

Notes for PX453, Advanced Quantum Theory (part II)

Nicholas d'Ambrumenil

Updated: April 5, 2024

Contents

Preface	i
1 Introducing QFT	2
1.1 Phonons and a 1D QFT	5
1.2 Time-Dependence	7
1.3 Summary and questions	8
2 QFT—Fermions	10
2.1 Non-relativistic QFT for electrons	13
2.2 Relativistic QFT for fermions	16
2.3 Normal ordering	18
2.4 Interactions	19
3 Interacting Electrons—Superconductivity	20
3.1 Many-body systems	20
3.2 Cooper instability	21
3.3 Mean Field Theory	24
3.4 BCS theory	25
3.5 Properties of the BCS state	28
3.6 Questions	30
4 Interacting Spins	32
4.1 Holstein-Primakoff bosons	34
4.2 Dimensionality	39
5 Density Matrices	41

<i>CONTENTS</i>	2
5.1 Properties of the Density Matrix	43
5.2 Sub-systems and Entanglement	45
5.3 Conclusion	48
A Continuum Limit of Lattice Oscillations	49
B Relativistic Quantum Fields—Bosons	53
B.1 Klein-Gordon field	53
B.2 Electromagnetic field	54
B.3 Lorentz and gauge invariance	56
C Attractive Interaction from Phonon Exchange	58
D Spin Interactions	60

Preface

This section of the module introduces some quantum theoretic techniques used across different areas of physics. The description of these techniques and their successes can become detailed very quickly and much of the detail belongs in graduate texts. However, with a little effort it is possible to understand how they might make sense.

As this module is also designed to support other more research-oriented modules in the fourth year, we will concentrate on ideas explored further or used in other modules. They are the idea of field operators and the density matrix. Quantum field theories (QFT) are important in condensed matter physics, nuclear physics, particle physics and statistical physics. They were first introduced in the 1920's by Dirac and others but it took a long time before they were developed into reliable tool for describing and predicting phenomena. We will introduce field operators and explain how they describe free bosons and free fermions. The only system involving interactions, that we will have time to look at, is that of a superconductor.

The density matrix is another old idea, also introduced in the 1920's, that has become increasingly important recently. von Neumann and Landau both introduced the idea of the density matrix in 1927. It is important for treating open systems, where the interaction of a subsystem with a larger system, which might be the environment, needs a quantum treatment. It is also important in quantum information theory.

Chapter 1

Introducing QFT

The message you may have picked up from the quantum mechanics (QM) you have studied until now is that something is not quite right. And some people may have suggested that quantum field theory (QFT) solves the problems. QFT has its own problems but does indeed solve many issues from QM. It allows us to explain and predict new things—the hugely successful Standard Model of particle physics is a QFT after all.

As you have just seen, Dirac's equation describes beautifully the behaviour of an electron but at the price of introducing an infinite sea of negative energy states. One way of rationalising these states is to assume that they are all filled. If we ignore any interaction between the electron in a positive energy state with this infinite sea of electrons occupying the negative energy states, we can often get away with thinking we are doing quantum mechanics—writing a wave-function for a single particle, or small number of particles, and ignoring the infinitely many states with negative energy. However, these extra particles are there. Interactions can lead to electron-positron creation and hence to a system with a changing number of particles. A state-function for such a system needs to be able to encompass such possible changes in the numbers of particles. This is one thing that QFT was invented to handle.

We also need to discuss how to quantise theories that are already field theories classically. An example that should be familiar is electromagnetism. The fields $\mathbf{E}(t, \mathbf{r})$ and $\mathbf{B}(t, \mathbf{r})$, or equivalently the corresponding four potential A with its contravariant components $A^\alpha = (\phi/c, \mathbf{A})$, are defined at all points in space-time. There are infinitely many points and hence infinitely many degrees of freedom to keep track of. Maxwell's equations determine the classical values of these fields. The quantised theory needs to describe how light particles, the photons with energy $\hbar\omega$, appear in the theory.

What is a Quantum Field Theory?

The boundary between quantum mechanics and quantum field theory is not always clear. Roughly speaking, mechanics describes systems with a finite number of degrees of freedom (DoF) while field theories describe systems with infinitely many DoF. This is true both classically and after quantisation. We will look first at a single oscillator then a 1D chain of ions, with length La where a is the lattice parameter. The oscillator has 1 DoF and the chain of ions has L DoF so both are systems described by mechanics. The continuum version of this lattice model described in Appendix A, and electromagnetism described in Appendix B, have infinitely many DoF and are therefore described by field theories. The

point about the Dirac equation is that it starts out as a wave equation (quantum mechanics) for a single electron but ends up requiring infinitely many electrons occupying the negative energy states, and so needs reformulating as a quantum field theory.

While the distinction between mechanics and field theory based on the number of DoF is conventional, many people refer to field operators for particles or excitations whether or not there are infinitely many of them and this is what we will do. The central properties of field operators are their commutation relations which are similar in the case of quantum mechanical systems and quantum field theories. These operators can change the number of particles in a system. They are the tool used to understand how to handle the question of the infinity of negative energy states in the Dirac equation mentioned above (and see Chapter 2). Working with field operators, rather than wave-functions, will be the theme of the first four chapters of these notes.

Field Operators in the Simple Harmonic Oscillator

Funnily enough you should have met precursors of field operators. These are the raising and lowering operators used to characterise the quantum states of a simple harmonic oscillator (SHO). The classical Hamiltonian for a particle of mass m , with coordinate x , and momentum p , moving in a quadratic potential ($V = kx^2/2$) in 1D is

$$H = \frac{p^2}{2m} + \frac{k}{2}x^2 \quad \omega_0^2 = k/m \quad \frac{p^2}{2m} + \frac{m\omega_0^2}{2}x^2. \quad (1.1)$$

We can write this as (so far the Hamiltonian is classical but it helps to put in \hbar as we know that the energy scale after quantising will be $\hbar\omega_0$)

$$H = \frac{\hbar\omega_0}{2} \left(\frac{p^2}{\hbar m \omega_0} + \frac{m\omega_0}{\hbar} x^2 \right) = \frac{\hbar\omega_0}{2} (p'^2 + x'^2). \quad (1.2)$$

We have introduced $p' = p/\sqrt{\hbar m \omega_0}$ and $x' = x\sqrt{m\omega_0/\hbar}$.

When we quantise the Hamiltonian, momentum and position become non-commuting operators \hat{p} and \hat{x} :

$$[\hat{x}, \hat{p}] = i\hbar \quad \text{or equivalently} \quad [\hat{x}', \hat{p}'] = i. \quad (1.3)$$

We have introduced the notation for a commutator $[\hat{A}, \hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A}$. We can factorise the Hamiltonian in 1.2 in terms of two Hermitian-conjugate operators, \hat{a} and \hat{a}^\dagger , as follows

$$\begin{aligned} \hat{H} &= \frac{\hbar\omega_0}{2} (\hat{p}'^2 + \hat{x}'^2) = \frac{\hbar\omega_0}{2} \left(\frac{(\hat{x}' + i\hat{p}')(\hat{x}' - i\hat{p}') + (\hat{x}' - i\hat{p}')(\hat{x}' + i\hat{p}')}{2} \right) \\ &= \frac{\hbar\omega_0}{2} (\hat{a}^\dagger \hat{a} + \hat{a} \hat{a}^\dagger) = \hbar\omega_0 \left(\hat{a}^\dagger \hat{a} + \frac{1}{2} [\hat{a}, \hat{a}^\dagger] \right) \\ &= \hbar\omega_0 \left(\hat{a}^\dagger \hat{a} + \frac{1}{2} \right), \quad \text{where } \hat{a} = \frac{(\hat{x}' + i\hat{p}')}{\sqrt{2}}, \quad \hat{a}^\dagger = \frac{(\hat{x}' - i\hat{p}')}{\sqrt{2}}. \end{aligned} \quad (1.4)$$

The position and momentum operators are

$$\hat{x}' = \frac{\hat{a} + \hat{a}^\dagger}{\sqrt{2}} \quad \text{and} \quad \hat{p}' = \frac{\hat{a} - \hat{a}^\dagger}{i\sqrt{2}}. \quad (1.5)$$

Both \hat{x}' and \hat{p}' are Hermitian as they should be.

The commutator

$$[\hat{a}, \hat{a}^\dagger] = 1. \quad (1.6)$$

Consider the harmonic oscillator in its ground state $|0\rangle$ with corresponding wavefunction $\psi_0 = (1/\sqrt{\pi}) \exp(-x'^2/2)$. This function is annihilated by \hat{a} :

$$\hat{a} \exp(-x'^2/2) = \frac{1}{\sqrt{2}} \left(x' + \frac{\partial}{\partial x'} \right) \exp(-x'^2/2) = \frac{1}{\sqrt{2}} (x' - x') \exp(-x'^2/2) = 0. \quad (1.7)$$

The operator $(\hat{a}^\dagger)^n$ generates n quanta of $\hbar\omega_0$:

$$\begin{aligned} \hat{H} (\hat{a}^\dagger)^n |0\rangle &= \hbar\omega_0 \left(\hat{a}^\dagger \hat{a} + \frac{1}{2} \right) (\hat{a}^\dagger)^n |0\rangle = \hbar\omega_0 \left(\hat{a}^\dagger (\hat{a}^\dagger)^n \hat{a} + \hat{a}^\dagger [\hat{a}, (\hat{a}^\dagger)^n] + \frac{(\hat{a}^\dagger)^n}{2} \right) |0\rangle \\ &= \hbar\omega_0 \left(\hat{a}^\dagger [\hat{a}, (\hat{a}^\dagger)^n] + \frac{(\hat{a}^\dagger)^n}{2} \right) |0\rangle. \end{aligned} \quad (1.8)$$

The commutator

$$\begin{aligned} [\hat{a}, (\hat{a}^\dagger)^n] &= \hat{a} (\hat{a}^\dagger)^n - \hat{a}^\dagger \hat{a} (\hat{a}^\dagger)^{n-1} + \hat{a}^\dagger \hat{a} (\hat{a}^\dagger)^{n-1} - (\hat{a}^\dagger)^2 \hat{a} (\hat{a}^\dagger)^{n-2} + \dots + (\hat{a}^\dagger)^{n-1} \hat{a} \hat{a}^\dagger - (\hat{a}^\dagger)^n \hat{a} \\ &= [\hat{a}, \hat{a}^\dagger] (\hat{a}^\dagger)^{n-1} + \hat{a}^\dagger [\hat{a}, \hat{a}^\dagger] (\hat{a}^\dagger)^{n-2} \dots (\hat{a}^\dagger)^{n-1} [\hat{a}, \hat{a}^\dagger] = n (\hat{a}^\dagger)^{n-1}. \end{aligned}$$

Inserting this result into 1.8 gives

$$\hat{H} (\hat{a}^\dagger)^n |0\rangle = \hbar\omega_0 \left(n + \frac{1}{2} \right) (\hat{a}^\dagger)^n |0\rangle. \quad (1.9)$$

This says that the state $(\hat{a}^\dagger)^n |0\rangle$ is an eigenstate of \hat{H} with eigenvalue $(n + 1/2)\hbar\omega_0$. In other words, the operator $(\hat{a}^\dagger)^n$ introduces n quanta of $\hbar\omega_0$. Here we have shown this only when $(\hat{a}^\dagger)^n$ acts on the ground state, but it is true whichever state it acts on.

We should normalise the state $(\hat{a}^\dagger)^n |0\rangle$. You should work through the algebra on the problems sheet to convince yourself of the following results

$$|n\rangle = \frac{1}{\sqrt{n!}} (\hat{a}^\dagger)^n |0\rangle \text{ is the normalised eigenstate of } \hat{H} \text{ with } n \text{ quanta} \quad (1.10)$$

$$\hat{a}^\dagger |n\rangle = \sqrt{n+1} |n+1\rangle, \quad \hat{a} |n\rangle = \sqrt{n} |n-1\rangle \quad (1.11)$$

$$\hat{a}^\dagger \hat{a} |n\rangle = n |n\rangle. \quad (1.12)$$

The conjugate operators, \hat{a}^\dagger and \hat{a} , are usually called ladder (or raising and lowering) operators in the context of the SHO. They will be called creation and annihilation operators in the context of QFT. This is to reflect the idea that \hat{a}^\dagger creates a quantum of energy $\hbar\omega$ while \hat{a} annihilates a quantum of energy. The operator combination, $\hat{a}^\dagger \hat{a}$, is a counting operator. It counts the number of quanta in the oscillator. The ground state $|0\rangle$ satisfies

$$\hat{a} |0\rangle = 0. \quad (1.13)$$

Relation to Quantum Field Theories

Equation 1.4 will turn out to be the basis of any QFT we can solve, at least for bosonic particles. In effect, there will be many modes into which we can introduce energy but the

Hamiltonian will end up with creation and annihilation operators exactly as in 1.4 but with a mode index l :

$$[\hat{a}_l, \hat{a}_{l'}^\dagger] = \delta_{l,l'}. \quad (1.14)$$

The operators \hat{a}_l^\dagger and \hat{a}_l will be the creation and annihilation operators for particles or quanta in modes denoted by l , where l might involve a wavevector, \mathbf{q} , and a polarisation index. The Hamiltonian for free bosons (which means non-interacting bosons) will be

$$\hat{H} = \sum_l \hbar \omega_l \left(\hat{a}_l^\dagger \hat{a}_l + \frac{1}{2} \right). \quad (1.15)$$

We will need to be careful to count modes correctly. If the system is a quantum mechanical one, the sum will be a sum over discrete labels l . If there are infinitely many modes, then the sum will become an integral over the quantum number(s) and involve a density of states for these modes. In translationally invariant systems the sum over l will usually include an integral over wave-vectors \mathbf{k} .

1.1 Phonons and a 1D QFT

A quantum mechanical model, which can be solved using creation and annihilation operators, describes the excitations of a chain of ions of mass m . The motion of the ions from their equilibrium position is restricted to 1D. Each ion in the chain will be at $la + x_l$, where $x_l \ll a$ is the displacement from equilibrium, which can depend on time. We are assuming that, in the ground state, the ions sit on a lattice with lattice constant a . We will identify the N 'th site in the lattice with the 0'th site (periodic boundary conditions). The momentum of each ion $p_i = m\dot{x}_i$. The Hamiltonian is

$$H = \sum_{l=0}^{N-1} \frac{p_l^2}{2m} + \sum_{l,l'=0}^{N-1} V(x_l - x_{l'}). \quad (1.16)$$

The interaction (here assumed to be a pairwise interaction) between ions could be computed from a realistic model using interionic potentials or first principles calculations (and this is often done in practice). However, to illustrate the use of creation and annihilation operators, we will assume that the energy involves nearest neighbours only:

$$\sum_{l,l'=0}^{N-1} V(|x_l - x_{l'}|) \rightarrow \frac{k}{2} \sum_{i=0}^{N-1} (x_i - x_{i+1})^2 = \frac{k}{2} \sum_{l=0}^{N-1} 2x_l^2 - x_l(x_{l-1} + x_{l+1}). \quad (1.17)$$

(When writing out the term $(x_l - x_{l+1})^2 = x_l^2 + x_{l+1}^2 - 2x_l x_{l+1}$, we have put one of the terms $x_l x_{l+1}$ with x_l^2 and one with x_{l+1}^2 in the expression on the right.) In 1.17 take $x_{-1} = x_{N-1}$.

As this system is periodic, working with Fourier decompositions should be helpful. We write $x_l = (1/\sqrt{N}) \sum_n u_n(t) e^{iq_n l a}$ and $p_n = (1/\sqrt{N}) \sum_l \pi_n(t) e^{iq_n l a}$, where $q_n = 2n\pi/Na$ for

integer $n \in \{-N/2, (N/2) - 1\}$, giving¹

$$\begin{aligned}
 \sum_l 2x_l^2 - x_l(x_{l-1} + x_{l+1}) &= \frac{1}{N} \sum_{l,n,n'} u_n u_{n'} (e^{i(q_{n'}+q_n)la} (2 - (e^{-iq_{n'}a} + e^{iq_{n'}a}))) \\
 &= \sum_{n,n'} u_n u_{n'} \delta_{n+n'} (2 - 2\cos(q_n a)) \\
 &= \sum_n u_{-n} u_n (2 - 2\cos(q_n a)) \quad \text{while} \\
 \sum_l \frac{1}{2m} p_l^2 &= \frac{1}{2m} \sum_n \pi_n \pi_{-n}.
 \end{aligned}$$

Inserting the Fourier decompositions into 1.16 gives

$$\begin{aligned}
 H &= \sum_n \frac{\pi_n \pi_{-n}}{2m} + \frac{4k \sin^2(q_n a/2)}{2} u_n u_{-n} = \sum_n \frac{\pi_n \pi_{-n}}{2m} + \frac{m\omega_n^2}{2} u_n u_{-n}, \\
 &= \sum_n \frac{1}{2} \hbar \omega_n (\pi'_n \pi'_{-n} + u'_n u'_{-n})
 \end{aligned} \tag{1.18}$$

where $\omega_n^2 = 4k \sin^2(q_n a/2)/m$ and we have introduced scaled quantities: $u'_n = u_n \sqrt{m\omega_n/\hbar}$ and $\pi'_n = \pi_n/\sqrt{\hbar m\omega_n}$.

We can write 1.18 as

$$H = \sum_n \frac{1}{2} \hbar \omega_n \frac{(u'_{-n} - i\pi'_n)(u'_n + i\pi'_{-n}) + (u'_{-n} + i\pi'_n)(u'_n - i\pi'_{-n})}{2}. \tag{1.19}$$

The requirement, that x_l and p_l are real, forces $u_n = u_n^*$ and $\pi_n = \pi_n^*$ (the coefficients of $e^{iq_n la}$ and $e^{-iq_n la}$ must be complex conjugates of each other). This means that $(u'_{-n} + i\pi'_n)$ is the conjugate of $(u'_n - i\pi'_{-n})$. Defining conjugate variables

$$a_n = \frac{u'_n + i\pi'_{-n}}{\sqrt{2}} \quad \text{and} \quad a_n^* = \frac{(u'_n)^* - i(\pi'_{-n})^*}{\sqrt{2}} = \frac{u'_{-n} - i\pi'_n}{\sqrt{2}}, \tag{1.20}$$

the Hamiltonian becomes

$$H = \sum_n \frac{1}{2} \hbar \omega_n (a_n a_n^* + a_n^* a_n). \tag{1.21}$$

(The term $(u'_{-n} + i\pi'_n)(u'_n - i\pi'_{-n}) \rightarrow a_{-n} a_{-n}^*$ and has been switched with the corresponding term with $n \rightarrow -n$ in the sum.)

The Hamiltonian 1.19 is equivalent to that of N harmonic oscillators (see 1.2). With

$$u'_n = \frac{a_n + a_{-n}^*}{\sqrt{2}} \quad \text{and} \quad \pi'_n = \frac{a_n - a_{-n}^*}{i\sqrt{2}}, \tag{1.22}$$

the general expressions for x_l and p_l are

$$\begin{aligned}
 x_l &= \frac{1}{\sqrt{N}} \sum_n \sqrt{\frac{\hbar}{2m\omega_n}} (a_n + a_{-n}^*) e^{iq_n la} = \frac{1}{\sqrt{N}} \sum_n \sqrt{\frac{\hbar}{2m\omega_n}} (a_n e^{iq_n la} + a_n^* e^{-iq_n la}) \\
 p_l &= \frac{-i}{\sqrt{N}} \sum_n \sqrt{\frac{\hbar m\omega_n}{2}} (a_n - a_{-n}^*) e^{iq_n la} = \frac{-i}{\sqrt{N}} \sum_n \sqrt{\frac{\hbar m\omega_n}{2}} (a_n e^{iq_n la} - a_n^* e^{-iq_n la})
 \end{aligned} \tag{1.23}$$

¹Care needs to be taken to distinguish between rings with an even number of sites, $n \in [-N/2, (N/2) - 1]$, or an odd number of sites, $n \in [-(N-1)/2, (N-1)/2]$.

Here the dummy variable n has been replaced by $-n$ for the term involving a_{-n}^* to obtain the obviously real forms on the right. The factors of $\sqrt{\hbar/m\omega_n}$ are the analogues of those relating x' and p' to x and p in 1.2, but scale individual modes labelled by n differently.

Everything in 1.21 is classical. We quantise by writing:

$$a_n \rightarrow \hat{a}_n \text{ and } a_n^* \rightarrow \hat{a}_n^\dagger, \quad \text{with } [\hat{a}_n, \hat{a}_{n'}^\dagger] = \delta_{n,n'}. \quad (1.24)$$

This is analogous to what we had in 1.4 but with a creation and annihilation operator for each mode now labelled by n . The form for the quantised Hamiltonian is

$$\hat{H} = \sum_n \hbar\omega_n \left(\hat{a}_n^\dagger \hat{a}_n + \frac{1}{2} \right). \quad (1.25)$$

The operators, \hat{a}_n^\dagger and \hat{a}_n are the creation and annihilation operators for what are called the phonon excitations of the 1D chain. For a finite system on a lattice, we are doing the same quantum mechanical calculations that were used to solve the simple harmonic oscillator. There are now N modes with frequencies given by ω_n . The operator $\hat{a}_n^\dagger \hat{a}_n$ is the counting operator for quanta (see Eq 1.12, with different modes labelled by n). We see that the energy of the system includes a sum over the phonon modes of the number of phonons in each mode times its energy $\hbar\omega_n$. Eq 1.25 is in the form we predicted in 1.15, with the mode index l given by the index n .

Imposing the commutation relations 1.24 gives the correct commutation relations between position and momentum:

$$[\hat{x}_l, \hat{p}_l] = i\hbar\delta_{ll}, \quad [\hat{x}_l, \hat{x}_{l'}] = 0 \quad \text{and} \quad [\hat{p}_l, \hat{p}_{l'}] = 0. \quad (1.26)$$

Note that imposing the result 1.26 is the canonical quantisation process. We have chosen to impose the corresponding commutator on the creation and annihilation operators to bring out the connection to the SHO.

The notation is compact and reduces all calculations to algebra (of the creation and annihilation operators). The phonon modes in this model are independent but, if the original potential function in 1.17 included higher order terms, the modes would be coupled. The relation of 1.24 to 1.23 is behind the notion of second quantisation. It is the amplitudes of the waves that have become operators. They generate states in the space of allowed states of the quantum system with well-defined numbers of energy quanta, $\hbar\omega_n$.

A continuum version of the phonon system is included in Appendix A and some relativistic fields are discussed in Appendix B.

1.2 Time-Dependence

The classical fields in 1.23 are time-dependent—the displacement is varying as a function of time. This means that u_n and π_n , and the a_n and a_n^* , are time-dependent. (In fact we know what the dependence is as the modes are those of simple harmonic oscillators, so $u_n \sim e^{-i\omega_n t}$.) When quantising the amplitudes in 1.24, there is a choice where to put the

time-dependence. In the Schrodinger picture the time-dependence belongs to the wave-function $\psi(t) = e^{-i\omega_0 t}\psi(0)$, while the operators are time-independent.

There is an alternative picture in which the wave-function or state-function is time-independent and the operators become time-dependent. This is called the Heisenberg picture. The derivation of the correct time-dependence of the operators \hat{a}_n and \hat{a}_n^\dagger is left as an exercise on the examples sheet. We find $\hat{a}_n \sim e^{-i\omega_n t}$ and $\hat{a}_n^\dagger \sim e^{i\omega_n t}$. It is not necessary to work in the Heisenberg picture but, in establishing QFTs, it is usual. It makes the notation easier.

Where does this apparent ambiguity come from? In quantum theory, what matter are matrix elements between states. The matrix element structure allows us to put the time-dependence where we want it. It can be associated with the state-function $|\psi(t)\rangle$, or with the operators \hat{O} . (In fact, when doing calculations it is often split between the two. Whatever makes the calculations easier to do safely is what people usually do.) The idea of this transformation is that, in the Schrodinger picture, the state-function satisfies

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = \hat{H} |\psi(t)\rangle \Rightarrow |\psi(t)\rangle = e^{-i\hat{H}t/\hbar} |\psi(0)\rangle \equiv \hat{U}(t) |\psi(0)\rangle. \quad (1.27)$$

When computing the matrix elements of an operator we can write

$$\langle \psi'(t) | \hat{O}_S | \psi(t) \rangle = \langle \psi'(0) | \hat{U}^\dagger(t) \hat{O}_S \hat{U}(t) | \psi(0) \rangle \equiv \langle \psi'(0) | \hat{O}_H(t) | \psi(0) \rangle. \quad (1.28)$$

Here $\hat{O}_H(t)$ is the operator in the Heisenberg representation. The matrix element (matrix elements define what is measurable) is the same in both pictures. The relation between the Heisenberg representation of an operator, $\hat{O}_H(t)$, and the Schrodinger representation, \hat{O}_S , is

$$\hat{O}_H(t) = \hat{U}^\dagger(t) \hat{O}_S \hat{U}(t). \quad (1.29)$$

1.3 Summary and questions

We have introduced the Hamiltonians for some free bosonic fields. The QFTs are essentially the same for these systems, namely 1.15. We have looked at the phonon system 1.25 and the field theoretic model of oscillations in Appendix A, see A.12. The KG field B.3 and electromagnetic waves B.10 are outlined in the Appendix B. There may be different dispersion relations and a different number of modes but the form of the Hamiltonians is similar in each case. The creation and annihilation operators describe how particles/quanta are created and annihilated in the system. The number operator, $\hat{a}_l^\dagger \hat{a}_l$, counts the number of particles or quanta that are added² in mode l . In the continuum cases, the operator $\hat{a}_p^\dagger \hat{a}_p$ gives a density or energy density operator in momentum space depending on how the creation and annihilation operators have been normalised. The fields are free in the sense that the Hamiltonian contains only terms involving the number (or density) operator. We also call fields free, if we can find a transformation that takes the Hamiltonian to this form.

An issue relating to QFT is illustrated perfectly by the 1D continuum theory of elastic waves. The field theory includes the limit $q \rightarrow \infty$ corresponding to infinitely many

²We will tend to refer to quanta in a wave-system in which $m_0 = 0$ and particles when $m_0 > 0$.

excitations with energies tending to infinity. Of course we know in the case of the 1D ionic chain that these are unphysical modes. Any mode with a wavelength on length scales less than the lattice spacing is not physical—on this length scale there is nothing that can oscillate! But what happens in the KG case or the case of the EM field? In both these cases there are also infinitely many modes ($p \rightarrow \infty$).

One approach to dealing with these modes is to assume that there must be a high frequency cutoff (as there clearly is in the case of the elastic waves) or equivalently argue that the theory is a long wavelength effective theory. We can ask what might set such a limit in other cases. Here's an argument to show that EM Theory cannot be right for arbitrarily large wavenumber. If a photon wavepacket has wavelengths less than λ_m , it can be localised in a spherical region of radius $\sim \lambda_m$. A wavepacket with only one photon on average would have energy $> hc/\lambda_m$ equivalent to a mass of $h/c\lambda_m$. If λ_m is less than the Schwarzschild radius for the one-photon wavepacket mass ($r_S = 2GM/c^2 \sim 2Gh/c^3$) a single photon will create its own black hole. At such short wavelengths the theory cannot be right.

Another worry is the role of zero-point energy, E_0 , which is infinite (the sum $\sum_q \hbar\omega_q/2$ is unbounded). One could argue that E_0 is not physical as it is not measurable within the QFT. All spectroscopic and thermodynamic quantities depend only on energy differences and so E_0 drops out. But, if this energy is there, it should couple to gravity and contribute to the stress-energy tensor.

Quantum theory was introduced by Planck to explain the absence of an ultra-violet divergence (the infinite number of short wavelength modes implied infinite radiation according to classical thermodynamics). We have ended up QFTs which contain similar problems—unphysical infinities.

Chapter 2

QFT—Fermions

We will “guess” how to construct field operators to describe fermions. It is possible to get a long way from first principles but this approach is probably better suited to a full length graduate module. It is covered in all the main textbooks on QFT. We will take the line that the justification is always empirical anyway—if it describes and predicts what we observe then that’s good enough. Good sources for more extensive treatments are [Eduardo Fradkin](#) and [David Tong](#).

We assume that there are field operators \hat{c}_λ and \hat{c}_λ^\dagger , which annihilate and create fermions in some mode or state labelled by λ . λ could include orbital and spin indices. If we take a simple one-state system (the rough equivalent of the harmonic oscillator for the boson), we obtain most of the answer. There are two states, the empty state or vacuum $|0\rangle$ and the occupied state $|1\rangle$:

$$\hat{c}^\dagger|0\rangle = |1\rangle, \quad \hat{c}|1\rangle = |0\rangle, \quad \text{while} \quad \hat{c}^\dagger|1\rangle = 0 \quad \hat{c}|0\rangle = 0. \quad (2.1)$$

The first two identities are as for bosons, see 1.11, but allowing only occupation $n = 0$ or $n = 1$. The third result encodes Pauli’s exclusion principle, which states that a single quantum state can contain no more than one fermion. These results can be written as follows

$$\{\hat{c}, \hat{c}^\dagger\} \equiv \hat{c}\hat{c}^\dagger + \hat{c}^\dagger\hat{c} = 1, \quad \{\hat{c}, \hat{c}\} = 0, \quad \{\hat{c}^\dagger, \hat{c}^\dagger\} = 0. \quad (2.2)$$

The quantity $\{\hat{A}, \hat{B}\}$ is called the anticommutator of \hat{A} and \hat{B} . The first result follows by considering the action of the anticommutator on the states of the system, which are $|0\rangle$ and $|1\rangle$. If the operator identity holds when acting on all states for the system (there are only two here) then it is true as an operator identity.

What’s left is the so-called spin-statistics theorem. The state-function for a system of indistinguishable fermions must be antisymmetric with respect to the interchange of any two particles. The idea is that, as the particles are indistinguishable, no measurement can distinguish between permutations of the coordinates of the indistinguishable particles, so $\psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_i, \dots, \mathbf{r}_j, \dots)$ and $\psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_j, \dots, \mathbf{r}_i, \dots)$ must describe the same state. Permuting the positions of two particle positions twice takes the wavefunction back to where it started. It follows that, under permutation of indistinguishable particles (here particles i and j), a state-function must satisfy

$$\psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_j, \dots, \mathbf{r}_i, \dots) = \pm \psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_i, \dots, \mathbf{r}_j, \dots). \quad (2.3)$$

The symmetric case (plus sign on rhs of 2.3) is correct for bosons. The antisymmetric case (minus sign in 2.3) applies in the case of fermions¹. It is possible, at least for free particles, to show that the only consistent quantisation of the Dirac equation describes fermions. However, we will take this to be an axiom.

If we allow for the fermion to have two spin states which we label \uparrow and \downarrow , the possible states of the single-orbital two spin state system are: $|0\rangle$ (the vacuum), $|\uparrow\rangle = \hat{c}_{\uparrow}^{\dagger}|0\rangle$, $|\downarrow\rangle = \hat{c}_{\downarrow}^{\dagger}|0\rangle$ and $|\uparrow\downarrow\rangle = \hat{c}_{\uparrow}^{\dagger}\hat{c}_{\downarrow}^{\dagger}|0\rangle$. The antisymmetric condition translates into the following anticommutation relations

$$\{\hat{c}_{\sigma}, \hat{c}_{\sigma'}^{\dagger}\} = \delta_{\sigma,\sigma'}, \quad \{\hat{c}_{\sigma}, \hat{c}_{\sigma'}\} = 0 = \{\hat{c}_{\sigma}^{\dagger}, \hat{c}_{\sigma'}^{\dagger}\}. \quad (2.4)$$

For example the antisymmetric condition for the two fermion state is

$$\hat{c}_{\downarrow}^{\dagger}\hat{c}_{\uparrow}^{\dagger}|0\rangle = -\hat{c}_{\uparrow}^{\dagger}\hat{c}_{\downarrow}^{\dagger}|0\rangle, \quad \text{which can be written} \quad (\hat{c}_{\downarrow}^{\dagger}\hat{c}_{\uparrow}^{\dagger} + \hat{c}_{\uparrow}^{\dagger}\hat{c}_{\downarrow}^{\dagger})|0\rangle = 0|0\rangle. \quad (2.5)$$

The first equation says that if we change the order, in which we populate the two spin states with the fermions, the state-function must change sign. The second equation is the first equation written in terms of the anticommutator of $\hat{c}_{\uparrow}^{\dagger}$ and $\hat{c}_{\downarrow}^{\dagger}$. You can check that the action of the operators, as they appear in the anticommutators, on all four states of the system is correct. If it is true for all states in a complete basis, then it is true as an operator identity.

The general result corresponding to 2.4 for fermionic states (not just the two-spin one-orbital case that we have considered) is

$$\{\hat{c}_{\lambda}, \hat{c}_{\lambda'}^{\dagger}\} = \delta_{\lambda,\lambda'}, \quad \{\hat{c}_{\lambda}, \hat{c}_{\lambda'}\} = 0 = \{\hat{c}_{\lambda}^{\dagger}, \hat{c}_{\lambda'}^{\dagger}\}. \quad (2.6)$$

The number operator for a single quantum state (here labelled by λ) is

$$\hat{n}(\lambda) = \hat{c}_{\lambda}^{\dagger}\hat{c}_{\lambda}. \quad (2.7)$$

The operator $\hat{n}(\lambda)$ has eigenvalues 0 or 1, as no single quantum state can be occupied by more than one fermion.

The index λ is in general some composite index. Examples we will mention are a Helium atom, free non-relativistic fermions described by Schrodinger's equation, and Dirac fermions. Before looking at some examples, let's ask what the wavefunctions generated by the $\hat{c}_{\lambda}^{\dagger}$ look like. Suppose λ characterises both an orbital in some atom and the spin state. The state $|\psi\rangle = \hat{c}_1^{\dagger}\hat{c}_2^{\dagger}|0\rangle$ corresponds to occupation of state χ_1 and state χ_2 by fermions. For example, χ_1 could be an s-orbital with spin \uparrow and χ_2 could be a p-orbital in an atom with spin \downarrow . The wavefunction cannot be $\chi_1(\mathbf{r}_1)\chi_2(\mathbf{r}_2)$ as this has particle 1 in state 1 and we cannot know which state a particular particle occupies. The wavefunction must be antisymmetric with respect to permutation of the particles. We must therefore add all permutations of the assignment of the particle coordinates and spins to the orbitals and spin states that are occupied, with a sign determined by how many permutations are needed to get to a given term from some reference state. (Any state can be the reference state as the overall phase of the wavefunction is not important.) We also need to normalise correctly.

¹The more precise argument takes account of what can happen when a particle encircles another. In the 3D this is equivalent to no change as the path can be shrunk to a point without crossing the other particle. This means that bosons and fermions are the only possibility. In 2D this is not true, a fact behind the phenomena of anyons and non-abelian statistics.

For the two state case the wavefunction corresponding to $|\psi\rangle = \hat{c}_1^\dagger \hat{c}_2^\dagger |0\rangle$ would be

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{\sqrt{2}} (\chi_1(\mathbf{r}_1)\chi_2(\mathbf{r}_2) - \chi_2(\mathbf{r}_1)\chi_1(\mathbf{r}_2)) = \frac{1}{\sqrt{2!}} \begin{vmatrix} \chi_1(\mathbf{r}_1) & \chi_1(\mathbf{r}_2) \\ \chi_2(\mathbf{r}_1) & \chi_2(\mathbf{r}_2) \end{vmatrix}. \quad (2.8)$$

Writing the antisymmetric combination in the two-particle case as a determinant generalises naturally to the N -particle case, $|\psi\rangle = \left(\prod_{i=1}^N \hat{c}_i^\dagger\right) |0\rangle$,

$$\psi(\mathbf{r}_1, \dots, \mathbf{r}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(\mathbf{r}_1) & \chi_1(\mathbf{r}_2) & \cdots & \chi_1(\mathbf{r}_N) \\ \chi_2(\mathbf{r}_1) & \chi_2(\mathbf{r}_2) & \cdots & \chi_2(\mathbf{r}_N) \\ \vdots & \vdots & \cdots & \vdots \\ \chi_N(\mathbf{r}_1) & \chi_N(\mathbf{r}_2) & \cdots & \chi_N(\mathbf{r}_N) \end{vmatrix}. \quad (2.9)$$

States of this type are called Slater determinants. The structure of the determinant means that, if any two or more of the states are the same, two (or more) rows of the matrix become identical and the determinant vanishes. This ensures that Fermi's exclusion principle is satisfied automatically for these many-particle fermionic wavefunctions. The spin-statistics theorem is also satisfied as exchanging any two rows (or columns) corresponds to permuting the corresponding orbitals (or particles) and leads to a change in sign.

Helium Atom

The single-particle states in a Coulomb potential are characterised by the principal quantum number $n = 1, 2, \dots$, the angular momentum $l = 0, 1, \dots, (n-1)$, and the projection of the angular momentum along some direction usually denoted by z , $m_z = l, (l-1), \dots, -l$. Each of these orbitals can be occupied by a particle with spin \uparrow , and with spin \downarrow . If we ignore the interaction between the two electrons, we would write the ground state as:

$$|\text{gs}\rangle_{\text{He}} = \hat{c}_{1,0,\uparrow}^\dagger \hat{c}_{1,0,\downarrow}^\dagger |0\rangle. \quad (2.10)$$

This describes a state with two electrons with opposite spin both occupying the 1s orbital. Here the labels denote, $n = 1$ and $l = 0$ (it is not strictly necessary to specify a value of l as this is automatically zero when $n = 1$). Let's look at what this means in terms of wavefunctions. The states in 2.8 are χ_1 , which has $n = 1$, $l = 0$ and $\sigma = \uparrow$ while χ_2 has $n = 1$, $l = 0$ and $\sigma = \downarrow$:

$$\phi(\mathbf{r}_1, \sigma_1, \mathbf{r}_2, \sigma_2) = \phi_{1,0}(\mathbf{r}_1)\phi_{1,0}(\mathbf{r}_2) \left(\frac{\uparrow_1 \downarrow_2 - \downarrow_1 \uparrow_2}{\sqrt{2}} \right). \quad (2.11)$$

In this case the wavefunction factorises into a spatial part, that is symmetric with respect to the interchange of the particles, and a spin part that is antisymmetric. (Note that not all states separate into spin and orbital parts like this.) The suffix in 2.10 is only there to emphasise that the 'vacuum', $|0\rangle$, is for a He atom. In future, we will assume that, unless discussed specifically, it is clear which is the relevant "vacuum".

An excited state could take an electron out of the $n = 0$ shell and put it in a higher shell. An example of an excited state (es) would be

$$\begin{aligned} |\text{es}\rangle_{\text{He}} &= \hat{c}_{2,1,-1\uparrow}^\dagger \hat{c}_{1,0,\uparrow}^\dagger |\text{gs}\rangle_{\text{He}} = \hat{c}_{2,1,-1\uparrow}^\dagger \hat{c}_{1,0,\uparrow}^\dagger \hat{c}_{1,0,\uparrow}^\dagger \hat{c}_{1,0,\downarrow}^\dagger |0\rangle_{\text{He}} \\ &= \hat{c}_{2,1,-1\uparrow}^\dagger \hat{c}_{1,0,\downarrow}^\dagger |0\rangle_{\text{He}}. \end{aligned} \quad (2.12)$$

The excited state has one electron in the $n = 1$ shell and another in the $n = 2$ with total angular momentum quantum number $l = 1$ and projection of the angular momentum along the z -axis of $m_z = -1$.

The Slater determinant states involve products of single-particle states. We have put electrons in one-particle orbitals, which we assume are the best solutions for the eigenstates of some effective one-particle Hamiltonian. This approach can be made to treat the interactions between the electrons but only on average. However, the motion of the two electrons can be correlated, in which case the Slater determinant states form a basis for the true eigenstates of the two-body interacting Hamiltonian (see below), but these eigenstates will be superpositions of Slater determinants given in 2.9.

Fock Space

The notion of working with states characterised by the occupation numbers of modes rather than explicit wavefunctions is sometimes referred to as second quantisation. The space, in which states with different or varying numbers of particles (bosons or fermions) are possible, is called the Fock space. Implicitly we have been identifying states via the occupation numbers of single-particle states. This is almost always what one does. While it is essential to define this space in QFT when particle numbers are changing, it is also sometimes helpful to work with the Fock space even when the particle number is not changing. One may want to compute matrix elements of operators at fixed particle number by allowing the particle number to change and then come back to its original value.

2.1 Non-relativistic QFT for electrons

The bosonic general Hamiltonian 1.15, suggests that the Hamiltonian for free non-interacting fermions governed by the Schrodinger equation should be

$$\hat{H} = \sum_{\sigma=\uparrow,\downarrow} \sum_{\mathbf{k}} E_{\mathbf{k}} \hat{c}_{\mathbf{k},\sigma}^\dagger \hat{c}_{\mathbf{k},\sigma}. \quad (2.13)$$

The label λ in 2.6 has become the momentum label \mathbf{k} and the spin label σ . The energy of the fermion is $E_{\mathbf{k}} = \hbar^2 k^2 / 2m$. The number operator for a single quantum state (here labelled by \mathbf{k}, σ), is (see 2.7)

$$\hat{n}_\sigma(\mathbf{k}) = \hat{c}_{\mathbf{k},\sigma}^\dagger \hat{c}_{\mathbf{k},\sigma}. \quad (2.14)$$

The Hamiltonian 2.13 sums the energies $E_{\mathbf{k}}$ of the occupied states ($\hat{n}_\sigma(\mathbf{k}) = 1$).

We can write the ground state of the free Fermi gas by occupying all the states with energies up to and including the Fermi energy E_F , corresponding to wavenumber k_F . We will call this state the Fermi Sea (FS):

$$|FS\rangle = \left(\prod_{|\mathbf{k}| < k_F} \prod_{\sigma=\uparrow,\downarrow} \hat{c}_{\mathbf{k},\sigma}^\dagger \right) |0\rangle. \quad (2.15)$$

This is a formalised way of writing down the ground state of a Fermi gas. The idea of filling states up to the Fermi level should be familiar from earlier year modules.

We can ask what are the operators which create and annihilate particles locally in space. They are

$$\hat{\psi}_\sigma^\dagger(\mathbf{x}) = \sum_{\mathbf{k}} \hat{c}_{\mathbf{k},\sigma}^\dagger e^{-i\mathbf{k}\cdot\mathbf{x}} \quad \text{and} \quad \hat{\psi}_\sigma(\mathbf{x}) = \sum_{\mathbf{k}} \hat{c}_{\mathbf{k},\sigma} e^{i\mathbf{k}\cdot\mathbf{x}}. \quad (2.16)$$

We should be careful about the notation of the sums. In 2.16 we are using the sum over wavevectors rather loosely to represent either a sum over a discrete set of wavevectors or an integral over wavevectors. If the system is finite, then we would normally impose periodic boundary conditions and the sum would be over the discrete values of \mathbf{k} consistent with the conditions. We would normalise the plane wave with the factor $1/\sqrt{V}$ (this is similar to the factor $1/\sqrt{L}$ in A.8). The same factor would go into the definition of the Fourier amplitude, $f(\mathbf{k})$:

$$f(\mathbf{r}) = \frac{1}{\sqrt{V}} \sum_{\mathbf{k}} f(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}} \quad \text{and} \quad f(\mathbf{k}) = \frac{1}{\sqrt{V}} \int d^3\mathbf{r} f(\mathbf{r}) e^{-i\mathbf{k}\cdot\mathbf{r}}. \quad (2.17)$$

If the system is defined over all space then $\frac{1}{\sqrt{V}} \sum_{\mathbf{k}} \cdots \rightarrow \int \frac{d^3\mathbf{k}}{(2\pi)^3} \cdots$ in the infinite system limit. The factors of 2π have been put in the inverse Fourier transforms but can also be split between this and the Fourier transform.

Their anticommutation relations are

$$\begin{aligned} \{\hat{\psi}_\sigma(\mathbf{x}), \hat{\psi}_{\sigma'}^\dagger(\mathbf{x}')\} &= \sum_{\mathbf{k}} \sum_{\mathbf{k}'} \{\hat{c}_{\mathbf{k},\sigma}, \hat{c}_{\mathbf{k}',\sigma'}^\dagger\} e^{i\mathbf{k}\cdot\mathbf{x}} e^{-i\mathbf{k}'\cdot\mathbf{x}'} \\ &= \sum_{\mathbf{k}} \sum_{\mathbf{k}'} \delta_{\sigma,\sigma'} \delta_{\mathbf{k},\mathbf{k}'} e^{i\mathbf{k}\cdot\mathbf{x}} e^{-i\mathbf{k}'\cdot\mathbf{x}'} = \delta_{\sigma,\sigma'} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')} \\ \text{putting in the normalisation} &\rightarrow \delta_{\sigma,\sigma'} \int \frac{d^3\mathbf{k}}{(2\pi)^3} e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')} \end{aligned} \quad (2.18)$$

$$= \delta_{\sigma,\sigma'} \delta^{(3)}(\mathbf{x} - \mathbf{x}'). \quad (2.19)$$

Here $\delta^{(3)}$ is the 3D delta function and we have used the infinite system normalisation. The result 2.19 shows that the anticommutator is not a number but a density (you have to integrate the δ -function over a volume to get a number). The operators $\hat{\psi}^\dagger(\mathbf{x})$ and $\hat{\psi}(\mathbf{x})$ create and annihilate fermions at the point \mathbf{x} while $|\mathbf{x}\rangle = \hat{\psi}^\dagger(\mathbf{x})|0\rangle$ describes one particle exactly at this point. The corresponding wavefunction satisfies $|\psi(\mathbf{x})|^2 = \delta^{(3)}(\mathbf{x})$, suggesting $\psi(\mathbf{x}) = \sqrt{\delta^{(3)}(\mathbf{x})}$. This is actually correct. However, you should see this more as a formal device than a wavefunction that is ever occupied (the fermion would have infinite kinetic energy).

The operator $\hat{\psi}_\sigma^\dagger(\mathbf{x})\hat{\psi}_\sigma(\mathbf{x})$, rather than being a number operator (see 2.7), is the density operator:

$$\hat{\rho}(\mathbf{x}) = \sum_{\sigma} \hat{\psi}_\sigma^\dagger(\mathbf{x})\hat{\psi}_\sigma(\mathbf{x}). \quad (2.20)$$

(If the anticommutator is a density as in 2.19, so must its two component terms be.) Its Fourier transform is

$$\hat{\rho}(q) = \frac{1}{\sqrt{V}} \sum_{\sigma} \sum_{\mathbf{k}} \hat{c}_{\mathbf{k}+\mathbf{q},\sigma}^\dagger \hat{c}_{\mathbf{k},\sigma}. \quad (2.21)$$

Added for Interest To derive the result in the form given, we will work in a finite system of volume V . Then

$$\begin{aligned}
 \hat{\rho}(q) &= \frac{1}{\sqrt{V}} \int_V d^3\mathbf{x} e^{i\mathbf{q}\cdot\mathbf{x}} \hat{\rho}(\mathbf{x}) = \frac{1}{\sqrt{V}} \sum_{\sigma} \int_V d^3\mathbf{x} e^{i\mathbf{q}\cdot\mathbf{x}} \hat{\psi}_{\sigma}^{\dagger}(\mathbf{x}) \hat{\psi}_{\sigma}(\mathbf{x}) \\
 &= \frac{1}{\sqrt{V}} \sum_{\sigma, \mathbf{k}, \mathbf{k}'} \int_V d^3\mathbf{x} e^{i\mathbf{q}\cdot\mathbf{x}} \hat{c}_{\mathbf{k}, \sigma}^{\dagger} \frac{e^{-i\mathbf{k}\cdot\mathbf{x}}}{\sqrt{V}} \hat{c}_{\mathbf{k}', \sigma} \frac{e^{i\mathbf{k}'\cdot\mathbf{x}}}{\sqrt{V}} \\
 &= \frac{1}{\sqrt{V}} \sum_{\sigma, \mathbf{k}, \mathbf{k}'} \hat{c}_{\mathbf{k}, \sigma}^{\dagger} \hat{c}_{\mathbf{k}', \sigma} \delta_{-\mathbf{k} + \mathbf{q} + \mathbf{k}', 0} = \frac{1}{\sqrt{V}} \sum_{\sigma, \mathbf{k}'} \hat{c}_{\mathbf{k}' + \mathbf{q}, \sigma}^{\dagger} \hat{c}_{\mathbf{k}', \sigma}
 \end{aligned} \tag{2.22}$$

(A factor of V is cancelled by the integral over V .)

We will look at the role of interactions later, but we will write here the operator corresponding to the Coulomb interaction. This is a term that needs to be added to the Hamiltonian (and is behind all the problems worth studying in materials). We can write the Coulomb interaction ($\epsilon_0 = 1$) in Fourier space:

$$V(\mathbf{x} - \mathbf{x}') = \frac{e^2}{4\pi|\mathbf{x} - \mathbf{x}'|} = \frac{1}{V} \sum_{\mathbf{q}} \frac{e^2}{|\mathbf{q}|^2} e^{i\mathbf{q}\cdot(\mathbf{x} - \mathbf{x}')} \equiv \frac{1}{V} \sum_{\mathbf{q}} v(\mathbf{q}) e^{i\mathbf{q}\cdot(\mathbf{x} - \mathbf{x}')} \tag{2.23}$$

where $v(\mathbf{q}) = e^2/|\mathbf{q}|^2$. The final result for the interaction part of the Hamiltonian is

$$\hat{V} = \frac{1}{2} \sum_{\mathbf{q}} v(\mathbf{q}) \hat{\rho}(\mathbf{q}) \hat{\rho}(-\mathbf{q}). \tag{2.24}$$

It is derived in the box below. This result is a natural one. The Coulomb interaction, actually any interaction in a translationally invariant system, can scatter an electron from state \mathbf{k} to a state $\mathbf{k} + \mathbf{q}$ (this is the effect of the operator $\hat{\rho}(\mathbf{q})$) if another electron is scattered from state \mathbf{k}' to a state $\mathbf{k}' - \mathbf{q}$ (the effect of the operator $\hat{\rho}(-\mathbf{q})$). This conserves total momentum. The matrix element for this scattering is $v(\mathbf{q})$. It is almost more important to think of things this way than to worry about the full derivation for the particular interaction, in this case the Coulomb interaction. Any interaction which is translationally invariant (shifting all particles by the same amount in real space maps the system to itself) will have the same structure for \hat{V} with details about the nature of the interaction encoded in $v(\mathbf{q})$.

Added for Interest Being able to derive 2.24 is not necessary for what we want to do and is included here as additional information. The result follows after inserting the expression

for the density operators at \mathbf{x} and \mathbf{x}' , see 2.20 (set $e^2/4\pi\epsilon_0 = 1$ for the time-being):

$$\begin{aligned}\hat{V} &= \frac{1}{2} \int d^3\mathbf{x} \int d^3\mathbf{x}' \frac{\hat{\rho}(\mathbf{x})\hat{\rho}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} = \frac{1}{2} \int d^3\mathbf{x} \int d^3\mathbf{x}' \frac{\hat{\psi}^\dagger(\mathbf{x})\hat{\psi}(\mathbf{x})\hat{\psi}^\dagger(\mathbf{x}')\hat{\psi}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} \\ &= \frac{1}{2} \int d^3\mathbf{x} \int d^3\mathbf{x}' \frac{1}{V} \sum_{\mathbf{q}} v(\mathbf{q}) \hat{\rho}(\mathbf{x})\hat{\rho}(\mathbf{x}') e^{i\mathbf{q}\cdot(\mathbf{x}-\mathbf{x}')} \\ &= \frac{1}{2} \sum_{\mathbf{q}} v(\mathbf{q}) \hat{\rho}(\mathbf{q})\hat{\rho}(-\mathbf{q}).\end{aligned}\tag{2.25}$$

The result is the one we wrote down in 2.24. The factor of $1/2$ is there to avoid double-counting of the interaction. The (missing) factor of $e^2/4\pi\epsilon_0$ is usually absorbed into $v(\mathbf{q})$.

There is one other comment about pesky normalisations used here. The form quoted for $v(\mathbf{q})$ in 2.24 is the form correct in infinite systems, not for finite systems. You can see that this is dimensionally correct. The Coulomb interaction varies like $1/r$ ie as inverse length. The integral of $1/q^2$ over 3D q -space also varies with length -1 . However the sum over discrete wavevectors of $1/q^2$ will vary with length as r^2 . This is brought down to r^{-1} , by the factor $1/V$ added in to 2.23.

In later chapters will be looking at interacting Fermi gases in the non-relativistic limit, *i.e.* based on the Schrodinger equation.

2.2 Relativistic QFT for fermions

Following 2.13 we should expect the Hamiltonian for Dirac fermions to be something like

$$\hat{H} = \sum_{\sigma=1,2} \sum_{\mathbf{p}} p_0 \left(\hat{a}_{\sigma,+}^\dagger(\mathbf{p}) \hat{a}_{\sigma,+}(\mathbf{p}) - \hat{a}_{\sigma,-}^\dagger(\mathbf{p}) \hat{a}_{\sigma,-}(\mathbf{p}) \right).\tag{2.26}$$

The operator $\hat{a}_{\sigma,s}(\mathbf{p})$ annihilates a fermion in a state with wavevector \mathbf{p} , with “spin” σ . The label, $s = \pm$, denotes whether the state has positive energy $p_0 = \sqrt{m^2 + \mathbf{p}^2}$ or negative energy $p_0 = -\sqrt{m^2 + \mathbf{p}^2}$. For the time-being, assume that the sum over \mathbf{p} is a sum over discrete momenta.

Dirac suggested that all the negative energy states were occupied in the ground state and defined a new vacuum state $|\text{vac}\rangle$ to be

$$|\text{vac}\rangle = \prod_{\sigma,\mathbf{p}} \hat{a}_{\sigma,-}^\dagger(\mathbf{p}) |0\rangle.\tag{2.27}$$

This recipe would not work for bosons. Essentially the equation itself “knows” that it can describe fermions and not bosons—the corresponding QFT is only possible using anticommutation relations between the creation and annihilation operators. (Trying to use commutation relations quickly leads to contradictions, see [Fradkin](#) for more on this.)

It is also canonical to introduce separate operators for the positive and negative energy states as follows:

$$\hat{b}_\sigma(\mathbf{p}) = \hat{a}_{\sigma,+}(\mathbf{p}) \quad \text{and} \quad \hat{d}_\sigma(\mathbf{p}) = \hat{a}_{\sigma,-}^\dagger(\mathbf{p}).\tag{2.28}$$

The corresponding operators, $\hat{b}_\sigma^\dagger(\mathbf{p})$, create particles (electrons for example) while the $\hat{d}_\sigma^\dagger(\mathbf{p})$ create antiparticles (positrons for example). The vacuum state for the QFT is the one satisfying

$$\hat{b}_\sigma(\mathbf{p})|\text{vac}\rangle = 0 = \hat{d}_\sigma(\mathbf{p})|\text{vac}\rangle \quad \forall \mathbf{p}, \forall \sigma. \quad (2.29)$$

This can be read as stating that the vacuum state of the QFT is the one containing no particles and no antiparticles.

The Hamiltonian 2.26 is

$$\begin{aligned} \hat{H} &= \sum_{\sigma=1,2} \sum_{\mathbf{p}} p_0 \left(\hat{b}_\sigma^\dagger(\mathbf{p}) \hat{b}_\sigma(\mathbf{p}) - \hat{d}_\sigma(\mathbf{p}) \hat{d}_\sigma^\dagger(\mathbf{p}) \right) \\ &= \sum_{\sigma=1,2} \sum_{\mathbf{p}} p_0 \left(\hat{b}_\sigma^\dagger(\mathbf{p}) \hat{b}_\sigma(\mathbf{p}) + \hat{d}_\sigma^\dagger(\mathbf{p}) \hat{d}_\sigma(\mathbf{p}) - 1 \right), \end{aligned} \quad (2.30)$$

where we have assumed the standard anticommutator 2.6 for fermions described by discrete labels (in this case $\lambda = \{\mathbf{p}, \sigma\}$). Note that the zero of energy is now negative. When we complete the sum over \mathbf{p} , it adds another infinity in to the mix.

When we define field operators as integrals over wavevectors \mathbf{p} , we need to be careful about the normalisation used—as we had to be when defining field operators for bosons, see for example B.18. We will take the Hamiltonian to be given by

$$\hat{H} = \int \frac{d^3p}{(2\pi)^3} \frac{m}{p_0} \sum_{\sigma=1,2} p_0 \left(\hat{b}_\sigma^\dagger(\mathbf{p}) \hat{b}_\sigma(\mathbf{p}) - \hat{d}_\sigma(\mathbf{p}) \hat{d}_\sigma^\dagger(\mathbf{p}) \right) \quad (2.31)$$

with the following anticommutation relations

$$\begin{aligned} \{\hat{b}_\sigma(\mathbf{p}), \hat{b}_{\sigma'}(\mathbf{p}')\} &= 0 = \{\hat{d}_\sigma(\mathbf{p}), \hat{d}_{\sigma'}(\mathbf{p}')\} \\ \{\hat{b}_\sigma(\mathbf{p}), \hat{d}_{\sigma'}(\mathbf{p}')\} &= 0 = \{\hat{b}_\sigma(\mathbf{p}), \hat{d}_{\sigma'}^\dagger(\mathbf{p}')\} \\ \{\hat{b}_\sigma(\mathbf{p}), \hat{b}_{\sigma'}^\dagger(\mathbf{p}')\} &= (2\pi)^3 \frac{p_0}{m} \delta^{(3)}(\mathbf{p} - \mathbf{p}') \delta_{\sigma\sigma'} = \{\hat{d}_\sigma(\mathbf{p}), \hat{d}_{\sigma'}^\dagger(\mathbf{p}')\}. \end{aligned} \quad (2.32)$$

The field operators, $\hat{\psi}_\alpha(x)$, carry an index α to denote the component of the Dirac spinors, $u^{(\sigma)}$ or $v^{(\sigma)}$, which are the wavefunctions you studied in the first half of the module. They are defined in the Heisenberg representation, so include time-dependence, with $p \cdot x = p_0 x_0 - \mathbf{p} \cdot \mathbf{x}$, as follows

$$\hat{\psi}_\alpha(x) = \int \frac{d^3p}{(2\pi)^3} \frac{m}{p_0} \sum_{\sigma=1,2} \left[\hat{b}_\sigma(\mathbf{p}) u_\alpha^{(\sigma)}(\mathbf{p}) e^{-ip \cdot x} + \hat{d}_\sigma^\dagger(\mathbf{p}) v_\alpha^{(\sigma)}(\mathbf{p}) e^{+ip \cdot x} \right]. \quad (2.33)$$

Here the spinors are given by

$$u^{(\sigma)}(\mathbf{p}) = \begin{pmatrix} \sqrt{\frac{p_0+m}{2m}} \phi^{(\sigma)} \\ \frac{\boldsymbol{\sigma} \cdot \mathbf{p}}{\sqrt{2m(p_0+m)}} \phi^{(\sigma)} \end{pmatrix} \quad \text{and} \quad v^{(\sigma)}(\mathbf{p}) = \begin{pmatrix} \frac{\boldsymbol{\sigma} \cdot \mathbf{p}}{\sqrt{2m(p_0+m)}} \phi^{(\sigma)} \\ \sqrt{\frac{p_0+m}{2m}} \phi^{(\sigma)} \end{pmatrix}, \quad (2.34)$$

with $\phi^{(1)} = (1, 0)^T$ and $\phi^{(2)} = (0, 1)^T$. The equal-time anticommutation relations for the creation and annihilation operators are

$$\begin{aligned} \{\hat{\psi}_\alpha(\mathbf{x}), \hat{\psi}_\beta(\mathbf{x}')\} &= 0 = \{\hat{\psi}_\alpha^\dagger(\mathbf{x}), \hat{\psi}_\beta^\dagger(\mathbf{x}')\} \\ \{\hat{\psi}_\alpha(\mathbf{x}), \hat{\psi}_\beta^\dagger(\mathbf{x}')\} &= \delta_{\alpha\beta} \delta^{(3)}(\mathbf{x} - \mathbf{x}'). \end{aligned} \quad (2.35)$$

Normally, setting up the Dirac QFT starts by quantising field operators $\hat{\psi}_\alpha$ by introducing the anticommutators in 2.35. It takes quite a long time to get to the Hamiltonian 2.31. At some point you should probably work through these derivations but it is not the main point of this part of the module. I have been following [Eduardo Fradkin](#), but there are plenty of others. Not everyone includes the factor m in the Hamiltonian (2.31) and sum over modes (2.33). Whatever normalisation you use for the mode operators and spinors, $\hat{b}_\sigma(\mathbf{p})$ and $u^{(\sigma)}(\mathbf{p})$, the correct anticommutation relations for the field operators 2.35 must hold.

The effect of the field operators can be thought of as follows. The operator $\hat{\psi}_\alpha^\dagger(\mathbf{x})$ creates a fermion of “type” α at the point x in space-time. α refers to the component of the Dirac spinor. The conjugate operator annihilates a fermion. As in the case of the bosonic KG field, the product $\hat{\psi}_\alpha^\dagger(\mathbf{x})\hat{\psi}_\alpha(\mathbf{x})$ is a number density.

2.3 Normal ordering

One thing we have mostly been doing, without drawing attention to it, is to put the operators in the Hamiltonian into what is called normal order. The idea is to commute (for bosonic fields) or anticommute (for fermionic fields) all annihilation operators to the right. An annihilation operator acting on the ground state or vacuum state will annihilate the state. For example in the case of B.3 we converted $\hat{a}_\mathbf{p}\hat{a}_\mathbf{p}^\dagger \rightarrow \hat{a}_\mathbf{p}^\dagger\hat{a}_\mathbf{p} + [\hat{a}_\mathbf{p}, \hat{a}_\mathbf{p}^\dagger]$. We did something similar for 2.30 using the anticommutator $\{\hat{d}_\sigma(\mathbf{p}), \hat{d}_\sigma^\dagger(\mathbf{p})\}$.

It is also common to separate any constants from the operator. For example, the KG Hamiltonian in B.3 might be written

$$\begin{aligned}\hat{H} &= :\hat{H}: + E_0 \quad \text{where} \\ :\hat{H}: &= \int \frac{d^3p}{(2\pi)^3} \hbar\omega_{\mathbf{p}} \hat{a}_\mathbf{p}^\dagger \hat{a}_\mathbf{p}.\end{aligned}\tag{2.36}$$

We can write the Hamiltonian for the Dirac QFT 2.31 in normal ordered form as

$$\begin{aligned}\hat{H} &= :\hat{H}: + E_0 \quad \text{where} \\ :\hat{H}: &= \int \frac{d^3p}{(2\pi)^3} \frac{m}{p_0} \sum_{\sigma=1,2} p_0 \left(\hat{b}_\sigma^\dagger(\mathbf{p}) \hat{b}_\sigma(\mathbf{p}) + \hat{d}_\sigma^\dagger(\mathbf{p}) \hat{d}_\sigma(\mathbf{p}) \right).\end{aligned}\tag{2.37}$$

The energy E_0 here is negative. Using the anticommutator we find

$$E_0 = -2V \int d^3p \sqrt{p^2 + m^2}.\tag{2.38}$$

The factor V is the volume of the system and is associated with the delta function

$$\delta(0) = \lim_{p \rightarrow 0} \delta^3(0) = \lim_{p \rightarrow 0} \int \frac{d^3x}{(2\pi)^3} e^{i\mathbf{p}\cdot\mathbf{x}} = \frac{V}{(2\pi)^3}.\tag{2.39}$$

The volume is assumed infinite, but is needed here to ensure that we keep track of the various infinities. Interest in this particular infinity is rare and we will not consider these issues further.

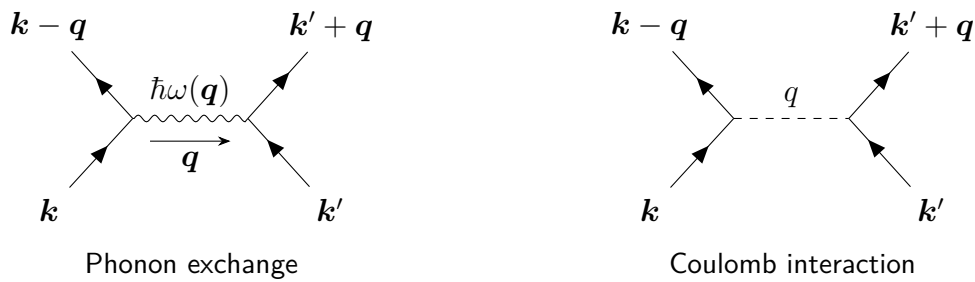
In almost all calculations, we work with the normal ordered operators, as we are interested in differences with respect to the ground state. Normal ordering is not quite as simple as commuting or anticommuting annihilation operators to the right. We need to know which operators annihilate the ground state (also called the vacuum state) and for this we need to know this state. The two questions are coupled. A simple case of this was that for Dirac QFT, where the operators annihilating the vacuum state are the $\hat{b}_\sigma(\mathbf{p})$ and the $\hat{d}_\sigma(\mathbf{p})$ - not the $\hat{a}_{s,\sigma}(\mathbf{p})$, see 2.28. Another case is the BCS theory of superconductivity which we look at next.

2.4 Interactions

You may be asking where are all the propagators and Feynman diagrams. Of course, we can define the propagators (called Green's functions in some branches of physics). However, the fields for which we have defined creation and annihilation operators are called free fields. There are no interactions yet. In essence there is nothing left to find, as we know the full spectrum of all modes and these do not interact with each other.

The formalism of Feynman diagrams, which you may have met in the context of the Standard Model, is needed to compute scattering amplitudes for various interactions. These interactions are the result of coupling between the matter fields, leptons or quarks for example, and bosonic fields such as photons.

We will look at the particular case of interactions in many-particle systems in condensed matter. These are the result of the exchange of particles, such as photons or phonons, between electrons. In the non-relativistic limit, the exchange of photons leads to the Coulomb interaction mentioned above. As the speed of light is so much larger than electron velocities, the Coulomb interaction can be treated as instantaneous without any need to treat the photons explicitly. In diagrammatic terms the phonon and photon exchange would look as below.



In the case of the phonon exchange, an electron with momentum \mathbf{k} emits a phonon (lattice oscillation) with wavevector \mathbf{q} and is scattered into a state with wavevector $\mathbf{k} - \mathbf{q}$. The phonon is absorbed by another electron with wavevector \mathbf{k}' . This exchange is behind the phenomenon of superconductivity which we consider in the next chapter (see also Appendix C).

Chapter 3

Interacting Electrons—Superconductivity

We will look at a system of interacting electrons. While solving for the eigenfunctions of the Hamiltonians of many-particle systems is not usually possible, many of the approximate treatments and some exact results for many-body systems are usually established using QFT. We will not explore all of this. Instead we will illustrate the power of QFT by looking at one of its greatest triumphs within condensed matter physics, namely the microscopic theory of superconductivity. QFT will allow us to work with a system which does not have a well-defined particle number and, rather surprisingly, this will lead to a relatively simple problem to solve.

QFT is not the answer to all problems. As just mentioned, it is not possible to construct many-body wavefunctions to describe the electronic structure of materials. However, it is possible to work only with the density of the electrons. Finding the ground state density can be shown to be equivalent to finding the eigenstates of an effective one-body Hamiltonian and not the many-body wavefunction. This is the approach of density functional theory (DFT). The many-body effects determine the effective one-body potential, which is of course not known in general (we would need the many-body wavefunction to find it). It can be approximated at each point by that for the homogeneous electron gas with the same density. Corrections, consistent with various sum rules, can be added empirically to account for variation of the density with position. The approach is amazingly successful and is widely used across much of computational materials science.

3.1 Many-body systems

The interaction between electrons in a metal is the Coulomb interaction. We can estimate orders of magnitude as follows. For a free electron gas with a number density n , we expect the average separation of electrons to be $r_s = \sqrt[3]{3/4\pi n}$. (The quantity r_s is the radius of a sphere containing one electron on average.) Normally, when discussing materials, we measure lengths in units of the Bohr radius $a = 4\pi\epsilon_0\hbar^2/me^2$. This is the length scale of electrons in an atom and expected to be the natural length scale for the outermost electrons in an atomic orbital and in a conduction or valence band.

The average kinetic energy of an electron in a free electron gas in 3D (using $k_F = (3\pi^2 n)^{1/3}$)

$$\begin{aligned}
 E_K &= \frac{3}{5} E_F = \frac{3}{5} \frac{\hbar^2 k_F^2}{2m} = \frac{3}{5} \frac{\hbar^2}{2m} (3\pi^2 n)^{2/3} \\
 &= \frac{3}{5} \frac{\hbar^2}{2m} \left(\frac{9\pi}{4} \right)^{2/3} \frac{1}{r_s^2} \\
 &\approx 1.1 \frac{\hbar^2}{m r_s^2}.
 \end{aligned} \tag{3.1}$$

We could have guessed the form of this answer—the Fermi wavenumber k_F has the dimensions of inverse length. The only length scale in a free electron gas is r_s , so the kinetic energy must be $\sim \hbar^2/(m r_s^2)$.

The average Coulomb interaction energy per particle should be close to $V_C = \frac{1}{2} e^2 / 4\pi\epsilon_0 r_s$. This is the interaction energy of two electrons a distance r_s apart. The factor of $1/2$ is there to avoid double counting (half the energy can be thought of as “belonging” to one and half to the other particle). The ratio of the potential energy to the kinetic energy is

$$\frac{V_C}{E_K} \approx \frac{1}{2} \left(\frac{1}{1.1} \frac{e^2}{4\pi\epsilon_0 r_s} \frac{m r_s^2}{\hbar^2} \right) \approx 0.45 \frac{r_s}{a}. \tag{3.2}$$

When estimating r_s for typical metals we should take the density of conduction band electrons, as these are the ones modelled by a free electron gas. For most metals r_s/a takes values between 1 and 6. This is always quite shocking the first time you meet it. We are modelling the metal as a free electron gas hoping that the interactions between electrons can be treated perturbatively afterwards somehow. Yet in almost all cases the typical interaction energy is comparable to (or even larger than) the typical kinetic energy. So how can we get away with this approach? A lot of work has gone into this question. Most studies have used the formalism of QFT, which makes it quite technical and long, and we will not cover it in this module. The gist of the argument is quite simple though. It is that, in a metal, the electron sea acting collectively screens the Coulomb interaction between individual particles so that its effect is less than the order of magnitude estimate above suggests it might be. There is also the ultimate argument: the method certainly works so let's get calculating.

For the non-relativistic case that we are considering 3.2 shows that interactions are dominant ($V/E_K \gg 1$) in the low density limit $r_s/a \gg 1$. In the high density limit the kinetic energy dominates. This is because the interaction energy scales as $1/r_s$ while the kinetic energy scales as $1/r_s^2$. Even though the interaction energies grow at high densities, the kinetic energy grows faster. Because the Coulomb energy does not contain \hbar , the low density behaviour and the ground state should be more classical, and indeed they are. At very low densities the free electron gas is expected to form a solid and its excitations are oscillations of the electrons about their equilibrium positions. This state is called the Wigner crystal.

3.2 Cooper instability

Superconductivity was discovered in 1911 and the microscopic theory, BCS theory, appeared in 1957. A microscopic theory took a long time coming! The idea that a microscopic theory

was possible might seem far-fetched in the first instance given that the transition temperatures were of the order of Kelvin, while the interaction energies were of the order of 10 eV or 110,000 K. How can one pick out an interaction effect “worth” 10K when overall interaction effects are 10^4 times bigger in general?

The insight that made a microscopic theory possible was that of the Cooper instability. It was known that there might be a residual attractive interaction between electrons due to phonon exchange. Experiments had shown that phonons were implicated somehow. Materials, containing different amounts of different isotopes for at least some of the ions, had different transition temperatures (the transition temperature varied as $M^{-1/2}$ where M was the mass of the ion). However, estimates showed that this interaction was weak and it was known that weak attractive interactions could not bind free electrons in 3D. Cooper realised that it was essential to take account of the presence of the filled Fermi sea (FS). Even if the direct interaction involving the electrons in the FS were not taken into account, the filled states do affect all other electrons (in this case the additional pair of electrons just outside the FS) by “blocking” the one-particle states.

In the free electron model, the orbitals for the free electrons will be $\psi_{\mathbf{k}}(\mathbf{r}) = V^{-1/2}e^{i\mathbf{k}\cdot\mathbf{r}}$, where V is the volume of the system and \mathbf{k} denotes a wavevector. Each electron will have a spin component $|\sigma\rangle = |\uparrow\rangle$ or $|\downarrow\rangle$. For electrons close to the Fermi energy, their energy is linear in the wavenumber difference with respect to the Fermi momentum, $\epsilon_{\mathbf{k}} = \epsilon_F + v_F\hbar(k - k_F)$. Here v_F is the Fermi velocity, $v = \partial\omega/\partial k$, and k_F is the Fermi wavenumber.

A bound state means a state with negative energy with respect to the filled Fermi sea. It will be expected to have zero momentum and total spin zero. (Zero momentum seems intuitive. Add momentum and you are likely to add kinetic energy to the system. Triplet spin pairing is also possible but, for the model interaction we will look at, does not give the lowest energy.) Total spin zero means the two electrons are in a spin-singlet state. This suggests looking for a two-particle state

$$\begin{aligned}\psi(\mathbf{r}_1, \mathbf{r}_2) &= \sum_{|\mathbf{k}| > k_F} \alpha_{\mathbf{k}} \psi_{\mathbf{k}}(\mathbf{r}_1) \psi_{-\mathbf{k}}(\mathbf{r}_2) \left(\frac{|\uparrow_1\rangle|\downarrow_2\rangle - |\downarrow_1\rangle|\uparrow_2\rangle}{\sqrt{2}} \right) \\ &= \frac{1}{V} \sum_{|\mathbf{k}| > k_F} \alpha_{\mathbf{k}} e^{i\mathbf{k}\cdot(\mathbf{r}_1 - \mathbf{r}_2)} |0\rangle_S \equiv \left(\sum_{|\mathbf{k}| > k_F} \alpha_{\mathbf{k}} \psi_{\mathbf{k}}^{(2)} \right) |0\rangle_S.\end{aligned}\quad (3.3)$$

We have denoted the spin singlet combination of the two spins as $|0\rangle_S$ (the idea being $S = S_1 + S_2 = 0$). We have also written the product of the two one-particle states, $\psi_{\mathbf{k}}^{(2)} = \psi_{\mathbf{k}}(\mathbf{r}_1) \psi_{-\mathbf{k}}(\mathbf{r}_2)$, as a convenient shorthand. The total wavefunction must be antisymmetric under interchange of the two particles. The spin singlet wavefunction is antisymmetric under interchanges of the particles, so that the spatial wavefunction, $\psi_{\mathbf{k}}^{(2)}$, must be symmetric in the spatial variables. This means $\alpha_{\mathbf{k}} = \alpha_{-\mathbf{k}}$. (See the discussion around 2.8 and 2.11.) The $\alpha_{\mathbf{k}}$ are coefficients (numbers not operators) and are what we need to find once we have specified the model Hamiltonian.

The model Hamiltonian, with a two-body interaction $U = U(\mathbf{r}_1 - \mathbf{r}_2)$, is

$$H = H_0 + U = H_0 + \sum_{\mathbf{q}} U_{\mathbf{q}} e^{i\mathbf{q}\cdot(\mathbf{r}_1 - \mathbf{r}_2)}, \quad (3.4)$$

where we have introduced a Fourier decomposition of the interaction potential. H_0 is the free electron part. A two-body state, $\psi_{\mathbf{k}'}^{(2)}$, is an eigenstate of H_0 with energy $2\epsilon_{\mathbf{k}'}$. This is because the spatial part involves a product of one-particle states with the same $|\mathbf{k}'|$. The effect of U on our two-body wavefunction (we will write only the spatial part as the interaction does not affect spin variables) is

$$\begin{aligned} U\psi_{\mathbf{k}'}^{(2)} &= \sum_{\mathbf{q}} U_{\mathbf{q}} e^{i\mathbf{q}\cdot(\mathbf{r}_1-\mathbf{r}_2)} \frac{1}{V} \alpha_{\mathbf{k}'} e^{i\mathbf{k}'\cdot(\mathbf{r}_1-\mathbf{r}_2)} \\ &= \sum_{\mathbf{q}} U_{\mathbf{q}} \psi_{\mathbf{k}'+\mathbf{q}}^{(2)} = \sum_{\mathbf{k}''} U_{\mathbf{k}''-\mathbf{k}'} \psi_{\mathbf{k}''}^{(2)}. \end{aligned} \quad (3.5)$$

This result says that, if we start with two electrons with wavevectors, \mathbf{k} and $-\mathbf{k}$, and hence zero net momentum, the interaction can scatter the electrons into another state with zero momentum. We denote the new wavevectors by \mathbf{k}'' and $-\mathbf{k}''$. The amplitude is given by the strength of the Fourier component of the interaction with wavevector, $\mathbf{q} = \mathbf{k}'' - \mathbf{k}'$. The argument also works for antiparticles or holes in this context. We could add two holes (remove two electrons). If a pair of electrons bind so would the pair of holes.

Without worrying quite how it got there (its origin is outlined in Appendix C), we will assume that the two-body interaction is an attractive interaction in an energy window close to the Fermi energy. As this interaction is known to be related to phonon-exchange, we will take the energy window to be set by $\hbar\omega_D$, where ω_D is taken as the highest frequency phonon in the system, usually called the Debye frequency. We will put

$$U_{\mathbf{k}'-\mathbf{k}''} = \begin{cases} -\frac{U}{V} & \text{if } |\epsilon_{\mathbf{k}'} - \epsilon_F| < \hbar\omega_D, |\epsilon_{\mathbf{k}''} - \epsilon_F| < \hbar\omega_D \\ 0 & \text{otherwise.} \end{cases} \quad (3.6)$$

Here $U > 0$. The factor of V (system volume) is associated with the Fourier series representation in 3.5. The Fourier amplitude $U_{\mathbf{q}} = \frac{1}{V} \int d^3\mathbf{r} U(\mathbf{r}) e^{-i\mathbf{q}\cdot\mathbf{r}}$ (making $U_{\mathbf{q}}$ an energy). This is an attractive interaction between two electrons that are in states within $\hbar\omega_D$ of the Fermi energy.

We are looking for an eigenstate of the Hamiltonian 3.4 with two more electrons than the filled Fermi sea, which has energy less than zero if the energies of the electrons are measured with respect to the Fermi energy. If there is such a bound state, it will tell us that the filled Fermi sea is not the ground state for our free electron gas—if two particles can bind then two more from within the Fermi sea can also bind and the energy will go down and so on.

An eigenstate will be defined by the choice of the $\alpha_{\mathbf{k}}$ in 3.3 satisfying

$$E \sum_{|\mathbf{k}'| > k_F} \alpha_{\mathbf{k}'} \psi_{\mathbf{k}'}^{(2)} = H \sum_{|\mathbf{k}'| > k_F} \alpha_{\mathbf{k}'} \psi_{\mathbf{k}'}^{(2)} \quad (3.7)$$

$$= \sum_{|\mathbf{k}'| > k_F} 2\epsilon_{\mathbf{k}'} \alpha_{\mathbf{k}'} \psi_{\mathbf{k}'}^{(2)} - \sum_{|\mathbf{k}'| > k_F} \sum_{\mathbf{k}''} \alpha_{\mathbf{k}'} \frac{U}{V} \psi_{\mathbf{k}''}^{(2)}. \quad (3.8)$$

In the sum on \mathbf{k}'' , we need $\epsilon_F < \epsilon_{\mathbf{k}''} < \epsilon_F + \hbar\omega_D$.

The $\psi_{\mathbf{k}}^{(2)}$ with different values of \mathbf{k} are orthogonal. Multiplying by $(\psi_{\mathbf{k}}^{(2)})^*$ and integrating

over both \mathbf{r}_1 and \mathbf{r}_2 gives

$$E \alpha_{\mathbf{k}} = 2\epsilon_{\mathbf{k}} \alpha_{\mathbf{k}} - \frac{U}{V} \sum_{\mathbf{k}''} \alpha_{\mathbf{k}''} \equiv 2\epsilon_{\mathbf{k}} \alpha_{\mathbf{k}} - \frac{U}{V} I, \quad \text{where } I = \sum_{\mathbf{k}''} \alpha_{\mathbf{k}''},$$

hence $\alpha_{\mathbf{k}} = -\frac{U}{V} \frac{I}{E - 2\epsilon_{\mathbf{k}}}.$ (3.9)

A bound state would mean $E = 2\epsilon_{\mathbf{F}} - \Delta$ with $\Delta > 0$ (or equivalently $E - 2\epsilon_{\mathbf{F}} < 0$). For weak attractive interactions, we can take the binding energy $\Delta \ll \hbar\omega_D$. Summing $\sum_{\mathbf{k}} \alpha_{\mathbf{k}} = I$, and introducing the density of states (DOS) per unit volume at the Fermi energy, $\nu_{\mathbf{F}}$, leads to a condition determining E :

$$\begin{aligned} 1 &= -\frac{U}{V} \sum_{\mathbf{k}''} \frac{1}{E - 2\epsilon_{\mathbf{k}''}} = \frac{U}{V} \int_{\epsilon_{\mathbf{F}}}^{\epsilon_{\mathbf{F}} + \hbar\omega_D} d\epsilon V \nu_{\mathbf{F}} \frac{1}{2\epsilon - E} \\ &= \frac{U\nu_{\mathbf{F}}}{2} \ln \left[\frac{2\epsilon_{\mathbf{F}} + \hbar\omega_D - E}{2\epsilon_{\mathbf{F}} - E} \right] = \frac{U\nu_{\mathbf{F}}}{2} \ln \left[\frac{\hbar\omega_D}{\Delta} \right]. \end{aligned} \quad (3.10)$$

The last expression follows as we are assuming $\Delta \ll \hbar\omega_D$. Note that the density of states ($V\nu_{\mathbf{F}}$) is extensive and proportional to volume while U is an energy times volume, see 3.6, which makes $U\nu_{\mathbf{F}}/2$ a number. Exponentiating 3.10 gives

$$\Delta = \hbar\omega_D e^{-2/U\nu_{\mathbf{F}}}. \quad (3.11)$$

The result 3.11 shows that, no matter how small the interaction is ($U\nu_{\mathbf{F}} \ll 1$), the energy $\Delta > 0$. A pair of electrons at the Fermi energy will always bind. The same calculation also applies to holes at the Fermi energy. In the presence of an arbitrarily small attractive interaction, the free Fermi gas is always unstable. Pairs of electrons and pairs of holes will bind and condense to form a new state. To describe the new state is what the BCS theory (named for Bardeen, Cooper and Schrieffer) attempts to do.

A comment on interactions in the free electron model. The Cooper instability (and others) are a warning about treating interacting electrons perturbatively. Models based on the free electron theory have been a hugely successful approach, but it is difficult to explain why a perturbation theoretic approach works. A perturbation theory normally implies that we are expanding about a non-interacting ($U = 0$) limit. We hope to find that an expansion converges in powers of the parameter $U/E_{\mathbf{F}}$. The result 3.11 shows that this is impossible. If the sign of U is negative (attractive interaction), however small in absolute strength, the free electron gas is unstable.

3.3 Mean Field Theory

In problems involving interacting fields, which in general cannot be solved exactly, we often approximate fields (these can be classical or quantum) by their average value. Sometimes, one particular expectation value can change from zero in one phase to non-zero in another phase, and acts as an order parameter for the second phase. The theory works explicitly with the dependence of this expectation value on parameters of the theory and quantities like

temperature and any applied fields. In quantum theory, operators are replaced by complex number (often referred to as a c-number) expectation values.

In the case below the product of (quantised) operators is of the form

$$\hat{c}_{\mathbf{k},\uparrow}^\dagger \hat{c}_{-\mathbf{k},\downarrow}^\dagger \hat{c}_{-\mathbf{k}',\downarrow} \hat{c}_{\mathbf{k}',\uparrow}. \quad (3.12)$$

The expectation values $\langle \hat{c}_{-\mathbf{k},\downarrow} \hat{c}_{\mathbf{k},\uparrow} \rangle = b_{\mathbf{k}}^*$ and $\langle \hat{c}_{\mathbf{k},\uparrow}^\dagger \hat{c}_{-\mathbf{k},\downarrow}^\dagger \rangle = b_{\mathbf{k}}$ will be assumed to be non-zero, ie $b_{\mathbf{k}} \neq 0$. This is an Ansatz or working hypothesis. We will need to check whether this non-zero expectation value is reproduced when we solve for the properties of the system within this approximation. This is called self-consistency. We will find that below a temperature, which we call T_c , this non-zero expectation value exists, and we will identify T_c with the superconducting ordering temperature.

The operator product is written in terms of the deviation of an operator from its expectation value

$$\begin{aligned} \hat{c}_{\mathbf{k},\uparrow}^\dagger \hat{c}_{-\mathbf{k},\downarrow}^\dagger \hat{c}_{-\mathbf{k}',\downarrow} \hat{c}_{\mathbf{k}',\uparrow} &= ((\hat{c}_{\mathbf{k},\uparrow}^\dagger \hat{c}_{-\mathbf{k},\downarrow}^\dagger - b_{\mathbf{k}}) + b_{\mathbf{k}}) ((\hat{c}_{-\mathbf{k}',\downarrow} \hat{c}_{\mathbf{k}',\uparrow} - b_{\mathbf{k}'}^*) + b_{\mathbf{k}'}^*) \\ &= (\hat{c}_{\mathbf{k},\uparrow}^\dagger \hat{c}_{-\mathbf{k},\downarrow}^\dagger - b_{\mathbf{k}})(\hat{c}_{-\mathbf{k}',\downarrow} \hat{c}_{\mathbf{k}',\uparrow} - b_{\mathbf{k}'}^*) \\ &\quad + (\hat{c}_{\mathbf{k},\uparrow}^\dagger \hat{c}_{-\mathbf{k},\downarrow}^\dagger - b_{\mathbf{k}})b_{\mathbf{k}'}^* + (\hat{c}_{-\mathbf{k}',\downarrow} \hat{c}_{\mathbf{k}',\uparrow} - b_{\mathbf{k}'}^*)b_{\mathbf{k}} + b_{\mathbf{k}}b_{\mathbf{k}'}^* \\ &\approx b_{\mathbf{k}'}^* \hat{c}_{\mathbf{k},\uparrow}^\dagger \hat{c}_{-\mathbf{k},\downarrow}^\dagger + b_{\mathbf{k}} \hat{c}_{-\mathbf{k}',\downarrow} \hat{c}_{\mathbf{k}',\uparrow} - b_{\mathbf{k}}b_{\mathbf{k}'}^*. \end{aligned} \quad (3.13)$$

In the last line we have discarded the term involving the product of the deviations of the operators from their expectation values (the first term on the previous line). The validity of any mean field approximation like this depends on the effects of these terms being small compared to those of the terms we keep. Unless absolute energies are required, we would normally also drop the last term $-b_{\mathbf{k}}^*b_{\mathbf{k}'}$, as this is a number and does not affect the values of the quantum fields. In the simplest case, the expectation value $b_{\mathbf{k}}$ will not depend on \mathbf{k} but only on k (see below).

As you have probably spotted, there is something radical with this particular mean field. The expectation value of an operator product that annihilates two particles $\langle \hat{c}_{-\mathbf{k},\downarrow} \hat{c}_{\mathbf{k},\uparrow} \rangle$ is non-zero. If the expectation value is taken with respect to a ground state wavefunction for a system with fixed particle number, it cannot be non-zero. In practice, there are ways of making the theory work at fixed particle number but they do not involve any particular new physics. Working with systems with variable particle number makes the sums easier and is fine provided we impose a constraint on the average total particle number as we do in statistical mechanics when working in the Grand Canonical Ensemble. We introduce a chemical potential, μ , which multiplies the total number of particles, \hat{N} , and we look for eigenstates of $\hat{H} - \mu\hat{N}$.

3.4 BCS theory

Main sources here are [John Chalker](#) and [Subir Sachdev](#).

The BCS theory (named after its originators Bardeen, Cooper and Schrieffer) is one of the great triumphs of QFT applied to many-particle systems. It is a theory in which the attractive part of the interparticle interaction is treated using mean field theory.

We will work with a weakly attractive interaction in the grand canonical ensemble. We consider the Hamiltonian

$$\hat{H} - \mu\hat{N} = \sum_{\mathbf{k},\sigma} \xi_{\mathbf{k}} \hat{c}_{\mathbf{k},\sigma}^\dagger \hat{c}_{\mathbf{k},\sigma} - \frac{U}{2V} \sum_{\mathbf{k},\mathbf{k}',\mathbf{q}}' \sum_{\sigma,\sigma'} \hat{c}_{\mathbf{k}+\mathbf{q},\sigma}^\dagger \hat{c}_{\mathbf{k}'-\mathbf{q},\sigma'}^\dagger \hat{c}_{\mathbf{k}',\sigma'} \hat{c}_{\mathbf{k},\sigma}. \quad (3.14)$$

The prime on the second sum denotes a restriction to the states within $\hbar\omega_D$ of the Fermi energy. This is because the attractive interaction is due to phonon exchange and so acts only for energies within $\hbar\omega_D$ of the Fermi energy. The quantity $\xi_{\mathbf{k}} = \epsilon_{\mathbf{k}} - \mu$, where μ is the chemical potential. It subtracts the term, $\sum_{\mathbf{k},\sigma} \mu \hat{c}_{\mathbf{k},\sigma}^\dagger \hat{c}_{\mathbf{k},\sigma} = \mu\hat{N}$, from the one-particle energies. It is needed to fix the average total number of particles. The factor of 1/2 in front of the interaction is necessary to avoid double-counting the pairwise interaction (as in 3.2). The factor of V is associated with the Fourier transform of an interaction $U_{\mathbf{q}} \sim (1/V) \int d^3\mathbf{r} U(\mathbf{r}) e^{-i\mathbf{q}\cdot\mathbf{r}}$. We will see at the end that the important parameter is $U\nu_F$ where ν_F is the density of states per unit volume.

We can't solve for the eigenstates of the Hamiltonian in 3.14. Nobody can. Instead we will use the insight of the Cooper instability to construct an approximate theory. We will assume that singlet pairs of particles form with zero net momentum and only consider interactions involving these pairs. We will look out for terms involving these pairs in the Hamiltonian:

$$\hat{H} - \mu\hat{N} = \sum_{\mathbf{k},\sigma} \xi_{\mathbf{k}} \hat{c}_{\mathbf{k},\sigma}^\dagger \hat{c}_{\mathbf{k},\sigma} - \frac{U}{V} \sum_{\mathbf{k},\mathbf{k}'}' \hat{c}_{\mathbf{k},\uparrow}^\dagger \hat{c}_{-\mathbf{k},\downarrow}^\dagger \hat{c}_{-\mathbf{k}',\downarrow} \hat{c}_{\mathbf{k}',\uparrow}. \quad (3.15)$$

We are assuming that all other terms do not alter the state we are looking for and can be dropped. The rationale for this is roughly that the single-particle spectrum explains the metallic state above the transition well and we are guessing that only the interaction involving the pairs matters. The factor of 1/2 in front of the interaction term has been cancelled by the sum over σ (we assume $\sigma' = -\sigma$ and summing over σ gives two copies of the same result).

We will assume non-zero averages (see section on Mean Field Theory above):

$$b_{\mathbf{k}} = \langle \hat{c}_{\mathbf{k},\uparrow}^\dagger \hat{c}_{-\mathbf{k},\downarrow}^\dagger \rangle, \quad b_{\mathbf{k}'}^* = \langle \hat{c}_{-\mathbf{k}',\downarrow} \hat{c}_{\mathbf{k}',\uparrow} \rangle, \quad (3.16)$$

and look for a self-consistent non-zero solution for these averages. The brackets mean take the expectation value in the ground state, which we do not know yet, but which we will identify at the end. This is what we mean by a self-consistency condition. At non-zero temperature the brackets indicate the expectation value with respect to the state of the system in thermal equilibrium.

Introducing the quantity Δ , which we will later identify with the order parameter $\Delta = \frac{U}{V} \sum_{\mathbf{k}'}' b_{\mathbf{k}'}$, and neglecting any constants, the mean field Hamiltonian is

$$\begin{aligned} \hat{H} - \mu\hat{N} &= \sum_{\mathbf{k},\sigma} \xi_{\mathbf{k}} \hat{c}_{\mathbf{k},\sigma}^\dagger \hat{c}_{\mathbf{k},\sigma} - \frac{U}{V} \sum_{\mathbf{k},\mathbf{k}'}' \left(b_{\mathbf{k}'}^* \hat{c}_{\mathbf{k},\uparrow}^\dagger \hat{c}_{-\mathbf{k},\downarrow}^\dagger + b_{\mathbf{k}} \hat{c}_{-\mathbf{k}',\downarrow} \hat{c}_{\mathbf{k}',\uparrow} \right) \\ &= \sum_{\mathbf{k},\sigma} \xi_{\mathbf{k}} \hat{c}_{\mathbf{k},\sigma}^\dagger \hat{c}_{\mathbf{k},\sigma} - \sum_{\mathbf{k}}' \left(\Delta^* \hat{c}_{\mathbf{k},\uparrow}^\dagger \hat{c}_{-\mathbf{k},\downarrow}^\dagger + \Delta \hat{c}_{-\mathbf{k},\downarrow} \hat{c}_{\mathbf{k},\uparrow} \right). \end{aligned} \quad (3.17)$$

We will assume $\xi_{\mathbf{k}} = \xi_k$ and write the Hamiltonian in a matrix form as follows

$$\hat{H} - \mu\hat{N} = \sum_{\mathbf{k}} \hat{H}_{\mathbf{k}} \quad \text{where}$$

$$\hat{H}_{\mathbf{k}} = \frac{1}{2} \begin{pmatrix} \hat{c}_{\mathbf{k},\uparrow}^\dagger & \hat{c}_{-\mathbf{k},\downarrow} & \hat{c}_{-\mathbf{k},\downarrow}^\dagger & \hat{c}_{\mathbf{k},\uparrow} \end{pmatrix} \begin{pmatrix} \xi_k & -\Delta^* & 0 & 0 \\ -\Delta & -\xi_k & 0 & 0 \\ 0 & 0 & \xi_k & \Delta^* \\ 0 & 0 & \Delta & -\xi_k \end{pmatrix} \begin{pmatrix} \hat{c}_{\mathbf{k},\uparrow} \\ \hat{c}_{-\mathbf{k},\downarrow}^\dagger \\ \hat{c}_{-\mathbf{k},\downarrow} \\ \hat{c}_{\mathbf{k},\uparrow}^\dagger \end{pmatrix} + \xi_k. \quad (3.18)$$

The constant at the end comes from anticommuting the creation and annihilation operators.

If we can diagonalise the matrix, we should be home and dry. However, it will only be of use if the transformed operators satisfy fermionic anticommutation relations. We will need this to characterise the ground and excited states (otherwise we would have a diagonal quadratic form with non-standard anticommutation relations and we would have to solve another problem—namely what do they mean?). The matrix is block diagonal which allows us to diagonalise each 2×2 sub-matrix separately.

We want to write

$$\begin{pmatrix} \hat{\gamma}_{\mathbf{k},1} \\ \hat{\gamma}_{\mathbf{k},2}^\dagger \end{pmatrix} = U \begin{pmatrix} \hat{c}_{\mathbf{k},\uparrow} \\ \hat{c}_{-\mathbf{k},\downarrow}^\dagger \end{pmatrix} \equiv \begin{pmatrix} u_{\mathbf{k}} & -v_{\mathbf{k}} \\ v_{\mathbf{k}}^* & u_{\mathbf{k}}^* \end{pmatrix} \begin{pmatrix} \hat{c}_{\mathbf{k},\uparrow} \\ \hat{c}_{-\mathbf{k},\downarrow}^\dagger \end{pmatrix}$$

$$\begin{pmatrix} \hat{c}_{\mathbf{k},\uparrow} \\ \hat{c}_{-\mathbf{k},\downarrow}^\dagger \end{pmatrix} = U^\dagger \begin{pmatrix} \hat{\gamma}_{\mathbf{k},1} \\ \hat{\gamma}_{\mathbf{k},2}^\dagger \end{pmatrix} \equiv \begin{pmatrix} u_{\mathbf{k}}^* & v_{\mathbf{k}} \\ -v_{\mathbf{k}}^* & u_{\mathbf{k}} \end{pmatrix} \begin{pmatrix} \hat{\gamma}_{\mathbf{k},1} \\ \hat{\gamma}_{\mathbf{k},2}^\dagger \end{pmatrix}. \quad (3.19)$$

We should check that the new operators $\hat{\gamma}_{\mathbf{k},\tau}$ can be made to satisfy fermionic anticommutation relations. This requires that

$$1 = \{\hat{\gamma}_{\mathbf{k},1}, \hat{\gamma}_{\mathbf{k},1}^\dagger\} = \{(u_{\mathbf{k}}\hat{c}_{\mathbf{k},\uparrow} - v_{\mathbf{k}}\hat{c}_{-\mathbf{k},\downarrow}^\dagger), (u_{\mathbf{k}}^*\hat{c}_{\mathbf{k},\uparrow}^\dagger - v_{\mathbf{k}}^*\hat{c}_{-\mathbf{k},\downarrow})\} = |u_{\mathbf{k}}|^2 + |v_{\mathbf{k}}|^2, \quad (3.20)$$

which can be satisfied by choosing $u_{\mathbf{k}} = \cos \theta_{\mathbf{k}} e^{i\phi_{\mathbf{k}}/2}$, $v_{\mathbf{k}} = \sin \theta_{\mathbf{k}} e^{-i\phi_{\mathbf{k}}/2}$.

The contribution to $\hat{H}_{\mathbf{k}}$ of the upper 2×2 matrix can be written in terms of the $\hat{\gamma}$ operators as

$$\frac{1}{2} \begin{pmatrix} \hat{\gamma}_{\mathbf{k},1}^\dagger & \hat{\gamma}_{\mathbf{k},2} \end{pmatrix} \begin{pmatrix} u_{\mathbf{k}} & -v_{\mathbf{k}} \\ v_{\mathbf{k}}^* & u_{\mathbf{k}}^* \end{pmatrix} \begin{pmatrix} \xi_k & -|\Delta|e^{-i\phi} \\ -|\Delta|e^{i\phi} & -\xi_k \end{pmatrix} \begin{pmatrix} u_{\mathbf{k}}^* & v_{\mathbf{k}} \\ -v_{\mathbf{k}}^* & u_{\mathbf{k}} \end{pmatrix} \begin{pmatrix} \hat{\gamma}_{\mathbf{k},1} \\ \hat{\gamma}_{\mathbf{k},2}^\dagger \end{pmatrix}.$$

The transformed Hamiltonian matrix (ie the product of the three matrices) is diagonal if the off-diagonal elements vanish. This happens if $+2\xi_k u_{\mathbf{k}} v_{\mathbf{k}} - |\Delta|e^{-i\phi} u_{\mathbf{k}}^2 + |\Delta|e^{i\phi} v_{\mathbf{k}}^2 = 0$, which is satisfied by choosing

$$\sin 2\theta_{\mathbf{k}} = \frac{|\Delta|}{\sqrt{\xi_k^2 + |\Delta|^2}}, \quad \cos 2\theta_{\mathbf{k}} = \frac{\xi_k}{\sqrt{\xi_k^2 + |\Delta|^2}}, \quad \phi_{\mathbf{k}} = \phi. \quad (3.21)$$

The angle $\theta_{\mathbf{k}}$ is defined such that $\theta_{\mathbf{k}} \in [0, \pi)$.

The eigenvalues are $E_{\mathbf{k}} = \pm \sqrt{\xi_k^2 + |\Delta|^2}$ (as they are for the lower 2×2 matrix). The full Hamiltonian becomes

$$\hat{H}_{\mathbf{k}} = \frac{1}{2} \begin{pmatrix} \hat{\gamma}_{\mathbf{k},1}^\dagger & \hat{\gamma}_{\mathbf{k},2} & \hat{\gamma}_{\mathbf{k},2}^\dagger & \hat{\gamma}_{\mathbf{k},1} \end{pmatrix} \begin{pmatrix} E_{\mathbf{k}} & 0 & 0 & 0 \\ 0 & -E_{\mathbf{k}} & 0 & 0 \\ 0 & 0 & E_{\mathbf{k}} & 0 \\ 0 & 0 & 0 & -E_{\mathbf{k}} \end{pmatrix} \begin{pmatrix} \hat{\gamma}_{\mathbf{k},1} \\ \hat{\gamma}_{\mathbf{k},2}^\dagger \\ \hat{\gamma}_{\mathbf{k},2} \\ \hat{\gamma}_{\mathbf{k},1}^\dagger \end{pmatrix} + \xi_k. \quad (3.22)$$

Rewriting the result 3.22 in the form of 3.17 gives

$$\hat{H} - \mu \hat{N} = \sum_{\mathbf{k}, \tau} E_{\mathbf{k}} \hat{\gamma}_{\mathbf{k}, \tau}^{\dagger} \hat{\gamma}_{\mathbf{k}, \tau} + \sum_{\mathbf{k}} (\xi_{\mathbf{k}} - E_{\mathbf{k}}). \quad (3.23)$$

The extra constant term on the right hand side appears, as we have used the anticommutation relation to write $\hat{\gamma}_{\mathbf{k}, \tau}^{\dagger} \hat{\gamma}_{\mathbf{k}, \tau}^{\dagger} = 1 - \hat{\gamma}_{\mathbf{k}, \tau}^{\dagger} \hat{\gamma}_{\mathbf{k}, \tau}$.

The parameter Δ should be computed self-consistently from the mean-field Hamiltonian 3.23. (The Hamiltonian assumes a non-zero Δ . Self-consistency means that the value of Δ determined from the treatment of the mean-field Hamiltonian matches that assumed to start with.) If there is a self-consistent non-zero solution for Δ the new state may be the answer. We have

$$\Delta = \frac{U}{V} \sum_{\mathbf{k}} b_{\mathbf{k}} = \frac{U}{V} \sum_{\mathbf{k}} \langle \hat{c}_{\mathbf{k}, \uparrow}^{\dagger} \hat{c}_{-\mathbf{k}, \downarrow}^{\dagger} \rangle = \frac{U}{V} \sum_{\mathbf{k}} u_{\mathbf{k}} v_{\mathbf{k}}^* \left(1 - \langle \hat{\gamma}_{\mathbf{k}, 1}^{\dagger} \hat{\gamma}_{\mathbf{k}, 1} \rangle - \langle \hat{\gamma}_{\mathbf{k}, 2}^{\dagger} \hat{\gamma}_{\mathbf{k}, 2} \rangle \right), \quad (3.24)$$

where we have used 3.19. The phase of $u_{\mathbf{k}}^* v_{\mathbf{k}}$ gives the phase of Δ . In the zero temperature case $\langle \hat{\gamma}_{\mathbf{k}, \tau}^{\dagger} \hat{\gamma}_{\mathbf{k}, \tau} \rangle = 0$. We also have $u_{\mathbf{k}}^* v_{\mathbf{k}} = |\Delta| / (2\sqrt{\xi_{\mathbf{k}}^2 + |\Delta|^2})$ from 3.21. Writing the sum over wavevectors \mathbf{k} , by and integral over energy using the density of states ν_F gives

$$1 = \frac{\nu_F U}{2} \int_{-\hbar\omega_D}^{\hbar\omega_D} d\xi \frac{1}{\sqrt{\xi^2 + |\Delta|^2}} = U \nu_F \sinh^{-1} \frac{\hbar\omega_D}{|\Delta|}. \quad (3.25)$$

Here ν_F is the density of states per unit volume. In the limit of weak interactions $U \nu_F \ll 1$, this has the solution

$$|\Delta| = 2\hbar\omega_D e^{-1/U\nu_F}. \quad (3.26)$$

The result 3.26 explained the isotope effect mentioned above. The transition temperature of a material scaled with the isotopic mass as $M^{-1/2}$. This is the dependence of prefactor ω_D (the phonon frequencies vary as $M^{-1/2}$ as we saw in 1.18).

Equation 3.23 is a free fermion Hamiltonian with single particle energies $E_{\mathbf{k}}$. (We used this fact to compute $\langle \hat{c}_{\mathbf{k}, \uparrow}^{\dagger} \hat{c}_{-\mathbf{k}, \downarrow}^{\dagger} \rangle$, when we wrote these in terms of the expectation of the $\hat{\gamma}$ operators in 3.24.) Once the mean field was introduced and products of four operators in 3.14 had been approximated by 3.17, we should have expected to be able to diagonalise the Hamiltonian like this—it was a quadratic form and these can be diagonalised. The only complication arising out of the use of operators was that we had to preserve the anticommutation relations.

3.5 Properties of the BCS state

The transformation, which we have used, is called a Bogoliubov transformation. It mixes particle creation and annihilation operators. If we look at the the limit $k \gg k_F$, we find $2\theta_{\mathbf{k}} \rightarrow 0$ and $\hat{\gamma}_{\mathbf{k}, 1}^{\dagger} \rightarrow \hat{c}_{\mathbf{k}, \uparrow}^{\dagger}$. In the limit $k \ll k_F$, we find $2\theta_{\mathbf{k}} \rightarrow \pi$ and $\hat{\gamma}_{\mathbf{k}, 1}^{\dagger} \rightarrow \hat{c}_{-\mathbf{k}, \downarrow}$. The excitations are particle-like far above the Fermi energy and hole-like far below the Fermi energy. For $\xi_{\mathbf{k}} = 0$, $\theta_{\mathbf{k}} = \pi/4$ and the excitations are equal mixtures of a particle and a hole. The term in the Hamiltonian, $(\Delta^* \hat{c}_{\mathbf{k}, \uparrow}^{\dagger} \hat{c}_{-\mathbf{k}, \downarrow}^{\dagger} + \Delta \hat{c}_{-\mathbf{k}, \downarrow} \hat{c}_{\mathbf{k}, \uparrow})$, when acting on an electron in

state \mathbf{k}, \uparrow , describes a process in which a single electron is replaced by a hole (and a pair of electrons is absorbed into the condensate). In the conjugate process, a hole is replaced by an electron (a pair of electrons is released from the condensate). These processes replace a particle by a hole and vice versa, so the true excitation of the MFT Hamiltonian becomes a linear superposition of a particle and a hole. The excitations at $\xi_{\mathbf{k}} = 0$ have energy $|\Delta|$, which is often called the energy gap.

Ground State

The ground state is the one annihilated by all $\hat{\gamma}_{\mathbf{k},\tau}$. We can find this ground state wavefunction in the basis of the original electronic states. The space of states acted on by the $\hat{\gamma}_{\mathbf{k},\tau}$ for given \mathbf{k} is

$$|\psi\rangle = p|0\rangle + q\hat{c}_{\mathbf{k},\uparrow}^\dagger|0\rangle + r\hat{c}_{-\mathbf{k},\downarrow}^\dagger|0\rangle + s\hat{c}_{\mathbf{k},\uparrow}^\dagger\hat{c}_{-\mathbf{k},\downarrow}^\dagger|0\rangle \quad (3.27)$$

where p, q, r, s are amplitudes. The BCS ground state must be annihilated by both $\hat{\gamma}_{\mathbf{k},1}$ and $\hat{\gamma}_{\mathbf{k},2}$. Substituting for the $\hat{\gamma}_{\mathbf{k},1}$ from 3.19, gives $q = 0$, $p \propto u_{\mathbf{k}}$ and $s \propto v_{\mathbf{k}}$. Requiring that $\hat{\gamma}_{\mathbf{k},2}$ also annihilates the state means $r = 0$. Taking the product over all \mathbf{k} gives the BCS ground state

$$|BCS\rangle = \prod_{\mathbf{k}} (u_{\mathbf{k}} + v_{\mathbf{k}}\hat{c}_{\mathbf{k},\uparrow}^\dagger\hat{c}_{-\mathbf{k},\downarrow}^\dagger)|0\rangle. \quad (3.28)$$

This wavefunction emphasises the nature of the pairing explicitly. In the original work, this wavefunction was introduced as a variational wavefunction and the total energy minimised to find the coefficients $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$. This gives the same results as the approach due to Bogoliubov, which we have outlined here and which appeared shortly after the BCS paper.

Density of States and Tunneling

The density of states, $n(E)$, for the excitations with energy $E_{\mathbf{k}}$ can be measured experimentally in tunneling experiments. Using $\xi = \sqrt{E^2 - |\Delta|^2}$, we find

$$n(E) = \frac{dN}{dE} = \frac{dN}{d\xi} \frac{d\xi}{dE} = \nu_F \frac{E}{\sqrt{E^2 - |\Delta|^2}} \theta(E - |\Delta|), \quad (3.29)$$

where the theta function means that there are no excitations with energies less than $|\Delta|$. Now imagine an electrode separated from a superconducting sample by a barrier which allows electrons to tunnel between the electrode and the sample. With no applied voltage the chemical potentials of the sample and the electrode align and there is no flow. When a voltage is applied there is a difference eV in the two chemical potentials. The net current, I , and differential conductance, dI/dV , should vary at $T = 0$ as

$$I(V) \sim \int_0^{eV} n(E) dE$$

$$\frac{dI}{dV} \sim n(eV) = \nu_F \frac{eV}{\sqrt{(eV)^2 - |\Delta|^2}} \theta(eV - |\Delta|). \quad (3.30)$$

The formula for $I(V)$ assumes that the tunneling matrix element between electronic states in the electrode is the same for all states within a few $|\Delta|$ of the chemical potentials. The density of states in the electrode is also assumed constant (equal to its value at the chemical potential) and that the only variable is the density of states in the superconductor. One would expect corrections to this formula to describe details of the system but the result 3.30 is remarkably good.

Non-zero Temperature

Many properties of the superconducting state can be computed using the Bogoliubov approach. To describe non-zero temperature we use the Fermi-Dirac distribution (note that $\sqrt{\xi^2 + |\Delta|^2}$ is always positive), $\langle \hat{\gamma}_{\mathbf{k},\tau}^\dagger \hat{\gamma}_{\mathbf{k},\tau} \rangle = 1/e^{(E_{\mathbf{k}} - \mu)/kT} + 1$), in 3.24. The corresponding result to 3.25 is

$$1 = \frac{U\nu_F}{2} \int_{-\hbar\omega_D}^{\hbar\omega_D} d\xi \frac{\tanh(\sqrt{\xi^2 + |\Delta|^2}/2kT)}{\sqrt{\xi^2 + |\Delta|^2}}. \quad (3.31)$$

The self-consistent solution can be found numerically but, in the limit $U\nu_F \ll 1$, it is possible to find analytically the maximum temperature T_c at which the superconducting state can exist:

$$\frac{2\Delta_0}{kT_c} = 2\pi e^{-\gamma} = 3.52775398 \dots, \quad (3.32)$$

where $\gamma = 0.5772 \dots$ is Euler's constant. This BCS prediction was quickly confirmed using optical experiments to measure the zero-temperature gap, Δ_0 .

Supercurrents

How does the BCS theory describe the response to em fields? The answer to this is more involved than one might expect. It requires inclusion of the coupling between the electrons and the magnetic vector potential, as should seem normal if you have studied the phenomenology of superconductivity and the London equations (these are mentioned in [Condensed Matter Physics II](#)). However, a rough idea of how the BCS state might carry a supercurrent runs as follows (this argument is usually attributed to Landau). We imagine shifting the Fermi volume an amount \mathbf{K} so that the system has a net momentum. (We might imagine switching on an electric field to set the current up and then switching it off leaving a non-zero supercurrent.) If we assume that the self-consistent equations of the spectrum go through as before, the system will not be able to create an excitation in the condensate unless $\xi_K \geq |\Delta|$. Then a particle or hole can be scattered across the Fermi surface with zero energy needed.

3.6 Questions

We should ask whether the theory is a good one. Actually we know it is, because it explains and predicts accurately much of what can be measured. Instead we should ask why this mean field theory (MFT) works so well. We should look at the typical contributions of the terms neglected in the MFT and compare these with the effects of the terms we have kept (see section 3.3). The “missing” contributions are of the type (angled brackets denote expectation values computed in the mean-field state)

$$< (\hat{c}_{\mathbf{k},\uparrow}^\dagger \hat{c}_{-\mathbf{k},\downarrow}^\dagger - b_{\mathbf{k}})(\hat{c}_{-\mathbf{k}',\downarrow} \hat{c}_{\mathbf{k}',\uparrow} - b_{\mathbf{k}}^*) > \equiv < \delta(\hat{c}_{\mathbf{k},\uparrow}^\dagger \hat{c}_{-\mathbf{k},\downarrow}^\dagger) \delta(\hat{c}_{-\mathbf{k}',\downarrow} \hat{c}_{\mathbf{k}',\uparrow}) >. \quad (3.33)$$

These terms characterise the effect of fluctuations about the mean field values $b_{\mathbf{k}}$, which is the reason for the symbol δ . (Normally the treatment of this problem is covered in books on second order phase transitions and critical phenomena. We will only summarise it here.)

The free energy owing to these fluctuations computed from the MFT, δF_{flucns} , needs to be compared with the computed free energy differences between the two phases δF_{eqm} . These

can be estimated to scale as

$$\delta F_{\text{flucns}} \sim U \frac{1}{R^2 \xi(T)} \quad \text{and} \quad \delta F_{\text{eqm}} \sim Ut. \quad (3.34)$$

Here $t = |T - T_C|/T_c$, R is the range of the interaction in real space and ξ is the correlation length for fluctuations where both R and ξ are in units of a microscopic length (usually the lattice parameter). If the interaction in Fourier space is restricted to energies within $\hbar\omega_D = \hbar v_F \Delta k$ of the Fermi energy, this corresponds to a range of interaction $R \sim 1/\Delta k \sim v_F/\omega_D$. At a second order transition, the correlation length for fluctuations diverges as $Rt^{-1/2}$. This is because the regions of the phase from the other side of the transition are larger the closer the system is to the transition and become infinite at the transition. The square root dependence is the mean field result.

The MFT should be good if, using $t = (R/\xi^2)$,

$$|\delta F_{\text{flucns}}| \ll \delta F_{\text{eqm}} \quad \Rightarrow \quad \xi \ll R^4. \quad (3.35)$$

The range of the interaction R is so large in superconductors that ξ has to become almost macroscopically large for the condition not to hold. The BCS theory is essentially always correct. If you want to know more on this, search for Ginzburg criterion. It is covered in Appendix II in “Fundamentals of the the Theory of Metals” by Abrikosov. A good source for the validity of MFT is Chapter 2 in “[Scaling and Renormalization in Statistical Physics](#)” by John Cardy.

There are many aspects not covered here—in fact almost everything! In particular, there is the origin of the attractive interaction. This is discussed in Appendix C. Even if there were a possible attractive interaction due to the exchange of phonons, how can it be strong enough to overcome the repulsive Coulomb interaction which is the strongest interaction in metals? Here, it is important to understand that the Coulomb interaction is screened by the other electrons and that the interaction induced by the exchange of phonons is retarded. Screening means that the Coulomb potential due to one electron attracts a cloud of positive charge (missing electrons) thereby reducing its effect on other electrons. Retarded means that the interaction is effectively delayed in time. The positively charged ions are perturbed by an electron. The phonon response occurs on a time-scale of $t_{ph} \sim 1/\omega_D$. Another electron is attracted to this distortion while the original electron has “moved on”. The two electrons end up being bound together but they are never too close to be forced away by the strong Coulomb repulsion.

Chapter 4

Interacting Spins

In this chapter, we will look at a system of spins on a lattice with nearest neighbour interactions. We will see that long-range order of any type is not possible in 1D at non-zero temperature. However, it is interesting to see why and how the field theoretic approach helps to see this.

For more on this topic, see [John Chalker](#).

We denote the spin operator in terms of its components $\hat{\mathbf{S}} = (\hat{S}^1, \hat{S}^2, \hat{S}^3)$. The algebra of these components is (assuming Einstein's summation convention)

$$[\hat{S}^i, \hat{S}^j] = i\hbar \epsilon_{ijk} \hat{S}^k, \quad \text{where} \quad \epsilon_{ijk} = \begin{cases} 1 & \text{for even permutation} \\ -1 & \text{for odd permutation} \\ 0 & \text{otherwise} \end{cases} \quad (4.1)$$

This is neither that of fermions nor of bosons. However, in a case where the spin is aligned with a given direction in the ground state, we can approximate these commutation relations by bosonic ones. In the spirit of mean field theory (see 3.3), we replace the right hand side by its expectation value in the ground state.

In the case of alignment of spins on neighbouring sites (ferromagnetic case), the resulting Hamiltonian can be solved by a Fourier decomposition as it was for excitations of ions on a lattice (phonons). In the antiferromagnetic case, where the exchange interaction favours anti-alignment of spins on neighbouring sites, the approximate Hamiltonian is similar to the BCS Hamiltonian. It can be solved the same way, by using a Bogoliubov transformation, but there is one crucial difference. The transformation that diagonalises the Hamiltonian needs to preserve bosonic commutation relations and not fermionic anti-commutation relations. In both cases, ferromagnetic and antiferromagnetic, the excitations of the system are called magnons.

The Heisenberg Hamiltonian, the microscopic origins of which are discussed in Appendix D, is

$$\hat{H} = \frac{J}{2} \sum_{\langle \mathbf{r}, \mathbf{r}' \rangle} \hat{\mathbf{S}}_{\mathbf{r}} \cdot \hat{\mathbf{S}}_{\mathbf{r}'} = \frac{J}{2} \sum_{\langle \mathbf{r}, \mathbf{r}' \rangle} \left(\frac{\hat{S}_{\mathbf{r}}^+ \hat{S}_{\mathbf{r}'}^- + \hat{S}_{\mathbf{r}}^- \hat{S}_{\mathbf{r}'}^+}{2} + \hat{S}_{\mathbf{r}}^z \hat{S}_{\mathbf{r}'}^z \right). \quad (4.2)$$

Here the angular brackets indicate that only neighbouring sites should be summed over and the factor of 1/2 is there because each term appears twice in the sum. After identifying a

z -direction, the Hamiltonian can also be written in terms of the spin raising and lowering operators. These raise and lower the projection of the angular momentum on the z -direction:

$$\hat{S}_r^\pm = \left(\hat{S}_r^x \pm i\hat{S}_r^y \right), \quad [\hat{S}_r^+, \hat{S}_r^-] = 2\hat{S}_r^z. \quad (4.3)$$

We are using the convention that the spin operators measure spin angular momentum in units of \hbar . We will take the spin quantum number to be S , ie $S_r^2 = S(S+1)$.

If the right hand side of the equation involving the commutator in 4.3 were a number, we could scale the raising and lowering operators so that the commutator looked like that for creation and annihilation operators for bosons 1.14. Our approximate treatment will replace this operator on the rhs by its expectation value. The approximation does not always work. Understanding why it sometimes works and sometimes does not work are interesting questions.

Ferromagnetic Case

If the exchange integral $J = -|J| < 0$, the spins will be predominantly aligned at low temperatures. The lowest energy state has all the spins aligned along one direction, which we will call the z -direction. We will call this state $|0\rangle_F$. We will assume that the number of nearest neighbours of each site, z , is even and the same for all sites. This would be the case for a simple cubic lattice in 3D or a square lattice in 2D.

We can see that the state with all spins aligned is an eigenstate by acting on the state with the Hamiltonian. The operators of the type $\hat{S}_r^+ \hat{S}_{r'}^-$ simply annihilate the state. This is because the raising operator (in this case acting on site r) annihilates the state as the spin already has maximum projection $S^z = S$. The operator $\hat{S}_r^z \hat{S}_{r'}^z$, acting on the state gives the value S^2 (as these spins are aligned with the z -direction $S^z = S$). Summing over all nearest neighbour pairs and multiplying by $-|J|/2$ gives an energy per site of $E_0 = -z|J|S^2/2$.

The band of single spin-flip states can also be found exactly for the ferromagnetic case. For the state

$$|\mathbf{r}\rangle = \hat{S}_r^- |0\rangle_F, \quad (4.4)$$

we find

$$(\hat{H} - NE_0) |\mathbf{r}\rangle = -\frac{|J|S}{2} \sum_{\langle \mathbf{r}' \rangle} |\mathbf{r}'\rangle + z|J|S |\mathbf{r}\rangle. \quad (4.5)$$

The operator $(\hat{H} - NE_0)$ will have eigenvalues which are the excitation energies. The sum over $\langle \mathbf{r}' \rangle$ is taken only over the z nearest neighbours of the site \mathbf{r} . The factor of S in the first term follows from the result $\hat{S}^- |S^z\rangle = \sqrt{S(S+1) - S^z(S^z-1)} |S^z-1\rangle$.

We will not go through the derivation in detail as the method is essentially the same as in the approximate treatment of the antiferromagnetic case we look at below. We form a Fourier decomposition of the states:

$$|\mathbf{r}\rangle = \frac{1}{\sqrt{N}} \sum_{\mathbf{q}} e^{i\mathbf{q}\cdot\mathbf{r}} |\mathbf{q}\rangle, \quad \text{and} \quad |\mathbf{q}\rangle = \frac{1}{\sqrt{N}} \sum_{\mathbf{r}} e^{-i\mathbf{q}\cdot\mathbf{r}} |\mathbf{r}\rangle, \quad (4.6)$$

where N is the number of lattice points in the system and the wavevectors \mathbf{q} are those satisfying the usually assumed periodic boundary conditions $e^{i\mathbf{q}\cdot\mathbf{R}} = 1$. The sum over wavevectors is restricted to those in the first Brillouin zone. Here \mathbf{R} is a vector which maps

equivalent points on opposite faces of the system. In crystals with a simple cubic symmetry, $z = 6$ and $\mathbf{R} = (N_x a, N_y a, N_z a)$, where a is the unit cell length and $N = N_x N_y N_z$.

If we act on the state $|\mathbf{q}\rangle$ with the Hamiltonian we find that it is an eigenstate. In the case of a cubic system

$$\begin{aligned}\hat{H}|\mathbf{q}\rangle &= +\frac{z}{2}|J|S|\mathbf{q}\rangle - |J|S(\cos q_x a + \cos q_y a + \cos q_z a)|\mathbf{q}\rangle \\ &= |J|S\left(\frac{z}{2} - \cos q_x a - \cos q_y a - \cos q_z a\right)|\mathbf{q}\rangle \equiv \epsilon(\mathbf{q})|\mathbf{q}\rangle.\end{aligned}\quad (4.7)$$

These excitations are called spin waves or magnons. They are magnetic analogs of the lattice oscillation modes called phonons. One difference is that the dispersion of these modes, $\lim_{q \rightarrow 0} \epsilon(\mathbf{q}) \rightarrow |J|S(q_x^2 + q_y^2 + q_z^2)a^2/2 = |J|Sq^2 a^2/2$. This is quadratic in q . For the phonon case the sound waves had a linear dispersion, $\omega_q = v|q|$ with v the speed of sound see paragraph around A.5.

It is not possible to go much further with exact eigenstates. If we try looking for states with two or more spin flips the calculation is complicated by the fact that there is an effective short range interaction between spin flips. This arises because the right hand side of 4.1 is an operator and not a number. The operators generating and annihilating the spin flips, the spin raising and lower operators, are not bosonic. The one magnon calculation does not work for the antiferromagnetic case either.

4.1 Holstein-Primakoff bosons

The approximation, which we will look at now, constructs a bosonic representation of the creation and annihilation of spin waves. We assume that the spin at a given site is aligned in the ground state with a given direction which we will call z . (This may not be the same direction on each site.) We then replace the operators on the right hand sides of the commutator in 4.3 with a number $\hat{S}_r^z \rightarrow S$ and write $[\hat{S}_r^+, \hat{S}_r^-] = 2S$. After rescaling the operators as follows

$$\begin{aligned}\hat{a}_r &= \frac{1}{\sqrt{2S}}\hat{S}_r^+ \quad \text{and} \quad \hat{a}_r^\dagger = \frac{1}{\sqrt{2S}}\hat{S}_r^-, \quad \text{we have} \\ \hat{S}_r^z &= S - \hat{a}_r^\dagger \hat{a}_r \quad \text{and} \quad [\hat{a}_r, \hat{a}_r^\dagger] = 1.\end{aligned}\quad (4.8)$$

The operators \hat{a}_r and \hat{a}_r^\dagger describe bosons. Eq. 4.8 indicates that the spin lowering operator creates a bosonic excitation. The boson is spin zero as there is no additional spin degree of freedom associated with the spin flip. However, the total spin of the system has been changed by 1 as the operator $\hat{a}_r \sim \hat{S}_r^+$. This turns out to be all we need to find an approximate excitation spectrum.

Before proceeding there are few words of warning. We have assumed that the spin at a given site is aligned along some direction in the ground state. This is a classical idea of a spin. Spins can be in linear combinations with spins at other sites for which this is not the case. The example of the two-site spin 1/2 system, mentioned in Appendix D, shows this. The ground state of this system is the spin singlet introduced in D.2:

$$|s_{12}\rangle = \frac{1}{\sqrt{2}}(|\uparrow_1 \downarrow_2\rangle - |\downarrow_1 \uparrow_2\rangle).$$

Here $\langle S_i^z \rangle = 0$ for both spins. There is no particular direction along which the spin is aligned and the use of the approximation 4.8 could not be made to work.

We should also worry that, for bosons, there is nothing to stop the application of the particle creation operator arbitrarily many times. We know that, if the total spin angular momentum on a site is S , and the spin is initially aligned with a direction (let's say the z -direction), we can only apply the spin lowering operator $2S$ times as there are only $(2S + 1)$ different states for a spin with $\hat{S} \cdot \hat{S} = S(S + 1)$. This particular problem can be solved by introducing a non-linear version of 4.8:

$$\hat{S}_r^+ = (2S)^{1/2} \left(1 - \frac{\hat{a}_r^\dagger \hat{a}_r}{2S} \right)^{1/2} \hat{a}_r \quad \text{and} \quad \hat{S}_r^- = (2S)^{1/2} \left(1 - \frac{\hat{a}_r^\dagger \hat{a}_r}{2S} \right)^{1/2} \hat{a}_r^\dagger. \quad (4.9)$$

We see that the mapping in 4.8 is obtained if we expand the factors, $(1 - \hat{a}_r^\dagger \hat{a}_r / 2S)^{1/2}$, in powers of the operator, $\hat{a}_r^\dagger \hat{a}_r$, and keep only the constant term. The transformation in 4.9 was proposed by Holstein and Primakoff.

Antiferromagnetic Case

When the exchange integral $J = |J|$, the lowest energy states will have spins on neighbouring sites anti-aligned. We will consider only a bipartite lattice. This is one which can be split into two sublattices, A and B, and all nearest neighbours of any site are on the opposite sublattice. A body centred cubic lattice would be an example.

For the bipartite lattice, the classical antiferromagnetic ground state has all spins as vectors pointing in one direction on sub-lattice A and pointing in the opposite direction on sublattice B. In this state all pairs of spins on nearest neighbour sites are anti-aligned. We will denote this state by $|0\rangle_{AF}$. It is also often called the Néel state.

To use the mapping as given in 4.8, we can rotate the axes by 180° about the y -axis on every site on the B lattice¹:

$$\begin{aligned} S^z &\rightarrow -S^z, & S^x &\rightarrow -S^x & \text{and} & S^y &\rightarrow S^y \\ S^+ &\rightarrow -S^- & S^- &\rightarrow -S^+ \end{aligned} \quad (4.10)$$

The magnetic unit cell is now twice as large and contains one site on the A sublattice and one on the B sublattice. As we are treating the sites on the A and B sublattices differently, we will denote the boson operators on the respective sublattices by \hat{a}_r and \hat{b}_r . Every link between neighbouring sites involves a spin on the A and a spin on the B sublattice.

We will go through the 1D case. It brings out the structure of the calculation nicely. (A 3D case is described in the box below.) The bosonic operators give rise to a change in the projection from $S^z = \pm 1$ on the A ($S^z = +1$) or ($S^z = -1$) on the B sub-lattice (we will take the even/odd numbered sites to be on the A/B sublattices). The Hamiltonian for this

¹We need to rotate the axes in order to preserve the commutation relations. Flipping the direction of z would not do this.

1D case is

$$\begin{aligned}
 \hat{H} &= \frac{J}{2} \sum_{\langle i,j \rangle}^N \hat{\mathbf{S}}_i \cdot \hat{\mathbf{S}}_j = \frac{J}{2} \sum_{\langle i,j \rangle}^N \left(\frac{\hat{S}_i^+ \hat{S}_j^- + \hat{S}_i^- \hat{S}_j^+}{2} + \hat{S}_i^z \hat{S}_j^z \right) \\
 &= -\frac{J}{2} \sum_{\langle i,j \rangle}^N S^2 + \frac{JS}{2} \sum_{\langle i,j \rangle}^N \left(\hat{a}_i^\dagger \hat{a}_i + \hat{b}_j^\dagger \hat{b}_j + \hat{a}_i \hat{b}_j + \hat{b}_j^\dagger \hat{a}_i^\dagger \right) \\
 (\hat{H} - NE_0) &= \frac{JS}{2} \sum_{\langle i,j \rangle}^N \left(\hat{a}_i^\dagger \hat{a}_i + \hat{b}_j^\dagger \hat{b}_j + \hat{a}_i \hat{b}_j + \hat{b}_j^\dagger \hat{a}_i^\dagger \right). \tag{4.11}
 \end{aligned}$$

Here the sum over j is over the sites $j = i \pm 1$. The energy per site in the presumed ground state $E_0 = -JS^2$. Now consider the Fourier representations

$$\hat{a}_l = \frac{1}{\sqrt{N/2}} \sum_q^{N/2} \hat{a}_q e^{iq l} \quad \text{and} \quad \hat{b}_{l'} = \frac{1}{\sqrt{N/2}} \sum_q^{N/2} \hat{b}_q e^{iq l'}. \tag{4.12}$$

Inserting these representations into 4.11 gives

$$\begin{aligned}
 (\hat{H} - NE_0) &= JS \sum_q \left(\hat{a}_q^\dagger \hat{a}_q + \hat{b}_{-q}^\dagger \hat{b}_{-q} + \gamma(q) \left(\hat{a}_q \hat{b}_{-q} + \hat{b}_{-q}^\dagger \hat{a}_q^\dagger \right) \right) = \frac{JS}{2} \sum_q \hat{H}_q \\
 \hat{H}_q &= \left(\hat{a}_q^\dagger, \hat{b}_{-q}, \hat{b}_{-q}^\dagger, \hat{a}_q \right) \begin{pmatrix} 1 & \gamma(q) & 0 & 0 \\ \gamma(q) & 1 & 0 & 0 \\ 0 & 0 & 1 & \gamma(q) \\ 0 & 0 & \gamma(q) & 1 \end{pmatrix} \begin{pmatrix} \hat{a}_q \\ \hat{b}_{-q}^\dagger \\ \hat{b}_{-q} \\ \hat{a}_q^\dagger \end{pmatrix} - 1 \tag{4.13}
 \end{aligned}$$

Here $\gamma(q) = \cos qa$ and a is the nearest neighbour distance. The constant comes from commuting operators.

The form of 4.13 is the bosonic equivalent of 3.18. We might expect to be able to diagonalise the Hamiltonian with a Bogoliubov transformation, which will mix creation and annihilation operators (we did this in the case of the BCS Hamiltonian). In this case, any such transformation should conserve the bosonic commutation relations. We can try (acting on one of the 2×2 sub-matrices for simplicity)

$$\begin{aligned}
 \begin{pmatrix} \hat{\alpha}_q \\ \hat{\beta}_q^\dagger \end{pmatrix} &= U \begin{pmatrix} \hat{a}_q \\ \hat{b}_{-q}^\dagger \end{pmatrix} \equiv \begin{pmatrix} u_q & v_q \\ v_q & u_q \end{pmatrix} \begin{pmatrix} \hat{a}_q \\ \hat{b}_{-q}^\dagger \end{pmatrix} \\
 \begin{pmatrix} \hat{a}_q \\ \hat{b}_{-q}^\dagger \end{pmatrix} &= U^\dagger \begin{pmatrix} \hat{\alpha}_q \\ \hat{\beta}_q^\dagger \end{pmatrix} \equiv \begin{pmatrix} u_q & v_q \\ v_q & u_q \end{pmatrix} \begin{pmatrix} \hat{\alpha}_q \\ \hat{\beta}_q^\dagger \end{pmatrix}. \tag{4.14}
 \end{aligned}$$

We have assumed here that the transformation matrix is real (this is fine for bosons not carrying any charge). The commutation relations need

$$1 = [\hat{\alpha}_q, \hat{\alpha}_q^\dagger] = u_q^2 - v_q^2$$

which can be satisfied if we set $u_q = \cosh \theta_q$ and $v_q = \sinh \theta_q$.

With transformation 4.14 we find

$$\hat{H}_q = \begin{pmatrix} \hat{\alpha}_q^\dagger & \hat{\beta}_q \end{pmatrix} \begin{pmatrix} u_q & v_q \\ v_q & u_q \end{pmatrix} \begin{pmatrix} 1 & \gamma(q) \\ \gamma(q) & 1 \end{pmatrix} \begin{pmatrix} u_q & v_q \\ v_q & u_q \end{pmatrix} \begin{pmatrix} \hat{\alpha}_q \\ \hat{\beta}_q^\dagger \end{pmatrix},$$

the matrix is diagonalised if $(u_q^2 + v_q^2)\gamma(q) + 2uv = \gamma_q \cosh 2\theta_q + \sinh 2\theta_q = 0$. This is satisfied if we choose

$$\tanh 2\theta_q = -\gamma(q), \quad \cosh 2\theta_q = \frac{1}{\sqrt{1 - \gamma(q)^2}} \quad \sinh 2\theta_q = \frac{-\gamma(q)}{\sqrt{1 - \gamma(q)^2}}. \quad (4.15)$$

This gives

$$\begin{aligned} \hat{H}_q &= \begin{pmatrix} \hat{\alpha}_q^\dagger & \hat{\beta}_q \end{pmatrix} \begin{pmatrix} u_q^2 + v_q^2 + 2u_q v_q \gamma(q) & 0 \\ 0 & u_q^2 + v_q^2 + 2u_q v_q \gamma(q) \end{pmatrix} \begin{pmatrix} \hat{\alpha}_q \\ \hat{\beta}_q^\dagger \end{pmatrix} \\ &= \begin{pmatrix} \hat{\alpha}_q^\dagger & \hat{\beta}_q \end{pmatrix} \begin{pmatrix} \epsilon(q) & 0 \\ 0 & \epsilon(q) \end{pmatrix} \begin{pmatrix} \hat{\alpha}_q \\ \hat{\beta}_q^\dagger \end{pmatrix}, \quad \text{where } \epsilon(q) = \sqrt{1 - \gamma(q)^2}. \end{aligned} \quad (4.16)$$

The quantities $JS\epsilon(q)$ are the energies of free bosonic excitations. These are spin wave excitations with dispersion given by $\epsilon(q)$. In the limit $q \rightarrow 0$, the spectrum is linear in q with energies varying as $JSqa$. This is in contrast with the ferromagnetic case where the dispersion for low energy excitations varied quadratically with q .

The result 4.16 extends to bipartite lattices in arbitrary dimensions. The box below treats the case for the BCC lattice in 3D. The scale JS changes to $JSz/2$ where z is the number of nearest neighbours of each site there are on the lattice, and $\gamma(q)$ becomes $\gamma(\mathbf{q})$ but still with $-1 \leq \gamma(\mathbf{q}) \leq 1$. The form of the dispersion term, $\gamma(q)$, will vary according to the particular bipartite lattice involved.

Higher Dimensions—BCC case

The Hamiltonian written in terms of the Holstein-Primakoff bosons becomes in higher dimensions

$$\begin{aligned} \hat{H} &= \frac{J}{2} \sum_{\langle \mathbf{r}, \mathbf{r}' \rangle} \hat{\mathbf{S}}_{\mathbf{r}} \cdot \hat{\mathbf{S}}_{\mathbf{r}'} = \frac{J}{2} \sum_{\langle \mathbf{r}, \mathbf{r}' \rangle} \left(\frac{\hat{S}_{\mathbf{r}}^+ \hat{S}_{\mathbf{r}'}^- + \hat{S}_{\mathbf{r}}^- \hat{S}_{\mathbf{r}'}^+}{2} + \hat{S}_{\mathbf{r}}^z \hat{S}_{\mathbf{r}'}^z \right) \\ &= -\frac{J}{2} \sum_{\langle \mathbf{r}, \mathbf{r}' \rangle} S^2 + \frac{JS}{2} \sum_{\langle \mathbf{r}, \mathbf{r}' \rangle} \left(\hat{a}_{\mathbf{r}}^\dagger \hat{a}_{\mathbf{r}} + \hat{b}_{\mathbf{r}'}^\dagger \hat{b}_{\mathbf{r}'} + \hat{a}_{\mathbf{r}} \hat{b}_{\mathbf{r}'} + \hat{b}_{\mathbf{r}'}^\dagger \hat{a}_{\mathbf{r}}^\dagger \right) \\ \left(\hat{H} - NE_0 \right) &= \frac{JS}{2} \sum_{\langle \mathbf{r}, \mathbf{r}' \rangle} \left(\hat{a}_{\mathbf{r}}^\dagger \hat{a}_{\mathbf{r}} + \hat{b}_{\mathbf{r}'}^\dagger \hat{b}_{\mathbf{r}'} + \hat{a}_{\mathbf{r}} \hat{b}_{\mathbf{r}'} + \hat{b}_{\mathbf{r}'}^\dagger \hat{a}_{\mathbf{r}}^\dagger \right). \end{aligned} \quad (4.17)$$

The energy per site in the presumed ground state $E_0 = -JS^2z/2$ (as it was in the ferromagnetic case). The term on the right hand side should describe the excitations.

We have assumed that the ground state has translational symmetry albeit with a unit cell which is twice as large (every other site on the bipartite lattice is equivalent). We should therefore expect that a Fourier decomposition should simplify things. The two Fourier

expansions are

$$\hat{a}_{\mathbf{r}} = \frac{1}{\sqrt{N'}} \sum_{\mathbf{q}} \hat{a}_{\mathbf{q}} e^{i\mathbf{q} \cdot \mathbf{r}} \quad \text{and} \quad \hat{b}_{\mathbf{r}'} = \frac{1}{\sqrt{N'}} \sum_{\mathbf{q}} \hat{b}_{\mathbf{q}} e^{i\mathbf{q} \cdot \mathbf{r}'}. \quad (4.18)$$

The allowed values of \mathbf{q} are those satisfying $e^{i\mathbf{q} \cdot \mathbf{R}'} = 1$. Here \mathbf{R}' is the vector mapping equivalent points on opposite faces as before. There are N' independent wavevectors satisfying this condition. As the smallest magnetic unit cell is larger than the smallest crystallographic one $N' \neq N$. In crystals with BCC symmetry $\mathbf{R}' = (aN'_x, aN'_y, aN'_z)$ and $N' = N'_x, N'_y, N'_z$ and $N' = N'_x N'_y N'_z$. Here a is the length of the unit cell. The BCC unit cell is usually taken as the unit cell used to describe the crystal structure. It has two ions in the unit cell, one at $(0, 0, 0)$ and one at $(a/2, a/2, a/2)$. Even though these two ions are equivalent in the crystal, this unit cell is usually preferred to the smaller unit cell primitive unit cell. However the BCC unit cell is the primitive unit cell of the magnetic structure.

If we insert the representations in 4.18 into 4.17 we find

$$\begin{aligned} (\hat{H} - NE_0) &= JS \frac{z}{2} \sum_{\mathbf{q}} \left(\hat{a}_{\mathbf{q}}^\dagger \hat{a}_{\mathbf{q}} + \hat{b}_{\mathbf{q}}^\dagger \hat{b}_{\mathbf{q}} + \gamma(\mathbf{q}) \left(\hat{a}_{\mathbf{q}} \hat{b}_{-\mathbf{q}} + \hat{b}_{-\mathbf{q}}^\dagger \hat{a}_{\mathbf{q}}^\dagger \right) \right), \quad \text{where} \\ \gamma(\mathbf{q}) &= \frac{1}{z} \sum_{\mathbf{x}} \cos \mathbf{q} \cdot \mathbf{x}. \end{aligned} \quad (4.19)$$

The sum over nearest neighbour links, \mathbf{x} . In the BCC lattice there are 8 nearest neighbour bonds, $(\pm a/2, \pm a/2, \pm a/2)$. This means Notice that $-1 \leq \gamma(\mathbf{q}) \leq 1$, with

$$\begin{aligned} \gamma(\mathbf{q}) &= \frac{2}{z} \sum_{\alpha=\pm 1} \sum_{\beta=\pm 1} \left(\cos \left(\frac{(q_x + \alpha q_y + \beta q_z)a}{2} \right) \right) \\ &\underset{\lim_{q \rightarrow 0}}{\approx} 1 - \frac{1}{z} q^2 a^2. \end{aligned} \quad (4.20)$$

This result of a quadratic dependence on q is common for antiferromagnetically ordered states on a bipartite lattice.

Predictions of the HP treatment

We can ask a little more about what the Holstein-Primakoff (HP) treatment predicts. We will see that it implies that the ground state is not antiferromagnetically ordered (even at zero temperature) in 1D but can be in higher dimensions. We will also see that there can never be true long range order at non-zero temperature in any 1D system.

First, let's look at the ground-state of the AF problem predicted by HP. It cannot be the classical Néel state as this is not annihilated by $\hat{\alpha}_q$, which is a linear superposition of creation and annihilation operators. If this acted on the Néel state, the state with all spins aligned with the alternating directions we have called z , the state will not be annihilated. We have that

$$\hat{\alpha}_q |\text{gs}\rangle = 0 \quad \forall q, \quad \text{while} \quad \hat{\alpha}_q |0\rangle = \left(u_q \hat{a}_q + v_q \hat{b}_q^\dagger \right) |0\rangle = v_q \hat{b}_q^\dagger |0\rangle \neq 0. \quad (4.21)$$

While it is possible to construct the ground state explicitly, we can study one interesting aspect of the new ground state quite easily as follows. We can estimate the reduction in the expected value of the spin on a given site. We would call this the result of quantum fluctuations. We can calculate the average reduction of the spin from its maximum value $S^z = S$ on the A sites (the calculation on the B sites is similar)

$$\begin{aligned}
\langle \Delta S^z \rangle &= -\frac{2}{N} \sum_q \langle \text{gs} | \hat{a}_q^\dagger \hat{a}_q | \text{gs} \rangle \\
&= -\frac{2}{N} \sum_q \langle \text{gs} | u_q^2 \hat{\alpha}_q^\dagger \hat{\alpha}_q + u_q v_q \hat{\alpha}_q \hat{\beta}_q + u_q v_q \hat{\alpha}_q^\dagger \hat{\beta}_q^\dagger + v_q^2 \hat{\beta}_q \hat{\beta}_q^\dagger | \text{gs} \rangle \\
&= -\frac{2}{N} \sum_q v_q^2 = \frac{1}{N} \sum_q \frac{1}{\sqrt{1 - \gamma(q)^2}} - 1.
\end{aligned} \tag{4.22}$$

In the limit $\gamma \rightarrow 0$, $1 - \gamma(q)^2 = (qa)^2$, which means the denominator vanishes as $1/q$. This also applies in higher dimensions (the prefactors may vary but the denominator always varies as $1/q$ for bipartite lattices in arbitrary dimension).

If the denominator vanishes as $q \rightarrow 0$, the sum may not be bounded. If so, the ordering must have been destroyed. Long-range ordering is only possible in dimensions less than 2. If the system is large, we can replace the sum by an integral over q . This gives in a system with dimension d , which has a density of states in q varying as q^{d-1} ,

$$\langle \Delta S^z \rangle = -\frac{1}{N} \sum_q \frac{1}{\sqrt{1 - \gamma(q)^2}} - 1 \sim - \int_{\text{BZ}} dq q^{d-1} \frac{1}{q}. \tag{4.23}$$

Here the integral over q is restricted to the values of q inside the first Brillouin Zone (BZ). For $d = 1$, the integral is unbounded. This implies that the spin cannot be assumed to align with any particular direction on any site. In two or higher dimensions it can be, and ordered antiferromagnets are found.

4.2 Dimensionality

The dimensionality of a system is important in determining whether there is the possibility of long-range order. In fact, one can show that long-range order at non-zero temperature is impossible in any 1D system with short-range interactions. We do not need to know the explicit form of the interaction.

A argument, attributed usually to Landau, goes as follows. Suppose long-range order exists in some system, for example the ordering of double bonds in a long molecule. Now imagine a defect in the order in a system of length N . If we represent the order using spin variables, the order might change from up to down:

$$\uparrow\uparrow\uparrow \dots \uparrow\downarrow\downarrow \dots \downarrow\downarrow$$

The middle of the line shows the domain wall as the system switches alignment direction.

Even if the system can relax and mix in other spin states there would always be a finite energy associated with this defect (we have assumed that the interaction is short-range,

which is equivalent to having a finite energy for the domain wall). We will take this energy to be ΔE . Where this defect is in the chain, it does not affect the total energy. It can be at any of the N sites in the chain (we would normally assume that the system has periodic boundary conditions to ensure that all sites are equivalent). The entropy of the system increases by $k_B \ln N$ by nucleating such a defect. The change in free energy, when we add this defect,

$$\Delta F = \Delta E - T\Delta S = \Delta E - k_B T \ln N.$$

In the thermodynamic limit $N \rightarrow \infty$, the free energy will always be lower if we introduced a defect breaking the long-range order. We can think of a length scale over which the order can persist which is where $\Delta F \approx 0$:

$$l \approx a e^{\Delta E/k_B T}. \quad (4.24)$$

The length scale in 4.24 helps understand why quasi-1D magnetic systems (these are systems which have much stronger interactions along chains inside a system than between chains) can have transitions into ordered states. The order along the chain develops over a distance, l , giving rise to blocks of aligned spins. These blocks interact with corresponding blocks in other chains. The interactions between the chains are effectively magnified by the factor l/a , and the system becomes 2- or 3-dimensional.

Mermin-Wagner Theorem

There is a theorem, due to Mermin and Wagner, which states that a continuous symmetry cannot be broken at non-zero temperature in systems in 2 or fewer dimensions. An example of this is the ferromagnetic system introduced above. The continuous symmetry is the choice of the direction z . The total moment of the ferromagnet could point in any direction. Yet we have chosen a particular direction and called it z . The change in the average value of S^z as a function of temperature can be computed if we assume that the spin waves are independent. (Actually they are not, although this would not affect the MW result.)

If the spin waves were free bosons, the occupation of the states at non-zero temperature T would be given by the Bose-Einstein distribution. The low energy spectrum we established was $\epsilon(\mathbf{q}) \approx |J|S q^2 a^2/2$. Considering only low temperatures, this is only significant for energies $\epsilon(\mathbf{q}) < k_B T$, this means (here $\beta = 1/k_B T$ and take $a = 1$):

$$\langle \Delta S^z \rangle = -\frac{1}{N} \sum_{\mathbf{q}} \frac{1}{e^{\beta \epsilon(\mathbf{q})} - 1} \sim - \int_0^{\sqrt{|J|S/k_B T}} d\mathbf{q} \frac{q^{d-1}}{\beta q^2}. \quad (4.25)$$

This integral is divergent for $d \leq 2$ (which is why we have not worried too much about where a should appear and factors of 2). The long-range magnetically ordered ferromagnetic state is unstable at any non-zero temperature ($\Delta S^z \rightarrow -\infty$) as stated by the MW theorem.

For further reading, see [John Chalker](#)

Chapter 5

Density Matrices

There are cases, other than that of particle creation and annihilation as a result of interactions, where systems cannot be described by a wavefunction. Every system is in contact with the rest of the Universe—even isolated atoms are an idealisation. They can exchange energy and information, and sometimes also particles, with outside systems. This is after all what happens in most experiments. We probe a system by perturbing it. We send in photons from outside to sense the states of individual ions or regions of the system or we apply electric and magnetic fields. If we didn't do this, we wouldn't know anything about them!

There are many good sources for extensive treatments on this topic. I have been using [Aram Harrow](#) and [Peter Young](#).

An example of another system, where we do not consider the whole system, is one in contact with a heat bath. The degrees of freedom of the heat bath are not treated explicitly but their effect determines properties of the quantum system of interest. We were implicitly working with this idea when we gave an expression for the heat capacity of a superconductor. We assumed that the average occupation of each eigenstate for the free fermion excitations of the mean-field theory was given by the Fermi-Dirac distribution, see 3.31. We said that (the quantity we looked at was the total energy but here we will take it to be any observable \hat{O})

$$\langle \hat{O} \rangle = \frac{1}{Z} \sum_i e^{-\beta E_i} \langle i | \hat{O} | i \rangle, \quad \text{where} \quad Z = \sum_i e^{-\beta E_i}, \quad (5.1)$$

and $\beta = 1/kT$. The states $|i\rangle$ are the normalised orthonormal eigenstates of the system of interest. We are assuming that we know these states and that they are all occupied according to the Boltzmann distribution.

The expression 5.1 is the right answer. We can write the result 5.1 by taking the Boltzmann factor inside and writing

$$\langle \hat{O} \rangle = \frac{1}{Z} \sum_i \langle i | e^{-\beta \hat{H}} \hat{O} | i \rangle \equiv \text{Tr} \frac{1}{Z} e^{-\beta \hat{H}} \hat{O}. \quad (5.2)$$

Here Tr denotes the trace. In a finite-dimensional space this means summing the diagonal matrix elements. We will assume that this idea generalises to an infinite-dimensional Hilbert space (it does). The advantage of writing things this way is that the trace of a matrix is

invariant under changes of orthonormal basis. If the states $|i\rangle$ are a complete orthonormal set of states, the value of the trace is the same. The expression 5.2 is therefore correct independent of whether the states $|i\rangle$ are the eigenstates of \hat{H} .

Writing the expectation value using 5.2 is at the expense of introducing a new operator which we call the density operator or density matrix. In 5.2 we have defined the density operator for thermal equilibrium

$$\hat{\rho}_{\text{eq}} = \frac{1}{Z} e^{-\beta \hat{H}} \quad \text{with} \quad Z = \text{Tr} e^{-\beta \hat{H}}. \quad (5.3)$$

What von Neumann and, independently, Landau showed in the 1920's, is that quantum theory can be reformulated in terms of a density operator.

Assume for the moment a complete orthonormal basis of states $|i\rangle$ (the non-orthogonal case brings up some interesting aspects). In a complete orthonormal basis the “resolution of the identity” is

$$1 = \sum_j |j\rangle \langle j|. \quad (5.4)$$

This operator has the effect of multiplying by 1. Acting with this operator on a state with expansion, $|\psi\rangle = \sum_i a_i |i\rangle$, we obtain $\sum_j |j\rangle \langle j| \psi\rangle = \sum_{ji} |j\rangle a_i \delta_{ji} = \sum_j |j\rangle a_j = |\psi\rangle$, i.e. we recover the original state $|\psi\rangle$. The thermal density matrix in this basis can be written (assuming the summation convention)

$$\hat{\rho}_{\text{eq}} = |j\rangle \langle j| \frac{e^{-\beta \hat{H}}}{Z} |i\rangle \langle i| \equiv |j\rangle (\hat{\rho}_{\text{eq}})_{ji} \langle i|. \quad (5.5)$$

(If the states $|i\rangle$ are the eigenstates of \hat{H} , $\rho_{ji} = e^{-\beta E_i} \delta_{ij} / Z$, and the density matrix is diagonal.)

Generally, we would expect to write

$$\hat{\rho} = |j\rangle \rho_{ji} \langle i| \quad \text{and} \quad \langle \hat{O} \rangle = \text{Tr} \hat{\rho} \hat{O} \quad (5.6)$$

for an arbitrary observable \hat{O} . Computing a single matrix element in the case of a wavefunction, namely the expectation value of an operator in a particular quantum state, has become a trace over the basis of the space. This is more general as there are cases where the system cannot be described by a single quantum state. We have looked at the case of thermal equilibrium but the approach applies in general.

We haven't made anything easier. The problem has been shifted from computing a wavefunction, which is a vector in a Hilbert space, to computing the density matrix. However, the density matrix allows us to consider problems not described by a wavefunction. Examples include a system in contact with a heat bath or systems that are in contact with other systems more generally (not necessarily a heat bath). More recently, there has been an increase in interest in open systems, for which the coupling to the outside universe should be considered explicitly. In the field of quantum information, qubits of interest can be subsystems of a larger system.

5.1 Properties of the Density Matrix

First, a word about names. The density matrix is not a density and not a matrix either but it's the name we all use. It should not be confused with charge and number density in systems of particles. It is an operator, which has matrix elements once a basis has been chosen.

We will start with a case that should be familiar from studying wavefunctions. If a system can be described by a wavefunction $|\psi\rangle$, the density matrix becomes

$$\hat{\rho} = |\psi\rangle\langle\psi| \quad \text{and} \quad \langle \hat{O} \rangle = \text{Tr } \hat{\rho} \hat{O} = \sum_i \langle i|\psi\rangle\langle\psi|\hat{O}|i\rangle = \langle\psi|\hat{O}\left(\sum_i |i\rangle\langle i|\right)|\psi\rangle = \langle\psi|\hat{O}|\psi\rangle, \quad (5.7)$$

where we have used the result 5.4. The density matrix gives the result we would expect for the expectation value of some operator, namely $\langle\psi|\hat{O}|\psi\rangle$.

For a system described by a wavefunction, the density matrix (or operator) is a projection operator:

$$\hat{\rho}^2 = |\psi\rangle\langle\psi|\psi\rangle\langle\psi| = |\psi\rangle\langle\psi| = \hat{\rho}. \quad (5.8)$$

Acting on an arbitrary state function, the density operator suppresses all components of the function that are orthogonal to the state $|\psi\rangle$.

Diagonal Form and Probabilities

If we insert the identity operator for \hat{O} in 5.6, we see that

$$\text{Tr } \hat{\rho} = 1. \quad (5.9)$$

This and the density matrix for thermal equilibrium suggest that, in some basis, all density matrices can be written

$$\hat{\rho} = \sum_i p_i |i\rangle\langle i|, \quad p_i \geq 0 \quad \text{and} \quad \sum_i p_i = 1. \quad (5.10)$$

We will take this as the required property of a density matrix. This implies that the operator must be Hermitian. (If in one basis the matrix is diagonal and the only non-zero matrix elements are probabilities, the matrix will be Hermitian after any unitary transformation.) If we diagonalise a density matrix the diagonal elements p_i give the probability that the system is in the state $|i\rangle$. In a non-orthogonal basis, we should be careful about what this means we can say about the states the system occupies (see below).

In the case of a wavefunction considered in 5.7 $\hat{\rho}$ is already diagonal. We take one of the basis states, $|n\rangle = |\psi\rangle$. The diagonal matrix elements (probabilities of occupation) are 1 for $i = n$ and zero for all $i \neq n$.

In the case of the equilibrium density matrix $\hat{\rho}_{eq}$ is again diagonal if we choose the basis states to be the eigenstates of \hat{H} :

$$\hat{\rho} = \hat{\rho}_{eq} = \sum_i |i\rangle \frac{e^{-\beta E_i}}{Z} \langle i|. \quad (5.11)$$

The diagonal matrix elements equal the probabilities in thermal equilibrium.

Equation of Motion for the Density Operator

The equation of motion for the density operator in the Schrodinger picture (see section 1.2) follows from the equation of motion for vectors $|\psi_i\rangle$. Write the density matrix in its diagonal form (we will assume that \hat{H} is not time-dependent), then

$$\begin{aligned}\frac{\partial \hat{\rho}}{\partial t} &= \left(\frac{\partial}{\partial t} \right) p_i |i\rangle \langle i| = p_i \frac{1}{i\hbar} \left(i\hbar \frac{\partial |i\rangle}{\partial t} \right) \langle i| - p_i |i\rangle \left(-i\hbar \frac{\partial \langle i|}{\partial t} \right) \frac{1}{i\hbar} \\ &= p_i \frac{1}{i\hbar} \hat{H} |i\rangle \langle i| - p_i |i\rangle \left(i\hbar \frac{\partial |i\rangle}{\partial t} \right)^\dagger \frac{1}{i\hbar} \\ &= \frac{1}{i\hbar} \left(\hat{H} p_i |i\rangle \langle i| - p_i |i\rangle \langle i| \hat{H} \right) = \frac{1}{i\hbar} [\hat{H}, \hat{\rho}].\end{aligned}$$

We have the result

$$i\hbar \frac{\partial \hat{\rho}}{\partial t} = [\hat{H}, \hat{\rho}]. \quad (5.12)$$

This is called the von Neumann equation.

Added for Interest

The von Neumann equation should not be confused with the one for the time-dependence of operators in the Heisenberg picture. We have not derived the Heisenberg operator equation but you may have met it or you can look it up in textbooks:

$$i\hbar \frac{\partial \hat{O}}{\partial t} = [\hat{O}, \hat{H}]. \quad (5.13)$$

For a time-independent Hamiltonian, in the Heisenberg picture the density operator is actually time-independent. The case of the single-wavefunction should make this clear:

$$\hat{\rho}_S = |\psi\rangle \langle \psi| = e^{-i\hat{H}t} |\psi(0)\rangle \langle \psi(0)| e^{+i\hat{H}t} \quad (5.14)$$

If we use the relation to move this operator into the Heisenberg representation Equation 1.29, we obtain

$$\hat{\rho}_H = \hat{U}^\dagger(t) \hat{\rho}_S \hat{U}(t) = |\psi(0)\rangle \langle \psi(0)|, \quad (5.15)$$

as $\hat{U}(t) = e^{-i\hat{H}t}$.

Pure and Mixed States

We distinguish between pure and mixed states. After diagonalising $\hat{\rho}$ in an orthonormal basis, the diagonal matrix elements can be considered as probabilities p_i . Hence $\text{Tr } \hat{\rho}^2 = \sum_i p_i^2 \leq \sum_i p_i = 1$. Equality occurs when $p_i = 1$ for one of the states and zero for all others. A state occurring with probability 1 is called a **pure** state. This was the case considered in 5.7. All states for which $\text{Tr } \hat{\rho}^2 < 1$ are called **mixed** states. An example of a system in a mixed state is one described by $\hat{\rho}_{\text{eq}}$.

One mixed state that illustrates the need to be careful how we interpret the density matrix in terms of probabilities is the following. Suppose that a two-state spin system with basis $|\uparrow\rangle$

and $|\downarrow\rangle$ is in state $|\uparrow\rangle$ with probability $2/3$ and in state $|\downarrow\rangle$ with probability $1/3$. The density operator and density matrix are

$$\hat{\rho} = \frac{2}{3}|\uparrow\rangle\langle\uparrow| + \frac{1}{3}|\downarrow\rangle\langle\downarrow|, \quad \rho_{ij} = \begin{pmatrix} \frac{2}{3} & 0 \\ 0 & \frac{1}{3} \end{pmatrix}. \quad (5.16)$$

Now consider the same system occupying

$$\begin{aligned} |\psi_1\rangle &= \sqrt{\frac{2}{3}}|\uparrow\rangle + \sqrt{\frac{1}{3}}|\downarrow\rangle, & \text{with probability } p_1 &= \frac{1}{2}, \\ |\psi_2\rangle &= \sqrt{\frac{2}{3}}|\uparrow\rangle - \sqrt{\frac{1}{3}}|\downarrow\rangle, & \text{with probability } p_2 &= \frac{1}{2}. \end{aligned} \quad (5.17)$$

The density operator is

$$\begin{aligned} \hat{\rho} &= \frac{1}{2}(|\psi_1\rangle\langle\psi_1| + |\psi_2\rangle\langle\psi_2|) \quad \text{giving} \\ \rho_{ij} &= \frac{1}{2} \begin{pmatrix} \frac{2}{3} & \frac{\sqrt{2}}{3} \\ \frac{\sqrt{2}}{3} & \frac{1}{3} \end{pmatrix} + \frac{1}{2} \begin{pmatrix} \frac{2}{3} & -\frac{\sqrt{2}}{3} \\ -\frac{\sqrt{2}}{3} & \frac{1}{3} \end{pmatrix} = \begin{pmatrix} \frac{2}{3} & 0 \\ 0 & \frac{1}{3} \end{pmatrix}. \end{aligned} \quad (5.18)$$

The system has the same density matrix for the two cases. The probabilities for the occupation of states are not uniquely determined by the density matrix. This should not be surprising. When a system is in a pure state $|\psi\rangle$, the system is completely defined and nothing is not known (that can be known) about the system. Once it is in a mixed state, the density matrix defines all that can be known and gives correctly the expectation values for measurements but there is not enough information to reconstruct a wavefunction for the system.

5.2 Sub-systems and Entanglement

Sometimes we may be interested in only part of a system and want to treat the rest of the system as its surroundings, in order to understand more about the full system. We might be able to compute the properties of a sub-system with its small number of degrees of freedom and treat its interactions with the rest of the field within some mean field theory. We would then insist that the mean field density matrix computed from the system matched its equivalent assumed for the whole system. Another example might be a case where we can measure the properties of a sub-part of the system directly. We should be able to define a density matrix only for the states that are in scope for the measurement. This indeed turns out to be the case and is what we consider here. We will look at a two-site system with the states of one of the sites being the sub-system.

A simple two-site system involving a $S = 1/2$ spin on each site has four linearly independent states. One choice of orthonormal basis is

$$|\sigma\tau\rangle = |\uparrow\uparrow\rangle, \quad |\uparrow\downarrow\rangle, \quad |\downarrow\uparrow\rangle, \quad \text{or} \quad |\downarrow\downarrow\rangle. \quad (5.19)$$

Technically $|\sigma\tau\rangle = |\sigma\rangle_1 \otimes |\tau\rangle_2$ where \otimes denotes what is called a tensor product. We have added the subscripts to emphasise that we are using σ to denote the spin variable on site 1

and τ to denote the spin variable on site 2. The idea of the tensor product (there is also the idea of the direct sum) is to extend the idea of multiplication (and addition) from numbers to vector spaces and to structures in abstract algebra in general. Mathematically the thinking behind these ideas emphasises the abstract nature of the relations and their universality. See [here](#) if you are interested in this, but be warned this is mathematics explained by a mathematician (at the board of course). In physics, we tend to be less interested in abstraction and more interested in particular systems and computing their properties. Normally we would start by choosing a basis for our sums that makes the calculations as easy as possible. This is explained [here](#) by a physicist (writing on a computer).

We will imagine combining the two simplest systems possible each of which can be in either of their two states. Both system 1 and system 2 have dimension 2 (states up and down) and the tensor product space has dimension $2 \times 2 = 4$ (there are four combinations of the states on site 1 and those on site 2). We take the states, $|\sigma \tau\rangle = |\sigma\rangle_1 \otimes |\tau\rangle_2$, as the basis states for our spin system. This notation means spin 1 is in state $|\sigma\rangle_1$ while spin 2 is in state $|\tau\rangle_2$. This is the only case we will look at. Spin systems are relatively simple and the tensor product nature of the states does not need much attention as there is no issue with permutation symmetries, because the individual spins are on different identifiable sites. To start with we will write the state $|\sigma\rangle_1 \otimes |\tau\rangle_2 = |\sigma\rangle_1 |\tau\rangle_2$.

If we were to measure an observable on lattice site 1, \hat{O}_1 , its expected value would be

$$\begin{aligned} \langle \hat{O}_1 \rangle &= \text{Tr} \hat{\rho} \hat{O}_1 = \sum_{\sigma} \sum_{\tau} {}_2\langle \tau | {}_1\langle \sigma | \hat{\rho} \hat{O}_1 | \sigma \rangle_1 | \tau \rangle_2 \\ &= \sum_{\sigma} {}_1\langle \sigma | \left(\sum_{\tau} {}_2\langle \tau | \hat{\rho} | \tau \rangle_2 \right) \hat{O}_1 | \sigma \rangle_1 \equiv \sum_{\sigma} {}_1\langle \sigma | \hat{\rho}_1 \hat{O}_1 | \sigma \rangle_1 \\ &= \text{Tr} \hat{\rho}_1 \hat{O}_1. \end{aligned} \quad (5.20)$$

We have used the fact that the operator, \hat{O}_1 , does not act on the spin on site 2.

The quantity in parentheses in the middle line of 5.20

$$\left(\sum_{\tau} {}_2\langle \tau | \hat{\rho} | \tau \rangle_2 \right) = \text{Tr}' \hat{\rho} \equiv \hat{\rho}_1 \quad (5.21)$$

is the partial trace (denoted by the prime) of the full density matrix over the degrees of freedom on site 2. It leaves the density matrix for the remaining sub-system, namely the spin on site 1.

In the case of a spin sub-system the algebra is simple, but the idea can be applied to quantum systems in general. If we are concerned with degrees of freedom in a sub-system (*i.e.* we are only going to measure quantities determined by the degrees of freedom in the sub-system), we imagine tracing over all other degrees of freedom to find the density matrix for the sub-system of interest. This is how von Neumann and Landau solved the conceptual problem of how to think about the properties of a system which is part of a larger system.

From now on, we will simplify the notation and write

$$|\sigma\rangle_1 |\tau\rangle_2 = |\sigma_1\rangle |\tau_2\rangle \quad \text{or often just} \quad |\sigma \tau\rangle.$$

Hopefully it will always be clear which sites the respective spin variables relate to.

Entanglement

When looking at a sub-system, one can ask how “entangled” with the rest of the larger system it is. The idea of entanglement also goes back to the early days of quantum theory. The example of the two-spin system is enough to illustrate the concept. We will look at two cases for which the density matrix for the full system is that of a pure state. The first is

$$\begin{aligned} |\psi\rangle &= |\uparrow_1\rangle|\downarrow_2\rangle \quad \text{for which} \quad \hat{\rho} = |\uparrow_1\rangle|\downarrow_2\rangle\langle\downarrow_2|\langle\uparrow_1| \\ \hat{\rho}_1 &= \sum_{\sigma_2} \langle\sigma_2| |\uparrow_1\rangle|\downarrow_2\rangle\langle\downarrow_2|\langle\uparrow_1| |\sigma_2\rangle = |\uparrow_1\rangle\langle\uparrow_1|. \end{aligned} \quad (5.22)$$

(The inner products are between bras and kets on the same site, for example $\langle\sigma_2|\downarrow_2\rangle = \delta_{\sigma_2\downarrow_2}$.)

The other pure state of the full system we will look at is $|\psi\rangle$ where:

$$\begin{aligned} |\psi\rangle &= \frac{1}{\sqrt{2}} (|\uparrow_1\rangle|\downarrow_2\rangle - |\downarrow_1\rangle|\uparrow_2\rangle) \quad \text{for which} \\ \hat{\rho} &= |\psi\rangle\langle\psi| = \frac{1}{2} (|\uparrow_1\rangle|\downarrow_2\rangle\langle\downarrow_2|\langle\uparrow_1| - |\downarrow_1\rangle|\uparrow_2\rangle\langle\downarrow_2|\langle\uparrow_1| \\ &\quad - |\uparrow_1\rangle|\downarrow_2\rangle\langle\uparrow_2|\langle\downarrow_1| + |\downarrow_1\rangle|\uparrow_2\rangle\langle\uparrow_2|\langle\downarrow_1|), \\ \hat{\rho}_1 &= \text{Tr}' \hat{\rho} = \frac{1}{2} |\uparrow_1\rangle\langle\uparrow_1| + \frac{1}{2} |\downarrow_1\rangle\langle\downarrow_1|. \end{aligned} \quad (5.23)$$

The two cases are different. The density matrix for site 1 in 5.22 is that of a pure state, as $\hat{\rho}_1$ has one eigenvalue equal to 1 (the other is zero). The case 5.23 is that of a mixed state. In the first case the full system is in what is called a product state of the spins on the two sites. The state of the spin on site 1 is independent or unentangled with that on site 2.

All states of the full system that are not product states are said to be entangled states. This means that the density matrix for the sub-system is that of a mixed state. If the subsystem has N states, the state is said to be maximally entangled (with the part of the system that has been traced over) if $p_i = 1/N$. In the example, the spin on site 1 has $N = 2$ and is maximally entangled with the spin on site 2. A simple test for whether a sub-system is entangled with the rest of the system is to check whether the square of the density matrix for the sub-system, $\hat{\rho}^2$, has trace less than one. If it does, there is entanglement.

Here is another state for our two site spin system:

$$|\psi\rangle = \frac{1}{2} (|\downarrow\downarrow\rangle + |\downarrow\uparrow\rangle - |\uparrow\downarrow\rangle - |\downarrow\downarrow\rangle), \quad (5.24)$$

which can be seen by eye to be a product state:

$$|\psi\rangle = \frac{1}{\sqrt{2}} (|\downarrow_1\rangle - |\uparrow_1\rangle) \frac{1}{\sqrt{2}} (|\downarrow_2\rangle + |\uparrow_2\rangle). \quad (5.25)$$

If we trace out site 2 we obtain

$$\begin{aligned} \hat{\rho}_1 &= \frac{1}{\sqrt{2}} (|\downarrow_1\rangle - |\uparrow_1\rangle) (\langle\downarrow_1| - \langle\uparrow_1|) \frac{1}{\sqrt{2}} \quad \text{or as a matrix} \\ (\rho_1)_{ij} &= \begin{pmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{pmatrix}. \end{aligned} \quad (5.26)$$

Here the basis of the states is $|1\rangle = |\uparrow\rangle$ and $|2\rangle = |\downarrow\rangle$. The density matrix has eigenvalues 1 and 0 and describes a pure state. As expected (given that the starting state was a product state), there is no entanglement between the two sites.

5.3 Conclusion

The density matrix approach is used widely to handle systems in contact with the outside world. This includes the field of quantum information, where the outside world is essential (we live in it and want to access answers generated or stored by our quantum machines) and a pest—all connection with the outside world leads to loss of information. Loss of information is via the decoherence (loss of phase information) as degrees of freedom in the outside world interact with those we want to use. We do not have control of these outside degrees of freedom and have to treat their effect statistically. For more on quantum computation there is the module [PX447 Quantum Computation and Simulation](#).

Another area of interest is that of open systems, where we are interested in describing explicitly how the dynamics of our sub-system are affected by different types of environment.

The density matrix method can also incorporate uncertainty we have as observers. An example of this might be as follows. Suppose someone has made a measurement on a system described by a wavefunction. According to quantum theory, after the measurement the system should be left in a state which is an eigenfunction of the operator corresponding to the observable. Now suppose we don't know what the result of the measurement was, even though we know a measurement has taken place. What is the expectation of any measurement we might make on the system? The density matrix reverts to the states allowed by the measurement with respective probabilities computed from the original wavefunction.

Appendix A

Continuum Limit of Lattice Oscillations

The phonon calculation looked at in Chapter 1 is a good microscopic theory of excitations of the lattice. It might seem counter-intuitive to look at a long wavelength version of the theory and then quantise it, but that is what we are going to do. It will illustrate the difference between working on a lattice with a finite number (N) degrees of freedom and working with a field defined at every point. Even on a finite interval of length L , there will be infinitely many degrees of freedom. (Roughly speaking a finite number of degrees of freedom is usually described as mechanics, while infinitely many degrees of freedom are described by a field theory. The phonon model in the main text is quantum mechanical, which we solved using a set of operators. The continuum model described here is a field theory—there are infinitely many degrees of freedom.)

In the continuum model corresponding to the model described by the Hamiltonian 1.16, we will use the coordinate x to denote where we are in the system and $\phi(x)$ to describe the displacement from equilibrium of the matter at point x . The mapping from the model of the ions to the continuum model is $la \rightarrow x$ and $x_l \rightarrow \phi(x)$. In writing the elastic energy, we approximate $(x_l - x_{l+1}) \approx -a\partial\phi(x)/\partial x$, while the elastic energy term

$$\begin{aligned} V &= \frac{k}{2} \sum_{l=0}^{N-1} (x_l - x_{l+1})^2 \approx \frac{k}{2} \sum_{l=0}^{N-1} a^2 \left(\frac{\partial\phi}{\partial x} \Big|_{la} \right)^2 \\ &\rightarrow \frac{ka^2}{2} \int_{x=0}^{Na} d\left(\frac{x}{a}\right) \left(\frac{\partial\phi}{\partial x} \right)^2 = \frac{B}{2} \int_0^L dx \left(\frac{\partial\phi}{\partial x} \right)^2, \end{aligned} \quad (\text{A.1})$$

where $L = Na$. The vertical bar, $|_{la}$, is there to indicate that we evaluate ϕ at $x = la$. $B = ka$ should be thought as an elastic modulus. The single right-pointing arrow denotes where we switch to the continuum model. The kinetic energy is 1.16

$$\sum_{l=0}^{N-1} \frac{p_l^2}{2m} = \sum_{l=0}^{N-1} m \frac{\dot{x}_l^2}{2} = \sum_{l=0}^{N-1} \rho a \frac{\dot{\phi}^2(la)}{2} \rightarrow \rho a \int_0^L d\left(\frac{x}{a}\right) \frac{\dot{\phi}^2(x)}{2} = \rho \int_0^L dx \frac{\dot{\phi}^2(x)}{2}. \quad (\text{A.2})$$

The density per unit length $\rho = m/a$.

We put the kinetic energy and potential energy together to form the Lagrangian $T - V$. As the dependent variable, $\phi(x, t)$, depends on two independent variables x and t , we write the

action in terms of the Lagrangian density \mathcal{L}

$$S = \int_0^L dt dx \mathcal{L} \quad (\text{A.3})$$

$$\mathcal{L} = \frac{1}{2} \left(\rho \dot{\phi}^2 - B \left(\frac{\partial \phi}{\partial x} \right)^2 \right). \quad (\text{A.4})$$

Here ρ is a mass per unit length. The action is the integral over time of the Lagrangian. In a field theory it is the integral over time and space of the Lagrangian density \mathcal{L} . We would normally specify the time interval and values for ϕ as a function of position at these times.

Extremising the action determines the classical equation of motion with the solution taking the field from its initial profile at t_0 to its final profile at t_f . The Euler-Lagrange equation for the action in A.4 is ($\phi_x = \partial\phi/\partial x$, etc)

$$\begin{aligned} 0 &= \frac{\partial \mathcal{L}}{\partial \phi} - \partial_x \left(\frac{\partial \mathcal{L}}{\partial \phi_x} \right) - \partial_t \left(\frac{\partial \mathcal{L}}{\partial \dot{\phi}} \right) = B \frac{\partial^2 \phi}{\partial x^2} - \rho \frac{\partial^2 \phi}{\partial t^2} \\ \Rightarrow \quad \frac{\partial^2 \phi}{\partial t^2} &= \frac{1}{v^2} \frac{\partial^2 \phi}{\partial x^2}, \quad v^2 = \frac{B}{\rho}. \end{aligned} \quad (\text{A.5})$$

Eq A.5 is the usual wave equation describing sound waves, which propagate at speed $v = \sqrt{B/\rho}$ (dispersion relation $\omega_q = v|q|$).

The action functional is the standard starting point for quantising field theories. You might ask why do we need to do this, given that we could deduce the momentum field by working from the p_i directly and defining a momentum density. The answer is that it is not always the case that the momentum density is given by $\rho \dot{\phi}$ (as it is here). Instead we need to deduce the conjugate momentum field in the usual way from the Lagrangian. Another reason is that, within the Lagrangian approach, time and position appear in almost the same way. This helps to make symmetries involving time and space explicit (Lorentz invariance).

Added for Interest A reason for working with the action S (and not Hamiltonians) is that a formulation of quantum theory, due to Feynman and others, quantises a system by computing the action over all classical paths and summing the factors $e^{iS/\hbar}$, to give the quantum amplitude for the system to propagate from the initial to final states specified in the limits of the integrals in A.4. This is the path integral formulation of quantum theory.

The path integral approach is often used to formulate theories. First, it has the advantage that Lorentz invariance can be included naturally. The Hamiltonian, as the generator of time-translations, appears to treat time differently from space. Another reason is that many quantum field theories contain infinities associated with interactions (and imposing Lorentz invariance) and establishing a Hamiltonian is not possible. However, at the level of the quantum amplitudes used in the path integral formulation, it is sometimes possible to arrive at an effective action that is free of divergences. This is called renormalisation. The conceptual idea is that high energy excitations can be “integrated out” so that they are no longer explicitly present. Their effect is included in amended interactions for

the remaining degrees of freedom. The resulting effective field theory may be free of divergences and allow the computation of quantum amplitudes for observable quantities.

The momentum density for the field ϕ and resulting Hamiltonian density are given by

$$\pi = \frac{\partial \mathcal{L}}{\partial \dot{\phi}} = \rho, \quad \mathcal{H} = \pi \dot{\phi} - \mathcal{L} = \frac{\pi^2}{2\rho} + B \left(\frac{\partial \phi}{\partial x} \right)^2. \quad (\text{A.6})$$

The Hamiltonian is then

$$H = \int_0^L dx \left(\frac{\pi^2}{2\rho} + B \left(\frac{\partial \phi}{\partial x} \right)^2 \right). \quad (\text{A.7})$$

If we put periodic boundary conditions onto the system, $\phi(0, t) = \phi(L, t)$ and $\pi(0, t) = \pi(L, t)$, it is again natural to use a Fourier decomposition:

$$\begin{aligned} \phi(x, t) &= \frac{1}{\sqrt{L}} \sum_{n=0}^{\infty} \phi_{q_n}(t) e^{iq_n x} + \phi_{-q_n}(t) e^{-iq_n x}, \\ \pi(x, t) &= \frac{1}{\sqrt{L}} \sum_{n=0}^{\infty} \pi_{q_n}(t) e^{iq_n x} + \pi_{-q_n}(t) e^{-iq_n x}. \end{aligned} \quad (\text{A.8})$$

Here $q_n = 2\pi n/L$. Inserting these expansions into A.7 gives the Hamiltonian as a Fourier series:

$$H = \sum_{n=0}^{\infty} \frac{1}{2\rho} \pi_{q_n} \pi_{-q_n} + B q^2 \phi_{q_n} \phi_{-q_n} = \sum_{n=0}^{\infty} \frac{\hbar \omega_q}{2} \left(\frac{\pi_{q_n} \pi_{-q_n}}{\hbar \rho \omega_{q_n}} + \frac{\rho \omega_{q_n} \phi_{q_n} \phi_{-q_n}}{\hbar} \right), \quad (\text{A.9})$$

where $\omega_{q_n} = v|q_n|$ (see after A.5). From now on we will drop the subscript n on q and remember that, for our periodic boundary conditions, $q = 2\pi n/L$.

Introducing scaled fields $\pi'_q = \pi_q / \sqrt{\hbar \rho \omega_q}$ and $\phi'_q = \sqrt{\rho \omega_q / \hbar} \phi_q$, Eq A.9 is in the form of 1.18, but with an infinite sum over wavevectors $q = 2\pi n/L$. Noting that, as we know that the field is real-valued, $\phi_{-q_n}(t) = \phi_{q_n}^*(t)$ and $\pi_{-q_n}(t) = \pi_{q_n}^*(t)$. This leads to the quantised Hamiltonian for the field theory, (using 1.22 and making the substitutions: $n \rightarrow q$, $x_l \rightarrow \phi(x, t)$, $\phi(x, t) \rightarrow \hat{\phi}(x, t)$):

$$\begin{aligned} \hat{\phi}(x, t) &= \sum_q \sqrt{\frac{\hbar}{2\rho\omega_q}} \hat{a}(q) e^{i(qx - \omega_q t)} + \sqrt{\frac{\hbar}{2\rho\omega_q}} \hat{a}^\dagger(q) e^{-i(qx - \omega_q t)} \\ \hat{\pi}(x, t) &= \sum_q (-i) \sqrt{\frac{\hbar}{2\rho\omega_q}} \hat{a}(q) e^{i(qx - \omega_q t)} + (-i) \sqrt{\frac{\hbar}{2\rho\omega_q}} \hat{a}^\dagger(q) e^{-i(qx - \omega_q t)}. \end{aligned} \quad (\text{A.10})$$

We have included the time-dependence with the creation and annihilation operators, see Section 1.2. (We know that the creation and annihilation operators have the time-dependence $\sim e^{\pm i\omega_a t}$ - see question sheet.) Introducing the commutators for the creation and annihilation operators (the \hat{a} and \hat{a}^\dagger) is equivalent to canonical quantisation which assumes the equal time commutator $[\hat{\phi}(x, t), \hat{\pi}(y, t)] = \delta(x - y)$.

The commutation relations are

$$[\hat{a}_q, \hat{a}_{q'}^\dagger] = \delta_{q, q'}. \quad (\text{A.11})$$

The Hamiltonian is again in the form of 1.15:

$$\hat{H} = \sum_q \hbar \omega_q \left(\hat{a}_q^\dagger \hat{a}_q + \frac{1}{2} \right), \quad (\text{A.12})$$

as expected.

The field theory gives the same answer as the microscopic theory in the regime (long wavelengths or small wavenumber) it was set up to describe. The allowed wavevectors are the same in the microscopic theory and field theory, provided that $|n| < L/2a$ (there is no upper limit to $|n|$ in the field theory, while there is in the microscopic theory). The frequency of oscillations in the field theory is $\omega_q = v|q|$, where $v = \sqrt{B/\rho}$. Using the definition of B , this means $v = \sqrt{ka/(m/a)} = a\sqrt{k/m}$. (B , ρ and ω_q are defined in A.1, A.2 and A.7.) The frequency of the oscillations in the microscopic theory is $\omega_n = 2\sqrt{k/m}|\sin q_n a/2|$. In the limit of long wavelengths, $qa \rightarrow 0$, $\omega_n \approx aq_n\sqrt{k/m} \rightarrow vq$ which coincides with the result from the field theory.

Why introduce a field theory, if we have a complete microscopic theory? Well, it's a useful model and, if we are only interested in long wavelengths, helps understand phenomena. Classically, of course, we almost always work in the long-wavelength limit (we do not 'see' individual ions). This is the field of elasticity and of great importance both in mathematics and engineering as well as physics. The long wavelength limit is completely defined by a system's elastic constants (B here). In 3D systems, the elastic constants are tensors and characterise the relation between the stress and strain which are both second rank tensors.

At the quantum level, there are other considerations. Again, if we are interested only in long wavelengths, it can help to work with models characterised by a small number of parameters (in this case B) which we fit to experiment. It is also often easier to work with effective theories of the continuum which can have continuous symmetries. The models can be more general in the sense that they can be used to describe other systems than the simple microscopic system we have derived it for.

Appendix B

Relativistic Quantum Fields—Bosons

The creation and annihilation operators generalise to the quantisation of classical fields involving bosons. In fact, we can guess what the corresponding Hamiltonians will be for a few systems and identify where the problems may be. Let's have a go.

The sources I have used have been mainly [Eduardo Fradkin](#) and [David Tong](#).

B.1 Klein-Gordon field

Once the dust settled in the development of quantum theory, it became clear that the KG field described bosons not electrons or other fermions. We will look only at the real-valued scalar field. The second quantised form for the Klein-Gordon equation in 3D (infinite domain) might be expected to be

$$\hat{H} = \frac{1}{2} \sum_{\mathbf{p}} \hbar \omega_{\mathbf{p}} (\hat{a}_{\mathbf{p}}^\dagger \hat{a}_{\mathbf{p}} + \hat{a}_{\mathbf{p}} \hat{a}_{\mathbf{p}}^\dagger), \quad (\text{B.1})$$

where

$$\hbar \omega_{\mathbf{p}} = \sqrt{p^2 c^2 + m_0^2 c^4} \quad \text{and} \quad [\hat{a}_{\mathbf{p}}, \hat{a}_{\mathbf{p}'}^\dagger] = \delta_{\mathbf{p}, \mathbf{p}'}. \quad (\text{B.2})$$

We should convert the sum over momenta to an integral over momenta. This would give

$$\hat{H} = \frac{1}{2} \int \frac{d^3 p}{(2\pi)^3} \hbar \omega_{\mathbf{p}} (\hat{a}_{\mathbf{p}}^\dagger \hat{a}_{\mathbf{p}} + \hat{a}_{\mathbf{p}} \hat{a}_{\mathbf{p}}^\dagger) = \int \frac{d^3 p}{(2\pi)^3} \hbar \omega_{\mathbf{p}} \left(\hat{a}_{\mathbf{p}}^\dagger \hat{a}_{\mathbf{p}} + \frac{1}{2} (2\pi)^3 \delta^{(3)}(0) \right). \quad (\text{B.3})$$

Here we have assumed the commutator is

$$[\hat{a}_{\mathbf{p}}, \hat{a}_{\mathbf{p}'}^\dagger] = (2\pi)^3 \delta^{(3)}(\mathbf{p} - \mathbf{p}'). \quad (\text{B.4})$$

The operator, $\hat{a}_{\mathbf{p}}^\dagger \hat{a}_{\mathbf{p}}$, is now a “density in momentum” operator (integrating over momentum gives a number).

Eqs B.3 and B.4 would be a natural generalisation of B.1 and B.2 to the case of continuous momenta—the sum over \mathbf{p} becomes an integral, the Kronecker delta becomes a delta function. The relation of the fields, $\hat{\phi}(\mathbf{x}, t)$ and $\hat{\pi}(\mathbf{x}, t)$, to the creation and annihilation

operators, with this normalisation for the commutation relations, is (using natural units, $\hbar = 1$, $c = 1$)

$$\begin{aligned}\hat{\phi}(\mathbf{x}, t) &= \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2\omega_p}} \left(\hat{a}_{\mathbf{p}} e^{i(\mathbf{p}\cdot\mathbf{x} - \omega_p t)} + \hat{a}_{\mathbf{p}}^\dagger e^{-i(\mathbf{p}\cdot\mathbf{x} - \omega_p t)} \right) \quad \text{and} \\ \hat{\pi}(\mathbf{x}, t) &= \int \frac{d^3p}{(2\pi)^3} (-i) \sqrt{\frac{\omega_p}{2}} \left(\hat{a}_{\mathbf{p}} e^{i(\mathbf{p}\cdot\mathbf{x} - \omega_p t)} - \hat{a}_{\mathbf{p}}^\dagger e^{-i(\mathbf{p}\cdot\mathbf{x} - \omega_p t)} \right).\end{aligned}\tag{B.5}$$

The ground-state of the system satisfies

$$\hat{a}_{\mathbf{p}}|0\rangle = 0 \quad \forall \mathbf{p}.\tag{B.6}$$

The factors of $\sqrt{\omega_p}$ in B.5 are factors needed to match the dimensions correctly ($\sqrt{\omega_p}\hat{\phi}(\mathbf{x}, t)$ and $\hat{\pi}(\mathbf{x}, t)/\sqrt{\omega_p}$ have the same dimensions as they should). Similar factors appeared in the scaling from $\hat{\phi} \rightarrow \hat{\phi}'$ in A.10. The operators in B.5 are time-dependent and we are using the Heisenberg representation.

This formulation is used (see eqn 2.84 here for example). However, it is not the most commonly used definition of the KG QFT. As we discuss below, most authors prefer to normalise the creation and annihilation operators differently, so that $\hat{a}_{\mathbf{p}}^\dagger \hat{a}_{\mathbf{p}}$ is not a density operator. This is to keep the theory explicitly Lorentz invariant. We will look at this after discussing briefly where photons are hiding in Maxwell's equations.

Added for Interest The real-valued KG field describes a wave system with a mass gap equal to $m_0 c^2$. It could be used to model the propagation of a neutral spin zero boson, like the Higgs boson or the neutral pion. The complex-valued KG field, $\phi(x, t) \neq \phi^*(x, t)$, when quantised, leads to a doublet of fields. There are two creation and two annihilation operators. One describes negatively charged particles and the other their antiparticles which are positively charged. The model could be used to model the propagation of the charged pions.

B.2 Electromagnetic field

Photons as particles need to be built into electromagnetism. The Lagrangian density for EM fields is (in the absence of sources, which means no charge or current densities)

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu}, \quad \text{where } F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu,\tag{B.7}$$

and $A_\mu = (\phi, -\mathbf{A})$ is the (covariant) four potential. The corresponding Euler-Lagrange equations are Maxwell's equations. We have taken $c = 1$. This suggests that A might become the quantum field (corresponding to ϕ in the other QFTs above).

The field A has four components. Naively, we might expect four different particles. On the other hand, we know from experiment (and from what we are told by our teachers) that photons are spin-1 particles and that there are only two polarisations. What is going on? First, there is no term in B.7 depending on $\partial_0 A_0$ and hence no momentum corresponding to A_0 . We say A_0 is not a dynamical variable. This means that it is fixed by \mathbf{A} .

Background The quantity $F_{\mu\nu}$ is a tensor, whose components are the electric and magnetic fields:

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E^1 & -E^2 & -E^3 \\ E^1 & 0 & -B^3 & B^2 \\ E^2 & B^3 & 0 & -B^1 \\ E^3 & -B^2 & B^1 & 0 \end{pmatrix}. \quad (\text{B.8})$$

The conjugate momentum to A^i is $\partial\mathcal{L}/\partial\dot{A}^i = E^i$. The momentum field will be \mathbf{E} .

We are now down to three independent components. Then we remember that the three components of \mathbf{A} are not wholly independent. Any gauge transformation, $\mathbf{A} \rightarrow \mathbf{A} + \nabla\chi$ where χ is an arbitrary function with well-defined gradient, leaves the physics unchanged. Variations in \mathbf{A} which are gauge transformation cannot describe excitations as they describe an unaltered physical state. After imposing gauge invariance (identifying all configurations that differ only by a change of gauge as equivalent) we can expect a reduction in the number of degrees of freedom. It turns out that we end up with just two independent polarisations. This still leaves us with what to do about the function χ . The usual approach is to choose a particular gauge. Here, we will adopt the Coulomb gauge condition:

$$\nabla \cdot \mathbf{A} = 0. \quad (\text{B.9})$$

We can guess the Hamiltonian for photons: ($\hbar = 1, c = 1, \Rightarrow \hbar\omega_{\mathbf{p}} = |\mathbf{p}| \equiv p_0$)

$$\hat{H} = \int \frac{d^3p}{(2\pi)^3} \sum_{s=1}^2 p_0 \left(\hat{a}_{\mathbf{p}}^{s\dagger} \hat{a}_{\mathbf{p}}^s + \frac{1}{2} (2\pi)^3 \delta^{(3)}(0) \right), \quad (\text{B.10})$$

with

$$[\hat{a}_{\mathbf{p}}^s, \hat{a}_{\mathbf{p}'}^{s'\dagger}] = (2\pi)^3 \delta^{s,s'} \delta(\mathbf{p} - \mathbf{p}'), \quad [\hat{a}_{\mathbf{p}}^s, \hat{a}_{\mathbf{p}'}^{s'}] = [\hat{a}_{\mathbf{p}}^{s\dagger}, \hat{a}_{\mathbf{p}'}^{s'\dagger}] = 0, \quad (\text{B.11})$$

$$\hat{a}_{\mathbf{p}}^s |0\rangle = 0, \quad \forall \mathbf{p}. \quad (\text{B.12})$$

We say that the operators, $\hat{a}_{\mathbf{p}}^{s\dagger}$ and $\hat{a}_{\mathbf{p}}^s$, create and annihilate photons with polarisation s . The vacuum (ground state) $|0\rangle$ is annihilated by all $\hat{a}_{\mathbf{p}}^s$. Again the “guess” B.10 is correct. The two polarisations, $s = 1, 2$, can be related to the projection of the spin of the photon, which is parallel or antiparallel to \mathbf{p} .

Added for Interest If the photons have spin one why are there not three polarisations? The idea, that angular momentum has $(2s + 1)$ different projections if the spin angular momentum quantum number is s , is associated with rotational invariance. Normally there is rotational invariance when looked at in the rest frame of a particle, which is how spin is defined. Photons have no rest frame. There is not complete rotational invariance, as the direction of travel of the photon is a special direction. If gauge-invariance is lost (Higgs mechanism), then the photon can have a mass and the third polarisation is recovered as an excitation with non-zero mass.

We also need to know how to construct the fields corresponding to \mathbf{A} and momentum π . The conjugate momentum field to \mathbf{A} is actually the electric field. We will state the results

for completeness sake: ($p_0 = |\mathbf{p}|$, $c = 1$)

$$\begin{aligned}\hat{\mathbf{A}}(\mathbf{x}, t) &= \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2p_0}} \sum_{s=1}^2 \boldsymbol{\epsilon}_p^s \left(\hat{a}_p^s e^{i(\mathbf{p}\cdot\mathbf{x}-p_0t)} + \hat{a}_p^{s\dagger} e^{-i(\mathbf{p}\cdot\mathbf{x}-p_0t)} \right) \quad \text{and} \\ \hat{\mathbf{E}}(\mathbf{x}, t) &= \int \frac{d^3p}{(2\pi)^3} (-i) \sqrt{\frac{p_0}{2}} \sum_{s=1}^2 \boldsymbol{\epsilon}_p^s \left(\hat{a}_p^s e^{i(\mathbf{p}\cdot\mathbf{x}-p_0t)} - \hat{a}_p^{s\dagger} e^{-i(\mathbf{p}\cdot\mathbf{x}-p_0t)} \right).\end{aligned}\tag{B.13}$$

The vectors $\boldsymbol{\epsilon}_p^s$ are normalised and orthogonal polarisation vectors. In the Coulomb gauge ($\nabla \cdot \mathbf{A} = 0$) it is clear that these polarisation vectors are perpendicular to \mathbf{p} as they should be—EM waves are transverse waves.

The operators in B.13 are time-dependent. They have been written in the Heisenberg representation.

The exact form of these equations will not be important for us here, but we should mention the question of Lorentz invariance. Both the KG and the EM fields must describe the physics correctly in all inertial frames. That does not mean that the form of all equations should be explicitly Lorentz invariant—have the same form before and after a Lorentz transformation—but the results should not depend on the choice of reference frame.

B.3 Lorentz and gauge invariance

It is usually better to work with Lorentz invariant forms. It reduces the risk of coming up with some apparent result which is just an artefact of working in one reference frame or another. Something similar applies in the case of gauge-invariance: the results for physical predictions should not depend on the choice of gauge.

Here we have used a form for the integrals in for example B.3 and B.10 in which the measure is not Lorentz invariant. An alternative is to write the integrals over the four momentum using a delta-function to impose the correct dispersion:

$$\int d^4p \frac{1}{(2\pi)^4} \delta^{(4)}(p_0^2 - \omega_p^2) \cdots = \int d^3\mathbf{p} \frac{1}{(2\pi)^3} \frac{1}{2\omega_p} \cdots.\tag{B.14}$$

Remember $\omega_p^2 = |\mathbf{p}|^2 + m_0^2$ with $m_0 = 0$ for the EM case. The factor $2\omega_p$ is there as the derivative of the argument of the delta-function for p_0 . This measure is Lorentz invariant. The argument of the 4D delta-function is the Lorentz invariant, $p_0^2 - p^2 c^2 - m_0^2 c^4$ (written in natural units with $c = 1$), while the Jacobian of a Lorentz transformation is always 1.

If the operator \hat{a}_p is scaled, $\hat{a}_p \rightarrow \sqrt{2\omega_p} \hat{a}_p$, with respect to the one introduced in B.3, the fields $\hat{\phi}$ in in B.5 are given by

$$\begin{aligned}\hat{\phi}(\mathbf{x}, t) &= \int \frac{d^3p}{(2\pi)^3} \frac{1}{2\omega_p} \left(a_p e^{i(\mathbf{p}\cdot\mathbf{x}-\omega_p t)} + a_p^\dagger e^{-i(\mathbf{p}\cdot\mathbf{x}-\omega_p t)} \right) \\ \hat{\pi}(\mathbf{x}, t) &= \int \frac{d^3p}{(2\pi)^3} (-i) \frac{1}{2\omega_p} \omega_p \left(a_p e^{i(\mathbf{p}\cdot\mathbf{x}-\omega_p t)} - a_p^\dagger e^{-i(\mathbf{p}\cdot\mathbf{x}-\omega_p t)} \right).\end{aligned}\tag{B.15}$$

These expressions include the Lorentz invariant measure. The commutator involving the scaled operators becomes

$$[\hat{a}_{\mathbf{p}}^s, \hat{a}_{\mathbf{p}'}^{s'\dagger}] = (2\pi)^3 2\omega_{\mathbf{p}} \delta^{s,s'} \delta(\mathbf{p} - \mathbf{p}'). \quad (\text{B.16})$$

With the same scaling of the photon creation and annihilation operators $\hat{a}_{\mathbf{p}}^s \rightarrow \sqrt{2p_0} \hat{a}_{\mathbf{p}}^s$, we obtain the more usual representation of the quantised EM field operators:

$$\begin{aligned} \hat{\mathbf{A}}(\mathbf{x}, t) &= \int \frac{d^3p}{(2\pi)^3} \frac{1}{2p_0} \sum_{s=1}^2 \boldsymbol{\epsilon}_{\mathbf{p}}^s \left(\hat{a}_{\mathbf{p}}^s e^{i(\mathbf{p}\cdot\mathbf{x} - p_0 t)} + \hat{a}_{\mathbf{p}}^{s\dagger} e^{-i(\mathbf{p}\cdot\mathbf{x} - p_0 t)} \right) \quad \text{and} \\ \hat{\mathbf{E}}(\mathbf{x}, t) &= \int \frac{d^3p}{(2\pi)^3} (-i) \frac{1}{2p_0} p_0 \sum_{s=1}^2 \boldsymbol{\epsilon}_{\mathbf{p}}^s \left(\hat{a}_{\mathbf{p}}^s e^{i(\mathbf{p}\cdot\mathbf{x} - p_0 t)} - \hat{a}_{\mathbf{p}}^{s\dagger} e^{-i(\mathbf{p}\cdot\mathbf{x} - p_0 t)} \right). \end{aligned} \quad (\text{B.17})$$

With this normalisation, the commutation relation B.11 and the Hamiltonian become

$$\begin{aligned} [\hat{a}_{\mathbf{p}}^s, \hat{a}_{\mathbf{p}'}^{s'\dagger}] &= (2\pi)^3 2p_0 \delta^{s,s'} \delta(\mathbf{p} - \mathbf{p}') \\ \hat{H} &= \int \frac{d^3p}{(2\pi)^3} \frac{1}{2p_0} \sum_{s=1}^2 p_0 \left(\hat{a}_{\mathbf{p}}^{s\dagger} \hat{a}_{\mathbf{p}}^s + \frac{1}{2} (2\pi)^3 \delta^{(3)}(0) \right). \end{aligned} \quad (\text{B.18})$$

Note the extra factor of $2p_0$ in the commutator, and $1/2p_0$ in the Hamiltonian, with respect to B.11 and B.10. This reflects the different normalisation of the creation and annihilation operators when working with the explicitly Lorentz invariant measure of B.14. In most textbooks, these factors of $1/2\omega_p$ or $1/2p_0$ are commonly found in the measure for formulations of relativistic QFTs. This is why they are presented here in yellow. But be careful to check the normalisation of the creation and annihilation operators before using any formulae.

Lorentz invariance also affects the gauge choice. In setting up B.10 we used the Coulomb gauge. This is of course not Lorentz invariant as a Lorentz transformation mixes time and space and the condition $\nabla \cdot \mathbf{A} = 0$ will contain time derivatives in other inertial frames. The theory can be set up (and often is) in the Lorenz gauge, $\partial_\mu A^\mu = 0$, which is manifestly gauge-invariant. In the Lorenz gauge, finding the two independent polarisations is more complicated than in the Coulomb gauge, while the propagator for photons is easier in the Lorenz gauge. In practice, it is more common to work in the Lorenz gauge.

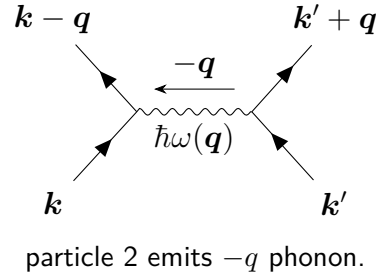
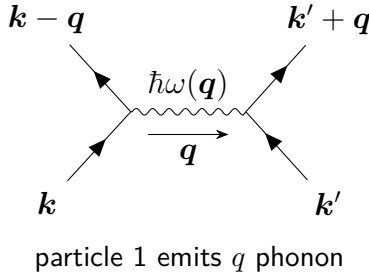
Appendix C

Attractive Interaction from Phonon Exchange

The effective interaction between electrons at the Fermi surface in a superconductor, see 3.14, was taken to be attractive. Where does this come from? Well, it was known that the transition temperature below which a material was superconducting, $T_c \propto M^{-1/2}$. Here M was the mass of the ions in the material. This mass could only appear if the frequency of oscillations of the ions (so-called phonons) was involved. (In 1.18 we found that $\omega_n \propto m^{-1/2}$ for a 1D chain of ions.) This dependence on the mass is generic for all lattice oscillations.

The following argument is based on the one in “Fundamentals of the the Theory of Metals” by Abrikosov.

Now consider the two processes shown schematically



In both cases, electron 1 is scattered to $\mathbf{k} - \mathbf{q}$ while electron 2 is scattered to $\mathbf{k}' + \mathbf{q}$. Energy conservation gives $\epsilon(\mathbf{k}) + \epsilon(\mathbf{k}') = \epsilon(\mathbf{k} - \mathbf{q}) + \epsilon(\mathbf{k}' + \mathbf{q})$. According to second order perturbation theory, the amplitude for the scattering processes are

$$1 : \frac{|U_q|^2}{\epsilon(\mathbf{k}) - \epsilon(\mathbf{k} - \mathbf{q}) - \hbar\omega(\mathbf{q})} \quad \text{and} \quad 2 : \frac{|U_q|^2}{\epsilon(\mathbf{k}') - \epsilon(\mathbf{k}' + \mathbf{q}) - \hbar\omega(\mathbf{q})}.$$

We are assuming that $\omega(\mathbf{q})$ and U_q are the same for \mathbf{q} and $-\mathbf{q}$.

The overall scattering amplitude (in second order) is the sum of these two terms. Note that because of energy conservation $\epsilon(\mathbf{k}) - \epsilon(\mathbf{k} - \mathbf{q}) = -\epsilon(\mathbf{k}') + \epsilon(\mathbf{k}' + \mathbf{q})$. This gives the scattering amplitude

$$-\frac{2|U_q|^2\hbar\omega(\mathbf{q})}{[\hbar\omega(\mathbf{q})]^2 - [\epsilon(\mathbf{k}) - \epsilon(\mathbf{k} - \mathbf{q})]^2}. \quad (\text{C.1})$$

Superconductivity involves the interaction between particles close to the Fermi energy, in states \mathbf{k} and $-\mathbf{k}$. The energy difference, $[\epsilon(\mathbf{k}) - \epsilon(\mathbf{k} - \mathbf{q})]$ should be small compared to $\hbar\omega(\mathbf{q})$ as the transition temperatures are generally much smaller than typical phonon frequencies. The amplitude is $\sim -|U_q|^2/\hbar\omega_q$ in such cases. This is independent of \mathbf{k} and is the same scattering amplitude one would obtain for an interaction $-U\delta(\mathbf{r} - \mathbf{r}')$ with $V > 0$, but remember the interaction is restricted to states close to the Fermi surface. Note that the matrix element U_q is between normalised wavefunctions of the type, $\psi_{\mathbf{k}} = \frac{e^{i\mathbf{k}\cdot\mathbf{r}}}{\sqrt{V}}$. The two factors of $\frac{1}{\sqrt{V}}$ are the origin of the factor $1/V$ in 3.6.

This interaction involves the exchange of phonons. As the density of states of the phonons is highest for large \mathbf{q} , and, as we are principally interested in phonon scattering across the Fermi surface, we should take the typical phonon energy to be that of large q phonons which is roughly the Debye frequency, $\hbar\omega_D$.

Finally, as mentioned in the main text, we think of the interaction as retarded. It acts on a time-scale $\Delta t \sim \omega_D^{-1}$, which is long on the scale of electronic interactions \hbar/ϵ_F .

Appendix D

Spin Interactions

You will sometimes see Hamiltonians involving spin operators. This is particularly the case in problems related to magnetism. The simplest of these is the nearest neighbour Heisenberg exchange interaction for spins on a lattice

$$\hat{H} = \frac{J}{2} \sum_{\langle i,j \rangle}^N \hat{\mathbf{S}}_i \cdot \hat{\mathbf{S}}_j. \quad (\text{D.1})$$

Normally we will assume periodic boundary conditions $\hat{\mathbf{S}}_{N+1} \equiv \hat{\mathbf{S}}_1$. The quantity J is called the exchange parameter. It gives the interaction between two spins (the factor of $1/2$ is there to avoid double-counting in the sum). The brackets $\langle \dots \rangle$ (not to be confused with $\langle \dots \rangle$ used to indicate expectation values) indicate that only sites, which are nearest neighbours on the lattice, are to be summed over. Models can also have unrestricted sums over lattice sites with the exchange integral, J_{ij} , dependent on position coordinates i and j . In interacting problems, the spin operators are usually defined in units of \hbar , so that the exchange parameter is an energy. Positive values of J in D.1 lead to anti-alignment of spins (antiferromagnetism) on neighbouring sites, $\langle \hat{\mathbf{S}}_i \cdot \hat{\mathbf{S}}_j \rangle < 0$, in the ground and low energy states. The case $J > 0$ leads to alignment of spins (ferromagnetism).

The eigenstates of two $s = 1/2$ spins interacting via an exchange interaction are the singlet, with total spin zero, $|s\rangle$, and the three states with total spin one:

$$\begin{aligned} |s_{12}\rangle &= \frac{1}{\sqrt{2}} (|\uparrow_1\downarrow_2\rangle - |\downarrow_1\uparrow_2\rangle) \\ |t_{12}\rangle &= \frac{1}{\sqrt{2}} (|\uparrow_1\downarrow_2\rangle + |\downarrow_1\uparrow_2\rangle), \quad |t_{12}\rangle_1 = |\uparrow_1\uparrow_2\rangle, \quad |t_{12}\rangle_{-1} = |\downarrow_1\downarrow_2\rangle. \end{aligned} \quad (\text{D.2})$$

The subscript ± 1 on $|t_{12}\rangle$ denotes that the projection on the z-direction of the total spin. The three triplet states are degenerate with energy $J/4$, while the singlet state has energy $-3J/4$.

The idea behind D.1 is that the spin on site i interacts with the spin on site j . These need not be single electron states nor spin- $1/2$ states. They could be multiple electron states with well-defined spin. An example might be a Gd^{3+} ion in some crystal or ligand complex. Its outer shell has 7 4f electrons. Hund's rules imply that it should have total spin $7/2$ (the maximum possible) and total angular momentum zero. (If the spins are all aligned, the spin

wavefunction is symmetric, while the occupation of the $2l + 1 = 7$ orbitals must be in an antisymmetric combination.) In fact, this large value of the total spin of ($S = 7/2$) with zero angular momentum ($L = 0$) makes Gd well-suited for use as an MRI image contrast agent. The large magnetic moment is preserved even when the ion is bonded to an organic ligand.

The interaction in D.1 usually has its origin in the Coulomb interaction. There are direct magnetic interactions but these are almost always smaller than indirect interactions arising out of the Coulomb effect. Generally, direct magnetic effects tend to be much smaller than electric ones when particles are moving at speeds $v \ll c$.

A model showing how an exchange interaction might come about considers a lattice with one available orbital per site. This orbital can be occupied by a spin up or spin down electron. There are then four states to consider at site i

$$|0\rangle, \quad |\uparrow_i\rangle = \hat{c}_{i\uparrow}^\dagger |0\rangle, \quad |\downarrow_i\rangle = \hat{c}_{i\downarrow}^\dagger |0\rangle, \quad |\uparrow_i\downarrow_i\rangle = \hat{c}_{i\uparrow}^\dagger \hat{c}_{i\downarrow}^\dagger |0\rangle. \quad (\text{D.3})$$

These states can be coupled to states on a neighbouring site, $i \pm 1$. We expect the Hamiltonian excluding any interaction effect \hat{H}_0 to have matrix elements

$$\epsilon_0 = \langle \uparrow_i | \hat{H}_0 | \uparrow_i \rangle = \langle \downarrow_i | \hat{H}_0 | \downarrow_i \rangle, \quad t = \langle \uparrow_i | \hat{H}_0 | \uparrow_{i+1} \rangle = \langle \uparrow_i | \hat{H}_0 | \uparrow_{i-1} \rangle. \quad (\text{D.4})$$

The matrix element t is usually called the hopping and describes the kinetic energy of electrons in the lattice. It can be taken to be real as here (unless there is a magnetic field applied). In second quantised notation, the non-interacting Hamiltonian is (assuming a 1D system for simplicity)

$$\hat{H}_0 = \text{constant} + \sum_{i,\sigma} t \hat{c}_{i\sigma}^\dagger \hat{c}_{i+1,\sigma} + h.c. \quad (\text{D.5})$$

Here the constant is $N\epsilon_0$ but it is not important and is usually dropped. The letters “h.c.” stand for hermitian conjugate (here the hermitian conjugate term is $t \hat{c}_{i+1}^\dagger \hat{c}_i$).

The Coulomb interaction will give rise to matrix elements between various states in the basis. In this model, there is only one orbital per site. The largest of the interaction matrix elements is expected to be the interaction between two electrons in the same orbital (the fourth state in the list D.3). We will take the energy of the doubly occupied state, $|\uparrow_i\downarrow_i\rangle$, to be $2\epsilon_0 + U$. If we ignore all other interaction terms we arrive at a model Hamiltonian

$$\hat{H} = \sum_{i,\sigma} t \left(\hat{c}_{i\sigma}^\dagger \hat{c}_{i+1,\sigma} + h.c. \right) + \sum_i^N U \hat{n}_{i\uparrow} \hat{n}_{i\downarrow}, \quad (\text{D.6})$$

where the product of the two number operators $\hat{n}_{i\uparrow} \hat{n}_{i\downarrow}$ is only non-zero when there are two electrons in the orbital.

Eq D.6 is the Hubbard Hamiltonian. It has not been solved analytically for arbitrary N except in the 1D case. Together with generalisations to include more than one orbital per site and additional interactions, it has become a standard model for discussing interaction effects in materials. The effects of the interactions will be more important when the orbitals are spatially compact leading to smaller values of t and larger values of U . This is the case in many transition metals and lanthanides. The 3d and 4f orbitals are the outer shell states but much of their high energy (large ϵ_0 , see D.4) is associated with angular momentum ($l = 2, 3$) not with the orbitals being far from the ionic nuclei.

In the case where there is one electron per site and the interaction energy $U \gg t$, one can expect all the sites to be occupied by 1 electron and no sites to be doubly occupied. Characterising the state of such a system reduces to specifying the spin degrees of freedom. Provided $U < \infty$, the possibility of doubly occupied sites has an effect on the orientation of the spins. Imagine adjacent sites (again think of 1D for simplicity) i and $i + 1$, and consider the cases where the two spins are aligned and where they are anti-aligned. The effect of the hopping term is different in the two cases:

$$\sum_{\sigma} \hat{c}_{i,\sigma}^{\dagger} \hat{c}_{i+1,\sigma} |\uparrow_i \uparrow_{i+1}\rangle = 0, \quad \text{while} \quad \sum_{\sigma} \hat{c}_{i,\sigma}^{\dagger} \hat{c}_{i+1,\sigma} |\uparrow_i \downarrow_{i+1}\rangle = |\uparrow_i \downarrow_i\rangle. \quad (\text{D.7})$$

In the non-aligned case the down spin electron hops from site $i + 1$ to site i .

We can find an effective spin Hamiltonian using second order perturbation theory. It is easiest to work with the singlet and triplet combinations, see D.2, and act on these with the Hamiltonian. You can check the singlet or triplet nature by action on the states with the spin raising operator, $\hat{S}^+ = \hat{S}_i^+ + \hat{S}_{i+1}^+$. This annihilates the singlet state, $|s_{12}\rangle$, and creates the $S_z = 1$ state when acting on the triplet state, $|t_{12}\rangle$. These states are given in D.2.

Acting with the hopping terms connecting sites i and $i + 1$ on the two states $|s\rangle$ and $|t\rangle$ gives:

$$\begin{aligned} \left(t \sum_{\sigma} \hat{c}_{i,\sigma}^{\dagger} \hat{c}_{i+1,\sigma} + \hat{c}_{i+1,\sigma}^{\dagger} \hat{c}_{i,\sigma} \right) |t_{i,i+1}\rangle &= 0 \\ \left(t \sum_{\sigma} \hat{c}_{i,\sigma}^{\dagger} \hat{c}_{i+1,\sigma} + \hat{c}_{i+1,\sigma}^{\dagger} \hat{c}_{i,\sigma} \right) |s_{i,i+1}\rangle &= t\sqrt{2} (|\uparrow_i \downarrow_i\rangle + |\uparrow_{i+1} \downarrow_{i+1}\rangle) \equiv 2t |d_{i,i+1}\rangle \end{aligned} \quad (\text{D.8})$$

(The state $|d_{i,i+1}\rangle$ denotes the normalised linear combination of the doubly occupied sites i and $i + 1$.) When operating with the fermionic creation and annihilation operators it is important to keep track of all anticommutations. The hopping term will take the state, $|\downarrow_i \uparrow_{i+1}\rangle$ to $|\downarrow_{i+1} \uparrow_i\rangle = -|\uparrow_{i+1} \downarrow_i\rangle$. This is why the triplet state is annihilated by the hopping. We should have expected this as the Hamiltonian is invariant under rotations in the spin space, and the effect on the $S_z = 0$ state in the triplet subspace must be the same as on the $S_z = \pm 1$ states.

The energy of the state on the right hand side of D.8 is $2\epsilon_0 + U$, which is U larger than the initial state with two singly occupied sites. We can use second order perturbation theory to compute the energies of the two states, $|s_{i,i+1}\rangle$ and $|t_{i,i+1}\rangle$. We are treating the interaction energy as the unperturbed part of the Hamiltonian and the hopping as the perturbation. This gives the energies

$$E_s^{(2)} = -\frac{|\langle s | \hat{H}_0 | d \rangle|^2}{E_d - E_s} = -4\frac{t^2}{U}, \quad E_t^{(2)} = 0. \quad (\text{D.9})$$

At the level of second order perturbation theory, the energy of the three states in the triplet subspace is $4t^2/U$ higher than that of the singlet state. Ignoring a constant term, this is equivalent to assuming that, in the subspace of singly-occupied states, the Hamiltonian for the system can be taken to be the exchange interaction given in D.1 with $J = 4t^2/U$.