^2 J[\sin(qa/2) - \sin(-qa/2)] \\ &= 2\hbar^2 J \sin(qa/2). \end{aligned}$$

The smallest excitation energy is obtained by taking $k = -\pi/(2a)$ and $k' = -\pi/(2a) + q$, resulting in

$$\epsilon_{\min}(q) = \hbar^2 J[\cos(qa - \pi/2) - \cos(-\pi/2)] = \hbar^2 J \sin(qa).$$

$\epsilon_{\max}(q)$ and $\epsilon_{\min}(q)$ can be expanded to the full range $q \in [0, 2\pi/a]$ by taking absolute values. The resulting continuum of excitations is shown in Fig. (b). Such two-particle continua are indeed observed in inelastic neutron scattering experiments on quasi one-dimensional quantum antiferromagnets, as shown in Fig. (c).

4.6.4 Spin-1 Haldane Spin Chain

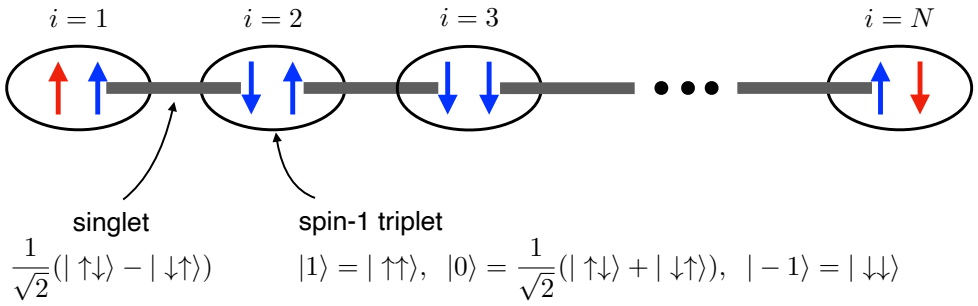
In the above, we were able to discuss the properties of the spin 1/2 chain by considering the simplified, XY spin chain. Similarly, a simplified version of the $S = 1$ chain enables us to appreciate some of its properties. This is known as the AKLT model (after Affleck, Kennedy, Lieb and Tasaki). The Hamiltonian is given by

$$\hat{\mathcal{H}} = J \sum_i \left\{ \hat{\mathbf{S}}_i \cdot \hat{\mathbf{S}}_{i+1} + \alpha (\hat{\mathbf{S}}_i \cdot \hat{\mathbf{S}}_{i+1})^2 \right\},$$

where the original spin-1 antiferromagnetic chain (the Haldane spin chain) is obtained for $\alpha = 0$. At first glance, the additional bi-quadratic term in the Hamiltonian seems to be

[4] This corresponds to exactly half filling, which reflects that there is on average the same number of up- and down spins in the system.

a further complication. However, as realised by Affleck, Kennedy, Lieb and Tasaki, the model is exactly solvable in the case $\alpha = 1/3$, using a cunning trick: The $S = 1$ spin on each site is represented by the triplet sector of two fractional $S = 1/2$ spins. It turns out that the groundstate of this model can be constructed by putting two of these fractional spins on neighbouring sites into a singlet state. This can be illustrated schematically as



The red and blue arrows represent a possible realisation of the spins after a projective measurement. In this particular example, site 1 is in the $m = 1$ triplet state, site 2 in the $m = 0$ state, site 3 in the $m = -1$ state and so on. It turns out that because of the constraint that the fractional spin-1/2 on neighbouring sites form singlets, the possible groundstates have a hidden antiferromagnetic order with an arbitrary number of $m = 0$ states inserted, e.g. $1, 0, -1, 0, 0, 0, 1, 0, -1, 1, 0, 0, 0, 0, 0, -1, \ldots$. This hidden order is an example of topological order. It is not characterised by a local order parameter like the staggered magnetisation, but instead by a non-local string order parameter.

An important signature of topological order is the presence of edge states. In the case of the AKLT chain these are the fractional S=1/2 spins at the ends of the chain (shown in red), which are not in a singlet configuration with a partner from a neighbouring site. The ground state of the AKLT model has an excitation gap to the first excited state, which can be understood as the energy required to break the singlet correlation between sites. The topological order is protected by the energy gap in the sense that small perturbation that do not result in a closing of the gap can't destroy the topological order. Numerical calculations show that the gap remains finite if α is reduced from $\alpha = 1/3$ to $\alpha = 0$. This proves that the Haldane spin chain is gapped and has topological order.

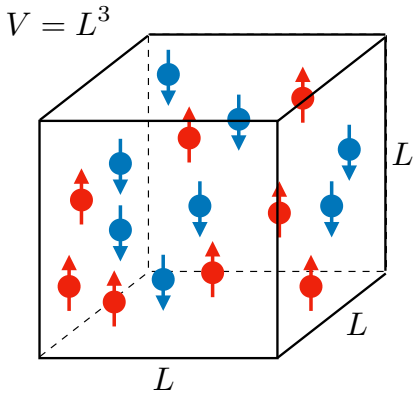
Chapter 5

The free Fermi gas

We will first review the *free Fermi gas* as a reference point and then consider a sequence of modifications to the non-interacting gas that lead to significant changes in physical behaviour. The modifications we discuss are the presence of *interactions* and the effects of a *lattice potential*.

The most important physical example of a system of fermions in condensed matter physics is the *electron gas in a metal*, but it is good to keep in mind as well examples from astrophysics (white dwarfs and neutron stars) and terrestrial low-temperature physics, in particular the case of liquid ^{3}He. The last example is particularly simple as it is a system that is rotationally symmetric and translationally invariant, because there exists no background lattice of neutralising ions.

Let us consider a gas of *non-interacting* fermions (e.g. electrons) with mass m in a cubic box ($d = 3$) with periodic boundary conditions.



The total number of particles is $N = N_\uparrow + N_\downarrow$ with $N_\uparrow = N_\downarrow$. The single-particle wavefunctions are given by the solutions of the single-particle Schrödinger equation

$$-\frac{\hbar^2}{2m}\nabla^2\psi_\sigma(\mathbf{r}) = \epsilon\psi_\sigma(\mathbf{r}) \quad (\sigma = \uparrow, \downarrow).$$

They are given by plane-wave states

$$\psi_\sigma(\mathbf{r}) = \frac{1}{\sqrt{V}} e^{i\mathbf{k}\cdot\mathbf{r}},$$

which are normalised, $\int_V d^3r\ \psi^*_\sigma(\mathbf{r})\psi_\sigma(\mathbf{r}) = 1$, and carry the energy

$$\epsilon_{\mathbf{k}} = \frac{\hbar^2\mathbf{k}^2}{2m}. \tag{5.1}$$

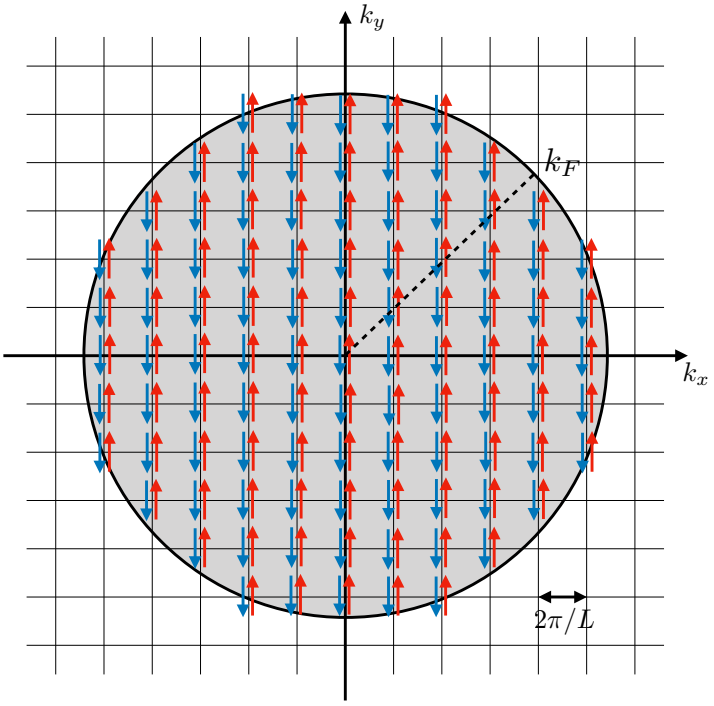
Because of the periodic boundary conditions the momenta can only take discrete values,

$$\psi_\sigma(\mathbf{r}) = \psi_\sigma(\mathbf{r} + L\hat{\mathbf{e}}_i) = \frac{1}{\sqrt{V}} e^{i\mathbf{k}\cdot\mathbf{r}} e^{ik_i L} = \psi_\sigma(\mathbf{r}) e^{ik_i L} \Rightarrow e^{ik_i L} = 1,$$

from which it follows that

$$k_i = n_i \frac{2\pi}{L} \quad \text{with} \quad n_i \in \mathbb{Z}. \tag{5.2}$$

Because of the *Pauli principle*, two fermions with the same spin cannot occupy the same momentum state. The groundstate of N fermions is therefore obtained by filling momentum states of increasing energy ('filling the Fermi sea').



The states are filled up to the Fermi momentum k_F, corresponding to the Fermi energy

$$\epsilon_F = \frac{\hbar^2 k_F^2}{2m}. \tag{5.3}$$

Note that in the thermodynamic limit, $L \to \infty$ (keeping the density fixed), $\Delta k = 2\pi/L \to 0$ and k_F becomes sharply defined. The density is given by

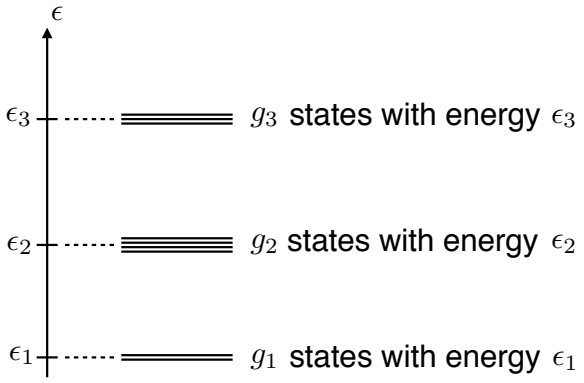
$$\begin{aligned}\rho &= \frac{N}{V} = \frac{N_\uparrow + N_\downarrow}{V} = \frac{2}{V}\sum_{\mathbf{k}}^{|\mathbf{k}|\le k_F} = \frac{2}{(2\pi)^3}\sum_{\mathbf{k}}^{|\mathbf{k}|\le k_F}(\Delta k)^3 \\ &\underset{L\to\infty}{\to} \frac{2}{(2\pi)^3}\int_{|\mathbf{k}|\le k_F} d^3k = \frac{2}{(2\pi)^3}\cdot\frac{4}{3}\pi k_F^3 = \frac{1}{3\pi^2}k_F^3. \end{aligned} \tag{5.4}$$

Using that $k_F = (3\pi^2\rho)^{1/3}$, we can express the Fermi energy (5.3) in terms of the density,

$$\epsilon_F = (3\pi^2)^{2/3}\frac{\hbar^2}{2m}\rho^{2/3}. \tag{5.5}$$

5.1 Fermi function and density of states

For the three-dimensional Fermi gas in a box with periodic boundary conditions we have 2 states with energy 0, 12 states with energy $\frac{\hbar^2(\Delta k)^2}{2m}$, etc. Let us consider a general situation of a system of fermions with discrete energy levels $\epsilon_1 < \epsilon_2 < \epsilon_2 < \ldots$ and assume that there exist g_1 states with energy ϵ_1, g_2 states with energy ϵ_2, and so on.



We want to calculate the *grand canonical partition function* Z_G and related *grand potential* $\Omega(V,T,\mu)$,

$$Z_G = \mathrm{Tr}\, e^{-\beta(\hat{H}-\mu\hat{N})} = e^{-\beta\Omega(V,T,\mu)}, \tag{5.6}$$

where $\beta = 1/(k_BT)$. The Hamiltonian and particle number operators are given by

$$\begin{aligned}\hat{H} &= \sum_i\sum_{\alpha_i=1}^{g_i}\epsilon_i c^\dagger_{i,\alpha_i}c_{i,\alpha_i}, \\ \hat{N} &= \sum_i\sum_{\alpha_i=1}^{g_i} c^\dagger_{i,\alpha_i}c_{i,\alpha_i}.\end{aligned}$$

Since $\hat{H}$ and $\hat{N}$ are diagonal in the occupation number operators, $\hat{n}_{i,\alpha_i} = c^\dagger_{i,\alpha_i}c_{i,\alpha_i}$, which can only take values 0 or 1, it is straightforward to compute the trace,

$$
\begin{aligned}
Z_G &= \mathrm{Tr}\, e^{-\beta \sum_i \sum_{\alpha_i=1}^{g_i} (\epsilon_i - \mu)\hat{n}_{i,\alpha_i}} = \sum_{\{n_{i,\alpha_i}\}} \prod_i \prod_{\alpha_i=1}^{g_i} e^{-\beta(\epsilon_i-\mu) n_{i,\alpha_i}} \\
&= \prod_i \prod_{\alpha_i=1}^{g_i} \left(e^{-\beta(\epsilon_i-\mu)} + 1\right) = \prod_i \left(e^{-\beta(\epsilon_i-\mu)} + 1\right)^{g_i} .
\end{aligned}
$$

From Z_G we can easily compute the potential

$$
\Omega = -\frac{1}{\beta} \ln Z_G = -\frac{1}{\beta} \sum_i g_i \ln\left(e^{-\beta(\epsilon_i-\mu)} + 1\right) . \tag{5.7}
$$

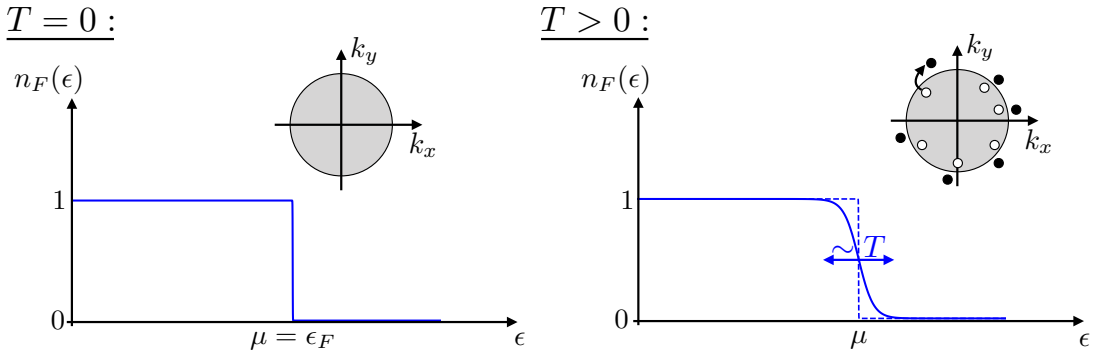
The average particle number is given by

$$
\begin{aligned}
N &= \langle \hat{N} \rangle = \frac{1}{Z_G} \mathrm{Tr}\, \hat{N} e^{-\beta(\hat{H}-\mu\hat{N})} = \frac{1}{\beta} \frac{\partial}{\partial \mu} \ln Z_G = -\left(\frac{\partial \Omega}{\partial \mu}\right)_{V,T} \\
&= \frac{1}{\beta} \sum_i g_i \frac{\beta e^{-\beta(\epsilon_i-\mu)}}{e^{-\beta(\epsilon_i-\mu)} + 1} = \sum_i g_i \frac{1}{e^{\beta(\epsilon_i-\mu)} + 1} = \sum_i g_i \, n_F(\epsilon_i).
\end{aligned} \tag{5.8}
$$

In the last step we have defined the Fermi function

$$
n_F(\epsilon) = \frac{1}{e^{\beta(\epsilon-\mu)} + 1}. \tag{5.9}
$$

From the expression for the average particle number, Eq. (5.8), it is evident that $n_F(\epsilon)$ is equal to the average occupation of a state with energy ϵ. As the temperature goes to zero ($\beta \to \infty$), $n_F(\epsilon) \to 1$ if $\epsilon < \mu$ and $n_F(\epsilon) \to 0$ if $\epsilon > \mu$. Hence the Fermi function becomes a step function at zero temperature, reflecting that only states with $\epsilon < \epsilon_F = \mu$ are occupied. In the context of the Fermi gas this is the groundstate of the filled Fermi sea. At finite temperature, $T > 0$, the Fermi function smears in an interval around μ of width $\sim T$. As we will see later, μ also slightly shifts ($\sim T^2$) away from ϵ_F. The smearing of the Fermi surface is a consequence of thermally excited particle-hole pairs.



For a continuous energy spectrum Eq. (5.8) takes the form

$$N = \int_0^\infty d\epsilon\, g(\epsilon) n_F(\epsilon), \tag{5.10}$$

where $g(\epsilon)$ is called the density of states (DOS) and $g(\epsilon)d\epsilon$ is equal to the number of states with energies between ϵ and $\epsilon + d\epsilon$. We conclude this section by computing the density of states $g(\epsilon)$ for the free Fermi gas in $d = 3$ spatial dimensions. We calculate the average particle number N explicitly by converting the momentum integral into an integral over energy and then obtain $g(\epsilon)$ from comparison with Eq. (5.10).

$$\begin{aligned} N &= \sum_{\mathbf{k},\sigma} n_F(\epsilon_\mathbf{k}) = 2\sum_{\mathbf{k}} n_F(\epsilon_\mathbf{k}) = 2\frac{V}{(2\pi)^3}\sum_{\mathbf{k}} (\Delta k)^3 n_F(\epsilon_\mathbf{k}) \\ &\overset{V\to\infty}{\simeq} 2\frac{V}{(2\pi)^3}\int d^3k\; n_F(\frac{\hbar^2k^2}{2m}) = \frac{V}{\pi^2}\int dk\; k^2 n_F(\frac{\hbar^2k^2}{2m}) \\ &= \frac{1}{2}\frac{V}{\pi^2}\frac{2m}{\hbar^2}\int_0^\infty d\epsilon \sqrt{\frac{2m}{\hbar^2}}\sqrt{\epsilon}\; n_F(\epsilon), \end{aligned} \tag{5.11}$$

from which it follows that

$$g(\epsilon) = \frac{V}{2\pi^2}\sqrt{\frac{2m}{\hbar^2}}^3\sqrt{\epsilon}. \tag{5.12}$$

Note that we have chosen to define $g(\epsilon)$ for the system as a whole rather than per unit volume, and so it is proportional to V. For the DOS per unit volume we will write $\nu(\epsilon) := g(\epsilon)/V$.

5.2 The Sommerfeld expansion

The Fermi temperatures in metals ($\epsilon_F = k_B T_F$) are typically of the order $T_F \sim 10^4 - 10^5$ K. We can therefore assume that

$$\frac{k_B T}{\epsilon_F} = \frac{T}{T_F} \ll 1, \tag{5.13}$$

and expand thermodynamic observables in terms of this small parameter. We need to calculate integrals of the form

$$I = \int_{-\infty}^\infty d\epsilon\; h(\epsilon) n_F(\epsilon),$$

where $n_F(\epsilon)$ is the Fermi function, Eq. (5.9), and $h(\epsilon)$ a function that is analytic at $\epsilon = \mu$. E.g., the expressions for the average particle number (5.10) and for the energy are of this form. The expansion in terms of the small parameter (5.13) is an expansion in terms of a small smearing of the Fermi step function. In a homework problem you will show that

$$\int_{-\infty}^\infty d\epsilon\; h(\epsilon) n_F(\epsilon) \approx \int_{-\infty}^\mu d\epsilon\; h(\epsilon) + \frac{\pi^2}{6}h'(\mu)(k_B T)^2 + \frac{7\pi^4}{360}h'''(\mu)(k_B T)^4. \tag{5.14}$$

This expansion is known as the *Sommerfeld expansion.*

We use the Sommerfeld expansion (5.14) to evaluate the integral for the particle number (5.10) up to order T^2. In order to obtain an integral from $-\infty$ to ∞ we can simply continue the DOS by setting $g(\epsilon) = 0$ for $\epsilon < 0$. We obtain:

$$\begin{aligned} N &= \int_0^\infty d\epsilon\, g(\epsilon) n_F(\epsilon) \approx \int_0^\mu d\epsilon\, g(\epsilon) + \frac{\pi^2}{6} g'(\mu)(k_B T)^2 \\ &\overset{g(\epsilon)\sim\sqrt{\epsilon}}{=} \frac{2}{3}\mu g(\mu) + \frac{\pi^2}{12}\frac{g(\mu)}{\mu}(k_B T)^2 = \frac{2}{3}\mu g(\mu)\left[1 + \frac{\pi^2}{8}\left(\frac{k_B T}{\mu}\right)^2\right] \\ &\approx \frac{2}{3}\mu g(\mu)\left[1 + \frac{\pi^2}{8}\left(\frac{k_B T}{\epsilon_F}\right)^2\right] \overset{(5.12)}{=} \frac{V}{3\pi^2}\sqrt{\frac{2m}{\hbar^2}}^3 \mu^{3/2}\left[1 + \frac{\pi^2}{8}\left(\frac{k_B T}{\epsilon_F}\right)^2\right] \end{aligned} \tag{5.15}$$

Solving this equation for the the chemical potential we obtain

$$\mu = \underbrace{(3\pi^2\rho)^{2/3}\frac{\hbar^2}{2m}}_{=\epsilon_F\ (5.5)}\left[1 + \frac{\pi^2}{8}\left(\frac{k_B T}{\epsilon_F}\right)^2\right]^{-2/3} \approx \epsilon_F\left[1 - \frac{\pi^2}{12}\left(\frac{k_B T}{\epsilon_F}\right)^2\right]. \tag{5.16}$$

At $T = 0$, $\mu = \epsilon_F$. Temperature leads to a small reduction of μ of order $(k_B T/\epsilon_F)^2$. For the energy,

$$E = \int_0^\infty d\epsilon\, \epsilon g(\epsilon) n_F(\epsilon), \tag{5.17}$$

we obtain from the Sommerfeld expansion (5.14) with $h(\epsilon) = \epsilon g(\epsilon) \sim \epsilon^{3/2}$ for $\epsilon > 0$ and $h(\epsilon) = 0$ for $\epsilon < 0$, keeping terms up to order T^2,

$$\begin{aligned} E &\approx \frac{2}{5}\mu^2 g(\mu) + \frac{\pi^2}{6}\frac{3}{2} g(\mu)(k_B T)^2 = \frac{2}{5}\mu^2 g(\mu)\left[1 + \frac{5\pi^2}{8}\left(\frac{k_B T}{\mu}\right)^2\right] \\ &\approx \frac{2}{5}\frac{V}{2\pi^2}\sqrt{\frac{2m}{\hbar^2}}^3 \mu^{5/2}\left[1 + \frac{5\pi^2}{8}\left(\frac{k_B T}{\epsilon_F}\right)^2\right] \\ &\overset{(5.16)}{=} \frac{2}{5}\frac{V}{2\pi^2}\sqrt{\frac{2m}{\hbar^2}}^3 \epsilon_F^{5/2}\left[1 - \frac{\pi^2}{12}\left(\frac{k_B T}{\epsilon_F}\right)^2\right]^{5/2}\left[1 + \frac{5\pi^2}{8}\left(\frac{k_B T}{\epsilon_F}\right)^2\right] \\ &\approx \frac{2}{5}\epsilon_F^2 g(\epsilon_F)\left[1 - \frac{5\pi^2}{24}\left(\frac{k_B T}{\epsilon_F}\right)^2\right]\left[1 + \frac{5\pi^2}{8}\left(\frac{k_B T}{\epsilon_F}\right)^2\right] \\ &\approx \frac{2}{5}\epsilon_F^2 g(\epsilon_F)\left[1 + \frac{5\pi^2}{12}\left(\frac{k_B T}{\epsilon_F}\right)^2\right] \\ &= E_0 + \frac{\pi^2}{6} g(\epsilon_F)(k_B T)^2. \end{aligned} \tag{5.18}$$

From the energy expression we obtain the specific heat

$$C_V = \left(\frac{\partial E}{\partial T}\right)_V = \frac{\pi^2}{3} g(\epsilon_F) k_B^2 T = \gamma T. \tag{5.19}$$

The linear specific heat coefficient

$$\gamma = \frac{\pi^2}{3} g(\epsilon_F) k_B^2 \tag{5.20}$$

is called the Sommerfeld coefficient.

5.3 Pauli (spin) susceptibility at $T = 0$

In this section we calculate the magnetic susceptibility or Pauli (spin) susceptibility,

$$\chi = \left.\frac{\partial M}{\partial B}\right|_{B=0}, \tag{5.21}$$

of the free electron gas at $T = 0$. Here B denotes the magnetic field and M the induced magnetisation. χ is a measure of how easy it is to magnetise the system. In the presence of a magnetic field, the Hamiltonian of the free electron gas becomes

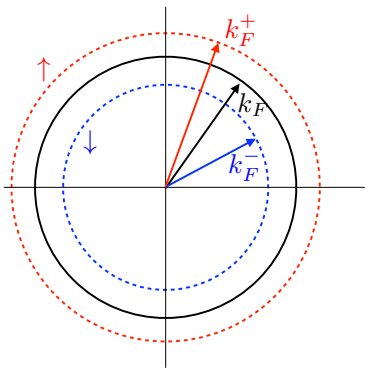
$$\hat{H} = \sum_{\mathbf{k}\sigma} \frac{\hbar^2 k^2}{2m} \underbrace{c^\dagger_{\mathbf{k}\sigma} c_{\mathbf{k}\sigma}}_{=\hat{n}_{\mathbf{k}\sigma}} - \frac{1}{2} g \mu_B B \sum_{\mathbf{k}} (\hat{n}_{\mathbf{k}\uparrow} - \hat{n}_{\mathbf{k}\downarrow}), \tag{5.22}$$

where the constants μ_B and g are the Bohr magneton and g-factor ($g = 2$ for electrons), respectively. The Hamiltonian is diagonal with electron dispersion

$$\epsilon_{\mathbf{k}\sigma} = \frac{\hbar^2 k^2}{2m} - \frac{1}{2} \sigma g \mu_B B, \tag{5.23}$$

where we have identified $\sigma \in \{\uparrow, \downarrow\} = \{+1, -1\}$. The opposite shifts of the dispersion of spin-up and spin-down electrons correspond to a splitting of the Fermi surfaces. From $\epsilon_{\mathbf{k}\sigma} = \mu = \epsilon_F$ we obtain

$$k_{F,\pm} = \sqrt{\frac{2m}{\hbar^2}} \sqrt{\epsilon_F \pm \frac{1}{2} g \mu_B B}. \tag{5.24}$$



The magnetisation is given by

$$\begin{aligned} M &= \frac{1}{2}g\mu_B(N_\uparrow - N_\downarrow) = \frac{1}{2}g\mu_B \sum_{\mathbf{k}} [n_F(\epsilon_{\mathbf{k}\uparrow}) - n_F(\epsilon_{\mathbf{k}\downarrow})] = \frac{1}{2}g\mu_B \frac{V}{(2\pi)^3}\frac{4}{3}\pi \left(k_{F,+}^3 - k_{F,-}^3\right) \\ &\overset{(5.24)}{=} \frac{V}{12\pi^2}g\mu_B \sqrt{\frac{2m}{\hbar^2}}^3 \left[(\epsilon_F + \frac{1}{2}g\mu_B B)^{3/2} - (\epsilon_F - \frac{1}{2}g\mu_B B)^{3/2}\right] \\ &\approx \frac{V}{12\pi^2}g\mu_B \sqrt{\frac{2m}{\hbar^2}}^3 \frac{3}{2}\sqrt{\epsilon_F}g\mu_B B \overset{(5.12)}{=} \frac{1}{4}(g\mu_B)^2 g(\epsilon_F)B. \end{aligned} \tag{5.25}$$

From this it follows that

$$\chi = \left.\frac{\partial M}{\partial B}\right|_{B=0} = \frac{1}{4}(g\mu_B)^2 g(\epsilon_F) \overset{g=2}{=} \mu_B^2 g(\epsilon_F) \tag{5.26}$$

The Sommerfeld coefficient γ (5.20) and the magnetic susceptibility χ are both proportional to the DOS at the Fermi level, $g(\epsilon_F)$. A useful quantity that is independent of the DOS is the Wilson ratio

$$R_W = \frac{\pi^2 k_B^2 \chi}{3\mu_B^2 \gamma}. \tag{5.27}$$

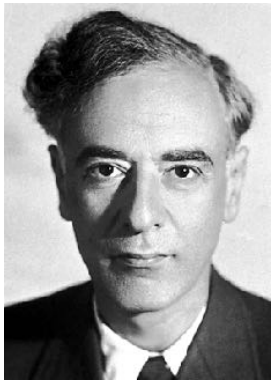
The coefficients are adjusted such that for the free Fermi gas $R_W = 1$. When we include *interactions* between electrons, both C_V and χ are modified, in principle differently. Deviations of R_W from 1 indicate strong electronic correlations. R_W may be $\gg 1$ in nearly ferromagnet metals.

Chapter 6

Landau theory of Fermi liquids

The thermal excitations in the free Fermi gas are given by particle-hole excitations around the Fermi surface in a momentum shell of relative width T/T_F. As a result, the heat capacity is linear in temperature and the Pauli susceptibility is constant. It is remarkable that the same behaviour is measured for electrons in metals and for liquid ^{3}He, since in these systems the scale for interactions is typically comparable with the Fermi energy and certainly much larger than the energy scale of excitations relevant for these physical properties.

The objective of the *Landau theory of Fermi liquids* is to understand why interactions have no qualitative effect, and to characterise their residual, quantitative consequences.



Lev Landau
1908 - 1968

Nobel Prize in Physics
(Theory of superfluidity)
1962

The central assumption of Fermi liquid theory is that, as long as we don't cross a phase transition, the ground state and the excitations evolve smoothly as we increase the interaction strength from zero to its physical value. This suggests that we can use the same set of quantum numbers as in the free gas, momentum and spin, to describe the low-energy excitations of the interacting system.

These excitations are described by elementary excitations, or *quasiparticles*, which are very similar to the usual electrons and holes in a free Fermi gas, despite strong interactions.

6.1 Lifetime of quasiparticles

A crucial assumption of Fermi-liquid theory is that there exists a well defined Fermi surface, with the radius k_F satisfying the same relation to the density,

$$\rho = \frac{N}{V} = \frac{k_F^3}{3\pi^2\hbar^3}$$

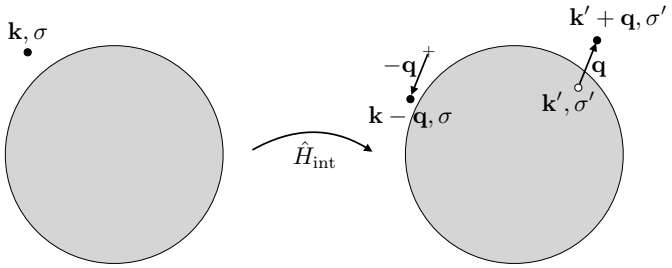
as in the free Fermi gas. We consider a quasiparticle with momentum $\mathbf{k}$ ($|\mathbf{k}| > k_F$) and spin σ, $c^\dagger_{\mathbf{k}\sigma}|\mathrm{FS}\rangle$. Interactions lead to *scattering processes* which give rise to a *finite lifetime* of the quasiparticle. The conventional density-density interactions are of the form

$$\hat{H}_{\mathrm{int}} = \frac{1}{2}\int d^3r \int d^3r'\, V(|\mathbf{r}-\mathbf{r}'|)\hat{n}(\mathbf{r})\hat{n}(\mathbf{r}'), \tag{6.1}$$

with $\hat{n}(\mathbf{r}) = \sum_{\sigma=\uparrow,\downarrow}\hat{n}_\sigma(\mathbf{r}) = \sum_\sigma c^\dagger_\sigma(\mathbf{r})c_\sigma(\mathbf{r})$. For the following considerations, the precise form of the interaction potential $V(r)$ is not important, it could e.g. be a screened Coulomb potential $V(r) = \frac{\lambda}{r}e^{-q_0 r}$ ($q_0 > 0$) with Fourier transform $\tilde{V}(q) = \frac{4\pi\lambda}{q^2+q_0^2}$ (homework). In momentum space, the interaction takes the form (homework)

$$\hat{H}_{\mathrm{int}} = \frac{1}{2}\sum_{\mathbf{k},\mathbf{k}',\mathbf{q}}\sum_{\sigma,\sigma'}\tilde{V}(\mathbf{q})c^\dagger_{\mathbf{k}-\mathbf{q},\sigma}c^\dagger_{\mathbf{k}'+\mathbf{q},\sigma'}c_{\mathbf{k}',\sigma'}c_{\mathbf{k},\sigma}. \tag{6.2}$$

Acting with $\hat{H}_{\mathrm{int}}$ on an initial quasiparticle state $|i\rangle = c^\dagger_{\mathbf{k}\sigma}|\mathrm{FS}\rangle$ ($|\mathbf{k}| > k_F$), we see that the quasiparticle loses energy by scattering another quasiparticle out of the Fermi sea, leaving behind a quasihole:



The inverse lifetime $\tau^{-1}_{\mathbf{k}}$ of a quasiparticle with momentum $\mathbf{k}$ is given by the *scattering rates* $w_{i\to f}$ summed over all final states $|f\rangle$. Using *Fermi's Golden rule*, we obtain

$$\tau^{-1}_{\mathbf{k}} = \sum_f w_{i\to f} = \frac{2\pi}{\hbar}\sum_f \left|\langle f|\hat{H}_{\mathrm{int}}|i\rangle\right|^2 \delta(E_f - E_i). \tag{6.3}$$

The delta function enforces conservation of energy, where

$$\begin{aligned} E_f &= E_{\mathrm{FS}} + \epsilon_{\mathbf{k}-\mathbf{q}} + \epsilon_{\mathbf{k}'+\mathbf{q}} - \epsilon_{\mathbf{k}'}, \\ E_i &= E_{\mathrm{FS}} + \epsilon_{\mathbf{k}}, \end{aligned}$$

with $E_{\rm FS}$ the energy of the fully occupied Fermi sea. Note that $\epsilon_{\mathbf{k}\sigma} = \epsilon_{\mathbf{k}} = \frac{\hbar^2 k^2}{2m}$. The summation over final states implies a sum (integral) over $\mathbf{k}'$, $\mathbf{q}$ and σ'. Inserting $\hat{H}_{\rm int}$ (6.2) into Eq. (6.3) we obtain

$$\tau_{\mathbf{k}}^{-1} = \frac{2\pi}{\hbar V^2} \sum_{\mathbf{k}',\mathbf{q},\sigma'} |\tilde{V}(\mathbf{q})|^2 \, n_F(\epsilon_{\mathbf{k}'}) \, [1-n_F(\epsilon_{\mathbf{k}'+\mathbf{q}})] \, [1-n_F(\epsilon_{\mathbf{k}-\mathbf{q}})] \, \delta(\epsilon_{\mathbf{k}-\mathbf{q}}-\epsilon_{\mathbf{k}}+\epsilon_{\mathbf{k}'+\mathbf{q}}-\epsilon_{\mathbf{k}'}). \quad (6.4)$$

The integrals can be performed analytically at $T = 0$ since the Fermi functions simply constrain the integration range and ensure that $\mathbf{k}'$ lies inside and $\mathbf{k}' + \mathbf{q}$ and $\mathbf{k} - \mathbf{q}$ outside the Fermi sea. The sum over σ' simply gives a factor of 2. A straightforward but tedious calculation (homework) gives

$$\tau_{\mathbf{k}}^{-1} \sim (\epsilon_{\mathbf{k}} - \epsilon_F)^2. \quad (6.5)$$

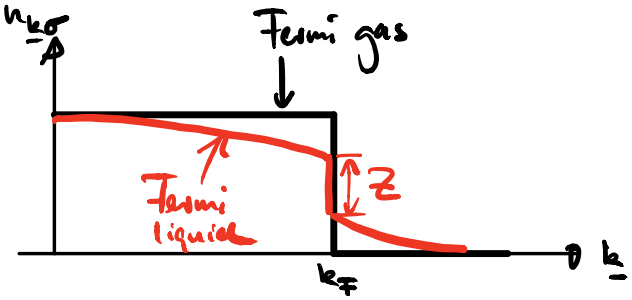
This result implies that the quasiparticles become longer and longer lived as the Fermi surface is approached. At zero temperature, the lifetime becomes infinite at the Fermi surface. Evaluating Eq. (6.4) at $T > 0$ one finds that temperature gives an additional contribution $\sim T^2$ to $\tau_{\mathbf{k}}^{-1}$.

6.2 Relation between bare electrons and quasiparticles

The action of an annihilation operator $c_{\mathbf{k}\sigma}$ of a *bare fermion* with wavevector $\mathbf{k}$ and spin σ of the groundstate $|\mathrm{GS}\rangle$ on the *interacting system* has an amplitude ($:= Z^{1/2}$) to generate a state $|\mathbf{k}\sigma\rangle$ containing a quasihole if $|\mathbf{k}| \leq k_F$. It also has an amplitude to generate superpositions of many excitations, which we denote as $|\mathrm{incoherent}\rangle$. We expect the incoherent processes to vary smoothly with $\mathbf{k}$.

$$c_{\mathbf{k}\sigma}|\mathrm{GS}\rangle \sim \begin{cases} Z^{1/2}|\mathbf{k}\sigma\rangle + |\mathrm{incoherent}\rangle & (|\mathbf{k}| \leq k_F) \\ |\mathrm{incoherent}\rangle & (|\mathbf{k}| > k_F) \end{cases} \quad (6.6)$$

At $T = 0$, the Fermi occupation number $n_{\mathbf{k}\sigma} = \langle \mathrm{GS}|c_{\mathbf{k}\sigma}^\dagger c_{\mathbf{k}\sigma}|\mathrm{GS}\rangle$ has a step of size Z at $|\mathbf{k}| = k_F$. This is known as the *Migdal discontinuity* and is the demonstration of the existence of a sharp Fermi surface.

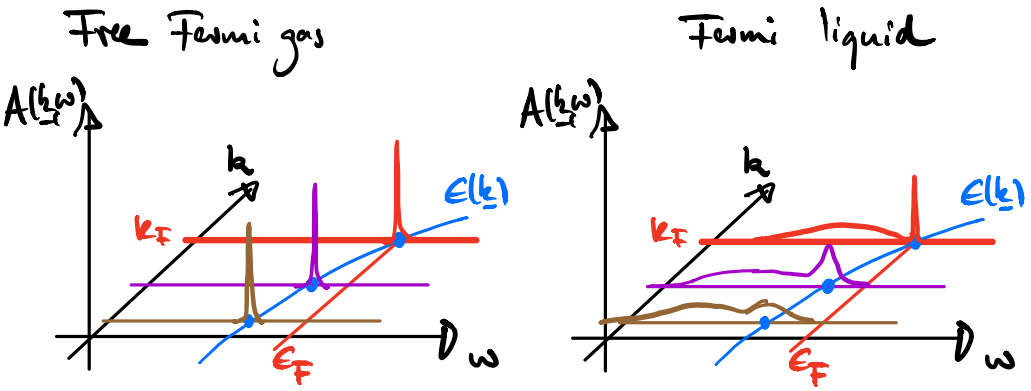


In the free fermion system $Z = 1$. The effect of interactions is to decrease Z and to give excitations (quasiparticles) an *effective mass* larger than that of the bare particles.

The existence of quasiparticles and an incoherent background can be seen in the spectral function $A(\mathbf{k},\omega)$. For $\omega > \epsilon_F$, $A(\mathbf{k},\omega)$ is equal to the probability of finding the system in the state with momentum $\mathbf{k}$ and energy ω after adding an electron. For $\omega < \epsilon_F$, $A(\mathbf{k},\omega)$ is equal to the probability of extracting an electron with energy ω and momentum $\mathbf{k}$. This is the process of (angle-resolved) photoemission (ARPES).

For the free electron gas, the system is made of bare electrons with a well defined relation between energy and momentum, $\omega = \epsilon_{\mathbf{k}}$, and hence

$$A(\mathbf{k},\omega) \simeq \delta(\omega - \epsilon_{\mathbf{k}}). \tag{6.7}$$



In the presence of interactions in a Fermi liquid there exists a quasiparticle which has a finite overlap $Z^{1/2} < 1$ with the bare electron. The remaining spectral weight is transferred to the incoherent background,

$$A(\mathbf{k},\omega) = A_{\rm qp}(\mathbf{k},\omega) + A_{\rm inc}(\mathbf{k},\omega). \tag{6.8}$$

At $\omega = \epsilon_F$ and $T = 0$ the lifetime of the quasiparticle is infinite and $A_{\rm qp}(\mathbf{k},\omega = \epsilon_F) \simeq Z\delta(\epsilon_F - \epsilon_{\mathbf{k}})$. Moving away from the Fermi surface, the lifetime of the quasiparticle is finite, $\tau_{\mathbf{k}} \sim |\epsilon_{\mathbf{k}} - \epsilon_F|^{-2}$, corresponding to a broadening of the quasiparticle peak. The peak can be approximated by a Lorentzian with width $\tau_{\mathbf{k}}^{-1}$.

6.3 Parametrising excitation energies

The key assumption of Fermi liquid theory is that the strongly interacting electron system renormalises into a gas of weakly interacting quasiparticles with dispersion

$$\epsilon_0(\mathbf{k}) = \frac{\hbar^2 k^2}{2m^*}. \tag{6.9}$$

Here, m^* is called the quasiparticle effective mass, which can be considerably larger than the bare electron mass m. E.g., in heavy fermion compounds such as $CeCu_6$ or $CeAl_3$ we observe $m^*/m \approx 10^3$. Since the quasiparticles are sharply defined only close to the Fermi surface, one often works with a linearised dispersion near k_F,

$$\epsilon_0(\mathbf{k}) - \epsilon_F \approx \frac{\hbar^2}{m^*} k_F (k - k_F). \tag{6.10}$$

Note that for simplicity, we assume that the Fermi surface of the Fermi liquid is spherical. Real materials often exhibit multiple bands with complicated Fermi surfaces. In this case, the effective mass can be different for the different bands and change as a function of momentum.

Having established the idea that excitations are of the same kind as in a free Fermi gas, and have sharply defined energies, it remains to discuss how these energies are influenced by residual interactions between quasiparticles. Suppose we somehow changed the distribution of quasiparticles,

$$n^{(0)}_{\mathbf{k}\sigma} \to n_{\mathbf{k}\sigma} = n^{(0)}_{\mathbf{k}\sigma} + \delta n_{\mathbf{k}\sigma} \tag{6.11}$$

The change $\delta n_{\mathbf{k}\sigma}$ could be due to an excitation or an external perturbation. The change of the total energy δE due to $\delta n_{\mathbf{k}\sigma}$ up to quadratic order is

$$\delta E = \sum_{\mathbf{k}\sigma} \frac{\hbar^2 k^2}{2m^*} \delta n_{\mathbf{k}\sigma} + \frac{1}{2} \sum_{\mathbf{k}\sigma, \mathbf{k}'\sigma'} f(\mathbf{k}\sigma, \mathbf{k}'\sigma') \delta n_{\mathbf{k}\sigma} \delta n_{\mathbf{k}'\sigma'}. \tag{6.12}$$

The function $f(\mathbf{k}\sigma, \mathbf{k}'\sigma')$ introduced phenomenologically in Eq. (6.12) is called the Landau interaction function. One can show that it is connected with the scattering amplitude of electrons. As quasiparticles interact with each other, the energy of each of them depends on the changes of the distribution function of the others. From $\delta E = \sum_{\mathbf{k}\sigma} \epsilon_\sigma(\mathbf{k}) \delta n_{\mathbf{k}\sigma}$ we obtain

$$\epsilon_\sigma(\mathbf{k}) = \frac{\delta(\delta E)}{\delta(\delta n_{\mathbf{k}\sigma})} = \frac{\hbar^2 k^2}{2m^*} + \sum_{\mathbf{k}'\sigma'} f(\mathbf{k}\sigma, \mathbf{k}'\sigma') \delta n_{\mathbf{k}'\sigma'}. \tag{6.13}$$

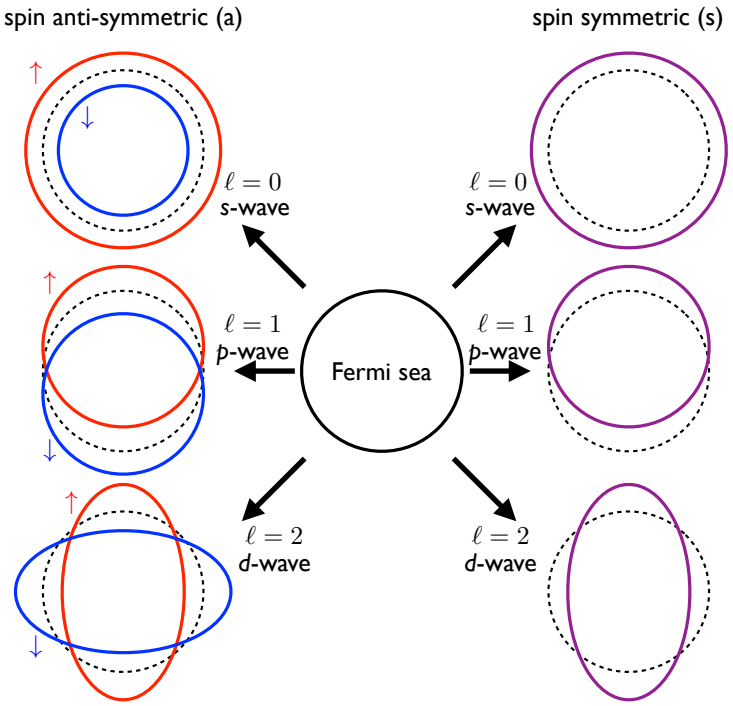
At this point the approach seems unpromising because the expansion coefficients involve not only a few fitting parameters but instead an unknown function $f(\mathbf{k}\sigma, \mathbf{k}'\sigma')$. We make things manageable by separating $\delta n_{\mathbf{k}\sigma}$ into spherical harmonics, and recognising that in most physical situations only the lowest two harmonics are relevant. In turn, for a spherical Fermi surface, only the zeroth and first harmonics of $f(\mathbf{k}\sigma, \mathbf{k}'\sigma')$ are important. For nonmagnetic systems, we can decompose $f(\mathbf{k}\sigma, \mathbf{k}'\sigma')$ into two terms, $f^{(s)}$, $f^{(a)}$,

$$
\begin{aligned}
f(\mathbf{k}\uparrow,\mathbf{k}'\uparrow) = f(\mathbf{k}\downarrow,\mathbf{k}'\downarrow) &= f^{(s)}(\mathbf{k},\mathbf{k}') + f^{(a)}(\mathbf{k},\mathbf{k}'),\\
f(\mathbf{k}\uparrow,\mathbf{k}'\downarrow) = f(\mathbf{k}\downarrow,\mathbf{k}'\uparrow) &= f^{(s)}(\mathbf{k},\mathbf{k}') - f^{(a)}(\mathbf{k},\mathbf{k}').
\end{aligned}
$$

Identifying $\{\uparrow,\downarrow\}$ with $\{+1,-1\}$ we can write the two equations in a compact form,

$$
f(\mathbf{k}\sigma,\mathbf{k}'\sigma') = f^{(s)}(\mathbf{k},\mathbf{k}') + \sigma\sigma' f^{(a)}(\mathbf{k},\mathbf{k}'). \tag{6.14}
$$

In the figure below, Fermi surface deformations in different angular momentum channel and symmetric or anti-symmetric in spin are shown.



Since the whole theory is valid only close to the Fermi surface, we can assume $|\mathbf{k}| \approx |\mathbf{k}'| \approx k_F$ and keep only the dependence on the angle θ between $\mathbf{k}$ and $\mathbf{k}'$. In this case, the expansion in terms of spherical harmonics reduces into an expansion in terms of simpler Legendre polynomials $P_\ell(x)$, $x = \cos\theta$,

$$
f^{(a,s)}(\mathbf{k},\mathbf{k}') = \sum_{\ell=0}^{\infty} f_\ell^{(a,s)} P_\ell(\cos\theta). \tag{6.15}
$$

The first Legendre polynomials are given by $P_0(x) = 1$, $P_1(x) = x$ and $P_2(x) = \frac{1}{2}(3x^2 - 1)$. From Eq. (6.12) it is clear that $f^{(a,s)}(\mathbf{k},\mathbf{k}')$ and hence $f_\ell^{(a,s)}$ carry units of energy. We therefore introduce dimensionless Landau parameters

$$F_\ell^{(a,s)} = g^*(\epsilon_F) f_\ell^{(a,s)}, \tag{6.16}$$

with

$$g^*(\epsilon_F) = \frac{V}{2\pi^2}\sqrt{\frac{2m^*}{\hbar^2}}^3 \sqrt{\epsilon_F} = \frac{V m^* k_F}{\pi^2 \hbar^2} \tag{6.17}$$

the DOS at the Fermi energy, where the asterisks denotes $m \to m^*$. Keeping only the first two harmonics we therefore obtain

$$f^{(a,s)}(\mathbf{k}, \mathbf{k}') \approx \frac{1}{g^*(\epsilon_F)} \left(F_0^{(a,s)} + F_1^{(a,s)} \cos\theta \right). \tag{6.18}$$

6.4 Measuring the Landau parameters

The isotropic Fermi liquid (spherical Fermi surface) is parametrised (to a good approximation) by a few dimensionless parameters,

$$\frac{m^*}{m}, F_0^{(s)}, F_0^{(a)}, \text{ and } F_1^{(s)}.$$

These are related to measurable quantities such as specific heat, magnetic susceptibility and compressibility.

Heat capacity:

Finite temperature generates an isotropic distribution (radially symmetric), $n_{\mathbf{k}\sigma} = n_{\mathbf{k}\sigma}^{(0)} + \delta n_{\mathbf{k}\sigma}$, in which there are the same number of quasiparticles and quasiholes and therefore

$$\sum_{\mathbf{k}} \delta n_{\mathbf{k}\sigma} = 0. \tag{6.19}$$

For any isotropic distribution $n_{\mathbf{k}\sigma}$, $\delta n_{\mathbf{k}\sigma}$ only contains an s-wave component ($\ell = 0$), coupling only to $F_0^{(s)}$ and $F_0^{(a)}$. In general, the orthogonality of Legendre polynomials,

$$\int_{-1}^{1} dx\ P_\ell(x) P_{\ell'}(x) = \int_0^\pi d\theta\ \sin\theta\ P_\ell(\cos\theta) P_{\ell'}(\cos\theta) = \frac{2}{2\ell+1}\delta_{\ell,\ell'},$$

implies that the ℓ-wave component of $\delta n_{\mathbf{k}\sigma}$ can only couple to $F_\ell^{(s)}$ and $F_\ell^{(a)}$. Starting from Eq. (6.13), we obtain

$$\begin{aligned}
\epsilon_\sigma(\mathbf{k}) &= \frac{\hbar^2 k^2}{2m^*} + \sum_{\mathbf{k}'\sigma'} f(\mathbf{k}\sigma, \mathbf{k}'\sigma') \delta n_{\mathbf{k}'\sigma'} = \frac{\hbar^2 k^2}{2m^*} + \sum_{\mathbf{k}'\sigma'} \left[f^{(s)}(\mathbf{k}, \mathbf{k}') + \sigma\sigma' f^{(a)}(\mathbf{k}, \mathbf{k}') \right] \delta n_{\mathbf{k}'\sigma'} \\
&\overset{\ell=0}{=} \frac{\hbar^2 k^2}{2m^*} + \frac{1}{g^*(\epsilon_F)} \sum_{\sigma'} \left(F_0^{(s)} + \sigma\sigma' F_0^{(a)} \right) \underbrace{\sum_{\mathbf{k}'} \delta n_{\mathbf{k}'\sigma'}}_{=0 \text{ (6.19)}} = \frac{\hbar^2 k^2}{2m^*}.
\end{aligned} \tag{6.20}$$

The interaction parameters do not contribute in this case. For this reason, interactions affect the heat capacity C_V only via the value of the *effective mass*,

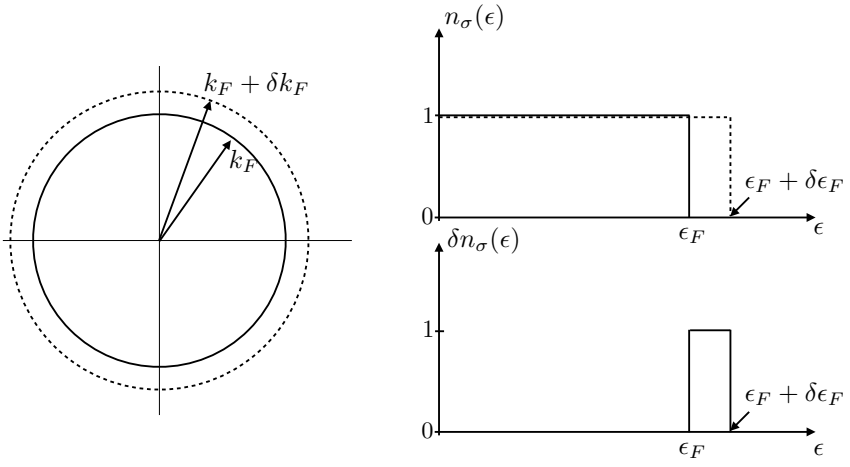
$$C_V = \frac{\pi^2}{3} k_B^2 g^*(\epsilon_F) T. \tag{6.21}$$

The ratio of the Sommerfeld coefficients of the Fermi liquid and the free Fermi gas is given by

$$\frac{\gamma}{\gamma_0} = \frac{g^*(\epsilon_F)}{g(\epsilon_F)} = \frac{m^*}{m}. \tag{6.22}$$

Compressibility:

A change in density can be represented by an isotropic, spin-independent $\delta n_{\mathbf{k}\sigma}$ ($\delta n_{\mathbf{k}\uparrow} = \delta n_{\mathbf{k}\downarrow}$). To compute the isothermal ($T \equiv 0$) compressibility, we change k_F to $k_F + \delta k_F$, ($\epsilon_F \to \epsilon_F + \delta\epsilon_F$).



Let $\delta n := \sum_{\mathbf{k}\sigma} \delta n_{\mathbf{k}\sigma}$. The resulting change in the total energy is given by

$$\begin{aligned}
\delta E &= \sum_{\mathbf{k}\sigma} \frac{\hbar^2 k^2}{2m^*} \delta n_{\mathbf{k}\sigma} + \frac{1}{2} \sum_{\mathbf{k}\sigma, \mathbf{k}'\sigma'} f(\mathbf{k}\sigma, \mathbf{k}'\sigma') \delta n_{\mathbf{k}\sigma} \delta n_{\mathbf{k}'\sigma'} \\
&\overset{\ell=0}{=} \sum_{\mathbf{k}\sigma} \frac{\hbar^2 k^2}{2m^*} \delta n_{\mathbf{k}\sigma} + \frac{1}{2g^*(\epsilon_F)} \sum_{\mathbf{k}\sigma, \mathbf{k}'\sigma'} \left(F_0^{(s)} + \sigma\sigma' F_0^{(a)} \right) \delta n_{\mathbf{k}\sigma} \delta n_{\mathbf{k}'\sigma'} \\
&= \sum_{\mathbf{k}\sigma} \frac{\hbar^2 k^2}{2m^*} \delta n_{\mathbf{k}\sigma} + \frac{F_0^{(s)}}{2g^*(\epsilon_F)} (\delta n)^2.
\end{aligned} \tag{6.23}$$

We need to evaluate the first term up to order $(\delta n)^2$. We first calculate it as a function of $\delta\epsilon_F$,

$$\begin{aligned}
\sum_{\mathbf{k}\sigma} \frac{\hbar^2 k^2}{2m^*} \delta n_{\mathbf{k}\sigma} &= \int_{\epsilon_F}^{\epsilon_F+\delta\epsilon_F} d\epsilon\, \epsilon g^*(\epsilon) = \frac{V}{2\pi^2}\sqrt{\frac{2m^*}{\hbar^2}}^3 \int_{\epsilon_F}^{\epsilon_F+\delta\epsilon_F} d\epsilon\, \epsilon^{3/2} \\
&= \frac{V}{2\pi^2}\sqrt{\frac{2m^*}{\hbar^2}}^3 \frac{2}{5}\left[(\epsilon_F+\delta\epsilon_F)^{5/2} - \epsilon_F^{5/2}\right] \\
&\approx \frac{V}{2\pi^2}\sqrt{\frac{2m^*}{\hbar^2}}^3 \left[\epsilon_F^{3/2}\delta\epsilon_F + \frac{3}{4}\epsilon_F^{1/2}(\delta\epsilon_F)^2\right] \\
&= g^*(\epsilon_F)\epsilon_F\delta\epsilon_F + \frac{3}{4}g^*(\epsilon_F)(\delta\epsilon_F)^2. \qquad (6.24)
\end{aligned}$$

In order to obtain a relation between $\delta\epsilon_F$ and δn we calculate δn in a similar way:

$$\begin{aligned}
\delta n = \sum_{\mathbf{k}\sigma} \delta n_{\mathbf{k}\sigma} &= \int_{\epsilon_F}^{\epsilon_F+\delta\epsilon_F} d\epsilon\, g^*(\epsilon) = \frac{V}{2\pi^2}\sqrt{\frac{2m^*}{\hbar^2}}^3 \int_{\epsilon_F}^{\epsilon_F+\delta\epsilon_F} d\epsilon\, \epsilon^{1/2} \\
&= = \frac{V}{2\pi^2}\sqrt{\frac{2m^*}{\hbar^2}}^3 \frac{2}{3}\left[(\epsilon_F+\delta\epsilon_F)^{3/2} - \epsilon_F^{3/2}\right] \\
&\approx \frac{V}{2\pi^2}\sqrt{\frac{2m^*}{\hbar^2}}^3 \left[\epsilon_F^{1/2}\delta\epsilon_F + \frac{1}{4}\epsilon_F^{-1/2}(\delta\epsilon_F)^2\right] \\
&= g^*(\epsilon_F)\delta\epsilon_F + \frac{1}{4}\frac{g^*(\epsilon_F)}{\epsilon_F}(\delta\epsilon_F)^2.
\end{aligned}$$

Solving this quadratic equation for $\delta\epsilon_F$ we obtain

$$\begin{aligned}
\delta\epsilon_F &= -2\epsilon_F \overset{+}{(-)} 2\epsilon_F \left(1 + \frac{\delta n}{\epsilon_F g^*(\epsilon_F)}\right)^{1/2} \\
&\approx \frac{\delta n}{g^*(\epsilon_F)} - \frac{1}{4}\frac{(\delta n)^2}{\epsilon_F g^*(\epsilon_F)^2},
\end{aligned}$$

which, after inserting into Eq. (6.24) yields

$$\sum_{\mathbf{k}\sigma} \frac{\hbar^2 k^2}{2m^*} \delta n_{\mathbf{k}\sigma} = \epsilon_F\delta n - \frac{(\delta n)^2}{4g^*(\epsilon_F)} + \frac{3(\delta n)^2}{4g^*(\epsilon_F)} = \epsilon_F\delta n + \frac{(\delta n)^2}{2g^*(\epsilon_F)}.$$

Inserting this into Eq. (6.23), we obtain the total energy change

$$\delta E = \epsilon_F\delta n + \frac{1+F_0^{(s)}}{2g^*(\epsilon_F)}(\delta n)^2. \qquad (6.25)$$

Using that the inverse compressibility is given by

$$\kappa^{-1} = -V\left.\frac{\partial p}{\partial V}\right|_{N,T} \overset{p=-(\partial E/\partial V)_N}{=} V\left.\frac{\partial^2 E}{\partial V^2}\right|_N, \qquad (6.26)$$

and that $\delta n \sim \delta V$, we obtain $\kappa \sim g^*(\epsilon_F)/[1+F_0^{(s)}]$. Hence the ratio of the compressibility of the Fermi liquid and the free Fermi gas is given by

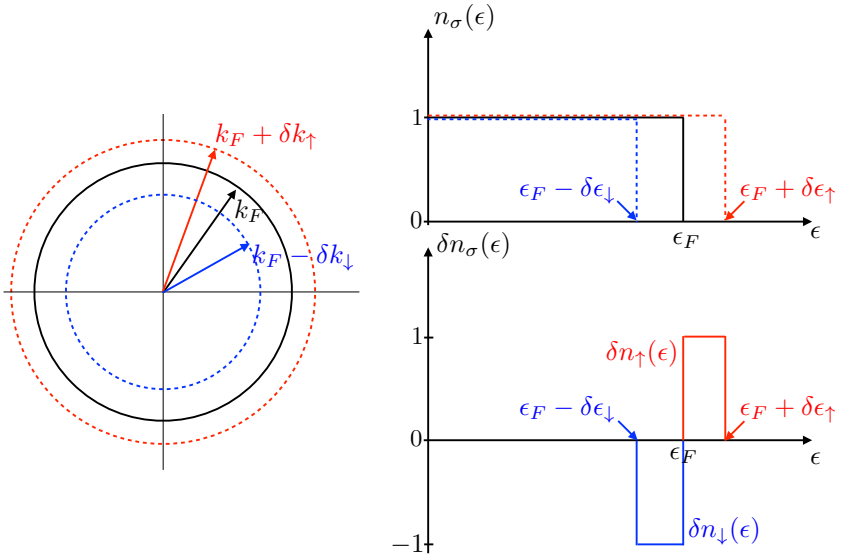
$$\frac{\kappa}{\kappa_0} = \frac{m^*/m}{1+F_0^{(s)}}. \tag{6.27}$$

Magnetic Pauli susceptibility:

We can probe the Landau parameter $F_0^{(a)}$ by measuring the Pauli spin susceptibility, since a magnetic Zeeman field B generates a spherically symmetric, spin antisymmetric change $\delta n_{\mathbf{k}\sigma}$ of the quasiparticle distribution. Let

$$\delta n_\sigma := \sum_{\mathbf{k}} \delta n_{\mathbf{k}\sigma} = \frac{1}{2}\int d\epsilon\, g^*(\epsilon)\delta n_\sigma(\epsilon),$$

with $\delta n_\uparrow = -\delta n_\downarrow$ (fixed density). This corresponds to a shift of the Fermi energies, $\epsilon_F \to \epsilon_F + \delta\epsilon_\uparrow$ for spin-up electrons and $\epsilon_F \to \epsilon_F - \delta\epsilon_\downarrow$ for spin-down electrons.



Note that since the DOS $g^*(\epsilon) \sim \sqrt{\epsilon}$ is not constant, $\delta n_\uparrow = -\delta n_\downarrow$ does *not* imply that $\delta\epsilon_\uparrow$ is equal to $\delta\epsilon_\downarrow$. The resulting magnetisation of the system is

$$M = \frac{1}{2}g\mu_B(\delta n_\uparrow - \delta n_\downarrow) = g\mu_B\delta n_\uparrow, \tag{6.28}$$

where g is the quasiparticle g-factor and μ_B the Bohr-magneton. The calculation of the change of total energy is very similar to the calculation we have done for the compressibility, and subject to a homework problem. The result is

$$\delta E = -g\mu_B B\delta n_\uparrow + \frac{2(1+F_0^{(a)})}{g^*(\epsilon_F)}(\delta n_\uparrow)^2. \tag{6.29}$$

The energy is minimised for $\delta n_\uparrow = \frac{1}{4}g\mu_B B\frac{g^*(\epsilon_F)}{1+F_0^{(a)}}$, giving rise to the magnetisation

$$M = g\mu_B \delta n_\uparrow = \frac{1}{4} g^2 \mu_B^2 B \frac{g^*(\epsilon_F)}{1 + F_0^{(a)}}, \tag{6.30}$$

from which we obtain the susceptibility

$$\chi = \left. \frac{\partial M}{\partial B} \right|_{B=0} = \mu_B^2 \frac{g^*(\epsilon_F)}{1 + F_0^{(a)}}, \tag{6.31}$$

where we have used $g = 2$ for electrons. The ratio of the susceptibilities of the Fermi liquid and the free Fermi gas is given by

$$\frac{\chi}{\chi_0} = \frac{m^*/m}{1 + F_0^{(a)}}. \tag{6.32}$$

A negative $F_0^{(a)}$ leads to an enhancement of χ. In the limit $F_0^{(a)} \to -1$, χ diverges, corresponding to an instability towards *ferromagnetic order*. This is sometimes referred to as the 'Stoner instability'.

We conclude this section by calculating the Wilson ratio of the Fermi liquid. Using Eqs. (6.22) and (6.32), we obtain

$$R_W = \frac{\pi^2 k_B^2}{3\mu_B^2} \frac{\chi}{\gamma} = \frac{\pi^2}{3\mu_B^2} \frac{\frac{m^*/m}{1+F_0^{(a)}} \chi_0}{\frac{m^*}{m} \gamma_0} = \frac{1}{1 + F_0^{(a)}}, \tag{6.33}$$

where in the last step we have used that the Wilson ratio of the free Fermi gas is equal to 1. The Wilson ratio becomes much larger than 1 close to a ferromagnetic instability.

Galilean invariance:

For translationally invariant Fermi liquids, e.g. liquid ^{3}He, there exists a simple relation between the quasiparticle effective mass, m^*/m, and the $\ell = 1$ Landau parameter $F_1^{(s)}$,

$$\frac{m^*}{m} = 1 + \frac{1}{3} F_1^{(s)}. \tag{6.34}$$

Note that this relation does not hold for quasi-electrons in a crystalline environment. Eq. (6.34) can be derived by considering a Galilean transformation to a frame moving at speed $\mathbf{v}$. The corresponding $\delta n_{\mathbf{k}\sigma}$ displaces the distribution of quasiparticles in a spin-symmetric way, coupling to $F_1^{(s)}$.

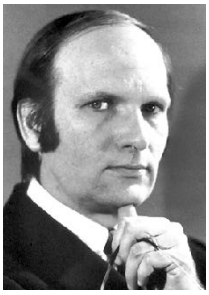
Chapter 7

BCS theory of superconductivity

BCS stands for Bardeen, Cooper and Schrieffer who published their microscopic theory for superconductivity in 1957 and received the Nobel prize in physics for it in 1972.

John Bardeen
1908 - 1991

Leon Neil Cooper
1930 -

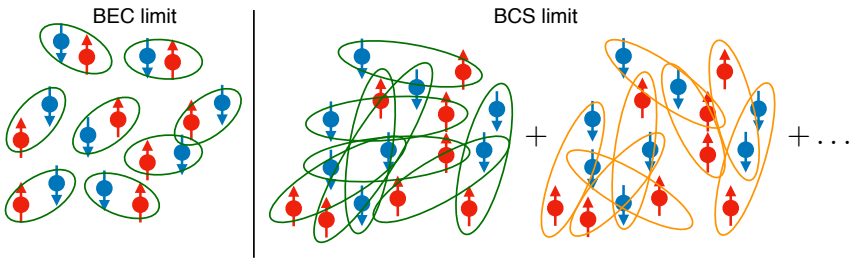


John Robert Schrieffer
1931 -

Nobel Prize in Physics
(BCS theory of
superconductivity)
1972

While the Fermi liquid is stable against weak repulsive interactions, *attractive* interactions by contrast lead to a qualitative change of the groundstate and excitation, no matter how weak the attraction is. The central idea of BCS theory is that electron-phonon coupling gives rise to such an attractive interaction that leads to the formation of bound electron pairs ('Cooper pairs'), which in a sense Bose condense.

This two-step picture is too naive however since in many systems, the size of Cooper pairs (the *coherence length*) is much larger than the average distance between pairs. For this reason, binding and condensation must be treated together in theory.



From a historical point of view, it is striking that it took more than 40 years between the experimental discovery of superconductivity by Kamerlingh-Onnes in 1911 (Nobel prize

in 1913) and the theoretical understanding due to Bardeen, Cooper and Schrieffer in 1957. This underlines what a revolutionary advance their theory represents.

7.1 Electron-phonon interaction

As an electron passes through the solid, due to its negative charge, it leaves behind a deformation trail, effecting the positions of the ion cores. These deformations are associated with an accumulation of positive charge, giving rise to an attractive potential for the other electrons. Let us derive this phonon-mediated, attractive electron-electron interaction from a simple toy model, $\hat{H} = \hat{H}_0 + \hat{H}_1$, written in terms of electron operators $c^\dagger_{\mathbf{k}\sigma}$, $c_{\mathbf{k}\sigma}$ and phonon operators $a^\dagger_{\mathbf{q}}$, $a_{\mathbf{q}}$,

$$\hat{H}_0 = \sum_{\mathbf{k}\sigma} \epsilon(\mathbf{k}) c^\dagger_{\mathbf{k}\sigma} c_{\mathbf{k}\sigma} + \hbar\omega_D \sum_{\mathbf{q}} a^\dagger_{\mathbf{q}} a_{\mathbf{q}}, \tag{7.1a}$$

$$\hat{H}_1 = \lambda \sum_{\mathbf{q}\mathbf{k}\sigma} \left(c^\dagger_{\mathbf{k}+\mathbf{q},\sigma} c_{\mathbf{k}\sigma} a_{\mathbf{q}} + c^\dagger_{\mathbf{k}-\mathbf{q},\sigma} c_{\mathbf{k}\sigma} a^\dagger_{\mathbf{q}} \right). \tag{7.1b}$$

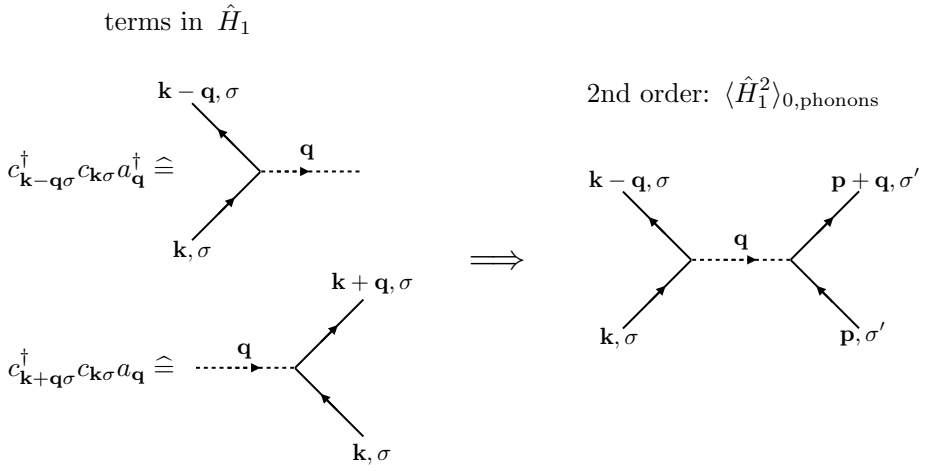
Here we have considered a single dispersionless phonon mode with frequency ω_D. This is known as the Einstein model which treats the solid as many individual, non-interacting quantum oscillators. $\hat{H}_1$ is the coupling between the electrons and the phonons. To be more precise, a coupling between the local electron density $\hat{n}_{\mathbf{r}} = \sum_\sigma c^\dagger_{\mathbf{r}\sigma} c_{\mathbf{r}\sigma}$ and the local lattice deformation $\delta\hat{x}_{\mathbf{r}} \sim a^\dagger_{\mathbf{r}} + a_{\mathbf{r}}$,

$$\begin{aligned}
\hat{H}_1 &\sim \sum_{\mathbf{r}} \hat{n}_{\mathbf{r}} \delta\hat{x} \sim \sum_{\mathbf{r}\sigma} c^\dagger_{\mathbf{r}\sigma} c_{\mathbf{r}\sigma} \left(a^\dagger_{\mathbf{r}} + a_{\mathbf{r}} \right) \\
&\sim \sum_{\mathbf{r}\sigma} \sum_{\mathbf{k}\mathbf{k}'\mathbf{q}} e^{i\mathbf{k}'\mathbf{r}} e^{-i\mathbf{k}\mathbf{r}} c^\dagger_{\mathbf{k}'\sigma} c_{\mathbf{k}\sigma} \left(e^{i\mathbf{q}\mathbf{r}} a^\dagger_{\mathbf{q}} + e^{-i\mathbf{q}\mathbf{r}} a_{\mathbf{q}} \right) \\
&\sim \sum_{\mathbf{k}\mathbf{k}'\mathbf{q}\sigma} c^\dagger_{\mathbf{k}'\sigma} c_{\mathbf{k}\sigma} \left[\delta(\mathbf{k}' - \mathbf{k} + \mathbf{q}) a^\dagger_{\mathbf{q}} + \delta(\mathbf{k}' - \mathbf{k} - \mathbf{q}) a_{\mathbf{q}} \right] \\
&= \sum_{\mathbf{k}\mathbf{q}\sigma} \left(c^\dagger_{\mathbf{k}-\mathbf{q}\sigma} c_{\mathbf{k}\sigma} a^\dagger_{\mathbf{q}} + c^\dagger_{\mathbf{k}+\mathbf{q}\sigma} c_{\mathbf{k}\sigma} a_{\mathbf{q}} \right),
\end{aligned}$$

in agreement with the electron-phonon Hamiltonian (7.1). We wish to focus on the electron system. To this end we eliminate the electron phonon coupling by means of a canonical transformation which we determine perturbatively. This is equivalent to treating the effect of $\hat{H}_1$ in 2nd order perturbation theory.
The resulting 2nd order diagram has the form of an electron-electron interaction,

$$\hat{H}_{\text{int}} = \sum_{\mathbf{k}\mathbf{p}\mathbf{q}} \sum_{\sigma\sigma'} V_{\mathbf{k}\mathbf{p}\mathbf{q}} c^\dagger_{\mathbf{p}+\mathbf{q}\sigma'} c_{\mathbf{p}\sigma'} c^\dagger_{\mathbf{k}-\mathbf{q}\sigma} c_{\mathbf{k}\sigma}. \tag{7.2}$$

To compute the matrix elements $V_{\mathbf{k}\mathbf{p}\mathbf{q}}$ we proceed with the canonical transformation,



$$\begin{aligned}\tilde{H} &= e^{-\hat{S}}\hat{H}e^{\hat{S}} = \hat{H} + [\hat{H},\hat{S}] + \frac{1}{2}[[\hat{H},\hat{S}],\hat{S}] + \dots \\ &= \hat{H}_0 + \hat{H}_1 + [\hat{H}_0,\hat{S}] + [\hat{H}_1,\hat{S}] + \frac{1}{2}[[\hat{H}_0,\hat{S}],\hat{S}] + \dots, \end{aligned} \tag{7.3}$$

and at leading order we fix $\hat{S}$ by demanding that

$$\hat{H}_1 + [\hat{H}_0,\hat{S}] = 0. \tag{7.4}$$

This way we have transformed away the linear $\hat{H}_1$ term. Inserting Eq. (7.4) into Eq. (7.3), we obtain

$$\tilde{H} \approx \hat{H}_0 + \frac{1}{2}[\hat{H}_1,\hat{S}] =: \hat{H}_0 + \hat{H}_{\text{int}}. \tag{7.5}$$

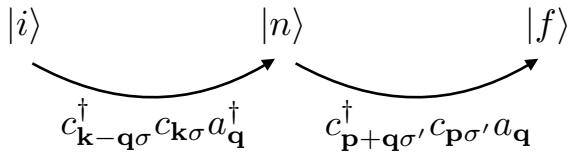
From Eq. (7.4) we can obtain a relation between the matrix elements of $\hat{S}$ and $\hat{H}_1$ in the basis $\{|n\rangle\}$ of eigenstates of $\hat{H}_0$ ($\hat{H}_0|n\rangle = E_n|n\rangle$),

$$\langle n|\hat{H}_1|m\rangle = -\langle n|\hat{H}_0\hat{S} - \hat{S}\hat{H}_0|m\rangle = (E_m - E_n)\langle n|\hat{S}|m\rangle. \tag{7.6}$$

We now use this relation to calculate the matrix elements of $\hat{H}_{\text{int}}$ for any given initial and final eigenstates of $\hat{H}_0$:

$$\begin{aligned}\langle f|\hat{H}_{\text{int}}|i\rangle &\overset{(7.5)}{=} \frac{1}{2}\langle f|\,[\hat{H}_1,\hat{S}]\,|i\rangle = \frac{1}{2}\langle f|\hat{H}_1\hat{S} - \hat{S}\hat{H}_1|i\rangle \\ &\overset{1=\sum_n |n\rangle\langle n|}{=} \frac{1}{2}\sum_n \Big(\langle f|\hat{H}_1|n\rangle\langle n|\hat{S}|i\rangle - \langle f|\hat{S}|n\rangle\langle n|\hat{H}_1|i\rangle\Big) \\ &\overset{(7.6)}{=} \frac{1}{2}\sum_n \langle f|\hat{H}_1|n\rangle\langle n|\hat{H}_1|i\rangle \left(\frac{1}{E_i - E_n} + \frac{1}{E_f - E_n}\right). \end{aligned} \tag{7.7}$$

From the matrix elements (7.7) and the definition of $\hat{H}_1$ (7.1) we can read of $\hat{H}_{\text{int}}$ in operator form. Let's start with an initial $|i\rangle$ without any phonons. In order to obtain a final state



$|f\rangle$ without phonons we first have to act with the $a^\dagger_{\mathbf{q}}$ term and then with the $a_{\mathbf{q}}$ term in $\hat{H}_1$,
The intermediate states $|n\rangle$ are states which contain one excited phonon. The corresponding energy changes are

$$\begin{aligned} E_n - E_i &= \epsilon(\mathbf{k}-\mathbf{q}) + \hbar\omega_D - \epsilon(\mathbf{k}), \\ E_f - E_i &= \epsilon(\mathbf{p}+\mathbf{q}) - \epsilon(\mathbf{p}) - \hbar\omega_D, \end{aligned}$$

from which we obtain

$$\begin{aligned} \hat{H}_{\rm int} &= \frac{1}{2}|\lambda|^2 \sum_{\mathbf{kpq}} \sum_{\sigma\sigma'} c^\dagger_{\mathbf{p}+\mathbf{q}\sigma'} c_{\mathbf{p}\sigma'} c^\dagger_{\mathbf{k}-\mathbf{q}\sigma} c_{\mathbf{k}\sigma} \\ &\quad \times \left(\frac{1}{\epsilon(\mathbf{k}) - \epsilon(\mathbf{k}-\mathbf{q}) - \hbar\omega_D} + \frac{1}{\epsilon(\mathbf{p}+\mathbf{q}) - \epsilon(\mathbf{p}) - \hbar\omega_D} \right). \end{aligned} \tag{7.8}$$

7.2 The Cooper problem

The phonon-mediated electron-electron interaction (7.8) is attractive for pairs of electrons with

$$|\epsilon(\mathbf{k}) - \epsilon(\mathbf{k} \pm \mathbf{q})| < \hbar\omega_D,$$

that is, for pairs within the Debye energy $\hbar\omega_D$ of the Fermi surface. It is however very weak.

This leaves us with a puzzle: for two particles moving in free space in three dimensions, an attractive interaction must exceed a critical strength to produce a bound state. How can a very weak attraction lead to superconductivity? The answer is that we have to consider a pair of quasiparticles moving not in free space, but above a filled Fermi sea. As we will show, Pauli exclusion facilitates binding!

Let us consider the wavefunction for such a pair of quasi-electrons. Since we want to write down a low-energy state, we set the centre-of-mass momentum to zero and choose the pair to be in a spin-singlet configuration. In this case the spatial wavefunction is symmetric, allowing us to take advantage of a local attractive interaction. To respect Pauli exclusion from a filled Fermi sea, we require the wavefunction to be built from momenta outside the Fermi surface. The general form of the pair wavefunction is then given by

$$\phi(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{\sqrt{2}} (\uparrow_1\downarrow_2 - \downarrow_1\uparrow_2) \sum_{\mathbf{k}}^{|\mathbf{k}|>k_F} g(\mathbf{k}) e^{i\mathbf{k}(\mathbf{r}_1 - \mathbf{r}_2)}. \tag{7.9}$$

The function $g(\mathbf{k})$ is obtained from requiring that $\phi(\mathbf{r}_1, \mathbf{r}_2)$ is a solution to the two-particle Schrödinger equation,

$$\left[-\frac{\hbar^2}{2m}(\nabla_1^2 + \nabla_2^2) + U(\mathbf{r}_1 - \mathbf{r}_2)\right] \phi(\mathbf{r}_1, \mathbf{r}_2) = E\phi(\mathbf{r}_1, \mathbf{r}_2), \tag{7.10}$$

where $U(\mathbf{r})$ denotes the interaction potential and E the pair energy. Inserting the Ansatz (7.9), we obtain

$$\sum_{\mathbf{k}}^{|\mathbf{k}|>k_F} g(\mathbf{k}) e^{i\mathbf{kr}} U(\mathbf{r}) = \sum_{\mathbf{k}}^{|\mathbf{k}|>k_F} \left[E - 2\epsilon(\mathbf{k})\right] g(\mathbf{k}) e^{i\mathbf{kr}},$$

with $\epsilon(\mathbf{k}) = \frac{\hbar^2 k^2}{2m}$ and $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$ the relative coordinate. We operate with $\frac{1}{V}\int d^3 r e^{-i\mathbf{qr}} \ldots$ on both sides, giving

$$\frac{1}{V} \sum_{\mathbf{k}}^{|\mathbf{k}|>k_F} g(\mathbf{k}) U_{\mathbf{k},\mathbf{q}} = \left[E - 2\epsilon(\mathbf{q})\right] g(\mathbf{q})$$

with $U_{\mathbf{k},\mathbf{q}} = \int d^3 r e^{i(\mathbf{k}-\mathbf{q})\mathbf{r}} U(\mathbf{r})$. We can understand the essentials in a simple way by taking

$$U_{\mathbf{k},\mathbf{q}} = \begin{cases} -U & \text{if} \quad \epsilon_F - \hbar\omega_D < \epsilon(\mathbf{k}), \epsilon(\mathbf{q}) < \epsilon_F + \hbar\omega_D \\ 0 & \text{otherwise} \end{cases} \tag{7.11}$$

For this simplified form of the interaction potential we obtain

$$\frac{U}{V} \sum_{\mathbf{k}}^{\epsilon_F < \epsilon(\mathbf{k}) < \epsilon_F + \hbar\omega_D} g(\mathbf{k}) = \left[2\epsilon(\mathbf{q}) - E\right] g(\mathbf{q}).$$

Since the left hand side of this equation is independent of $\mathbf{q}$ both sides are equal to a constant A and therefore

$$g(\mathbf{q}) = \frac{A}{2\epsilon(\mathbf{q}) - E}.$$

Summing this equation over momenta $\mathbf{q}$ with $\epsilon_F < \epsilon(\mathbf{q}) < \epsilon_F + \hbar\omega_D$, we obtain

$$A \sum_{\mathbf{q}}^{\epsilon_F < \epsilon(\mathbf{q}) < \epsilon_F + \hbar\omega_D} \frac{1}{2\epsilon(\mathbf{q}) - E} = \sum_{\mathbf{q}}^{\epsilon_F < \epsilon(\mathbf{q}) < \epsilon_F + \hbar\omega_D} g(\mathbf{q}) = \frac{V}{U} A,$$

and therefore the self-consistency equation

$$\frac{1}{U} = \frac{1}{V} \sum_{\mathbf{q}}^{\epsilon_F < \epsilon(\mathbf{q}) < \epsilon_F + \hbar\omega_D} \frac{1}{2\epsilon(\mathbf{q}) - E},$$

which determines the energy E of the pair. We calculate the momentum sum by converting it into an integral over energy and approximate the DOS in the small interval $[\epsilon_F, \epsilon_F + \hbar\omega_D]$ (small relative to ϵ_F) by a constant, $g(\epsilon)/(2V) \approx g(\epsilon_F)/(2V) =: \nu$ (DOS at Fermi energy per volume per spin),

$$\begin{aligned}\frac{1}{U} &= \frac{1}{2V}\int_{\epsilon_F}^{\epsilon_F+\hbar\omega_D}\frac{g(\epsilon)}{2\epsilon-E} \approx \nu\int_{\epsilon_F}^{\epsilon_F+\hbar\omega_D}\frac{1}{2\epsilon-E} \\ &= \nu\left[\frac{1}{2}\ln|2\epsilon-E|\right]_{\epsilon_F}^{\epsilon_F+\hbar\omega_D} = \frac{\rho}{2}\ln\frac{2(\epsilon_F+\hbar\omega_D)-E}{2\epsilon_F-E} \\ &= \frac{\nu}{2}\ln\left(1+\frac{2\hbar\omega_D}{2\epsilon_F-E}\right)\end{aligned}$$

At weak coupling (very small U), the argument of the logarithm is large and we can neglect the 1. After exponentiating the equation, we obtain an expression for the Cooper pair binding energy,

$$2\epsilon_F - E = 2\hbar\omega_D e^{-2/(\nu U)}. \tag{7.12}$$

This result shows that even for very small attraction U there exists a bound state: $2\epsilon_F - E > 0$ which means that the energy of two quasiparticles at the Fermi energy is higher than the energy E of the bound pair. The exponentially weak dependence on νU shows that the result is non-perturbative in the interaction.

7.3 The BCS wavefunction

Starting from the pair wavefunction of the form

$$\phi(\mathbf{r}_1,\mathbf{r}_2) = \frac{1}{\sqrt{2}}\left(\uparrow_1\downarrow_2 - \downarrow_1\uparrow_2\right) g(\mathbf{r}_1,\mathbf{r}_2),$$

$\phi(\mathbf{r}_2,\mathbf{r}_1) = -\phi(\mathbf{r}_1,\mathbf{r}_2)$, we can attempt to write down a wavefunction for $2N$ electrons in which pairs are Bose condensed, of the form

$$\Psi(\mathbf{r}_1,\ldots,\mathbf{r}_{2N}) = \mathcal{A}\prod_{i=1}^{N}\phi(\mathbf{r}_{2i-1},\mathbf{r}_{2i}), \tag{7.13}$$

where $\mathcal{A}$ denotes anti-symmetrisation. To illustrate this, let's consider the case of $N = 2$ pairs,

$$\begin{aligned}\Psi(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3,\mathbf{r}_4) &= \mathcal{A}\phi(\mathbf{r}_1,\mathbf{r}_2)\phi(\mathbf{r}_3,\mathbf{r}_4) \\ &= \phi(\mathbf{r}_1,\mathbf{r}_2)\phi(\mathbf{r}_3,\mathbf{r}_4) - \phi(\mathbf{r}_1,\mathbf{r}_3)\phi(\mathbf{r}_2,\mathbf{r}_4) + \phi(\mathbf{r}_1,\mathbf{r}_4)\phi(\mathbf{r}_2,\mathbf{r}_3).\end{aligned}$$

Note that $\Psi(\mathbf{r}_1,\ldots,\mathbf{r}_{2N})$ is an equal-weight superposition of all possible pairs, respecting the overall anti-symmetry of the wavefunction under the exchange of two fermions,

$$\Psi(\ldots,\mathbf{r}_i,\ldots,\mathbf{r}_j,\ldots) = -\Psi(\ldots,\mathbf{r}_j,\ldots,\mathbf{r}_i,\ldots). \tag{7.14}$$

The wavefunction (7.13) is much easier to write down in momentum space, using fermionic creation operators. For a single pair in a singlet configuration and with zero centre of mass momentum the wavefunction is given by

$$|\phi\rangle = \sum_{\mathbf{k}} g_{\mathbf{k}} c^\dagger_{\mathbf{k}\uparrow} c^\dagger_{-\mathbf{k}\downarrow} |0\rangle, \tag{7.15}$$

where $|0\rangle$ denotes the vacuum state. For the $2N$-particle state we therefore obtain

$$|\Psi\rangle = \prod_{i=1}^{N} \left(\sum_{\mathbf{k}_i} g_{\mathbf{k}_i} c^\dagger_{\mathbf{k}_i\uparrow} c^\dagger_{-\mathbf{k}_i\downarrow} \right) |0\rangle. \tag{7.16}$$

Expanding the product, only configurations with all $\mathbf{k}_i$'s different will survive because of Pauli exclusion. Moreover, the anti-commutator relations of Fermi creation operators guarantee the anti-symmetry of the wavefunction. The form of the wavefunction (7.16) is not very convenient, because the occupation of different orbitals is correlated to the constraint that exactly $2N$ electrons are present. The BCS wavefunction relaxes this constraint. This is similar to going from the canonical to the grand-canonical ensemble in statistical physics. The BCS wavefunction has the form

$$|\mathrm{BCS}\rangle = \prod_{\mathbf{k}} \left(u_{\mathbf{k}} + v_{\mathbf{k}} c^\dagger_{\mathbf{k}\uparrow} c^\dagger_{-\mathbf{k}\downarrow} \right) |0\rangle. \tag{7.17}$$

Here $u_{\mathbf{k}}$ is the amplitude for a pair of orbitals $(\mathbf{k}, -\mathbf{k})$ to be empty, whereas $v_{\mathbf{k}}$ is the amplitude for them to contain a Cooper pair. So we require

$$|u_{\mathbf{k}}|^2 + |v_{\mathbf{k}}|^2 = 1 \tag{7.18}$$

for normalisation, $\langle \mathrm{BCS}|\mathrm{BCS}\rangle = 1$, and

$$N = \langle \mathrm{BCS}|\hat{N}|\mathrm{BCS}\rangle = \sum_{\mathbf{k}} |v_{\mathbf{k}}|^2 \tag{7.19}$$

in a system containing $2N$ electrons on average (homework).

A filled Fermi sea simply has $v_{\mathbf{k}} = 1$ for all $\mathbf{k}$ inside the Fermi surface and $v_{\mathbf{k}} = 0$ outside. By contrast, in a state containing many Cooper pairs, we expect $v_{\mathbf{k}}$ to vary smoothly between 1 and 0 across a window of width $\hbar\omega_D$ of the Fermi surface. One approach is to find $v_{\mathbf{k}}$ *variationally* by minimising the grand canonical potential

$$\Omega = E - \mu N = \langle \mathrm{BCS}|\hat{H} - \mu\hat{N}|\mathrm{BCS}\rangle \tag{7.20}$$

over the wavefunctions parametrised by Eq. (7.17). This is subject to a homework problem. An alternative is a mean-field treatment of the attractive electron-electron interaction, as discussed in the next section.

7.4 BCS mean-field theory

Using the simplified paring interaction (7.11), our starting Hamiltonian is given by

$$\hat{H} = \sum_{\mathbf{k}\sigma} \epsilon_{\mathbf{k}} c^\dagger_{\mathbf{k}\sigma} c_{\mathbf{k}\sigma} - U \sum_{\mathbf{kq}}^{(*)} c^\dagger_{\mathbf{k}\uparrow} c^\dagger_{-\mathbf{k}\downarrow} c_{-\mathbf{q}\downarrow} c_{\mathbf{q}\uparrow}, \tag{7.21}$$

where $\sum_{\mathbf{kq}}^{(*)}$ is the sum over momenta $\mathbf{k}$, $\mathbf{q}$ with $\epsilon_F - \hbar\omega_D < \epsilon_\mathbf{k}, \epsilon_\mathbf{q} < \epsilon_F + \hbar\omega_D$. We wish to treat the interaction in a *mean-field approximation.*

Let us first consider an operator product $\hat{A}\cdot\hat{B}$ and rewrite the operators as their expectation value plus the operator that measures the deviation from the expectation value, e.g. $\hat{A} = \langle\hat{A}\rangle + \delta\hat{A}$. Assuming that $\delta\hat{A}$ and $\delta\hat{B}$ are small, we approximate

$$\begin{aligned}\hat{A}\cdot\hat{B} &= \left(\langle\hat{A}\rangle + \delta\hat{A}\right)\left(\langle\hat{B}\rangle + \delta\hat{B}\right)\\ &\approx \langle\hat{A}\rangle\langle\hat{B}\rangle + \langle\hat{A}\rangle\delta\hat{B} + \delta\hat{A}\langle\hat{B}\rangle\\ &\overset{\delta\hat{A}=\hat{A}-\langle\hat{A}\rangle}{=} \langle\hat{A}\rangle\hat{B} + \hat{B}\langle\hat{A}\rangle - \langle\hat{A}\rangle\langle\hat{B}\rangle. \end{aligned} \tag{7.22}$$

Using this mean-field decoupling for $\hat{A} = c^\dagger_{\mathbf{k}\uparrow}c^\dagger_{-\mathbf{k}\downarrow}$ and $\hat{B} = c_{-\mathbf{q}\downarrow}c_{\mathbf{q}\uparrow}$ and defining $\lambda_\mathbf{q} = \langle c_{-\mathbf{q}\downarrow}c_{\mathbf{q}\uparrow}\rangle$, we obtain

$$c^\dagger_{\mathbf{k}\uparrow}c^\dagger_{-\mathbf{k}\downarrow}c_{-\mathbf{q}\downarrow}c_{\mathbf{q}\uparrow} \approx \lambda^*_\mathbf{k}c_{-\mathbf{q}\downarrow}c_{\mathbf{q}\uparrow} + \lambda_\mathbf{q}c^\dagger_{\mathbf{k}\uparrow}c^\dagger_{-\mathbf{k}\downarrow} - \lambda^*_\mathbf{k}\lambda_\mathbf{q}. \tag{7.23}$$

Note that in the context of superconductivity this is the natural decoupling since $\lambda_\mathbf{k} = \langle c_{-\mathbf{k}\downarrow}c_{\mathbf{k}\uparrow}\rangle$ is directly linked to the formation of Cooper pairs. We define the superconducting order parameter as

$$\Delta = U\sum_\mathbf{k}^{(*)}\lambda_\mathbf{k} = U\sum_\mathbf{k}^{(*)}\langle c_{-\mathbf{k}\downarrow}c_{\mathbf{k}\uparrow}\rangle. \tag{7.24}$$

Using the mean-field approximation (7.23) we obtain the BCS mean-field Hamiltonian

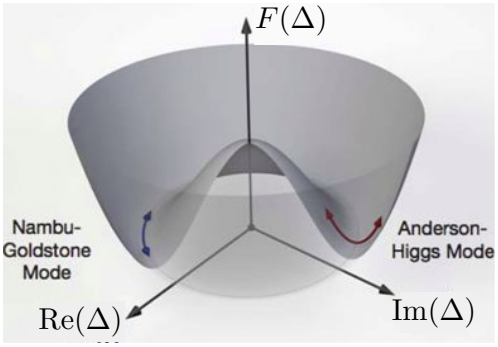
$$\hat{H}_\mathrm{mf} - \mu\hat{N} = \sum_{\mathbf{k}\sigma}\underbrace{(\epsilon_\mathbf{k}-\mu)}_{=:\xi_\mathbf{k}}c^\dagger_{\mathbf{k}\sigma}c_{\mathbf{k}\sigma} - \sum_\mathbf{k}^{(*)}\left(\Delta^* c_{-\mathbf{k}\downarrow}c_{\mathbf{k}\uparrow} + \Delta c^\dagger_{\mathbf{k}\uparrow}c^\dagger_{-\mathbf{k}\downarrow}\right) + \frac{|\Delta|^2}{U}. \tag{7.25}$$

For non-zero Δ, this Hamiltonian does not conserve the particle number, but it can be diagonalised by a fermionic *Bogoliubov transformation.* Note that the order parameter is *complex*, $\Delta = |\Delta|e^{i\phi}$. However, the free energy is independent of the phase ϕ. In the superconducting phase this continuous $U(1)$ symmetry is *spontaneously broken* and there exists a *Goldstone mode* corresponding to a change of the phase ϕ.

In the present calculation we can assume Δ to be real, without any loss of generality. Note that for a general complex order parameter we could use a Bogoliubov transformation with complex coefficients to transform away the phase of Δ. In the following, we assume $\Delta^* = \Delta$ and use the real-valued Bogoliubov transformation discussed in Section 2,

$$\begin{pmatrix} c_1 \\ c_2^\dagger \end{pmatrix} = \begin{pmatrix} u & v \\ -v & u \end{pmatrix}\begin{pmatrix} d_1 \\ d_2^\dagger \end{pmatrix},$$

where the Fermi anti-commutator relations are preserved if $u^2 + v^2 = 1$. We therefore make an ansatz



$$\begin{pmatrix} c_{\mathbf{k}\uparrow} \\ c^\dagger_{-\mathbf{k}\downarrow} \end{pmatrix} = \begin{pmatrix} u_\mathbf{k} & v_\mathbf{k} \\ -v_\mathbf{k} & u_\mathbf{k} \end{pmatrix} \begin{pmatrix} d_{\mathbf{k},1} \\ d^\dagger_{\mathbf{k},2} \end{pmatrix}, \tag{7.26}$$

with $u_\mathbf{k} = \cos\theta_\mathbf{k}$ and $v_\mathbf{k} = \sin\theta_\mathbf{k}$. Inserting the transformation (7.26) into the BCS mean-field Hamiltonian (7.25) and determining $\theta_\mathbf{k}$ such that the non-number-conserving terms vanish, we obtain (homework)

$$\cot(2\theta_\mathbf{k}) = \frac{\xi_\mathbf{k}}{\Delta} \tag{7.27}$$

and the diagonalised mean-field Hamiltonian

$$\hat{H}_\mathrm{mf} - \mu\hat{N} = \sum_{\mathbf{k}\alpha} \sqrt{\xi_\mathbf{k}^2 + \Delta^2}\, d^\dagger_{\mathbf{k}\alpha} d_{\mathbf{k}\alpha} + \sum_\mathbf{k} \left(\xi_\mathbf{k} - \sqrt{\xi_\mathbf{k}^2 + \Delta^2} \right) + \frac{\Delta^2}{U}. \tag{7.28}$$

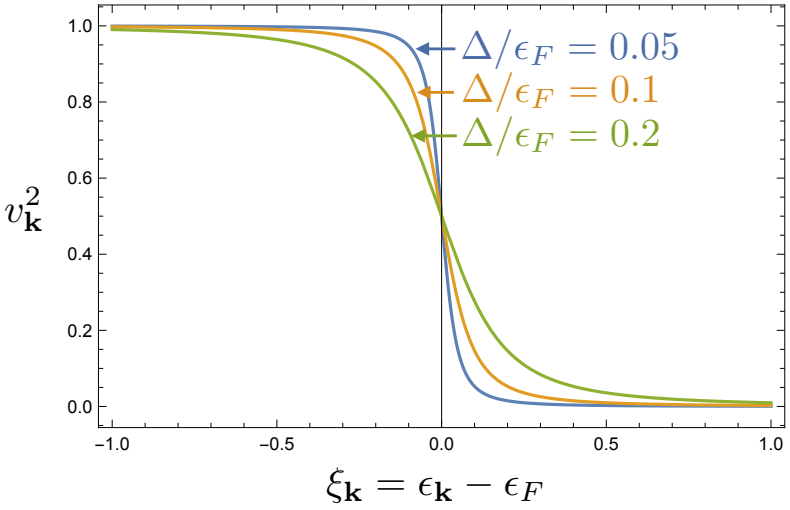
For completeness, we also give the matrix elements of the Bogoliubov transformation (7.26) that diagonalises the Hamiltonian:

$$u_\mathbf{k}^2 = \frac{1}{2}\left(1 + \frac{\xi_\mathbf{k}}{\sqrt{\xi_\mathbf{k}^2 + \Delta^2}}\right) \quad \text{and} \quad v_\mathbf{k}^2 = \frac{1}{2}\left(1 - \frac{\xi_\mathbf{k}}{\sqrt{\xi_\mathbf{k}^2 + \Delta^2}}\right) \tag{7.29}$$

Plotting $v_\mathbf{k}^2$, we find that it behaves exactly as we expect for the $v_\mathbf{k}$ we have introduced when we wrote down the ansatz for the BCS wavefunction (7.17). It is equal to the $T = 0$ Fermi step function for $\Delta = 0$ and smears around ϵ_F as we increase Δ. We will show that the ground state $|\Phi_0\rangle$ of the BCS mean-field Hamiltonian is indeed equal to $|\mathrm{BCS}\rangle = \prod_\mathbf{k} \left(u_\mathbf{k} + v_\mathbf{k} c^\dagger_{\mathbf{k}\uparrow} c^\dagger_{-\mathbf{k}\downarrow}\right)|0\rangle$ with $u_\mathbf{k}$ and $v_\mathbf{k}$ the coefficients of the Bogoliubov transformation (7.26).

The groundstate $|\Phi_0\rangle$ of the BCS mean-field Hamiltonian (7.28) is the state where no Bogoliubov quasiparticles (created and annihilated by $d^\dagger_{\mathbf{k}\alpha}$ and $d_{\mathbf{k}\alpha}$) are present. Creating such an excitation would cost an energy equal or larger than Δ. A state $|\Phi_0\rangle$ for which $d_{\mathbf{k}\alpha}|\Phi_0\rangle = 0$ for all $(\mathbf{k}, \alpha)$ can be easily written down,

$$|\Phi_0\rangle = \mathcal{N} \prod_{\mathbf{k}\alpha} d_{\mathbf{k}\alpha}|0\rangle = \mathcal{N} \prod_\mathbf{k} d_{\mathbf{k},2} d_{\mathbf{k},1}|0\rangle, \tag{7.30}$$



where $|0\rangle$ denotes the vacuum with no electrons present and $\mathcal{N}$ is a constant that ensures normalisation. Note that the above considerations hold if $|0\rangle$ is replaced by another wave-function $|\psi_0>$. However, only for $|\psi_0>=|0\rangle$ we recover the filled Fermi sea for $\Delta = 0$. From the inverse of the transformation (7.26),

$$\begin{pmatrix} d_{\mathbf{k},1} \\ d^\dagger_{\mathbf{k},2} \end{pmatrix} = \begin{pmatrix} u_{\mathbf{k}} & -v_{\mathbf{k}} \\ v_{\mathbf{k}} & u_{\mathbf{k}} \end{pmatrix} \begin{pmatrix} c_{\mathbf{k}\uparrow} \\ c^\dagger_{-\mathbf{k}\downarrow} \end{pmatrix},$$

we obtain

$$\begin{aligned} |\Phi_0\rangle &= \mathcal{N} \prod_{\mathbf{k}} \left(v_{\mathbf{k}} c^\dagger_{\mathbf{k}\uparrow} + u_{\mathbf{k}} c_{-\mathbf{k}\downarrow} \right) \left(u_{\mathbf{k}} c_{\mathbf{k}\uparrow} - v_{\mathbf{k}} c^\dagger_{-\mathbf{k}\downarrow} \right) |0\rangle \\ &= \mathcal{N} \prod_{\mathbf{k}} \left(-v_{\mathbf{k}}^2 c^\dagger_{\mathbf{k}\uparrow} c^\dagger_{-\mathbf{k}\downarrow} - u_{\mathbf{k}} v_{\mathbf{k}} \underbrace{c_{-\mathbf{k}\downarrow} c^\dagger_{-\mathbf{k}\downarrow}}_{=1-c^\dagger_{-\mathbf{k}\downarrow} c_{-\mathbf{k}\downarrow}} \right) |0\rangle \\ &= \mathcal{N} \prod_{\mathbf{k}} (-v_{\mathbf{k}}) \left(u_{\mathbf{k}} + v_{\mathbf{k}} c^\dagger_{\mathbf{k}\uparrow} c^\dagger_{-\mathbf{k}\downarrow} \right) |0\rangle \\ &= \prod_{\mathbf{k}} \left(u_{\mathbf{k}} + v_{\mathbf{k}} c^\dagger_{\mathbf{k}\uparrow} c^\dagger_{-\mathbf{k}\downarrow} \right) |0\rangle = |\mathrm{BCS}\rangle, \end{aligned} \tag{7.31}$$

where in the last step we have absorbed $\prod_{\mathbf{k}}(-v_{\mathbf{k}})$ into the normalisation and used that $|\mathrm{BCS}\rangle$ is already normalised because $u_{\mathbf{k}}^2 + v_{\mathbf{k}}^2 = 1$. This implies that $\mathcal{N}\prod_{\mathbf{k}}(-v_{\mathbf{k}})$ is simply a phase factor.

We have introduced the mean-field superconducting order parameter Δ 'by hand' and diagonalised the resulting Hamiltonian, but haven't answered the question whether it is energetically favourable for the system to become a superconductor ($\Delta > 0$). Instead of minimising the total grand-canonical free energy with respect to Δ, we can solve the self-consistency equation

$$\Delta = U \sum_{\mathbf{k}}^{(*)} \langle c_{-\mathbf{k}\downarrow} c_{\mathbf{k}\uparrow} \rangle,$$

where the expectation value on the r.h.s of the equation has to be taken with respect to the mean-field Hamiltonian which itself depends on Δ. In order to evaluate the integral, we first express the Fermi operators $c_{\mathbf{k}\sigma}$, $c^\dagger_{\mathbf{k}\sigma}$ in terms of the Bogoliubov quasiparticle operators,

$$\begin{aligned}
\Delta &= U \sum_{\mathbf{k}}^{(*)} \langle c_{-\mathbf{k}\downarrow} c_{\mathbf{k}\uparrow} \rangle \overset{(7.26)}{=} U \sum_{\mathbf{k}}^{(*)} \langle \Big(-v_{\mathbf{k}} d^\dagger_{\mathbf{k},1} + u_{\mathbf{k}} d_{\mathbf{k},2} \Big) \Big(u_{\mathbf{k}} d_{\mathbf{k},1} + v_{\mathbf{k}} d^\dagger_{\mathbf{k},2} \Big) \rangle \\
&= U \sum_{\mathbf{k}}^{(*)} \Big(- u_{\mathbf{k}} v_{\mathbf{k}} \langle d^\dagger_{\mathbf{k},1} d_{\mathbf{k},1} \rangle + u_{\mathbf{k}} v_{\mathbf{k}} \underbrace{\langle d_{\mathbf{k},2} d^\dagger_{\mathbf{k},2} \rangle}_{=1-\langle d^\dagger_{\mathbf{k},2} d_{\mathbf{k},2} \rangle} - v_{\mathbf{k}} \underbrace{\langle d^\dagger_{\mathbf{k},1} d^\dagger_{\mathbf{k},2} \rangle}_{=0} + u^2_{\mathbf{k}} \underbrace{\langle d_{\mathbf{k},2} d_{\mathbf{k},1} \rangle}_{=0} \Big) \\
&= U \sum_{\mathbf{k}}^{(*)} u_{\mathbf{k}} v_{\mathbf{k}} \left(1 - \sum_{\alpha} \langle d^\dagger_{\mathbf{k}\alpha} d_{\mathbf{k}\alpha} \rangle \right). \qquad (7.32)
\end{aligned}$$

Note that $\langle d_{\mathbf{k},2} d_{\mathbf{k},1} \rangle = \langle d^\dagger_{\mathbf{k},1} d^\dagger_{\mathbf{k},2} \rangle = 0$ because $\hat{H}_{\rm mf} - \mu \hat{N}$ is diagonal in the d-basis of Bogoliubov quasiparticles. At $T = 0$, $\langle d^\dagger_{\mathbf{k}\alpha} d_{\mathbf{k}\alpha} \rangle = 0$. Using that

$$u_{\mathbf{k}} v_{\mathbf{k}} = \cos\theta_{\mathbf{k}} \sin\theta_{\mathbf{k}} = \frac{1}{2} \sin(2\theta_{\mathbf{k}}) = \frac{1}{2} \frac{1}{\sqrt{1 + \cot^2(2\theta_{\mathbf{k}})}} \overset{(7.27)}{=} \frac{1}{2} \frac{\Delta}{\sqrt{\xi^2_{\mathbf{k}} + \Delta^2}},$$

we obtain the $T = 0$ self-consistency equation

$$\Delta = \frac{U}{2} \sum_{\mathbf{k}}^{(*)} \frac{\Delta}{\sqrt{\xi^2_{\mathbf{k}} + \Delta^2}}.$$

$\Delta = 0$ is always a solution of this equation. Any non-trivial solution $\Delta > 0$ must fulfil the equation

$$1 = \frac{U}{2} \sum_{\mathbf{k}}^{(*)} \frac{1}{\sqrt{\xi^2_{\mathbf{k}} + \Delta^2}}.$$

Approximating the DOS (per volume and per spin) over the interval $[\epsilon_F - \hbar\omega_D, \epsilon_F + \hbar\omega_D]$ by a constant $\nu = g(\epsilon_F)/(2V)$ and assuming that $\hbar\omega_D \gg \Delta$, we obtain

$$\begin{aligned}
1 &= \frac{U\nu}{2} \int_{-\hbar\omega_D}^{\hbar\omega_D} \frac{d\xi}{\sqrt{\xi^2 + \Delta^2}} \overset{x=\xi/\Delta}{=} \frac{U\nu}{2} \int_{-\hbar\omega_D/\Delta}^{\hbar\omega_D/\Delta} \frac{dx}{\sqrt{1+x^2}} = \frac{U\nu}{2} \ln\left(\sqrt{1+x^2} + x \right) \Big|_{-\hbar\omega_D/\Delta}^{\hbar\omega_D/\Delta} \\
&= \frac{U\nu}{2} \ln \left[\frac{\sqrt{1 + \left(\frac{\hbar\omega_D}{\Delta}\right)^2} + \frac{\hbar\omega_D}{\Delta}}{\sqrt{1 + \left(\frac{\hbar\omega_D}{\Delta}\right)^2} - \frac{\hbar\omega_D}{\Delta}} \right] = \frac{U\nu}{2} \ln \left[\frac{\sqrt{1 + \left(\frac{\Delta}{\hbar\omega_D}\right)^2} + 1}{\sqrt{1 + \left(\frac{\Delta}{\hbar\omega_D}\right)^2} - 1} \right] \approx U\nu \ln\left(\frac{2\hbar\omega_D}{\Delta} \right),
\end{aligned}$$

which leads to

$$\Delta = 2\hbar\omega_D e^{-1/(\nu U)}. \tag{7.33}$$

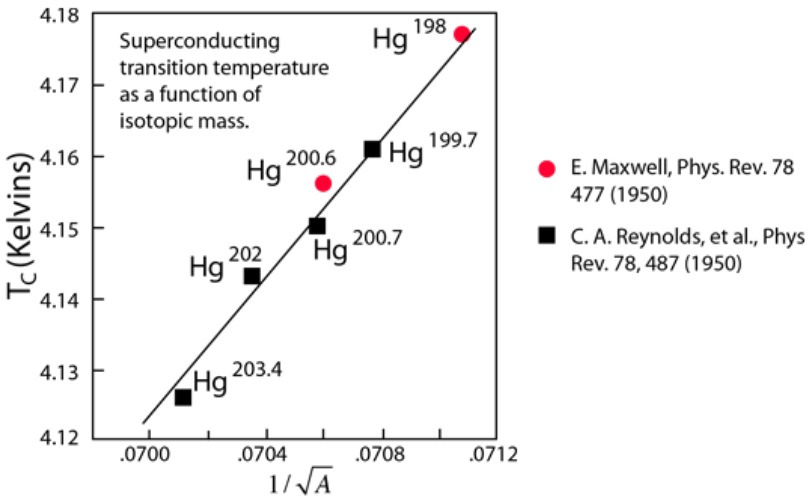
This result is of the same functional form as the pair-binding energy in the Cooper problem, showing once again that the superconducting instability is *non-perturbative* in νU. Regardless how weak the attractive interaction U, the ground state at $T = 0$ is a superconductor with a finite gap $\Delta(T = 0) > 0$, given by Eq. (7.33). At finite temperature, $\langle d^\dagger_{\mathbf{k}\alpha} d_{\mathbf{k}\alpha}\rangle$ is finite and determined from the quasiparticle energies using the Fermi distribution function. Δ decreases as temperature increases. Above a certain critical temperature T_c the only self-consistent solution is $\Delta = 0$. As $T \nearrow T_c$, the gap vanishes as

$$\Delta \overset{T \nearrow T_c}{\sim} (T_c - T)^{1/2}. \tag{7.34}$$

Such a square-root dependence of the order parameter close to the phase transition is generic to all mean-field theories. Since at weak coupling νU is the only parameter, the energy scales set by the zero temperature gap, $\Delta(T = 0)$ (7.33) and the thermal energy scale at the critical point, $k_B T_c$, have a universal relationship,

$$\frac{2\Delta(T = 0)}{k_B T_c} \approx 3.53, \tag{7.35}$$

which serves as a test for the theory.



One of the great successes of the BCS theory is the explanation of the isotope effect, which was first measured in mercury (Hg). Lowering the temperature, mercury solidifies and then becomes superconducting below about 4K. Using different isotopes of Hg, it is possible to change the mass of the ions in the lattice, without changing the electronic structure. In the case of Hg, isotopes with atomic masses between $A = 198$ and $A = 204$ (in units of u) are stable. The experiments show that the superconducting transition temperature T_c is proportional to $1/\sqrt{A}$. This was the first clear indication that phonons played a crucial role in the mechanism for superconductivity. BCS theory indeed shows that both $\Delta(T = 0)$ and T_c [see Eqs. (7.33) and (7.35)] are proportional to the Debye

frequency of phonons which is proportional to the inverse of the square root of the mass of lattice ions.

Chapter 8

Strong correlations

When the average energy scale of repulsive interactions between electrons becomes larger than the kinetic energy, electrons will have the *tendency to localise*. Materials and phenomena for which this factor plays an important role are at the centre of theoretical and experimental research. This interest was especially stimulated by the discovery of high-T_c superconductivity, in which strong electron correlations play a crucial role and weak-coupling BCS theory is no longer applicable.

Many interesting phenomena are connected with strong electron-electron interactions: metal-insulator transitions, magnetic and orbital ordering, heavy fermion behaviour, ... Real materials in which these phenomena occur are transition metal and rare earth compounds. In these materials, the electronic wavefunctions are rather localised, with a spatial extensions that is often smaller than the distance between lattice ions. This is because the electrons are tightly bound to the lattice ions. As a result, the amplitude for an electron to hop from one lattice site to another will be small. As the effective hopping determines the electronic bandwidth and their kinetic energy, we can have a situation in which the kinetic energy is much smaller than the electron interaction.

8.1 Tight-binding approximation

Since in the materials of interest, the electrons are tightly bound to the ions, their wavefunctions will be very similar to those in an isolated atom. In the *tight binding approximation* one expands the Hamiltonian in the basis of local atomic orbitals. In the absence of electron-electron interactions, the resulting tight-binding Hamiltonian describes the hopping of electrons from an atomic orbital β on site j to an atomic orbital α on site i,

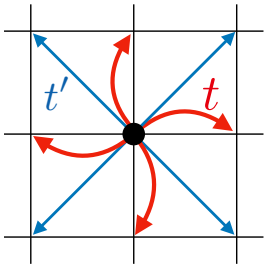
$$\hat{H} = -\sum_{ij}\sum_{\alpha\beta}\sum_{\sigma=\uparrow,\downarrow} t_{ij}^{\alpha\beta} c_{i\alpha\sigma}^{\dagger} c_{j\beta\sigma} + \sum_{i} \epsilon_{\alpha} \hat{n}_{i\alpha}, \tag{8.1}$$

with $\hat{n}_{i\alpha} = \hat{n}_{i\alpha\uparrow} + \hat{n}_{i\alpha\downarrow}$. The hopping amplitudes $t_{ij}^{\alpha\beta}$ depend on the orbitals involved and rapidly decay with the distance between sites i and j. The orbitals could be degenerate or have different energies ϵ_{α}. E.g., for an atom in a cubic crystal environment the five d-orbitals split into two sectors with different energies, the two e_g orbitals ($d_{x^2-y^2}$, d_{z^2}) and the three t_{2g} orbitals (d_{xy}, d_{yz}, d_{xz}).

For simplicity, we will from now on consider systems with one active orbital and include only nearest-neighbour (NN) and next-nearest-neighbour (NNN) hopping processes with amplitudes t and t',

$$\hat{H} = -t \sum_{\langle i,j \rangle} \sum_{\sigma} \left(c^{\dagger}_{i\sigma} c_{j\sigma} + \text{h.c.} \right) - t' \sum_{\langle\langle i,j \rangle\rangle} \sum_{\sigma} \left(c^{\dagger}_{i\sigma} c_{j\sigma} + \text{h.c.} \right). \tag{8.2}$$

Here $\sum_{\langle i,j \rangle}$ and $\sum_{\langle\langle i,j \rangle\rangle}$ denote sums over NN and NNN bonds, respectively, and h.c. stands for hermitian conjugate, $\left(c^{\dagger}_{i\sigma} c_{j\sigma} \right)^{\dagger} = c^{\dagger}_{j\sigma} c_{i\sigma}$.



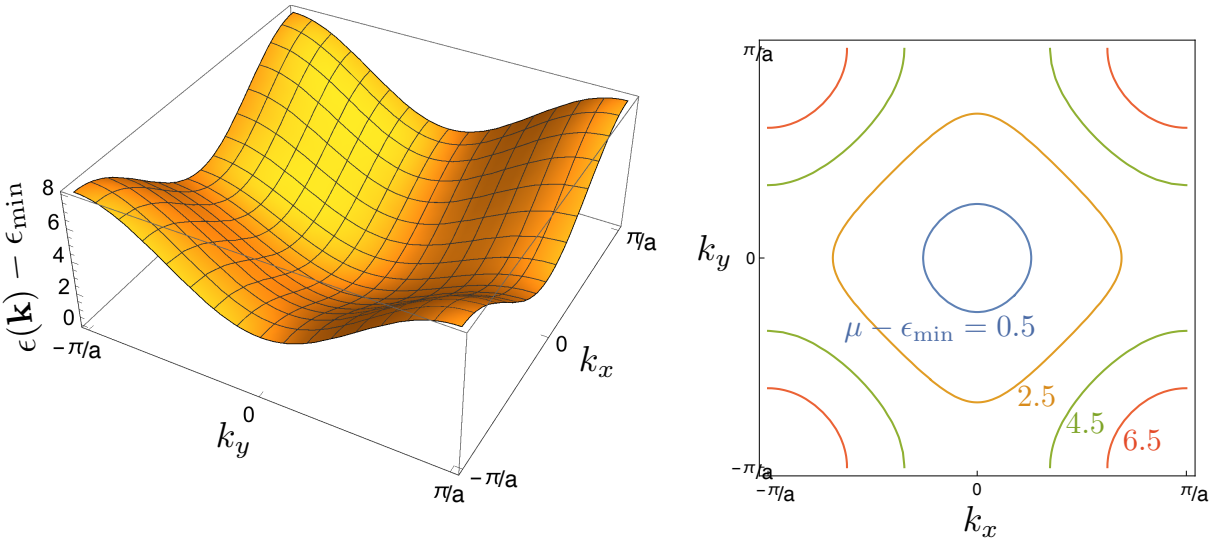
The tight-binding Hamiltonian (8.2) is a free-particle Hamiltonian, describing the motion of non-interacting electrons on a lattice. It can be diagonalised by transforming it to momentum space. For a square lattice we obtain:

$$\begin{aligned}
\hat{H} &= -\sum_{\mathbf{r}\sigma} \left[t \left(c^{\dagger}_{\mathbf{r}+a\hat{\mathbf{e}}_x,\sigma} + c^{\dagger}_{\mathbf{r}+a\hat{\mathbf{e}}_y,\sigma} \right) c_{\mathbf{r},\sigma} + t' \left(c^{\dagger}_{\mathbf{r}+a\hat{\mathbf{e}}_x+a\hat{\mathbf{e}}_y,\sigma} + c^{\dagger}_{\mathbf{r}-a\hat{\mathbf{e}}_x+a\hat{\mathbf{e}}_y,\sigma} \right) \right] + \text{h.c.} \\
&= -\sum_{\mathbf{r}\sigma} \frac{1}{N} \sum_{\mathbf{k}\mathbf{k}'} \left[t \left(e^{-i\mathbf{k}(\mathbf{r}+a\hat{\mathbf{e}}_x)} + e^{-i\mathbf{k}(\mathbf{r}+a\hat{\mathbf{e}}_y)} \right) + t' \left(e^{-i\mathbf{k}(\mathbf{r}+a\hat{\mathbf{e}}_x+a\hat{\mathbf{e}}_y)} + e^{-i\mathbf{k}(\mathbf{r}-a\hat{\mathbf{e}}_x+a\hat{\mathbf{e}}_y)} \right) \right] \\
&\qquad \times e^{i\mathbf{k}'\mathbf{r}} c^{\dagger}_{\mathbf{k}\sigma} c_{\mathbf{k}'\sigma} + \text{h.c.} \\
&= -\sum_{\mathbf{k}\sigma} \left[t \left(e^{-iak_x} + e^{-iak_y} \right) + t' \left(e^{-ia(k_x+k_y)} + e^{-ia(-k_x+k_y)} \right) \right] c^{\dagger}_{\mathbf{k}\sigma} c_{\mathbf{k}\sigma} + \text{h.c.} \\
&= \sum_{\mathbf{k}\sigma} \epsilon(\mathbf{k}) c^{\dagger}_{\mathbf{k}\sigma} c_{\mathbf{k}\sigma},
\end{aligned} \tag{8.3}$$

with the tight-binding dispersion

$$\epsilon(\mathbf{k}) = -2t \left[\cos(ak_x) + \cos(ak_y) \right] - 4t' \cos(ak_x) \cos(ak_y). \tag{8.4}$$

For $t > 0$ and $|t'| \ll t$, the bottom of the band is at $\mathbf{k} = 0$ with minimum energy $\epsilon_{\text{min}} = -4t - 4t'$. In the figure the dispersion and Fermi surfaces for different values of the chemical potential μ are shown for $t = 1.0$ and $t' = -0.1$.



Near the bottom of the band (small density of electrons) we can expand around $\mathbf{k} = 0$,

$$\begin{aligned}\epsilon(\mathbf{k}) - \epsilon_{\min} &\approx -2t\left(2 - \frac{1}{2}a^2k^2\right) - 4t'\left(1 - \frac{1}{2}a^2k_x^2\right)\left(1 - \frac{1}{2}a^2k_y^2\right) - (-4t - 4t') \\ &\approx (t + 2t')a^2k^2.\end{aligned}$$

This shows that the Fermi surface is well approximated by a circle at small filling. Close to the fully occupied band, we obtain an almost circular hole Fermi surface centred around $(\pi/a, \pi/a)$. The electronic bandwidth is equal to

$$W = \epsilon_{\max} - \epsilon_{\min} = \epsilon(\pi/a, \pi/a) - \epsilon(0, 0) = 4t - 4t' - (-4t - 4t') = 8t.$$

8.2 The Hubbard model

As we have discussed in the introduction of this chapter, the electronic bandwidth for tightly bound electrons is small. It is therefore important to include interactions, which are strong relative to the kinetic energy. The Coulomb repulsion is screened out over distances of a few lattice constants. In the Hubbard model one only includes the on-site repulsion U between electrons on the same site, usually referred to as the *Hubbard interaction*,

$$\hat{H} = -t\sum_{\langle i,j\rangle,\sigma}\left(c_{i\sigma}^\dagger c_{j\sigma} + \text{h.c.}\right) + U\sum_i \hat{n}_{i\uparrow}\hat{n}_{i\downarrow}, \tag{8.5}$$

where $\hat{n}_{i\sigma} = \sum_\sigma c_{i\sigma}^\dagger c_{i\sigma}$ is the number operator of electrons with spin σ on site i. Note that because of Pauli principle, two electrons with the same spin cannot occupy the same lattice site. We therefor only have to include the local repulsion between electrons with opposite spin. One might also include the repulsion between electrons on neighbouring lattice sites,

$V \sum_{\langle i,j \rangle} \hat{n}_i \hat{n}_j$ ($\hat{n}_i = \sum_\sigma \hat{n}_{i\sigma}$), or even longer ranged interactions. Such models are usually called *extended Hubbard model*. Other extensions include longer-ranged hopping, multiple orbitals, spin-orbit coupling, etc.

John Hubbard
1931 - 1980

In the following, we will stick to the simplest version of the Hubbard model, Eq. (8.5). It turns out that even this seemingly so simple model describes very rich physics and is at present far from being completely understood.

8.3 Mott insulators and antiferromagnetism

Phillip W. Anderson
1923 -

Sir Nevill Mott
1905 - 1996

John H. van Vleck
1899 - 1980

Nobel Prize in Physics
(theoretical investigation
of electronic structure
in magnetic and
disordered systems)
1977

There are two parameters that control the physical properties of the Hubbard model (8.5), the dimensionless ratio U/t and the electron concentration or band filling $n = N_{\text{electrons}}/N_{\text{sites}}$ ($0 \leq n \leq 2$). The often studied case of one electron per site ($n = 1$) corresponds to *half filling*.

The Hubbard model permits us quite naturally to describe two opposite limits: that of weakly interacting electrons ($U/t \ll 1$), and the case of strongly correlated electrons, $U/t \gg 1$. In the first one expects that the standard Fermi liquid picture will be valid.

Consider now the opposite case of strong interactions ($U \gg t$) at half filling ($n = 1$). In this limit, the system is in an *insulating* state since the gain in kinetic energy due to the hopping of electrons is much smaller than the energy cost for creating a doubly occupied site. This state is called a Mott insulator.
The state with exactly one electron per site is the ground state of the Hamiltonian at $t = 0$,

$$\hat{H}_0 = U \sum_i \hat{n}_{i\uparrow}\hat{n}_{i\downarrow}. \tag{8.6}$$

However, it is not a unique state because of a huge spin degeneracy. Each localised electron can have two spin orientations, $\uparrow$ or $\downarrow$. Hence, for a system with N sites, there are 2^N degenerate ground states. The hopping term

$$\hat{H}_1 = -t \sum_{\langle i,j\rangle,\sigma} \left(c^\dagger_{i\sigma} c_{j\sigma} + \text{h.c.} \right), \tag{8.7}$$

lifts this degeneracy and leads to an *antiferromagnetic* ground state. This can be understood by treating $\hat{H}_1$ in 2nd order perturbation theory. The virtual that start and end in the degenerate ground-state manifold of $\hat{H}_0$ involve a hopping of an electron from site i to site j, creating a doubly occupied state on site j, and then a hopping of one of the two electrons on site j back to site i. Because of Pauli principle this 2nd order process is only possible if the spins on i and j are antiparallel. It lowers the energy by $\sim t^2/U$, stabilising the antiferromagnetic arrangement.

We will show that the effective low-energy Hamiltonian at half filling and large U is indeed a local moment Heisenberg model with antiferromagnetic exchange couplings $J \sim t^2/U$. We use degenerate perturbation theory, treating the hopping $\hat{H}_1$ as a small perturbation to $\hat{H}_0$. We follow the same steps as in the electron-phonon problem, using a canonical transformation. Here we just use the general result for the matrix elements of the effective Hamiltonian,

$$\langle \text{fi}|\hat{H}_{\text{eff}}|\text{in}\rangle = \frac{1}{2}\sum_n \langle \text{fi}|\hat{H}_1|n\rangle\langle n|\hat{H}_1|\text{in}\rangle \left(\frac{1}{E_{\text{fi}} - E_n} - \frac{1}{E_n - E_{\text{in}}} \right), \tag{8.8}$$

where $|\text{in}\rangle$ and $|\text{fi}\rangle$ are initial and final eigenstates of $\hat{H}_0$ in the ground state manifold with one electron on each site. Each such 2nd order process involves only one particular bond $\langle i,j\rangle$ and the effective Hamiltonian is a sum

$$\hat{H}_{\text{eff}} = \sum_{\langle i,j\rangle} \hat{H}^{\langle i,j\rangle}_{\text{eff}}. \tag{8.9}$$

We can therefore derive $\hat{H}_{\rm eff}^{\langle i,j\rangle}$ on a given bond. Hopping is only possible if the electron spins on sites i and j are anti-parallel. The possible initial and final states are therefore

$$|\uparrow,\downarrow\rangle := c_{i\uparrow}^\dagger c_{j\downarrow}^\dagger|0\rangle, \tag{8.10a}$$
$$|\downarrow,\uparrow\rangle := c_{i\downarrow}^\dagger c_{j\uparrow}^\dagger|0\rangle. \tag{8.10b}$$

These are eigenstates of $\hat{H}_0$ with energy $E_{\rm in} = E_{\rm fi} = 0$ There are two possible intermediate states $\{|n\rangle\}$,

$$|\uparrow\downarrow,0\rangle := c_{i\uparrow}^\dagger c_{i\downarrow}^\dagger|0\rangle, \tag{8.11a}$$
$$|0,\uparrow\downarrow\rangle := c_{j\uparrow}^\dagger c_{j\downarrow}^\dagger|0\rangle, \tag{8.11b}$$

with energy $E_n = U$. We therefore obtain

$$\langle{\rm fi}|\hat{H}_{\rm eff}^{\langle i,j\rangle}|{\rm in}\rangle = -\frac{1}{U}\langle{\rm fi}|\hat{H}_1^{\langle i,j\rangle}\Big(|\uparrow\downarrow,0\rangle\langle\uparrow\downarrow,0| + |0,\uparrow\downarrow\rangle\langle 0,\uparrow\downarrow|\Big)\hat{H}_1^{\langle i,j\rangle}|{\rm in}\rangle, \tag{8.12}$$

with

$$\hat{H}_1^{\langle i,j\rangle} = -t\sum_\sigma\left(c_{i\sigma}^\dagger c_{j\sigma} + c_{j\sigma}^\dagger c_{i\sigma}\right).$$

We act with $\hat{H}_1^{\langle i,j\rangle}$ on the initial/final states, Eq. (8.10):

$$\begin{aligned}
\hat{H}_1^{\langle i,j\rangle}|\uparrow,\downarrow\rangle &= -t\left(c_{i\downarrow}^\dagger c_{j\downarrow} + c_{j\uparrow}^\dagger c_{i\uparrow}\right)c_{i\uparrow}^\dagger c_{j\downarrow}^\dagger|0\rangle \\
&= -t\left(-c_{i\downarrow}^\dagger c_{i\uparrow}^\dagger + c_{j\uparrow}^\dagger c_{j\downarrow}^\dagger\right)|0\rangle \\
&= -t\Big(|\uparrow\downarrow,0\rangle + |0,\uparrow\downarrow\rangle\Big), \qquad (8.13a)\\
\hat{H}_1^{\langle i,j\rangle}|\downarrow,\uparrow\rangle &= -t\left(c_{i\uparrow}^\dagger c_{j\uparrow} + c_{j\downarrow}^\dagger c_{i\downarrow}\right)c_{i\downarrow}^\dagger c_{j\uparrow}^\dagger|0\rangle \\
&= -t\left(-c_{i\uparrow}^\dagger c_{i\downarrow}^\dagger + c_{j\downarrow}^\dagger c_{j\uparrow}^\dagger\right)|0\rangle \\
&= t\Big(|\uparrow\downarrow,0\rangle + |0,\uparrow\downarrow\rangle\Big). \qquad (8.13b)
\end{aligned}$$

Using these results, we can evaluate the matrix elements (8.12),

$$\langle\uparrow,\downarrow|\hat{H}_{\rm eff}^{\langle i,j\rangle}|\uparrow,\downarrow\rangle = \langle\downarrow,\uparrow|\hat{H}_{\rm eff}^{\langle i,j\rangle}|\downarrow,\uparrow\rangle = -2t^2/U, \tag{8.14a}$$
$$\langle\uparrow,\downarrow|\hat{H}_{\rm eff}^{\langle i,j\rangle}|\downarrow,\uparrow\rangle = \langle\downarrow,\uparrow|\hat{H}_{\rm eff}^{\langle i,j\rangle}|\uparrow,\downarrow\rangle = 2t^2/U. \tag{8.14b}$$

From the matrix elements we obtain the effective Hamiltonian in operator form,

$$\begin{aligned}\hat{H}_{\rm eff}^{\langle i,j\rangle} &= \frac{2t^2}{U}\Big(|\uparrow,\downarrow\rangle\langle\downarrow,\uparrow| + |\downarrow,\uparrow\rangle\langle\uparrow,\downarrow| - |\uparrow,\downarrow\rangle\langle\uparrow,\downarrow| - |\downarrow,\uparrow\rangle\langle\downarrow,\uparrow|\Big)\\ &= \frac{2t^2}{U}\left(c_{i\uparrow}^\dagger c_{j\downarrow}^\dagger c_{j\uparrow}c_{i\downarrow} + c_{i\downarrow}^\dagger c_{j\uparrow}^\dagger c_{j\downarrow}c_{i\uparrow} - c_{i\uparrow}^\dagger c_{j\downarrow}^\dagger c_{j\downarrow}c_{i\uparrow} - c_{i\downarrow}^\dagger c_{j\uparrow}^\dagger c_{j\uparrow}c_{i\downarrow}\right)\\ &= \frac{2t^2}{U}\sum_\sigma\left(c_{i\sigma}^\dagger c_{i\overline{\sigma}}c_{j\overline{\sigma}}^\dagger c_{j\sigma} - \hat{n}_{i\sigma}\hat{n}_{j\overline{\sigma}}\right), \end{aligned} \tag{8.15}$$

where $\overline{\sigma} =\downarrow$ for $\sigma =\uparrow$ and $\overline{\sigma} =\uparrow$ for $\sigma =\downarrow$, or short $\overline{\sigma} = -\sigma$. We can introduce spin-$\frac{1}{2}$ operators

$$\hat{S}^\alpha = \frac{1}{2}(c_\uparrow^\dagger, c_\downarrow^\dagger)\sigma^\alpha\begin{pmatrix} c_\uparrow \\ c_\downarrow\end{pmatrix}, \tag{8.16}$$

on each site, where $\alpha = x, y, z$ and σ^α denote the spin-$\frac{1}{2}$ Pauli matrices,

$$\sigma^x = \begin{pmatrix} 0 & 1\\ 1 & 0\end{pmatrix}, \quad \sigma^y = \begin{pmatrix} 0 & -i\\ i & 0\end{pmatrix}, \quad \sigma^z = \begin{pmatrix} 1 & 0\\ 0 & -1\end{pmatrix}.$$

It can be shown (homework) that the operators defined in Eq. (8.16) indeed satisfy the spin commutator relations

$$\left[\hat{S}^\alpha, \hat{S}^\beta\right] = i\epsilon_{\alpha\beta\gamma}\hat{S}^\gamma.$$

As expected,

$$\hat{S}^z|\sigma\rangle = \frac{1}{2}\left(c_\uparrow^\dagger c_\uparrow - c_\downarrow^\dagger c_\downarrow\right)c_\sigma^\dagger|0\rangle = \frac{1}{2}\delta_{\sigma,\uparrow}|\uparrow\rangle - \frac{1}{2}\delta_{\sigma,\downarrow}|\downarrow\rangle = \frac{1}{2}\sigma|\sigma\rangle.$$

You will show in a homework problem, that in terms of the spin operators (8.16) the effective Hamiltonian (8.15) takes the form of a Heisenberg model,

$$\hat{H}_{\rm eff} = J\sum_{\langle i,j\rangle}\left(\hat{\mathbf{S}}_i\cdot\hat{\mathbf{S}}_j - \frac{1}{4}\right), \tag{8.17}$$

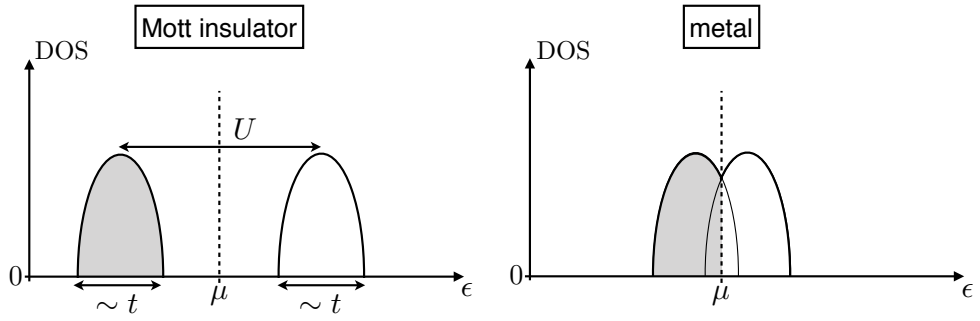
with antiferromagnetic superexchange $J = 4t^2/U$.

8.4 The Mott transition

It remains a heavily studied and highly non-trivial problem to understand what happens as U/t is reduced. For $U/t \gg 1$ the system is a Mott insulator. In the regime of weak interaction, $U/t \ll 1$, the system is a Fermi liquid with metallic behaviour. The transition between the metal and the insulator at a critical interaction strength $(U/t)_c$ is called the Mott transition.

Here we only present a cartoon picture for the case of half filling ($n = 1$). At large U/t, added electrons each cost an energy U The resulting doubly occupied sites (doublons) can move with an associated kinetic energy scale t. Holes are also mobile with kinetic energy t, but their creation does not involve the energy penalty U. At half filling we obtain a fully

occupied hole band and an empty electron band, both of width $\sim t$. The bands are split by U and symmetric around the chemical potential μ which is located in the middle of the energy gap. The DOS at the chemical potential is zero and the system is an insulator.



Reducing the Hubbard repulsion, the bands start to overlap at a critical value $U_c/t \sim 1$ and one expects a phase transition to metallic behaviour. In the Mott insulator there exists no Fermi surface. Approaching the Mott transition from the metallic side, we therefore expect the Migdal discontinuity Z to decrease and to vanish at the transition. The Mott transition is therefore linked to a breakdown of the quasiparticle picture, which makes it a very challenging theoretical problem.

8.5 Magnetic impurities in metals

The magnetic susceptibility of the free Fermi gas and of the Fermi liquid is constant at temperatures $T \ll T_F \simeq 10^4 - 10^5$K,

$$\begin{aligned} \chi^{(0)}_{\rm Pauli} &= \mu_B^2 g(\epsilon_F), \\ \chi_{\rm Pauli} &= \mu_B^2 \frac{g^*(\epsilon_F)}{1+F_0^{(a)}} = \frac{m^*/m}{1+F_0^{(a)}} \chi^{(0)}_{\rm Pauli}. \end{aligned}$$

Some metals, despite being expected to have a temperature-independent Pauli susceptibility, show Curie-like behaviour over a wide temperature range,

$$\chi(T) = \chi_{\rm Pauli} + \chi_{\rm Curie}(T) = \chi_{\rm Pauli} + \frac{A}{T+\theta}, \tag{8.18}$$

where A is a constant and θ the so-called *Curie-Weiss temperature*. A Curie susceptibility is characteristic for *local moments* and the observed behaviour is therefore an indication that there exist magnetic impurities in the metallic system. As a reminder, let us calculate the magnetic susceptibility of a single quantum spin. The Hamiltonian of the spin in a magnetic field $\mathbf{B}$ is

$$\hat{H} = -\mu_B B \hat{S}^z,$$

where without loss of generality we have chosen the field direction as our quantisation axis. For simplicity, let us assume $S = 1/2$. In this case there are only two eigenstates of $\hat{S}^z$, $m = \pm 1/2$, and the partition function is simply

$$Z = \mathrm{Tr} e^{-\beta \hat{H}} = e^{\frac{1}{2}\beta\mu_B B} + e^{-\frac{1}{2}\beta\mu_B B} = 2\cosh\left(\frac{\mu_B B}{2k_B T}\right).$$

Using that $Z = e^{-\beta F}$, we obtain the free energy

$$F = -k_B T \ln 2 - k_B T \ln\cosh\left(\frac{\mu_B B}{2k_B T}\right).$$

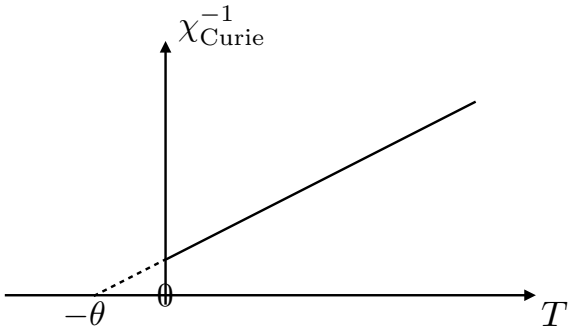
From this we can compute the magnetisation

$$\begin{aligned} M &= \mu_B \langle \hat{S}^z \rangle = \mu_B \frac{1}{Z} \mathrm{Tr}\ \hat{S}^z e^{-\beta\hat{H}} = k_B T \frac{\partial}{\partial B} \ln Z = -\frac{\partial F}{\partial B} \\ &= \frac{\mu_B}{2} \tanh\left(\frac{\mu_B B}{2k_B T}\right) \overset{B\to 0}{\simeq} \frac{\mu_B^2}{4k_B T} B. \end{aligned}$$

The resulting susceptibility is

$$\chi = \left.\frac{\partial M}{\partial B}\right|_{B=0} = \frac{\mu_B^2}{4k_B T}. \tag{8.19}$$

This is the Curie law. The susceptibility diverges as $T \to 0$ because at zero temperature, the spin will be completely aligned by an arbitrarily small magnetic field. In a real material there exists a finite concentration of magnetic impurities and interactions between the impurities. The Curie-Weiss temperature θ in Eq. (8.18) is introduced as a phenomenological scale which take account of the interactions between spins.



8.6 The Anderson impurity model

The question why impurities in metals become magnetic is non-trivial. A simple argument suggests that magnetic impurities should not be able to exist in metals. Assume that there exists an impurity level with energy ϵ_d. If $\epsilon_d > \epsilon_F$ the level will be empty, not giving rise to a magnetic moment. For $\epsilon_d < \epsilon_F$ one expects the level to be doubly occupied. The anti-symmetry of the wave function implies that the magnetic state is $\frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$, which is a singlet ($S = 0$) and so also non-magnetic. It can indeed be shown that a model that describes tunnelling between the conduction electrons and the impurity level,

$$\hat{H} = \underbrace{\sum_{\mathbf{k}\sigma} \epsilon(\mathbf{k}) c^\dagger_{\mathbf{k}\sigma} c_{\mathbf{k}\sigma}}_{\text{conduction sea}} + \underbrace{\epsilon_d \sum_\sigma d^\dagger_\sigma d_\sigma}_{\text{impurity level}} + \underbrace{\sum_{\mathbf{k}\sigma} \left(\lambda_{\mathbf{k}} c^\dagger_{\mathbf{k}\sigma} d_\sigma + \lambda^*_{\mathbf{k}} d^\dagger_\sigma c_{\mathbf{k}\sigma} \right)}_{\text{tunnelling}},$$

does not develop a Curie contribution to the susceptibility. The problem was resolved by P. W. Anderson in the 1950s. The impurity level is often a d- or f-orbital which have tightly compressed wave functions. As a result, the Coulomb energy cost U of doubly occupying the impurity level cannot be ignored. The Hamiltonian then becomes

$$\hat{H}_A = \sum_{\mathbf{k}\sigma} \epsilon(\mathbf{k}) c^\dagger_{\mathbf{k}\sigma} c_{\mathbf{k}\sigma} + \epsilon_d \sum_\sigma d^\dagger_\sigma d_\sigma + \lambda \sum_{\mathbf{k}\sigma} \left(c^\dagger_{\mathbf{k}\sigma} d_\sigma + d^\dagger_\sigma c_{\mathbf{k}\sigma} \right) + U d^\dagger_\uparrow d_\uparrow d^\dagger_\downarrow d_\downarrow, \tag{8.20}$$

where we have assumed that the tunnelling matrix element λ is independent of $\mathbf{k}$. This model is called the Anderson impurity model.

Similar to the on-site repulsion in the Hubbard model, a sufficiently strong repulsive interaction on the impurity can lead to the formation of a local moment. To understand this, let us first neglect the hybridisation λ. The state with a single electron on the impurity, which would be magnetic, is obtained by moving an electron from the Fermi surface (maximum energy gain) to the impurity level. Such a state has energy $E_{\rm FS} - \epsilon_F + \epsilon_d$, where $E_{\rm FS}$ is the energy of the filled Fermi sea. It is the ground state if its energy is smaller than $E_{\rm FS}$ and the energy $E_{\rm FS} - 2\epsilon_F + 2\epsilon_d + U$ of a state with a doubly occupied impurity level, from which it follows that

$$\epsilon_d < \epsilon_F < \epsilon_d + U. \tag{8.21}$$

This is the condition for local moment formation at zero temperature in the absence of hybridisation λ. For sufficient small λ we expect the local moment to be stable.

8.7 The Kondo model

We consider the regime $\epsilon_d < \epsilon_F < \epsilon_d + U$, corresponding to a ground-state manifold of

$$\hat{H}_0 = \sum_{\mathbf{k}\sigma} \epsilon(\mathbf{k}) c^\dagger_{\mathbf{k}\sigma} c_{\mathbf{k}\sigma} + \epsilon_d \sum_\sigma d^\dagger_\sigma d_\sigma + U d^\dagger_\uparrow d_\uparrow d^\dagger_\downarrow d_\downarrow \tag{8.22}$$

with one electron on the impurity site, giving rise to a local moment. We treat the hybridisation

$$\hat{H}_1 = \lambda \sum_{\mathbf{k}\sigma} \left(c^\dagger_{\mathbf{k}\sigma} d_\sigma + d^\dagger_\sigma c_{\mathbf{k}\sigma} \right) \tag{8.23}$$

as a perturbation to $\hat{H}_0$ and use degenerate second-order perturbation theory to derive an effective Hamiltonian that acts on the ground-state manifold. This effective Hamiltonian describes the coupling of a local impurity spin to the conduction electrons and is known as the Kondo Hamiltonian. The underlying canonical transformation is referred to as the Schrieffer-Wolf transformation.

Jun Kondo
1930 -

As we have derived in a previous lecture, the general expression for the matrix elements of the effective Hamiltonian is

$$\langle \mathrm{fi}|\hat{H}_{\mathrm{eff}}|\mathrm{in}\rangle = \frac{1}{2}\sum_n \langle \mathrm{fi}|\hat{H}_1|n\rangle\langle n|\hat{H}_1|\mathrm{in}\rangle \left(\frac{1}{E_{\mathrm{fi}}-E_n}+\frac{1}{E_{\mathrm{in}}-E_n}\right) \tag{8.24}$$

In the present case the initial and final states are eigenstates of $\hat{H}_0$ with one electron moved from the Fermi level to the impurity site. Such states have energy

$$E_0 = E_{\mathrm{fi}} = E_{\mathrm{in}} = E_{\mathrm{FS}} - \epsilon_F + \epsilon_d. \tag{8.25}$$

The intermediate states $\{|n\rangle\}$ are eigenstates of $\hat{H}_0$ with zero or two electrons moved to the impurity site, with energies

$$E_n^{(0)} = E_{\mathrm{FS}}, \tag{8.26a}$$
$$E_n^{(2)} = E_{\mathrm{FS}} - 2\epsilon_F + 2\epsilon_d + U. \tag{8.26b}$$

Introducing the projection operator $\hat{P}$ onto the ground-state manifold, we can write the effective Hamiltonian as

$$\hat{H}_{\mathrm{eff}} = \hat{P}\hat{H}_1(1-\hat{P})\frac{1}{E_0-\hat{H}_0}(1-\hat{P})\hat{H}_1\hat{P}. \tag{8.27}$$

Here

$$\hat{R} = \frac{1}{E_0-\hat{H}_0} = \sum_n \frac{|n\rangle\langle n|}{E_0-E_n}$$

is called the resolvent operator and $(1-\hat{P})$ ensures that division by zero does not occur. We obtain

$$
\begin{aligned}
\hat{H}_{\text{eff}} &= \lambda^2 \sum_{\mathbf{k}\mathbf{k}'\sigma\sigma'} \left(\frac{d^\dagger_\sigma c_{\mathbf{k}\sigma} c^\dagger_{\mathbf{k}'\sigma'} d_{\sigma'}}{E_0 - E_n^{(0)}} + \frac{c^\dagger_{\mathbf{k}\sigma} d_\sigma d^\dagger_{\sigma'} c_{\mathbf{k}'\sigma'}}{E_0 - E_n^{(2)}} \right) \\
&= \lambda^2 \sum_{\mathbf{k}\mathbf{k}'\sigma\sigma'} \left[\frac{d^\dagger_\sigma d_{\sigma'} \left(\delta_{\mathbf{k}\mathbf{k}'}\delta_{\sigma\sigma'} - c^\dagger_{\mathbf{k}'\sigma'} c_{\mathbf{k}\sigma} \right)}{\epsilon_d - \epsilon_F} + \frac{\left(\delta_{\sigma\sigma'} - d^\dagger_{\sigma'} d_\sigma \right) c^\dagger_{\mathbf{k}\sigma} c_{\mathbf{k}'\sigma'}}{\epsilon_F - \epsilon_d - U} \right] \\
&= \lambda^2 \left(\frac{1}{\epsilon_F - \epsilon_d} + \frac{1}{\epsilon_d + U - \epsilon_F} \right) \sum_{\mathbf{k}\mathbf{k}'\sigma\sigma'} d^\dagger_\sigma d_{\sigma'} c^\dagger_{\mathbf{k}'\sigma'} c_{\mathbf{k}\sigma} - \frac{\lambda^2}{\epsilon_d + U - \epsilon_F} \sum_{\mathbf{k}\mathbf{k}'\sigma} c^\dagger_{\mathbf{k}\sigma} c_{\mathbf{k}'\sigma} \\
&\quad \underbrace{- \frac{\lambda^2}{\epsilon_F - \epsilon_d} \sum_{\mathbf{k}} \overbrace{\sum_\sigma d^\dagger_\sigma d_\sigma}^{=1}}_{\text{const, drop}}
\end{aligned} \tag{8.28}
$$

We define the local moment spin operators

$$
\hat{S}^\alpha = \frac{1}{2} \left(d^\dagger_\uparrow, d^\dagger_\downarrow \right) \sigma^\alpha \begin{pmatrix} d_\uparrow \\ d_\downarrow \end{pmatrix} \tag{8.29}
$$

and the spin operator of the conduction electrons on the impurity site ($\mathbf{r} = 0$),

$$
\hat{s}^\alpha(0) = \frac{1}{2} \left(c^\dagger_\uparrow(0), c^\dagger_\downarrow(0) \right) \sigma^\alpha \begin{pmatrix} c_\uparrow(0) \\ c_\downarrow(0) \end{pmatrix} = \frac{1}{2} \sum_{\mathbf{k}\mathbf{k}'} \left(c^\dagger_{\mathbf{k}'\uparrow}, c^\dagger_{\mathbf{k}'\downarrow} \right) \sigma^\alpha \begin{pmatrix} c_{\mathbf{k}\uparrow} \\ c_{\mathbf{k}\downarrow} \end{pmatrix} \tag{8.30}
$$

We expect that the effective Hamiltonian does not prefer a particular orientation of the spins and is of the form $\hat{\mathbf{S}} \cdot \hat{\mathbf{s}}(0)$. We express the product of spin operators in terms of the fermionic creation and annihilation operators,

$$
\begin{aligned}
\hat{\mathbf{S}} \cdot \hat{\mathbf{s}}(0) &= \frac{1}{4} \left(d^\dagger_\uparrow d_\uparrow - d^\dagger_\downarrow d_\downarrow \right) \left(c^\dagger_\uparrow(0) c_\uparrow(0) - c^\dagger_\downarrow(0) c_\downarrow(0) \right) \\
&\quad + \frac{1}{4} \left(d^\dagger_\uparrow d_\downarrow + d^\dagger_\downarrow d_\uparrow \right) \left(c^\dagger_\uparrow(0) c_\downarrow(0) + c^\dagger_\downarrow(0) c_\uparrow(0) \right) \\
&\quad - \frac{1}{4} \left(d^\dagger_\uparrow d_\downarrow - d^\dagger_\downarrow d_\uparrow \right) \left(c^\dagger_\uparrow(0) c_\downarrow(0) - c^\dagger_\downarrow(0) c_\uparrow(0) \right)
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{4} \sum_\sigma d^\dagger_\sigma d_\sigma c^\dagger_\sigma(0) c_\sigma(0) - \frac{1}{4} \sum_\sigma d^\dagger_\sigma d_\sigma c^\dagger_{\overline{\sigma}}(0) c_{\overline{\sigma}}(0) + \frac{1}{2} \sum_\sigma d^\dagger_\sigma d_{\overline{\sigma}} c^\dagger_{\overline{\sigma}}(0) c_\sigma(0) \\
&= \frac{1}{2} \sum_{\sigma\sigma'} d^\dagger_\sigma d_{\sigma'} c^\dagger_{\sigma'}(0) c_\sigma(0) - \frac{1}{4} \sum_{\sigma\sigma'} d^\dagger_\sigma d_\sigma c^\dagger_{\sigma'}(0) c_{\sigma'}(0) \\
&= \frac{1}{2} \sum_{\mathbf{k}\mathbf{k}'\sigma\sigma'} d^\dagger_\sigma d_{\sigma'} c^\dagger_{\mathbf{k}'\sigma'} c_{\mathbf{k}\sigma} - \frac{1}{4} \sum_{\mathbf{k}\mathbf{k}'\sigma} c^\dagger_{\mathbf{k}'\sigma} c_{\mathbf{k}\sigma},
\end{aligned} \tag{8.31}
$$

where in the last step we have used that $\sum_\sigma d^\dagger_\sigma d_\sigma = 1$. Inserting this into the effective Hamiltonian (8.28), we obtain

$$\hat{H}_{\text{eff}} = J_K \hat{\mathbf{S}} \cdot \hat{\mathbf{s}}(0) + V \sum_{\mathbf{k}\mathbf{k}'\sigma} c^\dagger_{\mathbf{k}'\sigma} c_{\mathbf{k}\sigma}, \tag{8.32}$$

with

$$J_K = 2\lambda^2 \left(\frac{1}{\epsilon_F - \epsilon_d} + \frac{1}{\epsilon_d + U - \epsilon_F} \right) \quad \text{and} \quad V = \frac{1}{2}\lambda^2 \left(\frac{1}{\epsilon_F - \epsilon_d} - \frac{1}{\epsilon_d + U - \epsilon_F} \right). \tag{8.33}$$

The first term in the effective Hamiltonian (8.32) is the Kondo coupling between the local moment spin and the conduction electron spin at the impurity site $\mathbf{r} = 0$. Note that this coupling is *antiferromagnetic*, $J_K > 0$. The second term is a potential scattering term which we will neglect in the following. Including the unperturbed Hamiltonian $\hat{P}\hat{H}_0\hat{P}$, projected onto the ground-state manifold ($d^\dagger_\uparrow d_\uparrow d^\dagger_\downarrow d_\downarrow = 0$ and $\epsilon_d \sum_\sigma d^\dagger_\sigma d_\sigma = \epsilon_d$ an unimportant constant), we obtain the Kondo Hamiltonian

$$H_K = \sum_{\mathbf{k}\sigma} \epsilon(\mathbf{k}) c^\dagger_{\mathbf{k}\sigma} c_{\mathbf{k}\sigma} + J_K \hat{\mathbf{S}} \cdot \hat{\mathbf{s}}(0). \tag{8.34}$$

Kondo's original motivation for studying this model was to explain the resistivity minimum observed in many metals at low but finite temperature. Naively one expects that the resistivity $\rho(T)$ increases monotonously with temperature. In simple metals the scattering of electrons by phonons gives rise to a contribution $\rho_0(T) = AT^5$. In 1964, Kondo computed the scattering amplitude $A_{\mathbf{k},\mathbf{k}'}$ of conduction electrons by a local moment, treating the Kondo coupling in 2nd order perturbation theory. Here we only quote the final result,

$$A_{\mathbf{k},\mathbf{k}'} = J_K - \nu J_K^2 \ln \frac{k_B T}{\epsilon_F}, \tag{8.35}$$

where ν is the density of states per volume and spin at the Fermi surface. The scattering amplitude determines the inverse electron lifetime $\tau^{-1} \sim |A|^2$, which is proportional to the resistivity. Hence we obtain the resistivity contribution

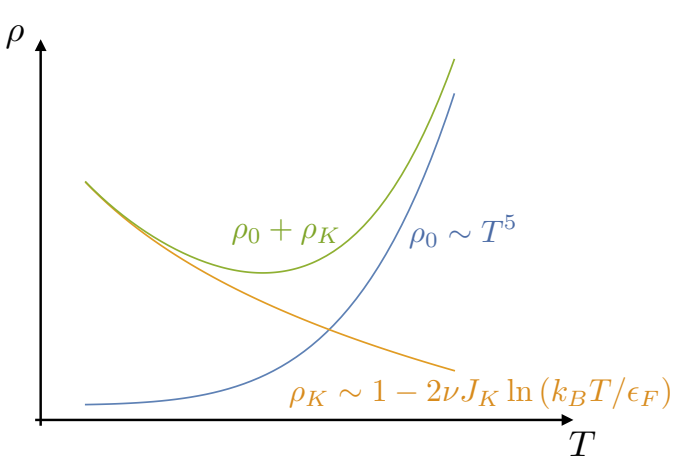
$$\rho_K(T) \approx B c J_K^2 \left(1 - 2\nu J_K \ln \frac{k_B T}{\epsilon_F} \right), \tag{8.36}$$

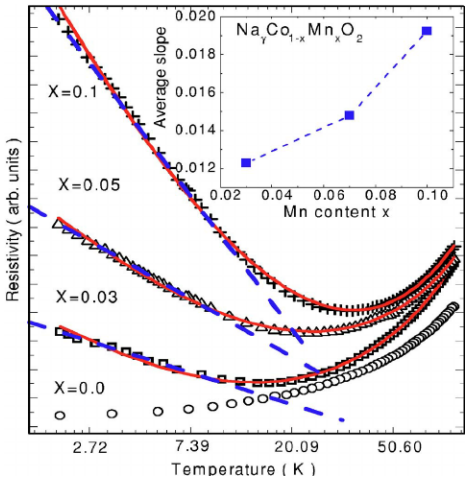
where c denotes the concentration of impurities and B a constant. $\rho_K(T)$ *increases* as temperature is decreased, giving rise to a minimum of $\rho(T) = \rho_0(T) + \rho_K(T)$,

$$\frac{d\rho}{dT} = 5AT^4 - 2Bc\nu J_K^3 T^{-1} = 0 \implies T_{\min} \sim c^{1/5}. \tag{8.37}$$

8.8 The Kondo problem

Disturbingly, the Kondo logarithm diverges as the temperature is lowered and hence the scattering amplitude (8.35) diverges logarithmically in the second order term as $T \to 0$. In fact divergencies are seen in all orders of the expansion and it can be shown that the scattering amplitude takes the form of a geometric sum





$$A_{\mathbf{k},\mathbf{k}'} = \frac{J_K}{1 + \nu J_K \ln \frac{k_B T}{\epsilon_F}}. \tag{8.38}$$

From this re-summed expression it is clear that the divergence occurs at a finite temperature T_K (the Kondo temperartue), determined by

$$1 + \nu J_K \ln \frac{k_B T}{\epsilon_F} = 0,$$

which gives

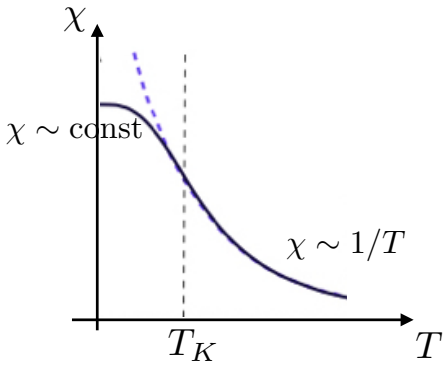
$$T_K = \frac{\epsilon_F}{k_B} e^{-1/(\nu J_K)}. \tag{8.39}$$

Interestingly, this exponential dependence is the same as in the BCS theory of superconductivity. Of course, in contrast to superconductivity, for a single impurity there cannot be a real phase transition. Instead, T_K gives a temperature scale at which there occurs a change of the behaviour of the system. The perturbation theory described above is only valid for $T \gg T_K$. What happens at temperatures below the Kondo temperature, $T < T_K$?

Experimentally, it is observed that below T_K the resistivity $\rho(T)$ saturates. There is also a crossover in the susceptibility χ from Curie-Weiss behaviour at $T \gg T_K$ to a constant value at $T \ll T_K$. This suggests that the local moments are screened below T_K and eventually disappear.

Due to the breakdown of perturbation theory with the appearance of a logarithmic singularity in the electron-local moment scattering amplitude, we must seek a different approach to analysing the single-impurity problem. A non-perturbative treatment was proposed by Varma and Yafet, starting from the Anderson impurity model (8.20) in the regime $U \to \infty$. In this limit, it is not possible to doubly occupy the impurity and the ground state at $T = 0$ is expected to be of the form

$$|\phi\rangle = \alpha_0 |\mathrm{FS}\rangle + \sum_{\mathbf{k}\sigma}^{|\mathbf{k}|<k_F} \alpha_{\mathbf{k}} d_\sigma^\dagger c_{\mathbf{k}\sigma} |\mathrm{FS}\rangle, \tag{8.40}$$



which is a superposition of the filled Fermi sea, empty impurity state with singlets formed between the impurity and the free electron state. We use $|\phi\rangle$ as a variational ansatz and compute the variational energy

$$E_V = E[\{\alpha_0, \alpha_{\mathbf{k}}\}] = \frac{\langle\phi|\hat{H}_A|\phi\rangle}{\langle\phi|\phi\rangle}, \tag{8.41}$$

with $\hat{H}_A$ the Anderson impurity Hamiltonian. Minimising E_V with respect to α_0 and $\alpha_{\mathbf{k}}$, Varma and Yafet demonstrated that the binding energy between the impurity and the conduction sea is equal to

$$\epsilon_B = E_V - E_{\text{FS}} - \epsilon_d \approx -\epsilon_F e^{-|\epsilon_d|/(2\lambda^2\nu)}. \tag{8.42}$$

Using that in the limit $U \to \infty$ the Kondo coupling (8.33) becomes

$$J_K = 2\lambda^2 \left(\frac{1}{\epsilon_F - \epsilon_d} + \frac{1}{\epsilon_d + U - \epsilon_F} \right) \overset{U\to\infty,\epsilon_F=0}{\longrightarrow} -2\frac{\lambda^2}{\epsilon_d} = 2\frac{\lambda^2}{|\epsilon_d|},$$

we can rewrite the expression for the binding energy as

$$|\epsilon_B| = \epsilon_F e^{-1/(\nu J_K)} \overset{(8.39)}{=} k_B T_K. \tag{8.43}$$

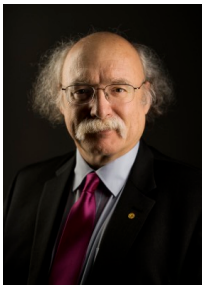
This suggests the following interpretation: the scattering between the conduction electrons and the impurity spin becomes significant at temperatures $T \sim T_K$, at which perturbation theory in J_K breaks down. This represents the crossover to a state where the impurity spin forms a bound state with the conduction sea. This non-magnetic bound state is called the Kondo singlet. The Varma-Yafet solution shows that at $T = 0$ the binding energy of the Kondo singlet is $k_B T_K$, and thus to break this bound state apart and recover the magnetic impurity it must require a temperature $T \sim T_K$.

Chapter 9

Topology in condensed matter (not examinable)

Topology plays an important role in condensed matter physics and is a very active area of research. The discoveries of the integer and fractional quantum Hall effects in the 1980's and of topological band insulators in the 2000's were landmarks in physics that enriched our view of electronic properties of solids. In a nutshell, these discoveries have taught us that quantum mechanical wavefunctions in crystalline solids may carry nontrivial topological invariants which have ramifications for the observable physics. The recent topological insulator revolution would never have happened without the groundbreaking theoretical work by Thouless, Haldane and Kosterlitz on topological phase transitions and topological phases of matter, for which they won the Nobel Prize in Physics in 2016.

David J. Thouless
1934 -



F. Duncan M. Haldane
1951 -

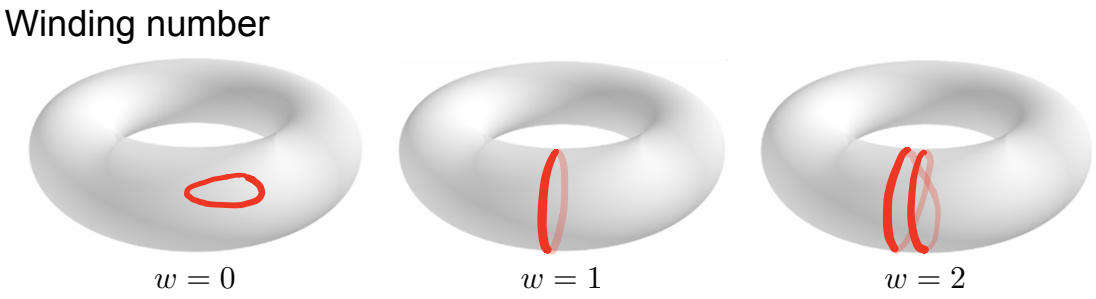
J. Michael Kosterlitz
1943 -

Nobel Prize in Physics
(theoretical discoveries of topological phase transitions and topological phases of matter)
2016

9.1 Topological invariants

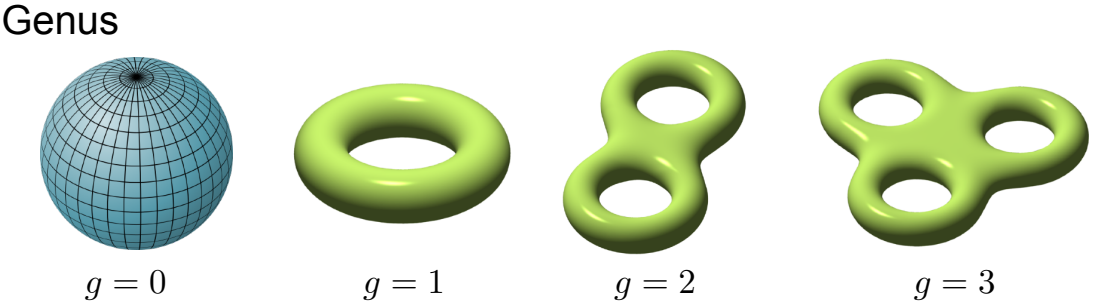
Topological order refers to properties of the system as a whole and is not related to a local order parameter such as the magnetisation. As a result, topological order is protected against small local perturbations of the system. Topological phases of matter are characterised by *topological invariants* rather than by spontaneous symmetry breaking. These invariants usually have a geometrical meaning, which we will illustrate in the following.

Consider for example a closed rubber string wrapped around a torus. The *winding number* w, which describes how often the rubber is wrapped around the torus, is a topological



invariant. Configurations with different w cannot be smoothly deformed into one another. Changing the winding number would require cutting and reconnecting the rubber. It is impossible to determine the winding number from an observation of a small segment of the rubber. Instead, one would have to look at the entire configuration or calculate an appropriate closed contour integral.

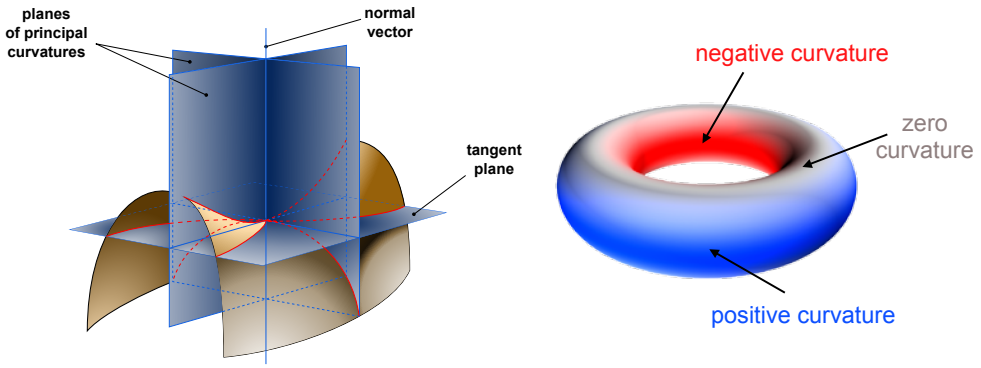
Another example of a topological invariant is the *genus* g of a manifold. The genus is equal to the number of holes or handles of the manifold. A sphere has genus $g = 0$, a torus or bagel $g = 1$, a pretzel $g = 2$ and so on. One cannot change the genus by smooth deformations of the object. It is possible however to smoothly deform a torus into a coffee mug without poking an addition hole into the object. This means that a torus and a coffee mug have the same genus and are topologically equivalent. Walking on the Earth, we are not able to tell if our planet is a sphere or a giant torus. Mountains and tales are just small ripples on the surface that do not tell us anything about topology. We would require an image from outer space or an appropriate surface integral over the entire Earth to work out its genus.



For a closed, orientable two-dimensional manifold M embedded in three dimensional space the genus is related to the *Euler characteristic* χ, which is an integral of the *Gaussian curvature* K over the manifold,

$$g = \frac{1}{2}(2 - \chi) = \frac{1}{2}\left(2 - \frac{1}{2\pi}\int_M K\, dA\right),$$

where dA denotes the surface area element of the manifold. This is a special case of the *Gauss-Bonet theorem*. The Gaussian curvature is equal to the product of the two principal curvatures, $K = \kappa_1\kappa_2$. Local minima or maxima are examples of points with $K > 0$, while saddle points (see picture) have $K < 0$.



Let us explicitly calculate g for a sphere and a torus. We parametrise the sphere of radius R by spherical coordinates,

$$\mathbf{r}(\theta,\phi) = R\begin{pmatrix}\sin\theta\cos\phi\\ \sin\theta\sin\phi\\ \cos\theta\end{pmatrix},$$

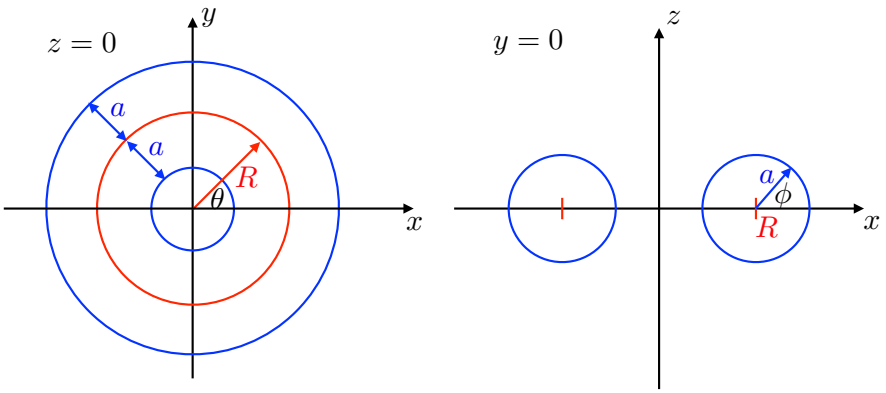
with $\theta \in [0,\pi]$ and $\phi \in [0,2\pi]$. The surface area element is $dA = R^2\sin\theta\, d\theta\, d\phi$ and the Gaussian curvature is constant, $K = 1/R^2$. We therefore obtain

$$g = \frac{1}{2}\left(2 - \frac{1}{2\pi}\int_0^{2\pi} d\phi \int_0^{\pi} \sin\theta d\theta\right) = \frac{1}{2}(2-2) = 0.$$

A possible parametrisation of a torus with radii R and a, $a < R$, is given by

$$\mathbf{r}(\theta,\phi) = \begin{pmatrix}(R + a\cos\phi)\cos\theta\\ (R + a\cos\phi)\sin\theta\\ a\sin\phi\end{pmatrix},$$

where $\theta \in [0,2\pi]$ and $\phi \in [0,2\pi]$.



For the surface area element we obtain

$$dA = \left\|\frac{\partial\mathbf{r}}{\partial\theta}\times\frac{\partial\mathbf{r}}{\partial\phi}\right\| d\theta\, d\phi = a(R + a\cos\phi)\, d\theta\, d\phi.$$

The Gaussian curvature K is a function of the angle ϕ. Here we just quote the result,

$$K = \frac{\cos\phi}{a(R + a\cos\phi)}.$$

This gives the genus

$$g = \frac{1}{2}\left(2 - \frac{1}{2\pi}\int_0^{2\pi} d\theta \int_0^{2\pi} d\phi\, \cos\phi\right) = \frac{1}{2}(2-0) = 1.$$

9.2 Integer Quantum Hall effect and Chern numbers

The simplest topological state of matter is realised in the integer quantum Hall effect. It is a quantum mechanical version of the Hall effect observed in a two-dimensional electron system at very low temperatures and in a strong magnetic field perpendicular to the two-dimensional system. If we drive a current through the system, the electrons are subject to the Lorentz force $F_L = q\mathbf{v} \times \mathbf{B}$ perpendicular to the current. This leads to the formation of cyclotron orbits and a build-up of positive and negative charges at opposite edges, giving rise to a Hall voltage V_H. In equilibrium, the forces from the resulting electric field and the applied magnetic field cancel each other. In the classical Hall effect, this leads to a Hall resistivity that is proportional to the magnetic field B. At low temperatures however, it is important to take into account quantum mechanical effects. As you have learned in your quantum mechanics lectures, a charged particle in a magnetic has quantised energy levels (Landau levels),

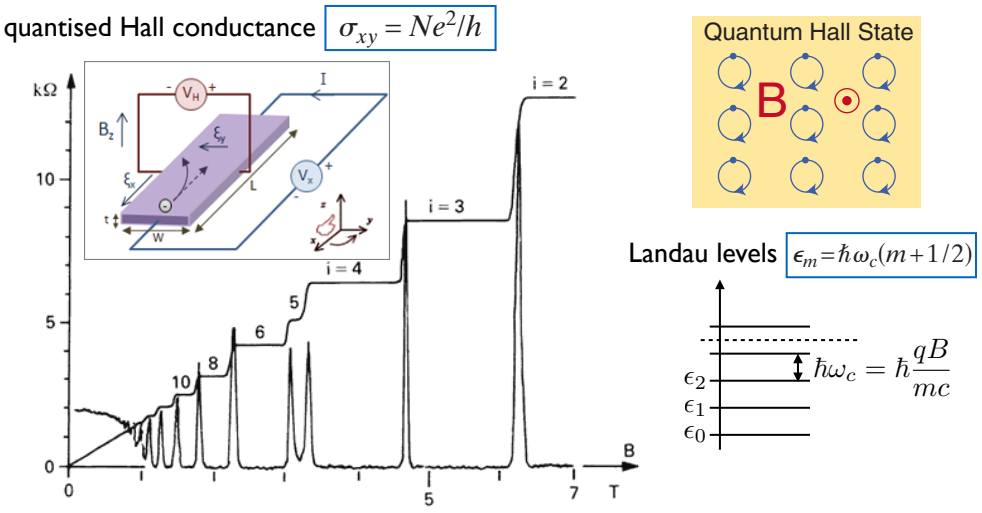
$$\epsilon_n = \hbar\omega_c\,(n + 1/2)\,, \tag{9.1}$$

where $\omega_c = \frac{qB}{mc}$ denotes the cyclotron frequency. With increasing field B, the spacing between energy levels increases. As a result, the number of filled Landau levels below the Fermi energy decreases in steps as B is increased. Based on approximate calculations, Ando, Matsumoto, and Uemura predicted in 1975 that, as a result, the Hall conductance is quantised. Subsequently, several experimental groups started to look for the effect and observed plateaux in the Hall conductance at low temperatures. In 1980, Klaus von Klitzing measured the Hall conductance of silicon-based samples (developed by Michael Pepper (UCL) and Gerhard Dorda) at the high magnetic field lab in Grenoble. His experiments demonstrated for the first time that the transverse Hall conductance is exactly quantised and equal to integer multiples of e^2/h,

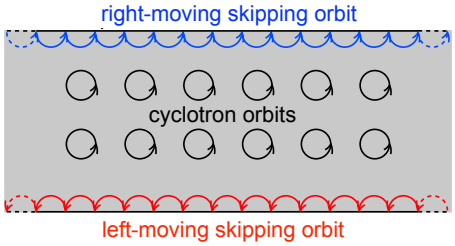
$$\sigma_{xy} = N\frac{e^2}{h}, \tag{9.2}$$

where e is the electron charge and h the Planck constant. For this finding, von Klitzing was awarded the 1985 Nobel Prize in Physics.

The quantisation of the Hall conductance has the important property of being exceedingly precise and robust against external perturbations. It is now possible to measure the quantum e^2/h with an accuracy of nearly one part in a billion. This astonishing accuracy has allowed for the definition of a new practical standard for electrical resistance, and since 1990, it is used in resistance calibrations worldwide



What is the reason for the exact quantisation and its robustness against perturbations such as disorder? In the bulk, the electrons are localised in orbits, and therefore form an insulating state. At the edges of the sample however, one expects *skipping orbits*, implying that electrons can move around the edges. This motion is unidirectional, where the chirality is set by the orientation of the magnetic field. These *edge modes* are protected against disorder since there is no channel for back-scattering.



It was realised by Thouless, Kohmoto, Nightingale, and Nijs in 1982, that the protection of the exact quantisation of the Hall conductance is routed in topology. They demonstrated that the integer quantum Hall states are characterised by a *topological invariant* which is defined in a very similar way to the Euler characteristic χ of a manifold. The TKNN topological invariant or *Chern number* ν is defined as an integral of the *Berry curvature*

$$\mathcal{F}(\mathbf{k}) = -i\epsilon_{ij}\langle\frac{\partial}{\partial k_i}\psi(\mathbf{k})|\frac{\partial}{\partial k_j}\psi(\mathbf{k})\rangle \tag{9.3}$$

of the electronic wavefunction $\psi(\mathbf{k})$ over the two-dimensional Brillouin zone,

$$\nu = \frac{1}{2\pi}\int_{BZ} d^2\mathbf{k}\,\mathcal{F}(\mathbf{k}). \tag{9.4}$$

Because of the periodic boundary condition in k_x and k_y the Brillouin zone is equivalent to a torus. The Chern number of the integer quantum Hall state is equal to the number of filled

Landau levels. Thouless, Kohmoto, Nightingale, and Nijs computed the Hall conductance using the Kubo formalism and found that the same integral (9.4) enters in the expression for σ_{xy}, proving that $\sigma_{xy} = \nu \frac{e^2}{h}$ and hence establishing the link between the exact quantisation and topology.

9.3 The Haldane model

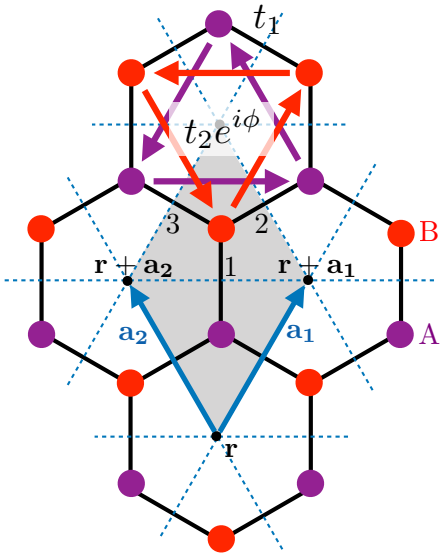
Haldane made an important contribution that established the link between topological properties of the insulating bulk, characterised by topological invariants, and the presence of gapless modes ('skipping orbits') that propagate around the edges of the system. This is often referred to as the *bulk-boundary correspondence*.

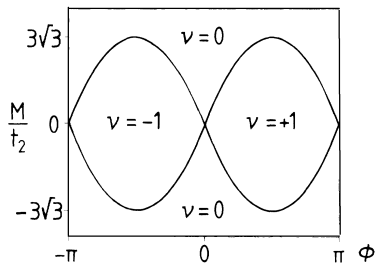
Haldane was interested in the question whether it was possible to obtain a quantum Hall state in the absence of an applied magnetic field. In 1988, he succeeded in constructing a simple model of spin-less, non-interacting fermions that indeed exhibits energy bands with non-zero Chern numbers, separated by an energy gap and located below and above the chemical potential. Such a state is a *topological Chern insulator*, exactly like the integer quantum Hall state.

In his model, now referred to as the *Haldane model*, he considered spin-less fermions on a honeycomb lattice that hop with amplitudes t_1 between neighbouring sites $\langle i, j\rangle$ and with *complex* amplitudes $t_2 \exp(i\phi_{ij})$ between next-nearest neighbour sites $\langle\langle i, j\rangle\rangle$,

$$\hat{H} = t_1 \sum_{\langle i,j\rangle} \left(c_i^\dagger c_j + \text{h.c.} \right) + \sum_{\langle\langle i,j\rangle\rangle} \left(t_2 e^{i\phi_{ij}} c_i^\dagger c_j + \text{h.c.} \right) + M \sum_i \epsilon_i c_i^\dagger c_i,$$

where $t_2 \geq 0$ and $\phi_{ij} = \phi$ if the 2nd-neighbour hopping is in an anti-clockwise direction and $\phi_{ij} = -\phi$ if it is in a clockwise direction. The last term in the Hamiltonian (9.5) is an on-site potential with $\epsilon_i = 1$ if the site i is on sub-lattice A and $\epsilon_i = -1$ if the site is on sub-lattice B.





F.D.M. Haldane (1988)

In momentum space the nearest-neighbour hopping takes the form

$$\begin{aligned} t_1 \sum_{\langle i,j \rangle} \left(c_i^\dagger c_j + \text{h.c.} \right) &= t_1 \sum_{\mathbf{r}} \left(c_A^\dagger(\mathbf{r}) c_B(\mathbf{r}) + c_A^\dagger(\mathbf{r}+\mathbf{a}_1) c_B(\mathbf{r}) + c_A^\dagger(\mathbf{r}+\mathbf{a}_2) c_B(\mathbf{r}) \right) + \text{h.c.} \\ &= t_1 \sum_{\mathbf{k}} \underbrace{\left(1 + e^{i\mathbf{k}\mathbf{a}_1} + e^{i\mathbf{k}\mathbf{a}_2}\right)}_{=:\gamma(\mathbf{k})} c_{\mathbf{k}A}^\dagger c_{\mathbf{k}B} + \text{h.c.} \\ &= t_1 \sum_{\mathbf{k}} \left(c_{\mathbf{k}A}^\dagger, c_{\mathbf{k}B}^\dagger \right) \begin{pmatrix} 0 & \gamma(\mathbf{k}) \\ \gamma^*(\mathbf{k}) & 0 \end{pmatrix} \begin{pmatrix} c_{\mathbf{k}A} \\ c_{\mathbf{k}B} \end{pmatrix}, \end{aligned}$$

The lattice vectors are given by $\mathbf{a}_1 = \left(\frac{\sqrt{3}}{2}, \frac{3}{2}\right)$ and $\mathbf{a}_2 = \left(-\frac{\sqrt{3}}{2}, \frac{3}{2}\right)$ where we set the lattice constant to $a = 1$. For simplicity, let us assume that the complex-hopping term is purely imaginary, $\phi = \pi/2$, $\exp(\pm i\phi) = \pm i$. In momentum space, we obtain

$$\begin{aligned} \sum_{\langle\langle i,j \rangle\rangle} \left(t_2 e^{i\phi_{ij}} c_i^\dagger c_j + \text{h.c.} \right) &= i t_2 \sum_{\mathbf{r}} \Big(c_A^\dagger(\mathbf{r}+\mathbf{a}_1-\mathbf{a}_2) c_A(\mathbf{r}) - c_A^\dagger(\mathbf{r}+\mathbf{a}_1) c_A(\mathbf{r}) \\ &\qquad + c_A^\dagger(\mathbf{r}+\mathbf{a}_2) c_A(\mathbf{r}) - c_B^\dagger(\mathbf{r}+\mathbf{a}_1-\mathbf{a}_2) c_B(\mathbf{r}) \\ &\qquad + c_B^\dagger(\mathbf{r}+\mathbf{a}_1) c_B(\mathbf{r}) - c_B^\dagger(\mathbf{r}+\mathbf{a}_2) c_B(\mathbf{r}) \Big) + \text{h.c.} \\ &= 2t_2 \sum_{\mathbf{k}} \left[\sin(\mathbf{k}\mathbf{a}_1) - \sin(\mathbf{k}\mathbf{a}_2) - \sin(\mathbf{k}(\mathbf{a}_1 - \mathbf{a}_2)) \right] \\ &\qquad \times \left(c_{\mathbf{k}A}^\dagger c_{\mathbf{k}A} - c_{\mathbf{k}B}^\dagger c_{\mathbf{k}B} \right) \\ &= t_2 \sum_{\mathbf{k}} \left(c_{\mathbf{k}A}^\dagger, c_{\mathbf{k}B}^\dagger \right) \begin{pmatrix} s(\mathbf{k}) & 0 \\ 0 & -s(\mathbf{k}) \end{pmatrix} \begin{pmatrix} c_{\mathbf{k}A} \\ c_{\mathbf{k}B} \end{pmatrix}, \end{aligned}$$

where in the last step we have defined $s(\mathbf{k}) = 2\left[\sin(\mathbf{k}\mathbf{a}_1) - \sin(\mathbf{k}\mathbf{a}_2) - \sin(\mathbf{k}(\mathbf{a}_1 - \mathbf{a}_2))\right]$. Finally, for the on-site potential we obtain

$$\begin{aligned} M \sum_i \epsilon_i c_i^\dagger c_i &= M \sum_{\mathbf{r}} \left(c_A^\dagger(\mathbf{r}) c_A(\mathbf{r}) - c_B^\dagger(\mathbf{r}) c_B(\mathbf{r}) \right) \\ &= \sum_{\mathbf{k}} \left(c_{\mathbf{k}A}^\dagger, c_{\mathbf{k}B}^\dagger \right) \begin{pmatrix} M & 0 \\ 0 & -M \end{pmatrix} \begin{pmatrix} c_{\mathbf{k}A} \\ c_{\mathbf{k}B} \end{pmatrix}. \end{aligned} \tag{9.5}$$

Combining the terms, the Hamiltonian can be written as

$$\begin{aligned} \hat{H} &= \sum_{\mathbf{k}} \left(c_{\mathbf{k}A}^\dagger, c_{\mathbf{k}B}^\dagger \right) \begin{pmatrix} d_z(\mathbf{k}) & d_x(\mathbf{k}) - i d_y(\mathbf{k}) \\ d_x(\mathbf{k}) + i d_y(\mathbf{k}) & -d_z(\mathbf{k}) \end{pmatrix} \begin{pmatrix} c_{\mathbf{k}A} \\ c_{\mathbf{k}B} \end{pmatrix} \\ &= \sum_{\mathbf{k}} \left(c_{\mathbf{k}A}^\dagger, c_{\mathbf{k}B}^\dagger \right) \left(\mathbf{d}(\mathbf{k}) \cdot \boldsymbol{\sigma} \right) \begin{pmatrix} c_{\mathbf{k}A} \\ c_{\mathbf{k}B} \end{pmatrix}, \end{aligned}$$

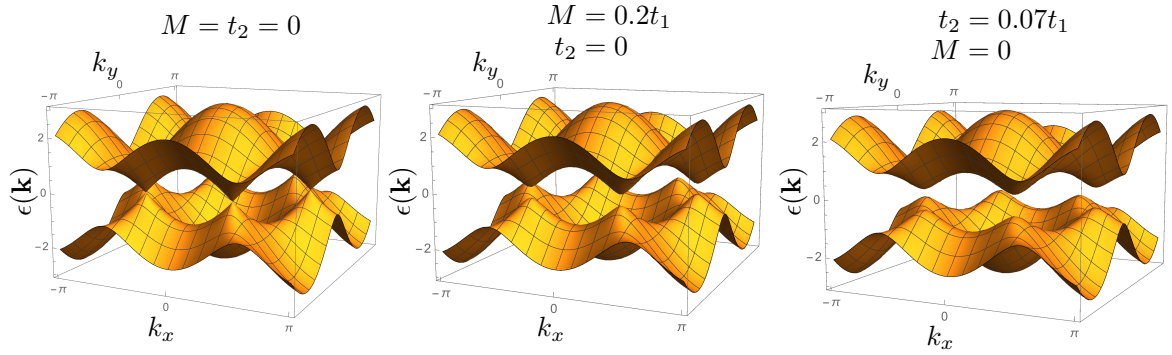
where $d_x(\mathbf{k}) = t_1 \text{Re}[\gamma(\mathbf{k})]$, $d_y(\mathbf{k}) = -t_1 \text{Im}[\gamma(\mathbf{k})]$ and $d_z(\mathbf{k}) = M + t_2 s(\mathbf{k})$. $\mathbf{d}(\mathbf{k})$ is a three-dimensional vector field and hence describes a mapping from the two-dimensional

Brillouin torus to a closed surface in three dimensional space. Diagonalising this two-by-two Hamiltonian, one obtains the dispersion

$$\epsilon(\mathbf{k}) = \pm\|\mathbf{d}(\mathbf{k})\| = \pm\sqrt{t_1^2|\gamma(\mathbf{k})|^2 + (M + t_2 s(\mathbf{k}))^2}. \tag{9.6}$$

For $M = t_2 = 0$ the dispersion vanishes *linearly* at the points $\mathbf{K}_\pm = \frac{2\pi}{3}\left(\pm\frac{1}{\sqrt{3}}, 1\right)$, forming so-called *Dirac cones*. The alternating potential M and the imaginary hopping t_2 both open a gap in the spectrum. In the presence of both terms, the gaps at $\mathbf{K}_\pm$ are given by

$$\Delta_\pm = |M \mp 3\sqrt{3}t_2|. \tag{9.7}$$



While the electron dispersions for the cases $M > 0$, $t_2 = 0$ and $M = 0$, $t_2 > 0$ look identical, the electronic wave functions and their topological properties are very different. In both gapped phases, $\epsilon(\mathbf{k})$ never reaches zero and the closed surface defined by all the vectors $\mathbf{d}(\mathbf{k})$ therefore does not include the origin. In this case one can show that the *Berry curvature* is equal to

$$\mathcal{F}(\mathbf{k}) = \frac{1}{2}\hat{\mathbf{d}}(\mathbf{k}) \cdot \left(\frac{\partial \hat{\mathbf{d}}}{\partial k_x} \times \frac{\partial \hat{\mathbf{d}}}{\partial k_y}\right), \tag{9.8}$$

where $\hat{\mathbf{d}}(\mathbf{k}) = \mathbf{d}(\mathbf{k})/\|\mathbf{d}(\mathbf{k})\|$. From the above equation, it is clear that the Berry curvature $\mathcal{F}(\mathbf{k})$ is equal to half the solid angle element of the mapping $\hat{\mathbf{d}}(\mathbf{k})$ and the integral of $\mathcal{F}(\mathbf{k})$ is therefore and integer multiple of 2π. The *Chern number*

$$\nu = \frac{1}{2\pi}\int_{BZ} d^2\mathbf{k}\, \mathcal{F}(\mathbf{k})$$

describes how many times $\hat{\mathbf{d}}(\mathbf{k})$ wraps around the sphere as a function of $\mathbf{k}$. Evaluating the integral for the Chern number we find

$$\nu = \begin{cases} 0 & \text{if} \quad 0 \leq 3\sqrt{3}t_2 < |M| \\ 1 & \text{if} \quad 0 \leq |M| < 3\sqrt{3}t_2 \end{cases} \tag{9.9}$$

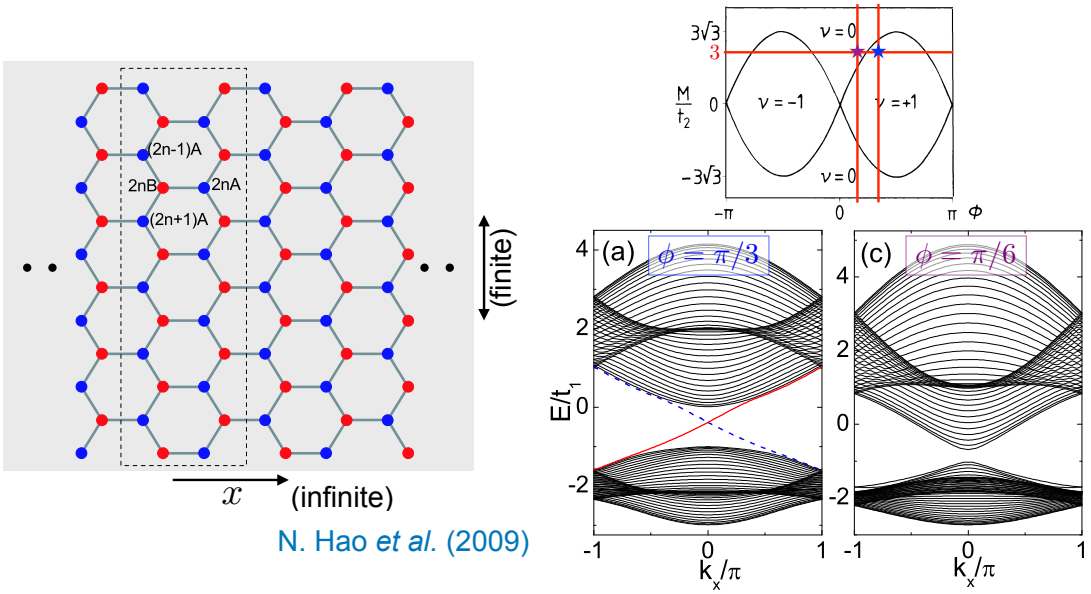
in agreement with Haldane's result. At the points $M = \pm 3\sqrt{3}t_2$, the gap closes at one of the Dirac points $\mathbf{K}_\pm$ and a transition between a conventional insulator ($\nu = 0$) and a topological Chern insulator ($\nu = 1$) occurs.

9.4 Bulk-boundary correspondence

Within the Haldane model, it is also possible to address the question whether there exist conducting edge modes if the gapped, insulating state has a non-trivial topology with non-zero Chern number. In order to establish this correspondence, one diagonalises the Hamiltonian of the Haldane model on a strip that is infinite along the x direction but has a finite width L_y along y. This system should be viewed as one-dimensional system with many lattice sites in the unit cells, giving rise to a large number of bands as a function of the continuous momentum k_x. Almost all of the bands have a trivial relation to the dispersion of the infinite, two-dimensional system: making the size along y finite, simply leads to bands corresponding to cuts through the two-dimensional dispersion with discrete momenta $k_y(j)$,

$$\epsilon_j(k_x) = \epsilon(k_x, k_y = j\frac{2\pi}{L_z}).$$

Since the two-dimensional system is an insulator, non of these bands will cross the Fermi energy. In addition, there could be non-trivial modes that are localised near the edges. If these modes are conducting they would have to cross the Fermi level and should therefore be clearly visible in the gap between the quasi-continua of bulk bands.



In the figure, the energy spectra of the strip system are shown for two different sets of parameters, one corresponding to a trivial insulating state ($\nu = 0$), and one for a topological Chern insulator with $\nu = 1$. As expected, only in the latter case there exist conducting modes that cross the Fermi level. An analysis of the eigenvectors shows that these modes propagate along the edges, the right-moving mode (red) along the upper edge of the strip and the left-moving mode (blue) along the lower edge.