Quantum Theory of Condensed Matter

John Chalker
Physics Department, Oxford University

2013

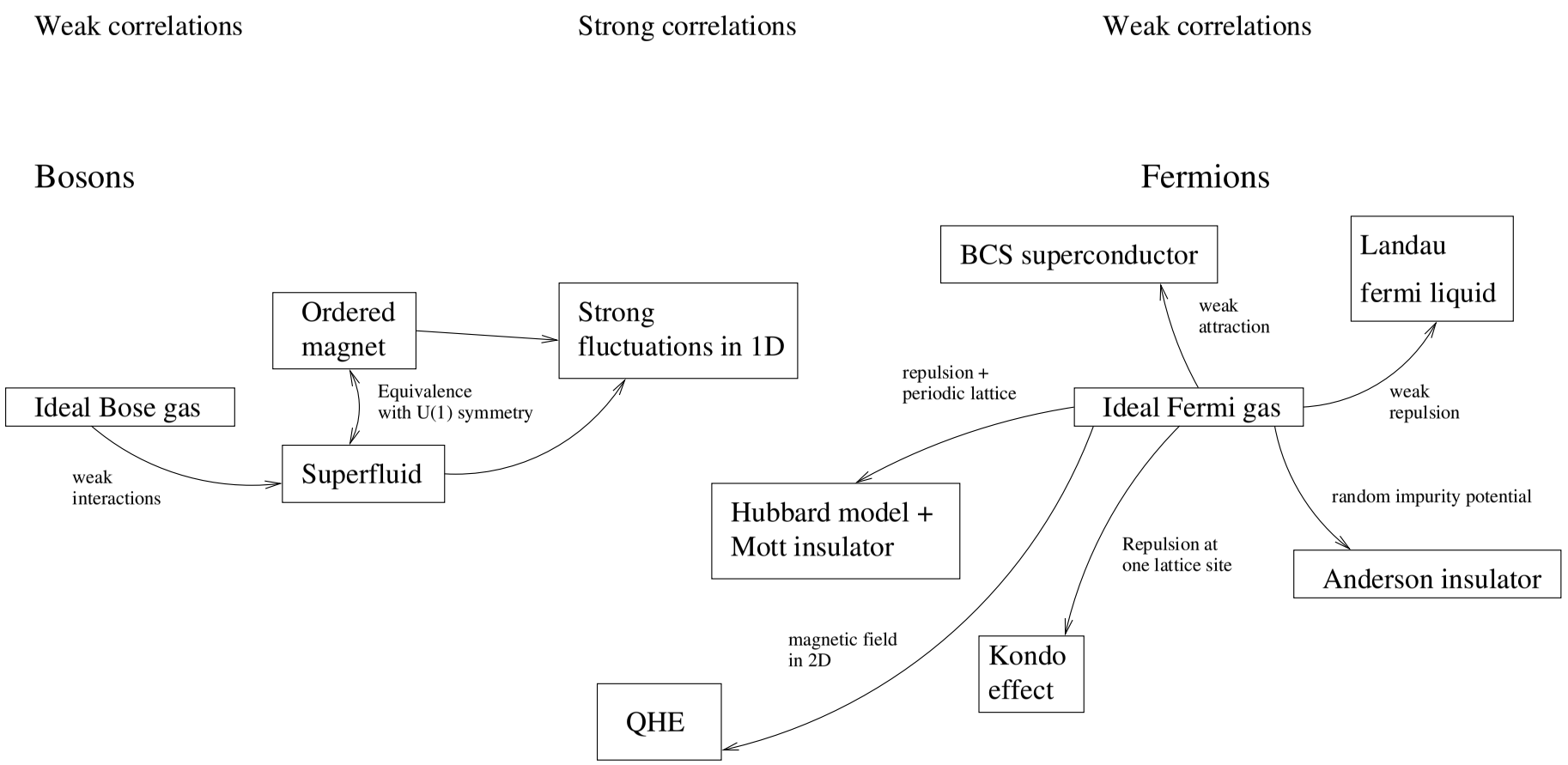
I aim to discuss a reasonably wide range of quantum-mechanical phenomena from condensed matter physics, with an emphasis mainly on physical ideas rather than mathematical formalism. The most important prerequisite is some understanding of second quantisation for fermions and bosons. There will be two problems classes in addition to the lectures.

Michaelmas Term 2013: Lectures in the Fisher Room, Dennis Wilkinson Building, Physics Department, on Wednesdays at 10:00 and Fridays at 11:00.

OUTLINE

- Overview
- Spin waves in magnetic insulators
- One-dimensional quantum magnets
- Superfluidity in a weakly interacting Bose gas
- Landau's theory of Fermi liquids
- BCS theory of superconductivity
- The Mott transition and the Hubbard model
- The Kondo effect
- Disordered conductors and Anderson localisation
- Anderson insulators
- The integer and fractional quantum Hall effects

Overall scheme



Bibliography

Background

N W Ashcroft and N D Mermin *Solid State Physics*, Holt-Sanders (1976). I assume familiarity with this material.

S-K Ma *Statistical Mechanics*, World Scientific (1985). Strongly recommended book at a level suitable for first year graduate students.

J-P Blaizot and G Ripka *Quantum Theory of Finite Systems*, MIT (1986). A very thorough treatment of second quantisation, canonical transformations and self-consistent field approximations.

Recent Graduate Texts

A. Altland and B. D. Simons *Quantum Field Theory in Condensed Matter Physics*, CUP (2006). An accessible introduction to the subject.

S. Sachdev *Quantum Phase Transitions*, CUP (1999).
An advanced survey of theoretical approaches to this subject.

H. Bruus and K. Flensberg *Many Body Quantum Theory in Condensed Matter Physics*, OUP (2004).
A detailed introduction to techniques and a discussion of topics of current interest, especially in connection with mesoscopic conductors and quantum dots.

X.-G. Wen *Quantum Field Theory of Many-Body Systems*, OUP (2004). An outline of basic material followed by an introduction to some advanced topics (topological order, the fractional quantum Hall effect, and spin liquids).

A. M. M. Tsvelik *Quantum Field Theory in Condensed Matter Physics*, CUP (1995). A concise survey of applications of field theory to condensed matter problems, especially in one dimension.

A. Auerbach *Interacting Electrons and Quantum Magnetism*, Springer (1994).
A reasonably gentle introduction to a range of current theoretical ideas.

General Texts

P W Anderson *Concepts in Solids*, Benjamin (1963). A classic introduction to solid state physics at a graduate level.

C Kittel *Quantum Theory of Solids*, Wiley (1963). [N.B. *not* the undergraduate text by the same author]. Includes most of the material covered in the first third of the lecture course.

D. Pines and P. Nozieres *The Theory of Quantum Liquids, Volume 1*, Addison Wesley (1989). A standard account of Fermi liquids.

P W Anderson *Basic Notions in Condensed Matter Physics*, Benjamin (1984). When first published was an advanced discussion of some of the most important ideas in the subject.

A. J. Leggett *Quantum Liquids*, OUP (2006). A clear and wide-ranging discussion of Bose condensation and Cooper pairing.

Electrons in disordered conductors.

N F Mott *Conduction in Non-Crystalline Materials* O.U.P. (1987) For a survey.

P. A. Lee and B. I. Altshuler, Physics Today **41** 36 (1988), and P.A. Lee and S-C Feng, Science **251** 633 (1991). Two accessible introductions to interference effects on transport in disordered media.

Y. Imry *Introduction to Mesoscopic Physics* O. U. P. (1997). Also an article in *Directions in Condensed Matter Physics*, Edited by G. Grinstein and G Mazenko, World Scientific (1986). A more advanced discussion, but in the same spirit as Lee and Altshuler's article.

Some consequences of electron-electron interactions.

N F Mott *Metal Insulator Transitions*, Taylor and Francis (1990), and A. C. Hewson *The Kondo Problem to Heavy Fermions* Cambridge (1993). Two complementary reviews.

The Quantum Hall Effect.

R. E. Prange and S. M. Girvin *The Quantum Hall Effect*, Springer (1990). A standard introduction, now rather old.

S. M. Girvin *The Quantum Hall Effect: Novel Excitations and Broken Symmetries* Lectures delivered at Ecole d'Ete Les Houches, July 1998; cond-mat/9907002
A more recent review of quantum Hall physics.

Green functions, response functions and perturbation theory

A. A. Abrikosov, L. P. Gorkov and I. E. Dzyaloshinski *Methods of quantum field theory in statistical physics*, Dover (1975). Still possibly the best starting point.

A. L. Fetter and J. D. Walecka *Quantum Theory of Many-Particle Systems*, McGraw-Hill, (1971); also available from Dover. A standard and straightforward introduction.

J. W. Negele and H. Orland *Quantum Many-Particle Systems*, Addison Wesley (1987). A modern treatment based on path integrals.

G. D. Mahan *Many-Particle Physics*, Plenum (1990).
A very detailed treatment of Green function techniques in many body theory.

Many-Particle Quantum Systems

1 Identical particles in quantum mechanics

Many-particle quantum systems are always made up of many *identical particles*, possibly of several different kinds. Symmetry under exchange of identical particles has very important consequences in quantum mechanics, and the formalism of many-particle quantum mechanics is designed to build these consequences properly into the theory. We start by reviewing these ideas.

Consider a system of N identical particles with coordinates $\mathbf{r}_1, \ldots \mathbf{r}_N$ described by a wavefunction $\psi(\mathbf{r}_1 \ldots \mathbf{r}_N)$. For illustration, suppose that the Hamiltonian has the form

$$\mathcal{H} = -\frac{\hbar^2}{2m}\sum_{i=1}^{N}\nabla_i^2 + \sum_{i=1}^{N}V(\mathbf{r}_i) + \sum_{i<j}U(\mathbf{r}_i - \mathbf{r}_j)\ .$$

Here there are three contributions to the energy: the kinetic energy of each particle (∇_i^2 operates on the coordinates $\mathbf{r}_i$); the one-body potential energy $V(\mathbf{r})$; and the two-particle interaction potential $U(\mathbf{r}_i - \mathbf{r}_j)$. To discuss symmetry under exchange of particles, we define the exchange operator $\mathcal{P}_{ij}$ via its action on wavefunctions:

$$\mathcal{P}_{ij}\psi(\ldots\mathbf{r}_i\ldots\mathbf{r}_j\ldots) = \psi(\ldots\mathbf{r}_j\ldots\mathbf{r}_i\ldots)\ .$$

Since $[\mathcal{H}, \mathcal{P}_{ij}] = 0$, we can find states that are simultaneous eigenstates of $\mathcal{H}$ and $\mathcal{P}_{ij}$. Moreover, a system that is initially in an eigenstate of $\mathcal{P}_{ij}$ will remain in one under time evolution with $\mathcal{H}$. For these reasons we examine the eigenvalues of $\mathcal{P}_{ij}$. Since $(\mathcal{P}_{ij})^2 = 1$, these are $+1$ and -1. Now, it is an observational fact (explained in relativistic quantum field theory by the spin-statistics theorem) that particles come in two kinds and that particles of a given kind are always associated with the same eigenvalue of the exchange operator: $+1$ for bosons and -1 for fermions.

1.1 Many particle basis states

In a discussion of many-particle quantum systems we should restrict ourselves to wavefunctions with the appropriate symmetry under particle exchange. We can do this by using a set of basis states that has the required symmetry. As a starting point, suppose that we have a complete, orthonormal set of single-particle states $\phi_1(\mathbf{r}), \phi_2(\mathbf{r})\ldots$. Next we would like to write down a wavefunction representing an N-particle system with one particle in state l_1, one in state l_2 and so on. The choice

$$\phi_{l_1}(\mathbf{r})\phi_{l_2}(\mathbf{r})\ldots\phi_{l_N}(\mathbf{r})$$

is unsatisfactory because for general $l_1, l_2\ldots$ it has no particular exchange symmetry. Instead we take

$$\psi(\mathbf{r}_1\ldots\mathbf{r}_N) = \mathcal{N}\sum_{\text{distinct}\ \ \text{perms.}}(\pm1)^P\phi_{k_1}(\mathbf{r}_1)\ldots\phi_{k_N}(\mathbf{r}_N)\ . \tag{1}$$

Several aspects of the notation in Eq. (1) require comment. The sign inside the brackets in $(\pm1)^P$ is $+1$ for bosons and -1 for fermions. The set of labels $\{k_1\ldots k_N\}$ is a permutation of the set $\{l_1\ldots l_N\}$. The permutation is called *even* if it can be produced by an even number of exchanges of adjacent pairs of labels, and is *odd* otherwise; the integer P is even or odd accordingly. The sum is over all *distinct* permutations of the labels. This means that if two or more of the labels l_n are the same, then permutations amongst equal labels do not appear as multiple contributions to the sum. Finally, $\mathcal{N}$ is a normalisation, which we determine next.

To normalise the wavefunction, we must evaluate

$$\int \mathrm{d}^d\mathbf{r}_1\ldots\int \mathrm{d}^d\mathbf{r}_N\,\psi^*(\mathbf{r}_1\ldots\mathbf{r}_N)\psi\ \ (\mathbf{r}_1\ldots\mathbf{r}_N)\ .$$

Substituting from Eq. (1), we obtain a double sum (over permutations $k_1\ldots k_N$ and $h_1\ldots h_N$) of terms of the form

$$\int \mathrm{d}^d\mathbf{r}_1\,\phi^*_{k_1}(\mathbf{r})\phi_{h_1}(\mathbf{r}_1)\ldots\int \mathrm{d}^d\mathbf{r}_N\,\phi^*_{k_N}(\mathbf{r})\phi_{h_N}(\mathbf{r}_1)\ .$$

These terms are zero unless $k_1 = h_1$, $k_2 = h_2$, and $\ldots k_N = h_N$, in which case they are unity. Therefore only the diagonal terms in the double sum contribute, and we have

$$\int \ldots \int |\psi|^2 = |\mathcal{N}|^2 \sum_{\text{dist. perms.}} (\pm 1)^{2P} = |\mathcal{N}|^2 \, \frac{N!}{n_1! \, n_2! \, \ldots}$$

where the n_1, $n_2 \ldots$ are the numbers of times that each distinct orbital appears in the set $\{l_1 \ldots l_N\}$, and the ratio of factorials is simply the number of distinct permutations. Hence we normalise the wavefunction to unity by taking

$$\mathcal{N} = \left(\frac{n_1! \; n_2! \; \ldots}{N!} \right)^{1/2} .$$

1.2 Slater determinants

For fermion wavefunctions we can get the correct signs by thinking of Eq. (1) as a determinant

$$\psi(\mathbf{r}_1 \ldots \mathbf{r}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_{l_1}(\mathbf{r_1}) & \ldots & \phi_{l_1}(\mathbf{r_N}) \\ & \ldots & \\ \phi_{l_N}(\mathbf{r_1}) & \ldots & \phi_{l_N}(\mathbf{r_N}) \end{vmatrix} . \tag{2}$$

Note that this determinant is zero either if two orbitals are the same ($l_i = l_j$) or if two coordinates coincide ($\mathbf{r}_i = \mathbf{r}_j$), so the Pauli exclusion principle is correctly built in. Note also that, since the sign of the determinant is changed if we exchange two adjacent rows, it is necessary to keep in mind a definite ordering convention for the single particle orbitals $\phi_l(\mathbf{r})$ to fix the phase of the wavefunction.

For bosons, we should use an object similar to a determinant, but having all terms combined with a positive sign: this is known as a *permanent*.

1.3 Occupation numbers

We can specify the basis states we have constructed by giving the number of particles n_l in each orbital l. Clearly, for fermions $n_l = 0$ or 1, while for bosons $n_l = 0, 1, \ldots$. These occupation numbers are used within Dirac notation as labels for a state: $|n_1, n_2, \ldots\rangle$.

1.4 Fock space

Combining states $|n_1, n_2, \ldots\rangle$ with all possible values of the occupation numbers, we have basis vectors for states with any number of particles. This vector space is known as *Fock space*. Using it, we can discuss processes in which particles are created or annihilated, as well as ones with fixed particle number, described by wavefunctions of the form $\psi(\mathbf{r}_1 \ldots \mathbf{r}_N)$.

1.5 The vacuum state

It is worth noting that one of the states in Fock space is the vacuum: the wavefunction for the quantum system when it contains no particles, written as $|0\rangle$. Clearly, in recognising this as a quantum state we have come some way from the notation of single-body and few-body quantum mechanics, with wavefunctions written as functions of particle coordinates. Of course, $|0\rangle$ is different from 0, and in particular $\langle 0|0\rangle = 1$.

1.6 Creation and annihilation operators

Many of the calculations we will want to do are carried out most efficiently by introducing creation operators, which add particles when they act to the right on states from Fock space. Their Hermitian conjugates are annihilation operators, which remove particles. Their definition rests on the set of single particle orbitals from which we built Fock space: $c_l^\dagger$ adds particles to the orbital $\phi_l(\mathbf{r})$. More formally, we define

$$c_{l_1}^\dagger c_{l_2}^\dagger \ldots c_{l_N}^\dagger |0\rangle \tag{3}$$

to be the state with coordinate wavefunction

$$\psi(\mathbf{r}_1, \ldots \mathbf{r}_N) = \frac{1}{\sqrt{N!}} \sum_{\text{all perms}} (\pm 1)^P \phi_{k_1}(\mathbf{r}_1) \ldots \phi_{k_N}(\mathbf{r}_N) = (n_1! \, n_2! \, \ldots)^{1/2} |n_1, n_2 \ldots\rangle \; . \tag{4}$$

A detail to note is that the sum in Eq. (4) is over all permutations, while that in Eq. (1) included only distinct permutations. The difference (which is significant only for bosons, since it is only for bosons that we can have $n_l > 1$), is the reason for the factor $(n_1!\, n_2!\, \ldots)^{1/2}$ appearing on the right of Eq. (4). This choice anticipates what is necessary in order for boson creation and annihilation operators to have convenient commutation relations.

Annihilation operators appear when we take the Hermitian conjugate of Eq. (3), obtaining $\langle 0|\, c_{l_N} \ldots c_{l_2} c_{l_1}$. Let's examine the effect of creation and annihilation operators when they act on various states. Since $c_l^\dagger|0\rangle$ is the state with coordinate wavefunction $\phi_l(\mathbf{r})$, we know that $\langle 0|c_l\, c_l^\dagger|0\rangle = 1$, but for any choice of the state $|\phi\rangle$ other than the vacuum, $c_l^\dagger|\phi\rangle$ contains more than one particle and hence $\langle 0|c_l\, c_l^\dagger|\phi\rangle = 0$. From this we can conclude that

$$c_l\, c_l^\dagger|0\rangle = |0\rangle\ ,$$

demonstrating that the effect of c_l is to remove a particle from the state $|n_l{=}1\rangle \equiv c_l^\dagger|0\rangle$. We also have for any $|\phi\rangle$ the inner products $\langle 0|c_l^\dagger|\phi\rangle = \langle\phi|c_l|0\rangle = 0$, and so we can conclude that

$$c_l|0\rangle = \langle 0|c_l^\dagger = 0\ .$$

1.7 Commutation and anticommutation relations

Recalling the factor of $(\pm 1)^P$ in Eq. (4), we have for any $|\phi\rangle$

$$c_l^\dagger c_m^\dagger|\phi\rangle = \pm c_m^\dagger c_l^\dagger|\phi\rangle\ ,$$

where the upper sign is for bosons and the lower one for fermions. From this we conclude that boson creation operators commute, and fermion creation operators anticommute: that is, for bosons

$$[c_l^\dagger, c_m^\dagger] = 0$$

and for fermions

$$\{c_l^\dagger, c_m^\dagger\} = 0\ ,$$

where we use the standard notation for an anticommutator of two operators A and B: $\{A, B\} = AB + BA$. Taking Hermitian conjugates of these two equations, we have for bosons

$$[c_l, c_m] = 0$$

and for fermions

$$\{c_l, c_m\} = 0\ .$$

Note for fermions we can conclude that $(c_l)^2{=}(c_l^\dagger)^2{=}0$, which illustrates again how the Pauli exclusion principle is built into our approach.

Finally, one can check that to reproduce the values of inner products of states appearing in Eq. (4), we require for bosons

$$[c_l\,, c_m^\dagger] = \delta_{lm}$$

and for fermions

$$\{c_l\,, c_m^\dagger\} = \delta_{lm}\ .$$

To illustrate the correctness of these relations, consider for a single boson orbital the value of $|[(c^\dagger)^n|0\rangle]|^2$. From Eq. (4) we have $|[(c^\dagger)^n|0\rangle]|^2 = n!$. Let's recover the same result by manipulating commutators: we have

$$\begin{aligned}
\langle 0|(c)^n(c^\dagger)^n|0\rangle &= \langle 0|(c)^{n-1}([c, c^\dagger] + c^\dagger c)(c^\dagger)^{n-1}|0\rangle \\
&= m\langle 0|(c)^{n-1}(c^\dagger)^{n-1}|0\rangle + \langle 0|c^\dagger(c)^{n-m}c^\dagger c(c)^m(c^\dagger)^{n-1}|0\rangle \\
&= n\langle 0|(c)^{n-1}(c^\dagger)^{n-1}|0\rangle + \langle 0|c^\dagger(c)^{n-1}(c^\dagger)^{n-1}|0\rangle \\
&= n(n-1)\ldots(n-l)\langle 0|(c^\dagger)^{n-l}(c)^{n-l}|0\rangle \\
&= n!\,\langle 0|0\rangle\ .
\end{aligned}$$

Of course, manipulations like these are familiar from the theory of raising and lowering operators for the harmonic oscillator.

1.8 Number operators

From Eq. (4) as the defining equation for the action of creation operators in Fock space we have

$$c_l^\dagger |n_1 \dots n_l \dots\rangle = (\pm 1)^{n_1+\dots+n_{l-1}} \sqrt{n_l+1}\, |n_1 \dots n_l + 1 \dots\rangle \ ,$$

or zero for fermions if n_l=1. Similarly, by considering the Hermitian conjugate of a similar equation, we have

$$c_l |n_1 \dots n_l \dots\rangle = (\pm 1)^{n_1+\dots+n_{l-1}} \sqrt{n_l}\, |n_1 \dots n_l - 1 \dots\rangle \ ,$$

or zero for both bosons and fermions if n_l=0. In this way we have

$$c_l^\dagger c_l \, |\dots n_l \dots\rangle = n_l |\dots n_l \dots\rangle$$

where the possible values of n_l are n_l=0, 1, 2 . . . for bosons and n_l=0, 1 for fermions. Thus the combination $c_l^\dagger c_l$, which we will also write as $\hat{n}_l$, is the number operator and counts particles in the orbital ϕ_l.

1.9 Transformations between bases

In the context of single-particle quantum mechanics it is often convenient to make transformations between different bases. Since we used a particular set of basis functions in our definition of creation and annihilation operators, we should understand what such transformations imply in operator language.

Suppose we have two complete, orthonormal sets of single-particle basis functions, $\{\phi_l(\mathbf{r})\}$ and $\{\rho_\alpha(\mathbf{r})\}$, which we also write as $\{|\phi_i\rangle\}$ and $\{|\rho_\alpha\rangle\}$. Then we can expand one in terms of the other, writing

$$\rho_\alpha(\mathbf{r}) = \sum_l \phi_l(\mathbf{r}) U_{l\alpha} \tag{5}$$

with $U_{l\alpha} = \langle \phi_l | \rho_\alpha \rangle$. Note that U is a unitary matrix, since

$$\begin{aligned} (\mathrm{U}\mathrm{U}^\dagger)_{ml} &= \sum_\alpha \langle \phi_m | \rho_\alpha \rangle \langle \rho_\alpha | \phi_l \rangle \\ &= \langle \phi_m | \phi_l \rangle \qquad \text{since} \qquad \sum_\alpha |\rho_\alpha\rangle\langle\rho_\alpha| = \mathbf{1} \\ &= \delta_{ml} \ . \end{aligned}$$

Now let $c_l^\dagger$ create a particle in orbital $\phi_l(\mathbf{r})$, and let $d_\alpha^\dagger$ create a particle in orbital $\rho_\alpha(\mathbf{r})$. We can read off from Eq. (5) an expression for $d_\alpha^\dagger$ in terms of $c_l^\dagger$:

$$d_\alpha^\dagger = \sum_l c_l^\dagger U_{l\alpha} \ .$$

From the Hermitian conjugate of this equation we also have

$$d_\alpha = \sum_l U_{l\alpha}^* c_l = \sum_l (\mathrm{U}^\dagger)_{\alpha l} c_l \ .$$

1.9.1 Effect of transformations on commutation relations

We should verify that such transformations preserve commutation relations. For example, suppose that c_l and $c_l^\dagger$ are fermion operators, obeying $\{c_l \, , c_m^\dagger\} = \delta_{lm}$. Then

$$\{d_\alpha, d_\beta^\dagger\} = \sum_{lm} U_{l\alpha}^* U_{m\beta} \, \{c_l \, , c_m^\dagger\} = (\mathrm{U}^\dagger \mathrm{U})_{\alpha\beta} = \delta_{\alpha\beta} \, .$$

Similarly, for boson operators commutation relations are preserved under unitary transformations.

1.10 General single-particle operators in second-quantised form

To continue our programme of formulating many-particle quantum mechanics in terms of creation and annihilation operators, we need to understand how to transcribe operators from coordinate representation or first-quantised form to so-called second-quantised form. In the first instance, we examine how to do this for one-body operators – those which involve the coordinates of one particle at a time. An example is the kinetic energy operator. Suppose in general that $A(\mathbf{r})$ represents such a quantity for a single-particle system. Then for a system of N particles in first-quantised notation we have

$$\hat{A} = \sum_{i=1}^{N} A(\mathbf{r}_i) \ .$$

We want to represent $\hat{A}$ using creation and annihilation operators. As a first step, we can characterise $A(\mathbf{r})$ by its matrix elements, writing

$$A_{lm} = \int \phi_l^*(\mathbf{r}) A(\mathbf{r}) \phi_m(\mathbf{r}) \mathrm{d}^d\mathbf{r} \ .$$

Then

$$A(\mathbf{r})\phi_m(\mathbf{r}) = \sum_l \phi_l(\mathbf{r}) A_{lm} \ . \tag{6}$$

The second-quantised representation is

$$\hat{A} = \sum_{pq} A_{pq} c_p^\dagger c_q \ . \tag{7}$$

To justify this, we should verify that reproduces the correct matrix elements between all states from the Fock space. We will simply check the action of $\hat{A}$ on single particles states. We have

$$\hat{A}|\phi_m\rangle = \sum_{pq} A_{pq} c_p^\dagger c_q c_m^\dagger |0\rangle \ .$$

Now, taking as an example bosons,

$$c_p^\dagger c_q c_m^\dagger |0\rangle = c_p^\dagger ([c_q, c_m^\dagger] + c_m^\dagger c_q)|0\rangle = c_p^\dagger \delta_{qm}|0\rangle$$

so

$$\hat{A}|\phi_m\rangle = \sum_p |\phi_p\rangle A_{pm} \ ,$$

reproducing Eq. (6), as required.

1.11 Two-particle operators in second-quantised form

Two-body operators depend on the coordinates of a pair of particles, an example being the two-body potential in an interacting system. Writing the operator in first-quantised form as $A(\mathbf{r}_1, \mathbf{r}_2)$, it has matrix elements which carry four labels:

$$A_{lmpq} = \int \phi_l^*(\mathbf{r}_1)\phi_m^*(\mathbf{r}_2) A(\mathbf{r}_1, \mathbf{r}_2) \phi_p(\mathbf{r}_2)\phi_q(\mathbf{r}_1) \mathrm{d}^d\mathbf{r}_1 \mathrm{d}^d\mathbf{r}_2 \ .$$

Its second-quantised form is

$$\hat{A} \equiv \sum_{ij} A(\mathbf{r}_i, \mathbf{r}_j) = \sum_{lmpq} A_{lmpq} c_l^\dagger c_m^\dagger c_p c_q \ . \tag{8}$$

Again, to justify this one should check matrix elements of the second-quantised form between all states in Fock space. We will content ourselves with matrix elements for two-particle states, evaluating

$$\langle A \rangle = \langle 0| c_y c_x \hat{A} c_a^\dagger c_b^\dagger |0\rangle$$

by two routes. In a first-quantised calculation with $\pm$ signs for bosons and fermions, we have

$$\begin{aligned}
\langle A \rangle &= \frac{1}{2} \int\int [\phi_x^*(\mathbf{r}_1)\phi_y^*(\mathbf{r}_2) \pm \phi_x^*(\mathbf{r}_2)\phi_y^*(\mathbf{r}_1)] \cdot [A(\mathbf{r}_1, \mathbf{r}_2) + A(\mathbf{r}_2, \mathbf{r}_1)] \cdot [\phi_a(\mathbf{r}_1)\phi_b(\mathbf{r}_2) \pm \phi_a(\mathbf{r}_2)\phi_b(\mathbf{r}_1)] \mathrm{d}^d\mathbf{r}_1 \mathrm{d}^d\mathbf{r}_2 \\
&= \frac{1}{2}[A_{xyba} \pm A_{xyab} + A_{yxab} \pm A_{yxba} + A_{xyba} \pm A_{xyab} + A_{yxab} \pm A_{yxba}] \\
&= (A_{xyba} + A_{yxab}) \pm (A_{xyab} + A_{yxba}) \ .
\end{aligned} \tag{9}$$

Using the proposed second-quantised form for $\hat{A}$, we have

$$\langle A\rangle = \sum_{lmpq} A_{lmpq}\langle 0|c_y c_x c_l^\dagger c_m^\dagger c_p c_q c_a^\dagger c_b^\dagger|0\rangle \ .$$

We can simplify the vacuum expectation value of products of creation and annihilation operators such as the one appearing here by using the appropriate commutation or anticommutation relation to move annihilation operators to the right, or creation operators to the left, whereupon acting on the vacuum they give zero. In particular

$$c_p c_q c_a^\dagger c_b^\dagger|0\rangle \ = (\delta_{aq}\delta_{bp} \pm \delta_{ap}\delta_{bq})|0\rangle$$

and

$$\langle 0|c_y c_x c_l^\dagger c_m^\dagger = \langle 0|(\delta_{ym}\delta_{xl} \pm \delta_{yl}\delta_{xm}) \ .$$

Combining these, we recover Eq. (9).

2 Diagonalisation of quadratic Hamiltonians

If a Hamiltonian is quadratic (or, more precisely, bilinear) in creation and annihilation operators we can diagonalise it, meaning we can reduce it to a form involving only number operators. This is an approach that applies directly to Hamiltonians for non-interacting systems, and also to Hamiltonians for interacting systems when interactions are treated within a mean field approximation.

2.1 Number-conserving quadratic Hamiltonians

Such Hamiltonians have the form

$$\mathcal{H} = \sum_{ij} H_{ij} a_i^\dagger a_j \ .$$

Note that in order for the operator $\mathcal{H}$ to be Hermitian, we require the matrix H to be Hermitian. Since the matrix H is Hermitian, it can be diagonalised by unitary transformation. Denote this unitary matrix by U and let the eigenvalues of H be ε_n. The same transformation applied to the creation and annihilation operators will diagonalise $\mathcal{H}$. The details of this procedure are as follows. Let

$$\alpha_l^\dagger = \sum_i a_i^\dagger U_{il} \ .$$

Inverting this, we have

$$\sum \alpha_l^\dagger (\mathrm{U}^\dagger)_{lj} = a_j^\dagger$$

and taking a Hermitian conjugate

$$\sum_l U_{jl}\alpha_l = a_j \ .$$

Substituting for $a^\dagger$'s and a's in terms of $\alpha^\dagger$'s and α's, we find

$$\mathcal{H} = \sum_{lm} \alpha_l^\dagger (\mathrm{U}^\dagger \mathrm{H U})_{lm}\alpha_m = \sum_n \varepsilon_n \alpha_n^\dagger \alpha_n \equiv \sum_n \varepsilon_n \hat{n}_n \ .$$

Thus the eigenstates of $\mathcal{H}$ are the occupation number eigenstates in the basis generated by the creation operators $\alpha_n^\dagger$.

2.2 Mixing creation and annihilation operators: Bogoliubov transformations

There are a number of physically important systems which, when treated approximately, have bilinear Hamiltonians that include terms with two creation operators, and others with two annihilation operators. Examples include superconductors, superfluids and antiferromagnets. These Hamiltonians can be diagonalised by what are known as *Bogoliubov transformations*, which mix creation and annihilation operators, but, as always, preserve commutation relations. We now illustrate these transformations, discussing fermions and bosons separately.

2.2.1 Fermions

Consider for fermion operators the Hamiltonian

$$\mathcal{H} = \epsilon(c_1^\dagger c_1 + c_2^\dagger c_2) + \lambda(c_1^\dagger c_2^\dagger + c_2 c_1) \ ,$$

which arises in the BCS theory of superconductivity. Note that λ must be real for $\mathcal{H}$ to be Hermitian (more generally, with complex λ the second term of $\mathcal{H}$ would read $\lambda c_1^\dagger c_2^\dagger + \lambda^* c_2 c_1$). Note as well the opposite ordering of labels in the terms $c_1^\dagger c_2^\dagger$ and $c_2 c_1$, which is also a requirement of Hermiticity.

The fermionic Bogoliubov transformation is

$$\begin{aligned} c_1^\dagger &= u d_1^\dagger + v d_2 \\ c_2^\dagger &= u d_2^\dagger - v d_1 \ , \end{aligned} \tag{10}$$

where u and v are c-numbers, which we can in fact take to be real, because we have restricted ourselves to real λ. The transformation is useful only if fermionic anticommutation relations apply to both sets of operators. Let us suppose they apply to the operators d and $d^\dagger$, and check the properties of the operators c and $c^\dagger$. The coefficients of the transformation have been chosen to ensure that $\{c_1^\dagger, c_2^\dagger\} = 0$, while

$$\{c_1^\dagger, c_1\} = u^2\{d_1^\dagger, d_1\} + v^2\{d_2^\dagger, d_2\}$$

and so we must require $u^2 + v^2 = 1$, suggesting the parameterisation $u = \cos\theta$, $v = \sin\theta$.

The remaining step is to substitute in $\mathcal{H}$ for $c^\dagger$ and c in terms of $d^\dagger$ and d, and pick θ so that terms in $d_1^\dagger d_2^\dagger + d_2 d_1$ have vanishing coefficient. The calculation is clearest when it is set out using matrix notation. First, we can write $\mathcal{H}$ as

$$\mathcal{H} = \frac{1}{2}\begin{pmatrix} c_1^\dagger & c_2 & c_2^\dagger & c_1 \end{pmatrix} \begin{pmatrix} \epsilon & \lambda & 0 & 0 \\ \lambda & -\epsilon & 0 & 0 \\ 0 & 0 & \epsilon & -\lambda \\ 0 & 0 & -\lambda & -\epsilon \end{pmatrix} \begin{pmatrix} c_1 \\ c_2^\dagger \\ c_2 \\ c_1^\dagger \end{pmatrix} + \epsilon$$

where we have used the anticommutator to make substitutions of the type $c^\dagger c = 1 - c\, c^\dagger$.

For conciseness, consider just the upper block

$$\begin{pmatrix} c_1^\dagger & c_2 \end{pmatrix} \begin{pmatrix} \epsilon & \lambda \\ \lambda & -\epsilon \end{pmatrix} \begin{pmatrix} c_1 \\ c_2^\dagger \end{pmatrix}$$

and write the Bogoliubov transformation also in matrix form as

$$\begin{pmatrix} c_1 \\ c_2^\dagger \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} d_1 \\ d_2^\dagger \end{pmatrix} \ .$$

We pick θ so that

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \epsilon & \lambda \\ \lambda & -\epsilon \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} \tilde{\epsilon} & 0 \\ 0 & -\tilde{\epsilon} \end{pmatrix} \ ,$$

where $\tilde{\epsilon} = \sqrt{\epsilon^2 + \lambda^2}$. Including the other 2×2 block of $\mathcal{H}$, we conclude that

$$\mathcal{H} = \tilde{\epsilon}(d_1^\dagger d_1 + d_2^\dagger d_2) + \epsilon - \tilde{\epsilon} \ .$$

2.2.2 Bosons

The Bogoliubov transformation for a bosonic system is similar in principle to what we have just set out, but different in detail. We are concerned with a Hamiltonian of the same form, but now written using boson creation and annihilation operators:

$$\mathcal{H} = \epsilon(c_1^\dagger c_1 + c_2^\dagger c_2) + \lambda(c_1^\dagger c_2^\dagger + c_2 c_1) \ .$$

We use a transformation of the form

$$\begin{aligned} c_1^\dagger &= u d_1^\dagger + v d_2 \\ c_2^\dagger &= u d_2^\dagger + v d_1 \ . \end{aligned}$$

Note that one sign has been chosen differently from its counterpart in Eq. (10) in order to ensure that bosonic commutation relations for the operators d and $d^\dagger$ imply the result $[c_1^\dagger, c_2^\dagger] = 0$. We also require

$$[c_1, c_1^\dagger] = u^2[d_1, d_1^\dagger] - v^2[d_2, d_2^\dagger] = 1$$

and hence $u^2 - v^2 = 1$. The bosonic Bogoliubov transformation may therefore be parameterised as $u = \cosh\theta$, $v = \sinh\theta$.

We can introduce matrix notation much as before (but note some crucial sign differences), with

$$\mathcal{H} = \frac{1}{2}\begin{pmatrix} c_1^\dagger & c_2 & c_2^\dagger & c_1 \end{pmatrix}\begin{pmatrix} \epsilon & \lambda & 0 & 0 \\ \lambda & \epsilon & 0 & 0 \\ 0 & 0 & \epsilon & \lambda \\ 0 & 0 & \lambda & \epsilon \end{pmatrix}\begin{pmatrix} c_1 \\ c_2^\dagger \\ c_2 \\ c_1^\dagger \end{pmatrix} - \epsilon\ ,$$

where for bosons we have used the commutator to write $c^\dagger c = c\,c^\dagger - 1$. Again, we focus on one 2×2 block

$$\begin{pmatrix} c_1^\dagger & c_2 \end{pmatrix}\begin{pmatrix} \epsilon & \lambda \\ \lambda & \epsilon \end{pmatrix}\begin{pmatrix} c_1 \\ c_2^\dagger \end{pmatrix}$$

and write the Bogoliubov transformation also in matrix form as

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

Substituting for c and $c^\dagger$ in terms of d and $d^\dagger$, this block of the Hamiltonian becomes

$$\begin{pmatrix} d_1^\dagger & d_2 \end{pmatrix}\begin{pmatrix} u & v \\ v & u \end{pmatrix}\begin{pmatrix} \epsilon & \lambda \\ \lambda & \epsilon \end{pmatrix}\begin{pmatrix} u & v \\ v & u \end{pmatrix}\begin{pmatrix} d_1 \\ d_2^\dagger \end{pmatrix}\ .$$

In the fermionic case the matrix transformation was simply an orthogonal rotation. Here it is not, and so we should examine it in more detail. We have

$$\begin{pmatrix} u & v \\ v & u \end{pmatrix}\begin{pmatrix} \epsilon & \lambda \\ \lambda & \epsilon \end{pmatrix}\begin{pmatrix} u & v \\ v & u \end{pmatrix} = \begin{pmatrix} \epsilon[u^2+v^2] + 2\lambda uv & 2\epsilon uv + \lambda[u^2+v^2] \\ 2\epsilon uv + \lambda[u^2+v^2] & \epsilon[u^2+v^2] + 2\lambda uv \end{pmatrix}\ .$$

It is useful to recall the double angle formulae $u^2 + v^2 = \cosh 2\theta$ and $2uv = \sinh 2\theta$. Then, setting $\tanh 2\theta = -\lambda/\epsilon$ we arrive at

$$\mathcal{H} = \tilde{\epsilon}(d_1^\dagger d_1 + d_2^\dagger d_2) - \epsilon + \tilde{\epsilon}\ .$$

with

$$\tilde{\epsilon} = \sqrt{\epsilon^2 - \lambda^2}. \tag{11}$$

Note that in the bosonic case the transformation requires $\epsilon > \lambda$: if this is not the case, $\mathcal{H}$ is not a Hamiltonian for normal mode oscillations about a stable equilibrium, but instead represents a system at an unstable equilibrium point.

2.3 Fourier transform conventions

We will use Fourier transforms extensively, because much of the time we will be considering systems that are translation-invariant, and the plane waves used in these transforms are eigenfunctions of translation operators. For convenience, we collect here some definitions. Although we are generally interested in the thermodynamic limit (the limit of infinite system size), it is usually clearest and cleanest to write transforms in the first instance for a finite system. In order to preserve translation invariance, we take this finite system to have periodic boundary conditions. Since some details differ, we consider lattice and continuum problems separately.

2.3.1 Lattice systems

Consider a three-dimensional Bravais lattice with basis vectors $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$. Lattice sites have coordinates

$$\mathbf{r} = l\mathbf{a} + m\mathbf{b} + n\mathbf{c} \tag{12}$$

with l,m and n integer. Periodic boundary conditions mean that $l + N_1 \equiv l$, $m + N_2 \equiv m$, and $n + N_3 \equiv n$, and the number of lattice sites is then $N = N_1 N_2 N_3$. In the usual way, reciprocal lattice vectors $\mathbf{G}_1$, $\mathbf{G}_2$ and $\mathbf{G}_3$ satisfy $\mathbf{G}_1 \cdot \mathbf{a} = 2\pi$, $\mathbf{G}_1 \cdot \mathbf{b} = \mathbf{G}_1 \cdot \mathbf{c} = 0$ and so on. Then the wave $\mathrm{e}^{\mathrm{i}\mathbf{kr}}$ satisfies periodic boundary conditions if

$$\mathbf{k} = 2\pi \left(\frac{n_1}{N_1}\mathbf{G}_1, \frac{n_2}{N_2}\mathbf{G}_2, \frac{n_3}{N_3}\mathbf{G}_3 \right) \tag{13}$$

with n_1, n_2 and n_3 integer. Note that we have N values of $\mathbf{k}$ in the Brillouin zone.

Let $c_{\mathbf{r}}^\dagger$ be a (boson or fermion) creation operator at the site $\mathbf{r}$. We define the Fourier transform and inverse transform by

$$c_{\mathbf{k}}^\dagger = \frac{1}{\sqrt{N}} \sum_{\mathbf{r}} \mathrm{e}^{-\mathrm{i}\mathbf{kr}} c_{\mathbf{r}}^\dagger \quad \text{and} \quad c_{\mathbf{r}}^\dagger = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} \mathrm{e}^{\mathrm{i}\mathbf{kr}} c_{\mathbf{k}}^\dagger \,. \tag{14}$$

There are several points to make here. First, one should check for consistency by substituting one expression into the other. Second, these definitions use the unitary $N \times N$ matrix U, which has elements

$$U_{\mathbf{kr}} = \frac{1}{\sqrt{N}} \mathrm{e}^{\mathrm{i}\mathbf{kr}} \,.$$

Third, if we consider time-dependence in the Heisenberg picture with a Hamiltonian $\mathcal{H} = \hbar\omega(\mathbf{k}) c_{\mathbf{k}}^\dagger c_{\mathbf{k}}$ we have

$$c_{\mathbf{k}}^\dagger(t) \equiv \mathrm{e}^{\mathrm{i}\mathcal{H}t} c_{\mathbf{k}}^\dagger \mathrm{e}^{-\mathrm{i}\mathcal{H}t} = c_{\mathbf{k}}^\dagger \mathrm{e}^{-\mathrm{i}\omega_{\mathbf{k}}t} \qquad \text{and} \qquad c_{\mathbf{r}}^\dagger(t) = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} \mathrm{e}^{\mathrm{i}[\mathbf{kr} - \omega(\mathbf{k})t]} c_{\mathbf{k}}^\dagger \,,$$

which has the usual traveling wave form.

2.3.2 Continuum systems

Consider a cube of side L and volume V with periodic boundary conditions. Take

$$\mathbf{k} = \frac{2\pi}{L} (n_1, n_2, n_3)$$

with n_1, n_2 and n_3 integer. Then the wavefunctions

$$\psi_{\mathbf{k}}(\mathbf{r}) = V^{-1/2} \mathrm{e}^{\mathrm{i}\mathbf{kr}}$$

form a normalised single-particle basis.

Let $c^\dagger(\mathbf{r})$ be a (boson or fermion) creation operator for a particle at the point $\mathbf{r}$. Then the creation operator for a particle in the state with wavefunction $\psi_{\mathbf{k}}(\mathbf{r})$ is

$$c_{\mathbf{k}}^\dagger = V^{-1/2} \int \mathrm{d}^3\mathbf{r}\, \mathrm{e}^{-\mathrm{i}\mathbf{kr}}\, c^\dagger(\mathbf{r}) \tag{15}$$

and the inverse transform is

$$c^\dagger(\mathbf{r}) = \frac{1}{\sqrt{V}} \sum_{\mathbf{k}} \mathrm{e}^{\mathrm{i}\mathbf{kr}} c_{\mathbf{k}}^\dagger \,. \tag{16}$$

2.3.3 Thermodynamic limit

In the thermodynamic limit sums on wavevectors can be replaced by integrals. On a lattice we have

$$N^{-1} \sum_{\mathbf{k}} \;\to\; \frac{1}{\Omega} \int_{\mathrm{BZ}} \mathrm{d}^d\mathbf{k} \,, \tag{17}$$

where the integral is over the Brillouin zone of volume Ω. In the continuum we have

$$V^{-1} \sum_{\mathbf{k}} \;\to\; (2\pi)^{-d} \int \mathrm{d}^d\mathbf{k} \,. \tag{18}$$

Quantum Magnets and the Bose Gas

3 The Heisenberg model

We move now to applying some of these ideas to the theory of magnetism. We will consider insulating magnets (as distinct from itinerant ones, in which the electrons involved in magnetism belong to a partially-filled conduction band that has a Fermi surface). We model an insulating magnetic material using spin operators to represent magnetic moments at the sites of a lattice. The formation of these magnetic moments can be understood in terms of the atomic physics of an isolated ion: we postpone discussion of this aspect to Section 6. Neighbouring magnetic moments are coupled by exchange interactions, and a simple model that captures this is the Heisenberg model. We take nearest neighbour interactions of strength J (where $J > 0$), and study the Hamiltonian

$$\mathcal{H} = \pm J \sum_{\langle \mathbf{r}\mathbf{r}' \rangle} \mathbf{S}_{\mathbf{r}} \cdot \mathbf{S}_{\mathbf{r}'} \equiv \pm J \sum_{\langle \mathbf{r}\mathbf{r}' \rangle} \left[S^z_{\mathbf{r}} S^z_{\mathbf{r}'} + \frac{1}{2} \left(S^+_{\mathbf{r}} S^-_{\mathbf{r}'} + S^-_{\mathbf{r}} S^+_{\mathbf{r}'} \right) \right] . \tag{19}$$

Here $\sum_{\langle \mathbf{r}\mathbf{r}' \rangle}$ denotes a sum over neighbouring pairs of sites on the lattice, with each pair counted once. With a negative sign in front of J, parallel spins have lower energy and the system is a ferromagnet, while with a positive sign antiparallel spins are favoured and the system is an antiferromagnet.

The three components of spin at site $\mathbf{r}$ are represented by operators $S^x_{\mathbf{r}}$, $S^y_{\mathbf{r}}$ and $S^z_{\mathbf{r}}$. Their commutation relations are the standard ones, and with $\hbar = 1$ take the form

$$[S^i_{\mathbf{r}_1}, S^j_{\mathbf{r}_2}] = \mathrm{i} \delta_{\mathbf{r}_1, \mathbf{r}_2} \epsilon_{ijk} S^k_{\mathbf{r}_1} .$$

We emphasise two points: first, the commutation relations are more complicated than those for creation and annihilation operators, since the commutator is itself another operator and not a number; and second, spin operators acting at different sites commute. We will also make use of spin raising and lowering operators, defined in the usual way as $S^+ = S^x + \mathrm{i}S^y$ and $S^- = S^x - \mathrm{i}S^y$.

4 Spin wave theory

4.1 Holstein Primakoff transformation

This transformation expresses spin operators in terms of boson operators. It provides an obvious way to build in the fact that spin operators at different sites commute. In a non-linear form it also reproduces exactly the commutation relations between two spin operators associated with the same site, but we will use a linearised version of the transformation which is approximate. At a single site we take the eigenvector of S^z with eigenvalue S to be the boson vacuum, and associate each unit reduction in S^z with the addition of a boson. Then

$$S^z = S - b^\dagger b \, .$$

From this we might guess $S^+ \propto b$ and $S^- \propto b^\dagger$. In an attempt to identify the proportionality constants we can compare the commutator $[S^+, S^-] = 2S^z$ with $[b, b^\dagger] = 1$. Since the commutator is an operator in the first case and a number in the second, our guessed proportionality cannot be exact, but within states for which $\langle S^z \rangle \approx S$ (meaning $\langle S^z \rangle - S \ll S$, which can be satisfied only if $S \gg 1$) we can take

$$S^+ \approx (2S)^{1/2} b \qquad \text{and} \qquad S^- \approx (2S)^{1/2} b^\dagger \, . \tag{20}$$

In an exact treatment, corrections to these expressions form a series in powers of $b^\dagger b / S$. The full expressions are

$$S^+ = (2S)^{1/2} \left(1 - \frac{b^\dagger b}{2S} \right)^{1/2} b \qquad \text{and} \qquad S^- = (2S)^{1/2} b^\dagger \left(1 - \frac{b^\dagger b}{2S} \right)^{1/2} . \tag{21}$$

4.2 Heisenberg ferromagnet

Consider the ferromagnetic Heisenberg model. If the spins were classical vectors, the ground state would be one in which all spins are aligned - say along the z-axis. We can define an equivalent quantum state $|0\rangle$, as one satisfying

$$S^z_{\mathbf{r}} |0\rangle = S |0\rangle$$

at every site in the lattice. This state is in fact an exact eigenstate of the Hamiltonian, Eq. (19) and is a ground state. Other, symmetry-related ground states are obtained by acting on this one with the total spin lowering operator, or with global spin rotation operators.

Next we would like to understand excitations from this ground state. Wavefunctions for the lowest branch of excitations can also be written down exactly, but to understand states with many excitations present we need to make approximations, and the Holstein-Primakoff transformation provides a convenient way to do so.

Using this transformation and omitting the higher order terms, the Hamiltonian may be rewritten approximately as

$$\mathcal{H} = -J \sum_{\langle \mathbf{r}\mathbf{r}' \rangle} S^2 - JS \sum_{\langle \mathbf{r}\mathbf{r}' \rangle} \left[b^\dagger_{\mathbf{r}} b_{\mathbf{r}'} + b^\dagger_{\mathbf{r}'} b_{\mathbf{r}} - b^\dagger_{\mathbf{r}} b_{\mathbf{r}} - b^\dagger_{\mathbf{r}'} b_{\mathbf{r}'} \right] . \tag{22}$$

Applying the approach of Section 2.1, we can diagonalise Eq. (22) by a unitary transformation of the creation and annihilation operators. In a translationally invariant system this is simply a Fourier transformation. Suppose the sites form a simple cubic lattice with unit spacing. Take the system to be a cube with side L and apply periodic boundary conditions. The number of lattice sites is then $N = L^3$ and allowed wavevectors are

$$\mathbf{k} = \frac{2\pi}{L}(l, m, n) \qquad \text{with} \quad l, m, n \quad \text{integer} \qquad \text{and} \quad 1 \leq l, m, n \leq L \, .$$

Boson operators in real space and reciprocal space are related by

$$b_{\mathbf{r}} = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} \mathrm{e}^{-\mathrm{i}\mathbf{k}\cdot\mathbf{r}} \, b_{\mathbf{k}} \qquad \text{and} \qquad b^\dagger_{\mathbf{r}} = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}} \, b^\dagger_{\mathbf{k}} \, .$$

We use these transformations, and introduce the notation $\mathbf{d}$ for vectors from a site to its nearest neighbours, and z for the coordination number of the lattice (the number of neighbours to a site: six for the simple cubic lattice), to obtain

$$\begin{aligned} \mathcal{H} &= -JS^2 N \frac{z}{2} - JS \sum_{\mathbf{r}\mathbf{d}} \sum_{\mathbf{k}\mathbf{q}} \frac{1}{N} \mathrm{e}^{\mathrm{i}\mathbf{r}\cdot(\mathbf{k}-\mathbf{q})} [\mathrm{e}^{\mathrm{i}\mathbf{d}\cdot\mathbf{q}} - 1] b^\dagger_{\mathbf{k}} b_{\mathbf{q}} \\ &= -JS^2 N \frac{z}{2} + \sum_{\mathbf{q}} \epsilon(\mathbf{q}) b^\dagger_{\mathbf{q}} b_{\mathbf{q}} \, , \end{aligned}$$

where

$$\epsilon(\mathbf{q}) = 2JS(3 - \cos q_x - \cos q_y - \cos q_z) \, .$$

In this way we have approximated the original Heisenberg Hamiltonian, involving spin operators, by one that is quadratic in boson creation and annihilation operators. By diagonalising this we obtain an approximate description of the low-lying excitations of the system as independent bosons. The most important feature of the result is the form of the dispersion a small wavevectors. For $q \ll 1$ we have $\epsilon(\mathbf{q}) = JSq^2 + \mathcal{O}(q^4)$, illustrating that excitations are gapless. This is expected because these excitations are Goldstone modes: they arise because the choice of ground state breaks the continuous symmetry of the Hamiltonian under spin rotations. The fact that dispersion is quadratic, and not linear as it is, for example for phonons, reflects broken time-reversal symmetry in the ground state of the ferromagnet.

4.3 Heisenberg antiferromagnet

We start again from the Heisenberg Hamiltonian, but now with antiferromagnetic interactions.

$$\mathcal{H} = J \sum_{\langle \mathbf{r}\mathbf{r}' \rangle} \mathbf{S}_{\mathbf{r}} \cdot \mathbf{S}_{\mathbf{r}'} = J \sum_{\langle \mathbf{r}\mathbf{r}' \rangle} \left[S^z_{\mathbf{r}} S^z_{\mathbf{r}'} + \frac{1}{2} \left(S^+_{\mathbf{r}} S^-_{\mathbf{r}'} + S^-_{\mathbf{r}} S^+_{\mathbf{r}'} \right) \right] . \tag{23}$$

We will only consider bipartite lattices: those for which the sites can be divided into two sets, in such a way that sites in one set have as their nearest neighbours only sites from the other set. The square lattice and the simple cubic lattice are examples, and we will treat the model in d dimensions on a hypercubic lattice. Approximating the quantum spins in the first instance as classical vectors, the exchange energy of a nearest neighbour pair is minimised when the two spins are antiparallel. For the lattice as a whole, the classical ground states are ones in which all spins on one sublattice have the same orientation, which is opposite to that of spins on the other

sublattice. This is a classical Néel state. The corresponding quantum state, taking the axis of orientation to be the z-axis, is defined by the property

$$S^z_{\mathbf{r}}|0\rangle = \pm S|0\rangle$$

with the sign positive on one sublattice, and negative on the other.

In contrast to the fully polarised ferromagnetic state, this is not an exact eigenstate of the Hamiltonian. We can see this by considering the action of the term $S^+_{\mathbf{r}}S^-_{\mathbf{r}'}$. If the site $\mathbf{r}$ is on the up sublattice and $\mathbf{r}'$ on the down sublattice, the operator simply annihilates $|0\rangle$. But if the sublattice assignments for the sites are the other way around, we generate a component in the resulting wavefunction that is different from $|0\rangle$: in this component the spin at site $\mathbf{r}$ has $S^z = S - 1$ and that at $\mathbf{r}'$ has $S^z = -(S-1)$.

To find out what the quantum ground state is, and to study excitations, we will again use the Holstein Primakoff transformation. Before we can do so, however, we need to adapt our spin coordinates to suit the classical Néel state. That is, we rotate axes in spin space for sites on the down sublattice, so that local z-axis is aligned with the spin direction in the classical Néel state. The required transformation is

$$S^z \to -S^z \qquad S^x \to -S^x \qquad S^y \to S^y\,.$$

As is necessary, this preserves the commutation relations, which inversion ($\mathbf{S} \to -\mathbf{S}$) would not do. After the transformation the Hamiltonian reads

$$\mathcal{H} = -J\sum_{\langle \mathbf{r}\mathbf{r}'\rangle}\left[S^z_{\mathbf{r}}S^z_{\mathbf{r}'} + \frac{1}{2}\left(S^+_{\mathbf{r}}S^+_{\mathbf{r}'} + S^-_{\mathbf{r}}S^-_{\mathbf{r}'}\right)\right]\,. \tag{24}$$

We use the Holstein Primakoff transformation denoting the boson annhiliation operator on sites from the up sublattice by $a_{\mathbf{r}}$ and those from the down sublattice by $b_{\mathbf{r}}$. (Note that Néel order means the magnetic unit cell has twice the volume of the chemical one.) Up to terms of quadratic order, we have

$$\mathcal{H} = -J\sum_{\langle \mathbf{r}\mathbf{r}'\rangle} S^2 + JS\sum_{\langle \mathbf{r}\mathbf{r}'\rangle}\left[a^\dagger_{\mathbf{r}}a_{\mathbf{r}} + b^\dagger_{\mathbf{r}'}b_{\mathbf{r}'} + a_{\mathbf{r}}b_{\mathbf{r}'} + b^\dagger_{\mathbf{r}'}a^\dagger_{\mathbf{r}}\right]\,. \tag{25}$$

Fourier transforming, this becomes

$$\mathcal{H} = -JS^2N\frac{z}{2} + JSd\sum_{\mathbf{k}}\left[a^\dagger_{\mathbf{k}}a_{\mathbf{k}} + b^\dagger_{-\mathbf{k}}b_{-\mathbf{k}} + \gamma(\mathbf{k})\left(a_{\mathbf{k}}b_{-\mathbf{k}} + b^\dagger_{-\mathbf{k}}a^\dagger_{\mathbf{k}}\right)\right] \tag{26}$$

where we have introduced the quantity

$$\gamma(\mathbf{k}) = \frac{1}{d}\sum_{\alpha=1}^{d}\cos(k_\alpha)\,,$$

which lies in the range $-1 \le \gamma(\mathbf{k}) \le 1$, and has the small k expansion $\gamma(\mathbf{k}) \approx 1 - k^2/2d$. To diagonalise the quadratic Hamiltonian of Eq. (26), we need to use the bosonic Bogoliubov transformation, as introduced in Section 2.2.2. We find

$$\mathcal{H} = -JS(S+1)N\frac{z}{2} + \sum_{\mathbf{k}}\epsilon(\mathbf{k})\left(\alpha^\dagger_{\mathbf{k}}\alpha_{\mathbf{k}} + \beta^\dagger_{-\mathbf{k}}\beta_{-\mathbf{k}} + 1\right) \tag{27}$$

with

$$a_{\mathbf{k}} = u_{\mathbf{k}}\alpha_{\mathbf{k}} - v_{\mathbf{k}}\beta^\dagger_{-\mathbf{k}} \qquad \text{and} \qquad b_{-\mathbf{k}} = u_{\mathbf{k}}\beta_{-\mathbf{k}} - v_{\mathbf{k}}\alpha^\dagger_{\mathbf{k}}$$

where

$$u_{\mathbf{k}} = \cosh(\theta_{\mathbf{k}})\,, \quad v_{\mathbf{k}} = \sinh(\theta_{\mathbf{k}})\,, \quad \text{and} \qquad \sinh(2\theta_{\mathbf{k}}) = \frac{\gamma(\mathbf{k})}{\sqrt{1-\gamma(\mathbf{k})^2}}\,.$$

The spinwave energy is

$$\epsilon(\mathbf{k}) = JSd(1-\gamma^2(\mathbf{k}))^{1/2}\,.$$

For small k the antiferromagnetic spinwave energy varies as $\epsilon(\mathbf{k}) \propto k$: a linear dependence on wavevector, in contrast to the quadratic variation for a ferromagnet, because the Néel state does not break time-reversal symmetry in a macroscopic sense (a symmetry of the state is time reversal, implying spin inversion, combined with exchange of sublattices).

4.4 Fluctuations and the order parameter

We can describe the fact that the ground states we have considered break spin rotation symmetry by using an order parameter. For the ferromagnet this is simply the magnetisation, and for the antiferromagnet it is the sublattice magnetisation. It is interesting to ask how the value of the order parameter is affected by fluctuations. Because the classical ferromagnetic state is also an exact quantum eigenstate of the Heisenberg Hamiltionian, there are no zero-point fluctuations in the ferromagnet, and in that case we will be interested in thermal fluctuations. On the other hand, we have seen that the classical Néel state is not an exact eigenstate, and so in this case quantum fluctuations are important as well.

4.4.1 Thermal fluctuations in a ferromagnet

The magnetisation (per site) is

$$M = \frac{1}{N}\sum_{\mathbf{r}} \langle S^z_{\mathbf{r}} \rangle \,.$$

Using the Holstein Primakoff transformation at leading order we have

$$M = S - \frac{1}{N}\sum_{\mathbf{k}} \langle b^\dagger_{\mathbf{k}} b_{\mathbf{k}} \rangle \equiv S - \Delta S \,. \tag{28}$$

Now, since the excitations are bosons with (like photons) no fixed number, the thermal average $\langle b^\dagger_{\mathbf{k}} b_{\mathbf{k}} \rangle$ is given by the Planck distribution. Using results from Section 2.3 to turn the sum on $\mathbf{k}$ into an integral, we have

$$\Delta S = \frac{1}{\Omega}\int_{\mathrm{BZ}} \mathrm{d}^d\mathbf{k} \frac{1}{\mathrm{e}^{\beta\epsilon(\mathbf{k})}-1} \,,$$

The most interesting aspects of this result are the generic ones, which emerge at low temperature. In that regime only low energy spinwaves are excited, and for these we can take the small wavevector form for their energy, finding

$$\Delta S \sim \frac{k_{\mathrm{B}}T}{J}\int_0^{\sqrt{k_{\mathrm{B}}T/J}} k^{d-1}\mathrm{d}k \, \frac{1}{k^2} \,.$$

The integral is divergent for $T > 0$ in $d = 1$ and $d = 2$, showing that low-range order is not possible in the Heisenberg ferromagnet in low dimensions (an illustration of the Mermin-Wagner theorem, which says that a continuous symmetry cannot be broken spontaneously at finite temperature for $d \leq 2$). In $d = 3$ we have $\Delta S \propto T^{3/2}$. The calculation of the spinwave contribution to the heat capacity is also interesting, but left as an exercise.

4.4.2 Quantum fluctuations in an antiferromagnet

The sublattice magnetisation on the 'up' sublattice is

$$S - \Delta S = S - \frac{1}{N}\sum_{\mathbf{r}} \langle a^\dagger_{\mathbf{k}} a_{\mathbf{k}} \rangle \,,$$

but now we need to relate $a_{\mathbf{k}}$ via the Bogoliubov transformation to the bosons that diagonalise the Hamiltonian. We find

$$\langle a^\dagger_{\mathbf{k}} a_{\mathbf{k}} \rangle = u^2_{\mathbf{k}} \langle \alpha^\dagger_{\mathbf{k}} \alpha_{\mathbf{k}} \rangle + v^2_{\mathbf{k}} [\langle \beta^\dagger_{-\mathbf{k}} \beta_{-\mathbf{k}} \rangle + 1] \,.$$

At zero temperature the boson occupation numbers are zero, and

$$\Delta S = \frac{1}{N}\sum_{\mathbf{k}} v^2_{\mathbf{k}} = \frac{1}{2\Omega}\int_{\mathrm{BZ}} \mathrm{d}^d\mathbf{k} \, [(1-\gamma^2(\mathbf{k}))^{-1/2} - 1] \,.$$

The most interesting question is to examine whether this integral converges. If it does, then for sufficiently large S the sublattice magnetisation is non-zero. But if it diverges, then our whole theoretical approach will collapse, because we started from the idea of a ground state with Néel order. A divergence can come only from points near the Brillouin zone center, where $\gamma(\mathbf{k})$ approaches 1. Expanding around this point, we have

$$\Delta S \sim \int k^{d-1}\mathrm{d}k \, \frac{1}{k} \,.$$

This integral is logarithmically divergent in one dimension, but convergent (since the Brillouin zone boundary sets an upper limit) in higher dimensions. We will see in the next section that one-dimensional antiferromagnets are particularly interesting, precisely because they have large quantum fluctuations.

5 Spin liquids and spinons in one-dimensional quantum magnets

As we have seen, quantum fluctuations melt antiferromagnetic order in one dimension. With some embellishments, the spinwave theory result for the sublattice magnetisation is

$$\langle S_r \rangle \approx S - \frac{1}{2\pi} \int_{\pi/L}^{\pi/a} \frac{\mathrm{d}k}{k} \,,$$

where we have considered a finite system of length L in order to have a small-k cut-off to the integral, and have introduced the lattice spacing a (previously set to one as the unit of length). Taking this approximate result seriously, $\langle S_r \rangle$ decreases with increasing L and reaches zero at a characteristic length, which we can identify as an estimate of the correlation length ξ in a quantum-disordered state. By this means we obtain $\xi \sim a\mathrm{e}^{2\pi S}$, which incidentally makes it clear that the spin size S has a dramatic influence on behaviour, and that quantum effects are most significant (ξ is shortest) if S is small.

In fact, this is only part of the story, although it is qualitatively correct for *integer* S. In that case, the finite correlation length goes hand in hand with an energy gap for excitations, known as the Haldane gap. By contrast, *half odd integer* spins, although disordered, are not characterised by a finite correlation length; instead they have correlations decaying with a power of separation, as we will see.

5.1 Spin one-half chain and transmutation of statistics in one dimension

For spin one-half the Holstein Primakoff transformation gives $S^z = \frac{1}{2} - b^\dagger b$, and since the eigenvalues of S^z are $\pm\frac{1}{2}$, we see that the allowed values of the boson number $b^\dagger b$ are 0 and 1. We can summarise this by saying that they are bosons with hard core interactions, which prevent more than one particle occupying the same site.

Now, hard core particles moving in one dimension can never alter their sequence. This means that our standard notions about symmetry of wavefunctions under particle exchange become an add-on to the theoretical description, and irrelevant to the dynamics. For that reason, it is possible to treat hard core bosons as spinless fermions. This is very useful, since for fermions we can do not need any interaction to prevent two particles occupying the same site: the Pauli exclusion principle ensures it, even in the absence of interactions. To put this idea to work, we need to understand in detail how to transform between the different operators used in the two descriptions.

5.1.1 Jordan-Wigner transformation

We want to transform from spin-half operators to a fermionic description. In spin language, operators S^z_n and $S^\pm_n$ obey the usual spin-half commutation relations. In the fermionic version, creation and annihilation operators $c^\dagger_n$ and c_n satisfy $\{c_n, c^\dagger_m\} = \delta_{nm}$. It is natural to set

$$S^z_m = c^\dagger_m c_m - 1/2 \equiv n_m - 1/2 \,.$$

This leads us to expect $S^+_n \propto c^\dagger_n$ and $S^-_n \propto c_n$. At any given site, everything works straightforwardly: $\{S^+_n, S^-_n\} = 1 = \{c_n, c^\dagger_n\}$. But for pairs of operators at different sites there is a problem: spin operators at different sites *commute*, while fermion operators *anticommute*.

The solution is provided by the Jordan-Wigner transformation, which reads

$$S^+_l = c^\dagger_l \mathrm{e}^{\mathrm{i}\pi \sum_{k<l} n_k} \qquad \text{and} \qquad S^-_l = \mathrm{e}^{-\mathrm{i}\pi \sum_{k<l} n_k} c_l \,, \tag{29}$$

where the factor $\mathrm{e}^{\mathrm{i}\pi \sum_{k<l} n_k}$, which depends on the total number of fermions on sites to the left of l, is termed a *Jordan-Wigner string*. To see that this transformation is indeed correct, consider first the relations

$$c^\dagger_m \mathrm{e}^{\mathrm{i}\pi n_m} = -\mathrm{e}^{\mathrm{i}\pi n_m} c^\dagger_m \qquad \text{and} \qquad c_m \mathrm{e}^{\mathrm{i}\pi n_m} = -\mathrm{e}^{\mathrm{i}\pi n_m} c_m \,, \tag{30}$$

which can be verified by comparing matrix elements of the left and right sides of each equation in the basis of fermion number eigenstates. Note also that

$$[c^\dagger_m, \mathrm{e}^{\mathrm{i}\pi n_l}] = 0$$

for $m \neq l$. Starting from $S^+_l S^+_m$ and substituting for the spin operators using Eq. (29), we hence obtain $S^+_l S^+_m = (-1)^2 S^+_m S^+_l$, where one factor of -1 comes as indicated in Eq. (30) and the other arises from exchanging the fermion operators $c^\dagger_l$ and $c^\dagger_m$. It is straightforward to check in a similar way that the Jordan-Wigner transformation also respects commutation of other pairs of the spin operators S^z, S^+ and S^- at different sites.

5.1.2 Application to spin chains

Consider a one-dimensional spin system with the Hamiltonian

$$\mathcal{H} = J \sum_n \left\{ \frac{1}{2} \left[s_n^+ s_{n+1}^- + s_n^- s_{n+1}^+ \right] + \Delta s_n^z s_{n+1}^z \right\} . \tag{31}$$

For $\Delta = 1$ this is the Heisenberg chain, and for $\Delta = 0$ it is an XY model. Using the Jordan-Wigner transformation we can re-write the Hamiltonian as

$$\mathcal{H} = \frac{J}{2} \sum_n \left\{ \left[c_n^\dagger c_{n+1} + c_{n+1}^\dagger c_n \right] + 2\Delta \left[(c_n^\dagger c_n - 1/2)(c_{n+1}^\dagger c_{n+1} - 1/2) \right] \right\} . \tag{32}$$

Clearly, the first term, arising from XY exchange, translates into fermion hopping, while the second term, arising from ZZ exchange, leads to interactions between fermions. Eigenstates can be found exactly for any Δ using the Bethe Ansatz, but the free fermion limit $\Delta = 0$ is the simplest by far, and we will consider only this case. Then we have a quadratic Hamiltonian, which with periodic boundary conditions is diagonalised by Fourier transform. (There in fact are some technical subtleties that arise when one combines the Jordan-Wigner transformation with periodic boundary conditions, but in the interests of brevity we will ignore these.) Writing

$$c_k = \frac{1}{\sqrt{N}} \sum_l \mathrm{e}^{\mathrm{i}kl} c_l$$

we have

$$\mathcal{H} = \frac{1}{N} \sum_k \epsilon(k) c_k^\dagger c_k \tag{33}$$

with $\epsilon(k) = J \cos k$. In the ground state, fermion orbitals k with $\epsilon(k) < 0$ are occupied and those with $\epsilon(k) > 0$ are empty.

5.1.3 Spin correlations

We will calculate the ground state correlation function $\langle S_0^z S_r^z \rangle$. Correlators of $S^\pm$ are much harder to evaluate because they involve the Jordan-Wigner strings. We have

$$\begin{aligned} \langle S_0^z S_r^z \rangle &= \langle c_0^\dagger c_0 c_r^\dagger c_r \rangle - \frac{1}{2} \langle c_0^\dagger c_0 + c_r^\dagger c_r \rangle + \frac{1}{4} \\ &= \frac{1}{N^2} \sum_{k_1,k_2,k_3,k_4} \langle c_{k_1}^\dagger c_{k_2} c_{k_3}^\dagger c_{k_4} \rangle \mathrm{e}^{\mathrm{i}r(k_4 - k_3)} - \frac{1}{4} . \end{aligned} \tag{34}$$

Contributions to $\langle c_{k_1}^\dagger c_{k_2} c_{k_3}^\dagger c_{k_4} \rangle$ are of two types:

(i) from $k_1 = k_2$ and $k_3 = k_4$ with both orbitals occupied: this cancels the term $-1/4$,

(ii) from $k_1 = k_4$ with the orbital occupied, and $k_2 = k_3$ with the orbital empty.

Thus, writing $n(k)$ for the occupation number of orbitals, we obtain

$$\begin{aligned} \langle S_0^z S_r^z \rangle &= \frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} \mathrm{d}k_1 \int_{-\pi}^{\pi} \mathrm{d}k_2 \, n(k_1)[1 - n(k_2)] \mathrm{e}^{\mathrm{i}r(k_1 - k_2)} \\ &= \frac{1}{2} \delta_{0,r} - \frac{1}{\pi^2 r^2} \sin^2 \left(\frac{\pi r}{2} \right) . \end{aligned} \tag{35}$$

We see that there are antiferromagnetic spin correlations (the correlation function is oscillatory) that decay as a power of separation — behaviour quite different to that in the classical Néel state.

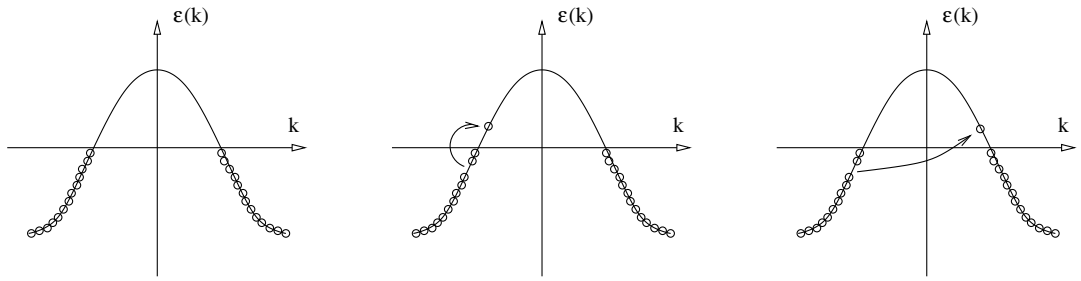


Figure 1: Ground state and excited states of spin one-half XY chain, in Jordan-Wigner fermion description

5.1.4 Excitations

Just as the ground state in this one-dimensional model is different from the ordered state we have discussed for higher-dimensional systems, so are excitations in one dimension are different from the spin waves we treated in higher dimensions. Note that the total number of Jordan-Wigner fermions, measured relative to a half-filled lattice, is proportional to the z-component of total spin. Excitations that do not change $S^z_{\rm tot}$ therefore involve a rearrangement of fermions in orbitals, without change in total fermion number. The simplest such excitation involves creating a particle-hole pair on top of the ground state. It can be characterised by its total momentum q, and low-lying states are of two types, with either $|q| \ll 1$ or $q \approx \pi$, as illustrated in Fig. 1. More generally, a given total excitation momentum q can be distributed between the particle and hole in a range of ways, so that a range of total energy is possible for a given momentum. This means that, instead of the sharp dispersion relation we found for spin waves, we have in this one-dimensional model a continuum of excitation energies – see Fig. 2.

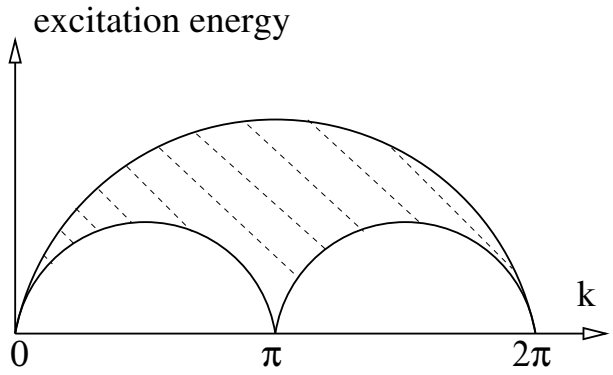


Figure 2: Range of possible energies for particle-hole excitations, as a function of total momentum.

5.2 Integer spin chains

Historically, many aspects of the behaviour of the spin-half chain were understood before higher spin versions (Bethe's exact solution of the spin-half Heisenberg model was published in 1931, although it took over 30 years before its physical interpretation was complete). It was Haldane's work in 1983 that showed there is qualitatively different behaviour in integer spin chains. We were able to discuss behaviour for spin one-half in a relatively simple fashion by treating the XY model, rather than the Heisenberg case. For spin one there is similarly a simplification, developed by AKLT (Affleck, Kennedy, Lieb and Tasaki). As even this simplified version is considerably more complicated than the free fermion problem arising from the spin-half XY model, we will discuss it only in a pictorial way.

The essential idea is to view the spin at each site in a spin-one chain as being a composite of two spin-half objects, taken in a symmetric combination. A wavefunction for the chain can be constructed by forming singlets across each bond in such a way that at each site, one spin-half is paired with the site to the left, and the other with the site to the right, as sketched in Fig. 3. Such a state is an exact ground state for a special Hamiltonian that has both Heisenberg $(\mathbf{S}_n \cdot \mathbf{S}_{n+1})$ and biquadratic $([\mathbf{S}_n \cdot \mathbf{S}_{n+1}]^2)$ exchange with suitably chosen relative strengths.

To see this, consider – for the wavefunction we have described – possible values of the total spin $S^{\rm tot}_{n,n+1}$ of two neighbouring sites, n and $n+1$. This total spin is built from four spin-half objects, two of which are in a singlet state. It may therefore take the values 0 or 1, but cannot take the value 2. Such a state is annihilated by the projection operator $\mathcal{P}_2(\mathbf{S}_n + \mathbf{S}_{n+1})$ onto $S^{\rm tot}_{n,n+1} = 2$. It is therefore a zero-energy eigenstate of the Hamiltonian

$$\mathcal{H} = J \sum_n \mathcal{P}_2(\mathbf{S}_n + \mathbf{S}_{n+1}) \tag{36}$$

and is a ground state for antiferromagnetic exchange ($J > 0$), since $\mathcal{H}$ in this case is a sum of non-negative terms. The projection operator can be written explicitly as

$$\mathcal{P}(\mathbf{L}) = |\mathbf{S}_n + \mathbf{S}_{n+1}|^2(|\mathbf{S}_n + \mathbf{S}_{n+1}|^2 - 2),$$

and expansion of this expression yields Heisenberg and biquadratic terms as discussed.

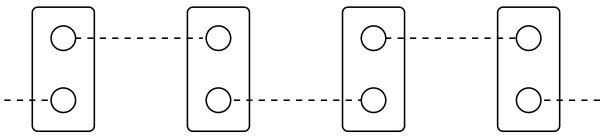


Figure 3: Schematic representation of the AKLT wavefunction. Boxes represent sites of the spin chain, and small circles represent spin one-half objects that together form spin one degrees of freedom. Dashed lines indicate that spin one-half objects from adjacent sites are in singlet states.

It is plausible and true (though the proof takes some work) that this wavefunction has only short range spin correlations. Note that if we wished to construct a similar state for spin one-half, we would be forced to break translation symmetry, because with just a single spin-half object at each site, we can form singlets only across alternate bonds.

6 Weakly interacting Bose gas

As a final example of a system of bosons, we treat excitations in a Bose gas with repulsive interactions between particles, using an approximation that is accurate if interactions are weak. There is good reason for wanting to understand this problem in connection with the phenomenon of superfluidity: the flow of Bose liquids without viscosity below a transition temperature, as first observed below 2.1 K in liquid ^{4}He. Indeed, an argument due to Landau connects the existence of superfluidity with the form of the excitation spectrum, and we summarise this argument next.

6.1 Critical superfluid velocity: Landau argument

Consider superfluid of mass M flowing with velocity $\mathbf{v}$, and examine whether friction can arise by generation of excitations, characterised by a wavevector $\mathbf{k}$ and an energy $\epsilon(k)$. Suppose production of one such excitation reduces the bulk velocity to $\mathbf{v} - \Delta\mathbf{v}$. From conservation of momentum

$$M\mathbf{v} = M\mathbf{v} - M\Delta\mathbf{v} + \hbar\mathbf{k}$$

and from conservation of energy

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

From these conditions we find at large M that $\mathbf{k}$, $\mathbf{v}$ and $\epsilon(k)$ should satisfy $\hbar\mathbf{k}\cdot\mathbf{v} = \epsilon(k)$. The left hand side of this equation can be made arbitrarily close to zero by choosing $\mathbf{k}$ to be almost perpendicular to $\mathbf{k}$, but it has a maximum for a given k, obtained by taking $\mathbf{k}$ parallel to $\mathbf{v}$. If $\hbar kv < \epsilon(k)$ for all k then the equality cannot be satisfied and frictional processes of this type are forbidden. This suggests that there should be a critical velocity $v_{\rm c}$ for superfluid flow, given by $v_{\rm c} = \min_k[\epsilon(k)/k]$. For $v_{\rm c}$ to be non-zero, we require a real, interacting Bose liquid to behave quite differently from the non-interacting gas, since without interactions the excitation energies are just those of individual particles, giving $\epsilon(k) = \hbar^2k^2/2m$ for bosons of mass m, and hence $v_{\rm c} = 0$. Reassuringly, we will find from the following calculation that interactions have the required effect. For completeness, we should note also that while a critical velocity of the magnitude these arguments suggest is observed in appropriate experiments, in others there can be additional sources of friction that lead to much lower values of $v_{\rm c}$.

6.2 Model for weakly interacting bosons

There are two contributions to the Hamiltonian of an interacting Bose gas: the single particle kinetic energy $\mathcal{H}_{\rm KE}$ and the interparticle potential energy $\mathcal{H}_{\rm int}$. We introduce boson creation and annihilation operators for plane wave states in a box with side L, as in Section 2.3. Then

$$\mathcal{H}_{\rm KE} = \sum_{\mathbf{k}} \frac{\hbar^2 k^2}{2m} c_{\mathbf{k}}^\dagger c_{\mathbf{k}} \; .$$

Short range repulsive interactions of strength parameterised by u are represented in first-quantised form by

$$\mathcal{H}_{\rm int} = \frac{u}{2} \sum_{i \neq j} \delta(\mathbf{r}_i - \mathbf{r}_j) \; .$$

Using Eq. (8) this can be written as

$$\mathcal{H}_{\rm int} = \frac{u}{2L^3} \sum_{\mathbf{kpq}} c_{\mathbf{k}}^\dagger c_{\mathbf{p}}^\dagger c_{\mathbf{q}} c_{\mathbf{k+p-q}} \; .$$

With this, our model is complete, with a Hamiltonian $\mathcal{H} = \mathcal{H}_{\rm KE} + \mathcal{H}_{\rm int}$.

6.3 Approximate diagonalisation of Hamiltonian

In order to apply the techniques set out in Section 2.1 we should approximate $\mathcal{H}$ by a quadratic Hamiltonian. The approach to take is suggested by recalling the ground state of the non-interacting Bose gas, in which all particles occupy the $\mathbf{k} = \mathbf{0}$ state. It is natural to suppose that the occupation of this orbital remains macroscopic for small u, so that the ground state expectation value $\langle c_{\mathbf{0}}^\dagger c_{\mathbf{0}} \rangle$ takes a value N_0 which is of the same order as N, the total number of particles. In this case we can approximate the operators $c_{\mathbf{0}}^\dagger$ and $c_{\mathbf{0}}$ by the c-number $\sqrt{N_0}$ and expand $\mathcal{H}$ in decreasing powers of N_0. We find

$$\mathcal{H}_{\rm int} = \frac{u N_0^2}{2L^3} + \frac{u N_0}{2L^3} \sum_{\mathbf{k} \neq \mathbf{0}} \left[2 c_{\mathbf{k}}^\dagger c_{\mathbf{k}} + 2 c_{-\mathbf{k}}^\dagger c_{-\mathbf{k}} + c_{\mathbf{k}}^\dagger c_{-\mathbf{k}}^\dagger + c_{\mathbf{k}} c_{-\mathbf{k}} \right] + \mathcal{O}([N_0]^0) \; .$$

At this stage N_0 is unknown, but we can write an operator expression for it, as

$$N_0 = N - \sum_{\mathbf{k} \neq \mathbf{0}} c_{\mathbf{k}}^\dagger c_{\mathbf{k}} \; .$$

It is also useful to introduce notation for the average number density $\rho = N/L^3$. Substituting for N_0 we obtain

$$\mathcal{H}_{\rm int} = \frac{u\rho}{2} N + \frac{u\rho}{2} \sum_{\mathbf{k} \neq \mathbf{0}} \left[c_{\mathbf{k}}^\dagger c_{\mathbf{k}} + c_{-\mathbf{k}}^\dagger c_{-\mathbf{k}} + c_{\mathbf{k}}^\dagger c_{-\mathbf{k}}^\dagger + c_{\mathbf{k}} c_{-\mathbf{k}} \right] + \mathcal{O}([N_0]^0)$$

and hence

$$\mathcal{H} = \frac{u\rho}{2} N + \frac{1}{2} \sum_{\mathbf{k} \neq \mathbf{0}} \left[E(k) \left(c_{\mathbf{k}}^\dagger c_{\mathbf{k}} + c_{-\mathbf{k}}^\dagger c_{-\mathbf{k}} \right) + u\rho \left(c_{\mathbf{k}}^\dagger c_{-\mathbf{k}}^\dagger + c_{\mathbf{k}} c_{-\mathbf{k}} \right) \right] + \ldots \tag{37}$$

with

$$E(k) = \frac{\hbar^2 k^2}{2m} + u\rho \; .$$

At this order we have a quadratic Hamiltonian, which we can diagonalise using the Bogoliubov transformation for bosons set out in Section 2.2.2. From Eq. (11), we find that the dispersion relation for excitations in the Bose gas is

$$\epsilon(k) = \left[\left(\frac{\hbar^2 k^2}{2m} + u\rho \right)^2 - (u\rho)^2 \right]^{1/2} \; .$$

At large k ($\hbar^2 k^2 / 2m \gg u\rho$), this reduces to the dispersion relation for free particles, but in the opposite limit it has the form

$$\epsilon(k) \simeq \hbar v k \qquad \text{with} \quad v = \sqrt{\frac{u\rho}{m}} \; .$$

In this way we obtain a critical velocity for superfluid flow, which is proportional to the interaction strength u, illustrating how interactions can lead to behaviour quite different from that in a non-interacting system.

7 Landau theory of Fermi liquids

We now switch our attention to systems of fermions. Our reference point is the free Fermi gas, and in this and later sections we will consider a sequence of modifications to the non-interacting gas that lead to increasingly significant changes in physical behaviour. The most important physical example of a system of fermions in condensed matter physics is of course the electron gas in a metal, but it is good to keep in mind as well both the astrophysical examples of white dwarf stars and neutron stars, and from terrestrial low-temperature physics, the case of liquid ^{3}He. This last system is particularly simple in the sense that it is translationally and rotationally symmetric, there being (in contrast to the case of metals) no background lattice of neutralising ions.

Let's recall some of the distinctive properties of the free Fermi gas at temperatures low compared to the Fermi temperature $T_{\rm F}$. The heat capacity is linear in temperature, and the Pauli susceptibility is constant, both being suppressed by a factor of $T/T_{\rm F}$ compared to their values in a non-degenerate system, as a result of the Pauli exclusion principle. It is remarkable that the same behaviour is measured for electrons in metals and for liquid ^{3}He, since in these systems the scale for interaction energies is typically comparable with the Fermi energy and certainly much larger than the energy scale of the excitations relevant for these physical properties. The objective of the Landau theory of Fermi liquids is to understand why interactions have no qualitative effect, and to characterise their residual, quantitative consequences.

Before starting our discussion of Fermi liquid theory, it is interesting to consider in a little more detail how we can characterise the strength of interactions for an electron gas in a uniform neutralising background. This (the jellium model) is a simple situation because it is parameterised by a single quantity: the number density n. This quantity sets both the electron spacing ($\propto n^{-1/3}$) and the Fermi wavevector ($\propto n^{1/3}$), and we can write the ratio of Coulomb to kinetic energies as

$$\frac{\text{Coulomb energy}}{\text{kinetic energy}} \propto \frac{(e^2 n^{1/3}/4\pi\epsilon_0)}{(\hbar^2 n^{2/3}/m)} = \frac{1}{a_o n^{1/3}}$$

where $a_0 = 4\pi\hbar^2\epsilon_0/me^2$ is the Bohr radius. The conventionally used parameter is in fact $r_{\rm s}$, the radius in units of a_0 of a sphere containing one electron, so that $4\pi r_{\rm s}^3/3 = 1/(na_0^3)$, and for typical metals $r_{\rm s}$ lies in the range $1.8 - 6$. Clearly, if Coulomb interactions dominate, $r_{\rm s}$ is large, and if kinetic energy dominates, $r_{\rm s}$ is small. It is at first a surprise to see that interactions are, relatively speaking, weak in the high density limit – the point is that although Coulomb energies grow with increasing density, the kinetic energy grows faster.

We can ask what happens if $r_{\rm s}$ is very large, so that Coulomb repulsion overwhelms the kinetic energy. This constitutes a classical limit, since the zero-point motion is then negligible, and it is straightforward to see that the ground state should involve a crystalline arrangement of electrons, to minimise potential energy. This state is called the Wigner crystal, and the electron gas is known from quantum Monte Carlo simulations to have a first-order transition from a Fermi liquid phase to the Wigner crystal at a critical value $r_{\rm s}^* \sim 100$.

Returning to our main theme of Landau theory, the central assumption is expressed as a statement about the behaviour of the ground state and long-lived excitations if the interaction strength is varied from zero to its physical value: we suppose that the ground state and excitations evolve smoothly. This means we assume that excitations in the interacting system can be labelled using the same set of quantum numbers (wavevector and spin) as in the ideal Fermi gas. It also means that we assume there are no ground-state phase transitions for interaction strengths in this range, and so would fail if we were to pass into a Wigner crystal.

7.1 Lifetime of excitations

For the ideal Fermi gas, states with particle or hole excitations are exact eigenstates, and so have infinite lifetime. This is not the case in the interacting system, and it is important to understand what determines the finite value of excitation lifetimes here. A key argument due to Migdal addresses this issue. Consider the scattering rate between the initial and final states sketched in Fig. 4, in which an initially isolated quasiparticle looses energy by scattering a fermion out of the Fermi sea, leaving a hole behind.

In a calculation of the rate for this process, we should sum over all final states. We can specify the final state in our example via the energies and momenta of the two quasiparticles, since those of the hole are then fixed by conservation of total energy and momentum. Since the energies of the final state quasiparticles cannot exceed that of the initial particle, and since the quasiparticles must lie outside the Fermi sea, the final state sum is highly constrained if the initial quasiparticle energy $\epsilon_{\mathbf{k}}$ is close to the chemical potential μ, yielding a rate that varies as $(\epsilon_{\mathbf{k}} - \mu)^2$ at zero temperature. This is an important conclusion: because the scattering rate vanishes more rapidly than the excitation energy $\epsilon_{\mathbf{k}} - \mu$ as the Fermi surface is approached, the energy of low-lying quasiparticles is

Figure 4: Left: initial state with a single quasiparticle excitation above a filled Fermi sea. Right: final state with two quasiparticles and a quasihole

(in the limit) sharply defined. That is to say, the idea of a sharp Fermi surface is self-consistent, because of the restrictions on scattering processes imposed by Pauli exclusion. Finite temperature sets a lifetime for quasiparticles at μ proportional to T^2.

7.2 Relation between bare fermions and quasiparticles

The action on the ground state wavefunction $|0\rangle$ for the interacting system of an annihilation operator $c_{\mathbf{k}\sigma}$ for a bare fermion with wavevector $\mathbf{k}$ and spin σ has an amplitude to generate a state $|\mathbf{k}\sigma\rangle$ containing a quasihole with these quantum numbers if $|\mathbf{k}| < k_{\rm F}$. It also has an amplitude to generate superpositions of many excitations, which we denote as $|\mathrm{incoherent}\rangle$, and we expect the amplitude for such processes to vary smoothly with $\mathbf{k}$. Denoting the amplitude for creation of a quasiparticle by $Z^{1/2}$, we can summarise these ideas by writing

$$c_{\mathbf{k}\sigma}|0\rangle \sim \begin{cases} Z^{1/2}|\mathbf{k}\sigma\rangle + |\mathrm{incoherent}\rangle & |\mathbf{k}| < k_{\rm F} \\ |\mathrm{incoherent}\rangle & |\mathbf{k}| > k_{\rm F} \end{cases} \tag{38}$$

Hence the dependence of $\langle 0|c^\dagger_{\mathbf{k}\sigma}c_{\mathbf{k}\sigma}|0\rangle$ on $|\mathbf{k}|$ has a step of size Z at $|\mathbf{k}| = k_{\rm F}$. This is known as the Migdal discontinuity and is a demonstration of the existence of a sharp Fermi surface. In the free fermion system $Z = 1$; the effect of interactions is to decrease Z and to give excitations an effective mass larger than that of the bare particles.

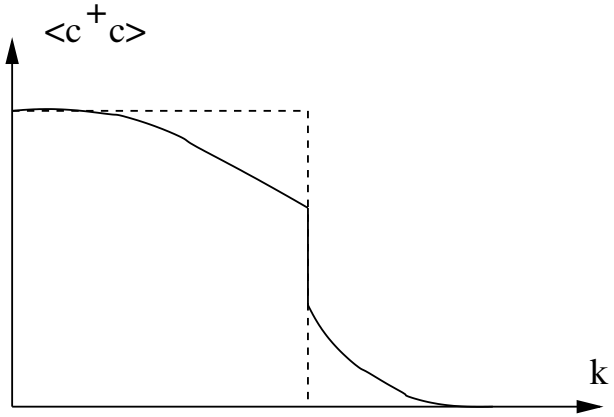


Figure 5: Relation between bare fermions and quasiparticles: dependence of occupation number on $|\mathbf{k}|$ in the ground states of a free Fermi gas (dashed line) and in interacting Fermi liquid (full line).

7.3 Parameterising excitation energies

Having established the idea that excitations are of the same kind as in a free gas, and have sharply defined energies, it remains to discuss how these energies are influenced by interactions. We specify the state of the system in terms of the occupation number $n_{\mathbf{k}\sigma}$ for quasiparticles with wavevector $\mathbf{k}$ and spin σ, and write the energies of these quasiparticles as $\epsilon_{\mathbf{k}\sigma}$. We separate ground state and excitation contributions by writing $n_{\mathbf{k}\sigma} = n^0_{\mathbf{k}\sigma} + \delta n_{\mathbf{k}\sigma}$, where $n^0_{\mathbf{k}\sigma} = 1$ if in the ground state $\epsilon_{\mathbf{k}\sigma} < \mu$ and is zero otherwise. We expect the energy $\epsilon_{\mathbf{k}\sigma}$ of a given quasiparticle to depend on the occupation $\delta n_{\mathbf{q}\sigma'}$ of all other excitations: the idea of Landau theory is to represent this dependence

using the first terms in a Taylor series, as

$$\epsilon_{\mathbf{k}\sigma} - \mu = \frac{\hbar^2 k_{\rm F}}{m^*}(k - k_{\rm F}) + \sum_{\mathbf{q}\sigma'} f(\mathbf{k}\sigma, \mathbf{q}\sigma')\delta n_{\mathbf{q}\sigma'} \,. \tag{39}$$

Here the zeroth order term in δn is assumed linear in the radial deviation $k - k_{\rm F}$ from the Fermi surface, and its magnitude is characterised by an effective mass m^*, while the first order terms involve the Landau f-parameters. Note that for physically important excited states, both the relevant values of $k - k_{\rm F}$ and the fraction of non-zero $\delta n_{\mathbf{q}\sigma'}$ are small, so that the two terms retained in Eq. (39) are comparable in magnitude, and parametrically larger that the neglected higher order terms. At this stage, the approach seems unpromising, because the expansion coefficients involve not simply a few fitting parameters but instead an unknown function $f(\mathbf{k}\sigma, \mathbf{q}\sigma')$. We make things manageable by separating $\delta n_{\mathbf{q}\sigma'}$ into spherical harmonics, and recognising that only the lowest two harmonics are generated in situations of physical interest. In turn, and assuming a spherically symmetric Fermi surface, only the zeroth and first harmonics of $f(\mathbf{k}\sigma, \mathbf{q}\sigma')$ are important, and symmetrising also in spin labels we are left with just three significant Landau parameters. Together with the effective mass they characterise interaction effects.

In more detail, we expect $f(\mathbf{k}\sigma, \mathbf{q}\sigma')$ to depend (for $\mathbf{k}$ and $\mathbf{q}$ close to the Fermi surface) only on the angle θ between these wavevectors, and so (suppressing spin labels) we write

$$f(\mathbf{k}\sigma, \mathbf{q}\sigma') = \sum_l f_l P_l(\cos\theta)$$

with $P_l(\cos\theta)$ the Legendre polynomials. Similarly, we write

$$f(\mathbf{k}\uparrow, \mathbf{q}\uparrow) = f(\mathbf{k}\downarrow, \mathbf{q}\downarrow) = f^{\rm s}_{\mathbf{kq}} + f^{\rm a}_{\mathbf{kq}} \qquad \text{and} \qquad f(\mathbf{k}\uparrow, \mathbf{q}\downarrow) = f(\mathbf{k}\downarrow, \mathbf{q}\uparrow) = f^{\rm s}_{\mathbf{kq}} - f^{\rm a}_{\mathbf{kq}} \,,$$

and finally we use the density of states at the Fermi surface $\nu(E_{\rm F})$ to form dimensionless combinations $F = \nu(E_{\rm F})f$. The Fermi liquid is then parameterised by

$$F^{\rm s}_0 \qquad F^{\rm a}_0 \qquad F^{\rm s}_1 \quad \text{and} \quad m^*$$

and of these only three are independent, because m^* and $F^{\rm s}_1$ are related.

7.4 Measuring Landau parameters

To understand the physical significance of these parameters, we should consider the situations in which each of them becomes important, by examining different ways of exciting the Fermi liquid.

7.4.1 Heat capacity

Finite temperature generates a distribution of excitations in which there are equal numbers of quasiparticles and quasiholes, so that the density integrated over the radial component of wavevector vanishes:

$$\int {\rm d}k \, \delta n_{\mathbf{k}\sigma} = 0 \,.$$

For this reason interactions affect the heat capacity $C_{\rm V}$ only via the value of effective mass, and

$$C_{\rm V} = \frac{\pi^2}{3} k_{\rm B}^2 \nu(E_{\rm F}) T = \frac{k_{\rm F} k_{\rm B}^2}{3\hbar^2} m^* T \,.$$

7.4.2 Compressibility

An increase in density can be represented as an isotropic, spin-independent $\delta n_{\mathbf{k}\sigma}$. Let

$$\delta n = \sum_{\mathbf{k}\sigma} \delta n_{\mathbf{k}\sigma} \,.$$

The resulting change in the total energy of the system is

$$\begin{aligned} \delta E &= \frac{\hbar^2 k_{\rm F}}{m^*} \sum_{\mathbf{k}\sigma} (k - k_{\rm F}) \delta n_{\mathbf{k}\sigma} + \frac{1}{2} \sum_{\mathbf{k}\sigma, \mathbf{q}\sigma'} f(\mathbf{k}\sigma, \mathbf{q}\sigma') \, \delta n_{\mathbf{k}\sigma} \, \delta n_{\mathbf{q}\sigma'} \\ &= E_{\rm F}^{\rm old} \delta n + \frac{1}{2}(E_{\rm F}^{\rm new} - E_{\rm F}^{\rm old}) \delta n + (f^{\rm s}_0 + f^{\rm a}_0) \left(\frac{\delta n}{2}\right)^2 + (f^{\rm s}_0 - f^{\rm a}_0) \left(\frac{\delta n}{2}\right)^2 . \end{aligned} \tag{40}$$

Now, we also have the relation

$$\delta n = (E_{\rm F}^{\rm new} - E_{\rm F}^{\rm old})\nu(E_{\rm F})$$

and so the change in energy as a result of a volume change is

$$\delta E = E_{\rm F}^{\rm old}\delta n + \frac{1}{2\nu(E_{\rm F})}[1 + F_0^{\rm s}](\delta n)^2 \,.$$

From the energy change we can obtain the compressibility κ, since this quantity, the pressure p, the volume V and the energy E of a system are related by

$$p = -\frac{\partial E}{\partial V} \qquad \text{and} \qquad \kappa^{-1} = -V\frac{\partial p}{\partial V}$$

giving

$$\kappa = \frac{\nu(E_{\rm F})}{V} \cdot \frac{1}{1 + F_0^{\rm s}} \,.$$

In this result, the first factor is the contribution from the degeneracy pressure of free fermions (note that we have chosen to define $\nu(E_{\rm F})$ for the system as a whole rather then per unit volume, and so it is proportional to V), while the second factor represents the influence of interactions between quasiparticles, which reduce the compressibility if they are repulsive ($F_0^{\rm s} > 0$), as one would expect.

7.4.3 Susceptibility

We can probe the Landau parameter $F_0^{\rm a}$ by considering a measurement of the Pauli susceptibility, since a Zeeman field generates a spherically symmetric distribution of quasiparticles with opposite signs of $\delta n_{\mathbf{k}\sigma}$ for spins orientated parallel or antiparallel to the Zeeman field of strength H.

Let

$$\delta n_\uparrow \equiv \sum_{\mathbf{k}} \delta n_{\mathbf{k}\uparrow} = -\delta n_\downarrow \equiv \sum_{\mathbf{k}} \delta n_{\mathbf{k}\downarrow} \,.$$

Then the magnetisation of the system (writing g for the g-factor of the quasiparticles) is

$$M = \frac{1}{2} g\mu_{\rm B}(\delta n_\uparrow - \delta n_\downarrow) = g\mu_{\rm B}\delta n_\uparrow$$

and the change in total energy, consisting of Zeeman, kinetic and interaction terms, is

$$\begin{aligned} \delta E &= -g\mu_0\mu_{\rm B}H\delta n_\uparrow + \frac{\hbar^2 k_{\rm F}}{m^*}\sum_{\mathbf{k}\sigma}(k - k_{\rm F})\delta n_{\mathbf{k}\sigma} + \frac{1}{2}\sum_{\mathbf{k}\sigma,\mathbf{q}\sigma'} f(\mathbf{k}\sigma, \mathbf{q}\sigma')\,\delta n_{\mathbf{k}\sigma}\delta n_{\mathbf{q}\sigma'} \\ &= -g\mu_0\mu_{\rm B}H\delta n_\uparrow + \frac{2}{\nu(E_{\rm F})}(\delta n_\uparrow)^2 + \frac{2F_0^{\rm a}}{\nu(E_{\rm F})}(\delta n_\uparrow)^2 \,. \end{aligned}$$

Minimising with respect to $\delta n_\uparrow$ yields the equilibrium value of the magnetisation and the susceptibility

$$\chi = \frac{\partial M}{\partial H} = \frac{\mu_{\rm B}^2\mu_0\nu(E_{\rm F})}{1 + F_0^{\rm a}} \,.$$

In this expression the numerator is the free fermion result modified by replacing bare mass with effective mass, while the denominator includes the influence of interactions between quasiparticles. Note that an attractive interaction between quasiparticles with the same spin leads to a negative value for $F_0^{\rm a}$ and an enhancement of χ. In the limit$F_0^{\rm a} \to -1$ this produces an instability towards ferromagnetic order.

7.4.4 Galilean invariance

The requirement of Galilean invariance leads to the relation

$$\frac{m^*}{m} = 1 + \frac{F_1^{\rm s}}{3}$$

and the derivation of this result is set as Question 1 on Problem Sheet 2.

7.4.5 Fermi liquid parameters for ^{3}He

It is interesting to see the measured values of the Landau parameters for ^{3}He, shown in the table below. Note that the effective mass is greatly enhanced compared to the bare mass, and that this enhancement increases with increasing density. Note also that the liquid is quite close to a ferromagnetic instability.

	m^*/m	$F_1^{\rm s}$	$F_0^{\rm s}$	$F_0^{\rm a}$
Low pressure (0.3 atmospheres)	3.1	6.3	10.8	-0.67
High pressure (27 atmospheres)	5.8	14.4	75.6	-0.72

Table 1: Fermi liquid parameters for liquid ^{3}He (from Pines and Nozieres, *The Theory of Quantum Liquids*).

8 BCS theory of superconductivity

We have seen that the Fermi liquid is stable to weak repulsive interactions, in the sense that excitations retain their character though their energy is modified. Attractive interactions by contrast lead to a qualitative change in the ground state and low-temperature properties, no matter how weak they are.

The central idea of BCS theory is that electron-phonon interactions lead to the formation of bound pairs of electrons, known as Cooper pairs, which in a sense Bose condense. However, the characteristic size of Cooper pairs – the coherence length – is much larger than their separation, so binding and condensation must be treated together in the theory. The same fact also leads to a simplification: since each Cooper pair interacts with many others, mean field theory is a good approximation.

From a historical perspective, it is striking how long the interval was between the experimental discovery of superconductivity, by Onnes in 1911, and the theoretical understanding due to Bardeen, Cooper and Schrieffer in 1957: this serves to underline what a revolutionary advance their treatment of a cooperative quantum phenomenon represents.

8.1 Electron-phonon interactions

Experiments on the isotope effect showed that phonons are central to superconductivity. In the ideal case, for different isotopes of the same superconductor the energy scales represented by the critical temperature $T_{\rm c}$ and the critical field $H_{\rm c}$ vary with isotope mass like phonon frequencies, as (ionic mass)$^{-1/2}$. For a pair of electrons that are close in energy, phonon exchange generates an attractive interaction that beats the obvious screened Coulomb repulsion.

To derive this effective interaction we start from the Hamiltonian $\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_1$ written in terms of electron operators $c^\dagger_{\mathbf{k}}$ and $c_{\mathbf{k}}$, and phonon operators $a^\dagger_{\mathbf{q}}$ and $a_{\mathbf{q}}$ as

$$\mathcal{H}_0 = \sum_{\mathbf{k}} \epsilon(\mathbf{k}) c^\dagger_{\mathbf{k}} c_{\mathbf{k}} + \hbar\omega \sum_{\mathbf{q}} a^\dagger_{\mathbf{q}} a_{\mathbf{q}} \qquad \text{and} \qquad \mathcal{H}_1 = \sum_{\mathbf{kq}} \left(M c^\dagger_{\mathbf{k+q}} c_{\mathbf{k}} a_{\mathbf{q}} + \text{h.c.} \right) .$$

Here $\epsilon(\mathbf{k})$ is the electron dispersion relation and the phonons are represented as Einstein oscillators, all with frequency ω; the electron-phonon coupling is represented by the matrix element M; and we have omitted spin labels, though they will be crucial later.

We wish to focus on the electron system. To this end we eliminate the electron-phonon coupling by means of a canonical transformation, which we determine perturbatively. We write

$$\tilde{\mathcal{H}} = \mathrm{e}^{-S} \mathcal{H} \mathrm{e}^{S} = \mathcal{H} + [\mathcal{H}, S] + \frac{1}{2}[[\mathcal{H}, S], S] + \ldots$$

and at leading order we fix S simply by setting

$$\mathcal{H}_1 + [\mathcal{H}_0, S] = 0 \,, \tag{41}$$

yielding

$$\tilde{\mathcal{H}} = \mathcal{H}_0 + \frac{1}{2}[\mathcal{H}_1, S] \equiv \mathcal{H}_0 + \mathcal{H}_{\rm int} \,. \tag{42}$$

To find S explicitly, we take matrix elements of Eq. (41) in eigenstates of $\mathcal{H}_0$, which satisfy $\mathcal{H}_0|n\rangle = E_n|n\rangle$. In this way we obtain

$$\langle n|S|m\rangle = \frac{\langle n|\mathcal{H}_1|m\rangle}{E_m - E_n} \qquad \text{and} \qquad \langle f|\mathcal{H}_{\rm int}|i\rangle = \frac{1}{2}\sum_v \langle f|\mathcal{H}_1|v\rangle\langle v|\mathcal{H}_1|i\rangle \left(\frac{1}{E_i - E_v} + \frac{1}{E_f - E_v}\right).$$

From its matrix elements we can read off $\mathcal{H}_{\rm int}$ in operator form, as

$$\mathcal{H}_{\rm int} = \frac{1}{2}\sum_{\mathbf{pkq}} c^\dagger_{\mathbf{p}+\mathbf{q}} c_{\mathbf{p}} c^\dagger_{\mathbf{k}-\mathbf{q}} c_{\mathbf{k}} |M|^2 \left(\frac{1}{\epsilon(\mathbf{k}) - \epsilon(\mathbf{k}-\mathbf{q}) - \hbar\omega} + \frac{1}{\epsilon(\mathbf{p}+\mathbf{q}) - \epsilon(\mathbf{p}) - \hbar\omega}\right).$$

8.2 The Cooper problem

This is an attractive interaction for pairs of electrons with $|\epsilon(\mathbf{k}) - \epsilon(\mathbf{k}-\mathbf{q})| < \hbar\omega$: that is, for pairs within the Debye energy $\hbar\omega_{\rm D}$ of the Fermi surface. It is, however, typically very weak. This leads us to a puzzle: for two particles moving in free space in three dimensions, an attractive interaction must exceed a critical strength to produce a bound state. So how can a weak attraction generate superconductivity? The Cooper problem takes us one step towards answering this question: we consider a pair of particles moving not in free space, but above a filled Fermi sea, and will find that Pauli exclusion facilitates binding.

Consider the wavefunction for this pair of particles. We want to write a low-energy state, and so we set the centre-of-mass momentum to zero. To take advantage of a local attractive interaction, we choose the pair to be in a spin-singlet state so that the spatial wavefunction is symmetric. And to respect Pauli exclusion from a filled Fermi sea, we require the wavefunction to be built from orbitals outside the Fermi surface. The general form is then

$$\psi(\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2) = (\uparrow_1\downarrow_2 - \downarrow_1\uparrow_2) \cdot \sum_{|\mathbf{k}|>k_{\rm F}} g_{\mathbf{k}} {\rm e}^{{\rm i}\mathbf{k}\cdot(\mathbf{r}_1 - \mathbf{r}_2)}$$

with $g_{\mathbf{k}}$ determined by requiring this to be a solution to the two-particle Schrödinger equation. Writing the pair energy as E and the pair potential as $U(\mathbf{r}_1 - \mathbf{r}_2)$, we have

$$\sum_{|\mathbf{k}|>k_{\rm F}} g_{\mathbf{k}}\, {\rm e}^{{\rm i}\mathbf{k}\cdot(\mathbf{r}_1-\mathbf{r}_2)} U(\mathbf{r}_1 - \mathbf{r}_2) = \sum_{|\mathbf{k}|>k_{\rm F}} (E - 2\epsilon(\mathbf{k})) g_{\mathbf{k}}\, {\rm e}^{{\rm i}\mathbf{k}\cdot(\mathbf{r}_1-\mathbf{r}_2)}\,. \tag{43}$$

To solve this Schrödinger equation, we operate on both sides with

$$\frac{1}{V}\int {\rm d}^d(\mathbf{r}_1 - \mathbf{r}_2) \ldots {\rm e}^{-{\rm i}\mathbf{q}\cdot(\mathbf{r}_1-\mathbf{r}_2)}$$

(where V is the system volume) and introduce the notation $\int {\rm d}^d\mathbf{r}\, U(\mathbf{r}) {\rm e}^{{\rm i}(\mathbf{k}-\mathbf{q})\cdot\mathbf{r}} = U_{\mathbf{kq}}$. Then Eq. (43) becomes

$$\frac{1}{V}\sum_{|\mathbf{k}|>k_{\rm F}} g_{\mathbf{k}} U_{\mathbf{kq}} = (E - 2\epsilon(\mathbf{q})) g_{\mathbf{q}}\,.$$

We can understand the essentials in a simple way by taking

$$U_{\mathbf{kq}} = \begin{cases} -U & \text{if } \epsilon(\mathbf{k}) \text{ and } \epsilon(\mathbf{q}) \text{ are within } \hbar\omega_{\rm D} \text{ of } E_{\rm F} \\ 0 & \text{otherwise}\,. \end{cases} \tag{44}$$

Then

$$\frac{U}{V}\sum_{\mathbf{k}}{}' g_{\mathbf{k}} = (2\epsilon(\mathbf{q}) - E) g_{\mathbf{q}}$$

where $\sum'_{\mathbf{k}}$ is a sum over states within $\hbar\omega_{\rm D}$ of $E_{\rm F}$.

$$\frac{1}{V}\sum_{\epsilon=E_{\rm F}}^{\epsilon=E_{\rm F}+\hbar\omega_{\rm D}} \frac{1}{2\epsilon - E} = \frac{1}{U}\,.$$

With a constant density of states ρ per unit volume, this yields

$$\int_{E_{\rm F}}^{E_{\rm F}+\hbar\omega_{\rm D}} {\rm d}\epsilon\, \frac{\rho}{2\epsilon - E} = \frac{\rho}{2}\ln\left[\frac{2(E_{\rm F} + \hbar\omega_{\rm D}) - E}{2E_{\rm F} - E}\right] = \frac{1}{U}\,.$$

There is a bound state for any positive U, and at weak coupling the binding energy is

$$2E_{\rm F} - E = 2\hbar\omega_{\rm D} {\rm e}^{-2/\rho U}\,.$$

Strikingly, this form is non-perturbative in ρU.

8.3 The BCS wavefunction

Given a wavefunction for a pair of electrons of the form we have been considering

$$\phi(\mathbf{r}_1, \mathbf{r}_2) = (\uparrow_1\downarrow_2 - \downarrow_1\uparrow_2) g(\mathbf{r}_1 - \mathbf{r}_2)$$

we can attempt to write a wavefunction for $2N$ electrons in which pairs are Bose condensed, of the form

$$\psi(\mathbf{r}_1 \ldots \mathbf{r}_{2N}) = \mathcal{A} \prod_{n=1}^{n=N} \phi(\mathbf{r}_{2n}, \mathbf{r}_{2n+1})$$

where $\mathcal{A}$ denotes antisymmetrisation. It is initially unclear whether Pauli exclusion allows this – whether a non-zero wavefunction survives after the operation of $\mathcal{A}$ when N is large. That it does is clearer when we write both these states in k-space and using operator notation, as

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

and

$$|\psi\rangle = \prod_{n=1}^{N} \left(\sum_{\mathbf{k}_n} g_{\mathbf{k}_n} c^\dagger_{\mathbf{k}_n\uparrow} c^\dagger_{-\mathbf{k}_n\downarrow} \right) |0\rangle \tag{45}$$

since when we expand the product, terms with all $\mathbf{k}_n$'s different will survive. The form of the wavefunction in Eq. (45) is not very convenient, because the occupation of different orbitals is correlated through the constraint that exactly $2N$ electrons are present. The BCS wavefunction relaxes this condition (passing to it is analogous to going from the cannonical to the grand canonical distributions in statistical mechanics). It has the form

$$|\mathrm{BCS}\rangle = \prod_{\mathbf{k}} (u_{\mathbf{k}} + v_{\mathbf{k}} c^\dagger_{\mathbf{k}\uparrow} c^\dagger_{-\mathbf{k}\downarrow}) |0\rangle \,. \tag{46}$$

Here $u_{\mathbf{k}}$ is the amplitude for a pair of orbitals to be empty, and $v_{\mathbf{k}}$ is the amplitude for them to contain a Cooper pair, so we require $|u_{\mathbf{k}}|^2 + |v_{\mathbf{k}}|^2 = 1$ for normalisation, and $\sum_{\mathbf{k}} |v_{\mathbf{k}}|^2 = N$ in a system containing $2N$ electrons on average.

In fact, a wavefunction of the form of Eq. (46) can represent a variety of states, depending on the choice of coefficients $v_{\mathbf{k}}$. A filled Fermi sea simply has $v_{\mathbf{k}} = 1$ for $\mathbf{k}$ inside the Fermi surface and $v_{\mathbf{k}} = 0$ outside. By contrast, in a state containing many Cooper pairs, we expect $v_{\mathbf{k}}$ to vary smoothly between 1 and 0 across a window of width $\hbar\omega_{\mathrm{D}}$ around the Fermi energy. One approach to finding $v_{\mathbf{k}}$ is variational, and the subject of Question 2 on Problem Sheet 2. An alternative is a mean field treatment of the Hamiltonian, as described next.

8.4 Mean field theory

Using the simplified pairing interaction of Eq. (44) the Hamiltonian we are concerned with is

$$\mathcal{H} = \sum_{\mathbf{k}\sigma} \epsilon_{\mathbf{k}} c^\dagger_{\mathbf{k}\sigma} c_{\mathbf{k}\sigma} - U \sum_{\mathbf{kq}}{}' c^\dagger_{\mathbf{k}\uparrow} c^\dagger_{-\mathbf{k}\downarrow} c_{-\mathbf{q}\downarrow} c_{\mathbf{q}\uparrow}$$

where $\sum'_{\mathbf{kq}}$ is a sum on states having energies $\epsilon_{\mathbf{k}}$, $\epsilon_{\mathbf{q}}$ within $\hbar\omega_{\mathrm{D}}$ of E_{F}.

We now wish to treat the interaction term in this Hamiltonian using a mean field approximation, in order to reduce it to a quadratic form. To this end we let $b_{\mathbf{k}} = \langle c^\dagger_{\mathbf{k}\uparrow} c^\dagger_{-\mathbf{k}\downarrow} \rangle$ and write

$$c^\dagger_{\mathbf{k}\uparrow} c^\dagger_{-\mathbf{k}\downarrow} = b_{\mathbf{k}} + \left(c^\dagger_{\mathbf{k}\uparrow} c^\dagger_{-\mathbf{k}\downarrow} - b_{\mathbf{k}} \right)$$

with the idea that the sum over $\mathbf{k}$ of the term in round brackets is small, and so need be taken only to first order. Then with $\xi_{\mathbf{k}} \equiv \epsilon_{\mathbf{k}} - \mu$ we have

$$\mathcal{H} - \mu N = \sum_{\mathbf{k}\sigma} \xi_{\mathbf{k}} c^\dagger_{\mathbf{k}\sigma} c_{\mathbf{k}\sigma} - U \sum_{\mathbf{kq}}{}' \left(b_{\mathbf{k}} c_{-\mathbf{q}\downarrow} c_{\mathbf{q}\uparrow} + b^*_{\mathbf{q}} c^\dagger_{\mathbf{k}\uparrow} c^\dagger_{-\mathbf{k}\downarrow} - b_{\mathbf{k}} b^*_{\mathbf{q}} \right) + \text{fluctuations}.$$

Dropping the fluctuation term and setting $U\sum_{\mathbf{k}}' b_{\mathbf{k}} \equiv \Delta$, this is

$$\mathcal{H} - \mu N = \sum_{\mathbf{k}\sigma} \xi_{\mathbf{k}} (c^\dagger_{\mathbf{k}\uparrow} c_{\mathbf{k}\uparrow} + c^\dagger_{-\mathbf{k}\downarrow} c_{-\mathbf{k}\downarrow}) - \sum_{\mathbf{k}}{}' \left(\Delta\, c_{-\mathbf{k}\downarrow} c_{\mathbf{k}\uparrow} + \Delta^* c^\dagger_{\mathbf{k}\uparrow} c^\dagger_{-\mathbf{k}\downarrow} \right) + \frac{|\Delta|^2}{U} \,. \tag{47}$$

This Hamiltonian is diagonalised by the fermionic Bogoliubov transformation, as described in Section 2.2.1. Setting

$$c^\dagger_{\mathbf{k}\uparrow} = u_{\mathbf{k}} \gamma^\dagger_{\mathbf{k}0} + v^*_{\mathbf{k}} \gamma_{\mathbf{k}1} \quad \text{and} \quad c^\dagger_{-\mathbf{k}\downarrow} = u_{\mathbf{k}} \gamma^\dagger_{\mathbf{k}1} - v^*_{\mathbf{k}} \gamma_{\mathbf{k}0}$$

we require $|u_{\mathbf{k}}|^2 + |v_{\mathbf{k}}|^2 = 1$ to preserve anticommutation relations, and $2\xi_{\mathbf{k}} u_{\mathbf{k}} v_{\mathbf{k}} + \Delta v^2_{\mathbf{k}} - \Delta^* u^2_{\mathbf{k}} = 0$ to eliminate number-changing terms from the transformed Hamiltonian. Writing $\Delta = \mathrm{e}^{\mathrm{i}\phi}|\Delta|$ these conditions are met by

$$u_{\mathbf{k}} = \mathrm{e}^{\mathrm{i}\phi/2} \cos\theta_{\mathbf{k}}, \qquad v_{\mathbf{k}} = \mathrm{e}^{-\mathrm{i}\phi/2} \sin\theta_{\mathbf{k}} \quad \text{and} \quad \cot 2\theta_{\mathbf{k}} = \frac{\xi_{\mathbf{k}}}{|\Delta|} \,.$$

Note that for $\xi_{\mathbf{k}} \gg |\Delta|$, $\theta_{\mathbf{k}} \to 0$ and so $c^\dagger_{\mathbf{k}\uparrow} \sim \gamma^\dagger_{\mathbf{k}0}$ and $c^\dagger_{-\mathbf{k}\downarrow} \sim \gamma^\dagger_{\mathbf{k}1}$. Conversely, for $\xi_{\mathbf{k}} \ll -|\Delta|$, $\theta_{\mathbf{k}} \to \pi/2$ and so $c^\dagger_{\mathbf{k}\uparrow} \sim \gamma_{\mathbf{k}1}$ and $c^\dagger_{-\mathbf{k}\downarrow} \sim \gamma_{\mathbf{k}0}$. The γ-particles thus interpolate between electrons and holes, and at the Fermi energy, where $\xi_{\mathbf{k}} = 0$ and $\theta_{\mathbf{k}} = \pi/4$, have equal electron and hole content.

After Bogoliubov transformation, the Hamiltonian is

$$\mathcal{H} - \mu N = \sum_{\mathbf{k}\sigma} \sqrt{\xi^2_{\mathbf{k}} + |\Delta|^2} \left(\gamma^\dagger_{\mathbf{k}0} \gamma_{\mathbf{k}0} + \gamma^\dagger_{\mathbf{k}1} \gamma_{\mathbf{k}1} \right) + \sum_{\mathbf{k}} \left(\xi_{\mathbf{k}} - \sqrt{\xi^2_{\mathbf{k}} + |\Delta|^2} \right) + \frac{|\Delta|^2}{U} \,. \tag{48}$$

Crucially, the value of $|\Delta|$ must be determined self-consistently. We have

$$|\Delta| = U \sum_{\mathbf{k}}{}' |\langle c^\dagger_{\mathbf{k}\uparrow} c^\dagger_{-\mathbf{k}\downarrow} \rangle| = U \sum_{\mathbf{k}}{}' u_{\mathbf{k}} v^*_{\mathbf{k}} \left(1 - \langle \gamma^\dagger_{\mathbf{k}0} \gamma_{\mathbf{k}0} \rangle - \langle \gamma^\dagger_{\mathbf{k}1} \gamma_{\mathbf{k}1} \rangle \right) \,. \tag{49}$$

At zero temperature $\langle \gamma^\dagger_{\mathbf{k}0} \gamma_{\mathbf{k}0} \rangle = \langle \gamma^\dagger_{\mathbf{k}1} \gamma_{\mathbf{k}1} \rangle = 0$. Then with a constant density of states ρ we have

$$1 = \frac{\rho U}{2} \int_{-\hbar\omega_{\mathrm{D}}}^{-\hbar\omega_{\mathrm{D}}} \frac{\mathrm{d}\xi}{\sqrt{\xi^2 + |\Delta|^2}} \,.$$

For $\hbar\omega_{\mathrm{D}} \gg |\Delta|$ this gives $1 \approx \rho U \ln(\hbar\omega_{\mathrm{D}}/|\Delta|)$ and we find

$$|\Delta| = 2\hbar\omega_{\mathrm{D}} \mathrm{e}^{-1/\rho U}$$

as we did for the pair binding energy in the Cooper problem.

At finite temperature $\langle \gamma^\dagger_{\mathbf{k}0} \gamma_{\mathbf{k}0} \rangle$ and $\langle \gamma^\dagger_{\mathbf{k}1} \gamma_{\mathbf{k}1} \rangle$ are determined from the quasiparticle energies using the Fermi distribution, and the gap $|\Delta|$ decreases as temperature decreases. Above a critical temperature T_{c}, the only solution to the self-consistency condition is $\Delta = 0$. Moreover, since at weak coupling ρU is the only parameter, the energy scales set by the zero-temperature gap and the thermal energy at the critical point have a universal relationship that serves as a test of the theory:

$$\frac{2|\Delta(T{=}0)|}{k_{\mathrm{B}} T_{\mathrm{c}}} = 3.53 \,.$$

The form of the quasiparticle density of states, probed by tunneling spectroscopy, provides another experimental test of these ideas.

9 The Mott transition and the Hubbard model

We now consider the combined consequences for electrons in a solid of electron-electron interactions and the background ionic lattice. Our most important conclusion will be that a new type of insulator is possible, in addition to the band insulator familiar from single-particle theory. Specifically, while in a solid without electron-electron interactions we have a insulator (in the absence of band overlap) when the number of electrons per unit cell is even, we shall see that strong correlations in a half-filled band with an *odd* number of electrons per unit cell can generate a new state known as a Mott insulator.

9.1 The Hubbard model

In the standard theoretical description of this phenomenon we use a tight-binding model

$$\mathcal{H}_0 = -t \sum_{\langle \mathbf{r}\mathbf{r}'\rangle\sigma} c^\dagger_{\mathbf{r}\,\sigma} c_{\mathbf{r}'\sigma}$$

and add only intra-site Coulomb interactions, of the form

$$\mathcal{H}_{\rm I} = U \sum_{\mathbf{r}} n_{\mathbf{r}\uparrow} n_{\mathbf{r}\downarrow} \tag{50}$$

where $n_{\mathbf{r}\sigma} = c^\dagger_{\mathbf{r}\sigma} c_{\mathbf{r}\sigma}$. The Hubbard model Hamiltonian is then

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_{\rm I}\,. \tag{51}$$

We would like to understand its phase diagram as a function of U/t, band-filling and temperature. At small U and above one dimension we expect Fermi liquid behaviour, from the same arguments that led to Landau theory, but at large U with one electron per site we find an insulator, by the following argument. In this regime there is a set of low energy states in which every site is singly occupied. Charge motion requires generation of empty and doubly occupied sites, and so is energetically prohibited.

9.1.1 Relation to the Heisenberg model

At large U and half filling, the set of low-lying states has a degeneracy of 2^N in a system of N lattice sites, arising from the choices of spin orientation at each site, in the limit $t/U \to 0$. Virtual hopping lifts this degeneracy and gives rise to Heisenberg exchange, as follows. Consider a pair of sites and treat the effects of $\mathcal{H}_0$ using perturbation theory. We can label eigenstates of the unperturbed Hamiltonian $\mathcal{H}_{\rm I}$ as $|\uparrow,\uparrow\rangle$, $|\uparrow,\downarrow\rangle$ and so on. The leading contributions to their energies are at second order, and are

$$\delta E_{\uparrow,\uparrow} = 0 \qquad \text{and} \qquad \delta E_{\uparrow,\downarrow} = -\frac{2t^2}{U} \qquad \text{from} \qquad \frac{|\langle\uparrow,\downarrow|\mathcal{H}_0|\uparrow\downarrow,0\rangle|^2}{E_{\uparrow,\downarrow} - E_{\uparrow\downarrow,0}} \quad \text{and similar.}$$

Compare these with the energies of eigenstates of the spin Hamiltonian $\mathcal{H} = J\mathbf{s}_1\cdot\mathbf{s}_2$, which are $E_{\rm singlet} = -3J/4$ and $E_{\rm triplet} = J/4$, and hence split by J. Noting that $|\uparrow,\uparrow\rangle$ is a triplet state, while $|\uparrow,\downarrow\rangle = 2^{-1/2}[|{\rm singlet}\rangle + |{\rm triplet}\rangle]$, we recognise that the Hubbard model with a half-filled band and large U reduces at low energy to the antiferromagnetic Heisenberg model. We read off the exchange strength, as

$$J = \frac{4t^2}{U}\,.$$

9.2 Mott transition

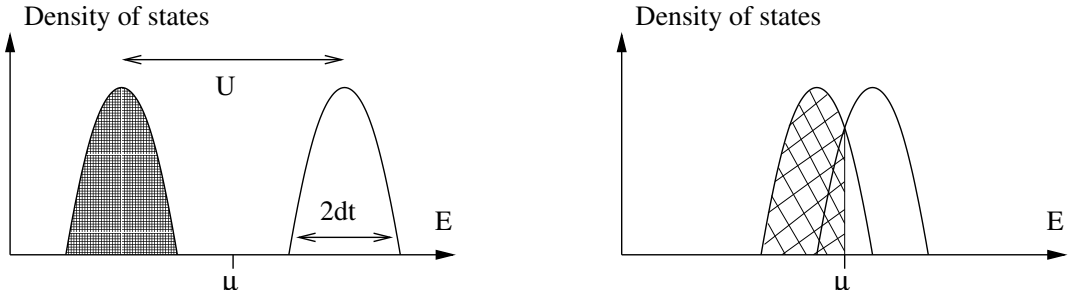


Figure 6: Spectral function expected for Hubbard model at half filling, showing upper and lower Hubbard bands: (left) in the insulating phase at large U/t; (right) in the metallic phase at small U/t.

It is an important and heavily-studied problem to understand behviour at half-filling as U/t is reduced. We will restrict ourselves to some cartoons. At large U added electrons each cost energy U as they necessarily hop

between doubly occupied sites with an associated kinetic energy scale t. Holes are also mobile with kinetic energy t, but their creation does not involve the energy penalty U. We therefore expect a spectral density at large U as shown in Fig. 6: at half-filling a fully occupied hole band and an empty electron band, both of width $\mathcal{O}(t)$, are split by U and sit symmetrically in energy either side of the chemical potential. Reducing the Hubbard repulsion, these bands overlap at a critical value $U_{\rm c} \sim 2dt$ and one expects a phase transition to metallic behaviour. As we increase U starting from the metallic side, we expect the Migdal discontinuity Z (see Fig. 5) to decrease, reaching zero at the transition. Indeed, deep in the Mott insulator, it is easy to see that there is no sign of the Fermi surface, since the ground state occupation is simply

$$\langle c^\dagger_{\mathbf{k}} c_{\mathbf{k}} \rangle = \frac{1}{N} \sum_{\mathbf{r}\mathbf{r}'} \langle c^\dagger_{\mathbf{r}} c_{\mathbf{r}'} \rangle \mathrm{e}^{\mathrm{i}(\mathbf{r}'-\mathbf{r})\cdot\mathbf{k}} = \frac{1}{2}$$

for all $\mathbf{k}$, independent of wavevector.

10 The Kondo effect

We now consider the consequences of large Hubbard repulsion acting just at one site of a lattice, to represent an impurity atom at that point. It is a surprise to find that this situation, apparently so close to a free-particle one, should generate a subtle many-body problem, involving what is known as the Kondo effect.

Experimentally, the resistance of many metals and alloys decreases with decreasing temperature as inelastic scattering processes are suppressed. Alloys containing dilute magnetic moments, such as a low concentration of Mn or Fe in Cu, are an exception: they show a resistivity minimum as a function of temperature and below it an increase in resistivity with decreasing temperature. An explanation of the minimum was given by Kondo in 1964, but it took a decade before the nature of the ground state was properly understood, and it was only in 1980 that a model for the phenomenon was solved exactly.

10.1 Model

The Kondo Hamiltonian describes a conduction band of independent electrons interacting via exchange with a single local moment on the impurity site, chosen to lie at the origin. Then

$$\mathcal{H} = \sum_{\mathbf{k}\sigma} \epsilon_{\mathbf{k}} c^\dagger_{\mathbf{k}\sigma} c_{\mathbf{k}\sigma} + \mathcal{H}_1 \qquad \text{with} \qquad \mathcal{H}_1 = J\, \mathbf{S}\cdot\mathbf{s}(0)\,. \tag{52}$$

Here $\mathbf{S}$ represents the local moment and $\mathbf{s}(0)$ is the spin density of the conduction electrons at the impurity site. We can write the exchange interaction as

$$\begin{aligned} J\,\mathbf{S}\cdot\mathbf{s}(0) &= J\left[S^z s^z(0) + \frac{1}{2}\Big(S^+ s^-(0) + S^- s^+(0)\Big)\right] \\ &= \frac{J}{V}\sum_{\mathbf{k}\mathbf{q}} \left[S^z\left(c^\dagger_{\mathbf{k}\uparrow} c_{\mathbf{q}\uparrow} - c^\dagger_{\mathbf{k}\downarrow} c_{\mathbf{q}\downarrow}\right) + \left(S^+ c^\dagger_{\mathbf{k}\downarrow} c_{\mathbf{q}\uparrow} + S^- c^\dagger_{\mathbf{k}\uparrow} c_{\mathbf{q}\downarrow}\right)\right] \end{aligned} \tag{53}$$

10.2 Scattering amplitude

The impurity spin mediates interactions between the conduction electrons, since the state of the impurity spin at a given time depends on previous scattering events. To see the consequences, we will calculate the scattering amplitude for an electron, taking as our example an initial state $\mathbf{k}_i\uparrow$ and a final state $\mathbf{k}_f\uparrow$, and working to second order in J, which will give the scattering rate to $\mathcal{O}(J^3)$.

At first order, the amplitude (left in the form of an operator on the impurity spin) is

$$\langle \mathbf{k}_f\uparrow | \mathcal{H}_1 | \mathbf{k}_i\uparrow\rangle = \frac{JS^z}{V}\,.$$

From this we can find the scattering rate at $\mathcal{O}(J^2)$, which is independent of temperature in the absence of a Zeeman field.

At second order we require

$$\sum_v \langle \mathbf{k}_f\uparrow | \mathcal{H}_1 | v\rangle \frac{1}{\epsilon - \epsilon_v} \langle v | \mathcal{H}_1 | \mathbf{k}_i\uparrow\rangle$$

where $\epsilon = \epsilon_{\mathbf{k}_i} = \epsilon_{\mathbf{k}_f}$ is the energy of the initial and final states, and $|v\rangle$ is an intermediate state with energy ϵ_v. The interesting contributions arise from $|v\rangle$ in which the impurity spin is flipped. These involve

$$S^+ c^\dagger_{\mathbf{q}\downarrow} c_{\mathbf{k}_i\uparrow} \frac{1}{\epsilon - [\epsilon + \epsilon_{\mathbf{k}_f} - \epsilon_{\mathbf{q}}]} S^- c^\dagger_{\mathbf{k}_f\uparrow} c_{\mathbf{q}\downarrow} \qquad \text{and} \qquad S^- c^\dagger_{\mathbf{k}_f\uparrow} c_{\mathbf{q}\downarrow} \frac{1}{\epsilon - [\epsilon - \epsilon_{\mathbf{k}_i} + \epsilon_{\mathbf{q}}]} S^+ c^\dagger_{\mathbf{q}\downarrow} c_{\mathbf{k}_i\uparrow} \,.$$

Now take a thermal average over scattering processes into the intermediate state, by making the replacements

$$c^\dagger_{\mathbf{q}\downarrow} c_{\mathbf{q}\downarrow} \to \langle c^\dagger_{\mathbf{q}\downarrow} c_{\mathbf{q}\downarrow} \rangle = f(\epsilon_{\mathbf{q}}) \qquad \text{and} \qquad c_{\mathbf{q}\downarrow} c^\dagger_{\mathbf{q}\downarrow} \to \langle c_{\mathbf{q}\downarrow} c^\dagger_{\mathbf{q}\downarrow} \rangle = 1 - f(\epsilon_{\mathbf{q}})$$

where $f(\epsilon) = [\mathrm{e}^{-\beta(\epsilon-\mu)} + 1]^{-1}$ is the Fermi function. Combining terms (and noting a negative sign arising from exchanging the order of $c_{\mathbf{k}_i\uparrow}$ and $c^\dagger_{\mathbf{k}_f\uparrow}$ in the first term), we then have contributions to the amplitude of

$$\begin{aligned} \frac{J^2}{V} \sum_{\mathbf{q}} \frac{1}{\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{q}}} \Big\{ S^+ S^- f(\epsilon_{\mathbf{q}}) \quad &+ \quad S^- S^+ [1 - f(\epsilon_{\mathbf{q}})] \Big\} \\ &= \quad S^- S^+ \frac{J^2}{V} \sum_{\mathbf{q}} \frac{1}{\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{q}}} + [S^+, S^-] \frac{J^2}{V} \sum_{\mathbf{q}} \frac{f(\epsilon_{\mathbf{q}})}{\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{q}}} \,. \end{aligned} \tag{54}$$

The factor $[\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{q}}]^{-1}$ in the sums on the right hand side is divergent at the energy $\epsilon_{\mathbf{k}}$ of the particle whose scattering amplitude we are calculating, which in the case of most interest is the Fermi energy. For the first sum, this divergence is not important, since contributions from above and below the Fermi energy cancel to leave a finite result. But in the second term the Fermi function limits this cancellation and we obtain

$$[S^+, S^-] \frac{1}{V} \sum_{\mathbf{q}} \frac{f(\epsilon_{\mathbf{q}})}{\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{q}}} = 2 S^z \rho \ln(D/k_{\mathrm{B}}T) \,,$$

where D is the energy width of the conduction band, which sets a lower limit on the sum, and ρ is the density of states in energy. To calculate the scattering rate we square the amplitude: combining our first and second order results, we find a result proportional to

$$J^2 + 4J^3 \rho \ln(D/k_{\mathrm{B}}T)$$

which grows with decreasing temperature. Combined with a phonon contribution to the electron scattering rate, this yields a minimum in the resistance as a function of temperature for alloys with magnetic impurities, which is the essence of the Kondo effect. The temperature scale at which the logarithmic term becomes important is the Kondo temperature $T_{\mathrm{K}} = D\mathrm{e}^{-1/2\rho J}$. For the physically relevant regime in which $J\rho$ is small, $T_{\mathrm{K}} \ll D$.

At temperatures $T \ll T_{\mathrm{K}}$ perturbation theory is not helpful. Instead we would like to understand the ground state for the coupled system consisting of the impurity spin and the Fermi gas. It turns out that this is just the same as one would guess by considering the strong coupling limit in which $J\rho$ is large: for antiferromagnetic J the impurity spin binds a conduction electron into a singlet state. This is characterised by a Pauli susceptibility (as distinct from the Curie susceptibility of a free impurity spin) and by a finite scattering rate for the remaining conduction electrons from the bound complex.

11 Disordered conductors and Anderson localisation

The theory of electron energy bands in solids gives an understanding of how the different properties of conductors and insulators arise. This picture, based on single-particle quantum mechanics for electrons in a periodic lattice, can fail when interactions are strong, as we have seen in our discussion of Mott insulators. It can also fail if disorder, in the form of deviations from a perfectly periodic lattice potential, is sufficiently large. This type of insulator is known as an Anderson insulator.

11.1 Anderson localisation

Consider electrons in impure or amorphous metals or doped semiconductors as examples of waves in a disordered medium. At *weak* disorder the main consequence of impurities is a finite elastic mean free path ℓ_{el} for electrons. This terminology emphasises a scattering description, but we can also discuss eigenstates in a large, finite system. At weak disorder an individual eigenstate is a superposition of waves with similar energies. Locally this superposition generates a plane wave, but the amplitude and phase of the wave vary on the scale of ℓ_{el}, which in this regime is large compared to the other relevant lengthscale, the Fermi wavelength λ_{F}.

The question arises: what happens if disorder is increased until what is known as the Ioffe-Regel limit, $\ell_{\rm el} \sim \lambda_{\rm F}$? At strong disorder (and also at weak disorder, in low dimensional systems), the nature of eigenfunctions in fact changes. In place of extended states, with the probability density of each individual state spread through the entire sample, eigenstates are spatially localised, with each electron trapped in its own localisation volume, of linear size ξ.

It is useful to compare this situation with other, more familiar eigenstate problems. An individual localised state is similar to the impurity state bound to a donor or acceptor in a semiconductor, in the sense that both states are spatially localised. There is, however an important difference in that the donor or acceptor levels are isolated in energy from other states, since they lie in the band gap of the semiconductor. By contrast in an Anderson insulator, we have a band of localised states, and some are very close in energy though far apart in space.

11.2 Stability of localised states

An immediate question is whether such a situation can be stable. In general, states that are close in energy are easily mixed by a perturbation, and if we mix two localised states with widely separated centres, the resulting superpositions are much less localised than the initial states. Suppose that the mixing is due to a small change δV in the potential felt by electrons, arising for example because of a change in the disorder realisation. The consequent change $\delta\psi_i$ in an eigenfunction with energy E_i is at first order in perturbation theory

$$\delta\psi_i = \sum_k \frac{\langle\psi_k|\delta V|\psi_i\rangle}{E_i - E_k}\psi_k$$

and the fact that states nearby in energy are easily mixed is represented here by the possibility of small denominators $E_i - E_k$. More specifically, with density of states ρ per unit energy and volume, the smallest value of $E_i - E_k$ arising from states k that have their localisation centres within a distance r of the state i is $(\rho r^d)^{-1}$. We are saved from divergent contributions to the sum at large r because for localised states the matrix element in the numerator fall off even faster: we expect

$$\langle\psi_k|\delta V|\psi_i\rangle \sim {\rm e}^{-r/\xi}\,.$$

In this way we see that the idea of a dense set of localised states is at least self-consistent.

11.3 Mobility edge

The strength of disorder is characterised by the amplitude of fluctuations in the potential felt by electrons, which should be compared with their kinetic energy to obtain a dimensionless measure. For that reason, in a given system electrons at low energy are in strong relative disorder, and those are high energy are at weak relative disorder. In consequence, if there is a transition in the nature of eigenstates as a function of disorder strength, there will be a division in the properties of states as a function of energy, with localised states at low energy and extended states at high energy. The dividing energy is called the mobility edge. Localised states can be characterised by their localisation length ξ and extended states by the value of the conductivity σ, and we expect the behaviour shown in Fig. 7.

11.4 Scaling theory

The change in behaviour from extended to localised states as a function of energy or disorder is (in a non-thermodynamic sense) a phase transition. The field-theoretic description (in terms of a non-linear sigma model) is quite involved, but there is also a simple phenomenological approach based on scaling ideas, known as one-parameter scaling theory. It has some useful predictive power and also offers a good setting for thinking about renormalisation group ideas.

In general, scaling approaches involve asking how physical properties of a system change with lengthscale. Here we focus on the conductance $G(L)$ of a (hyper)-cubic sample in d-dimensions, of linear size L. We measure this in units of the conductance quantum, e^2/h, writing

$$g(L) = \frac{G(L)}{e^2/h}\,.$$

Now suppose we combine 2^d such hypercubes to make a larger hypercube, of size $2L$. The hypothesis of scaling theory is that the conductance of this final cube depends only on the conductances of the initial cubes, and not on

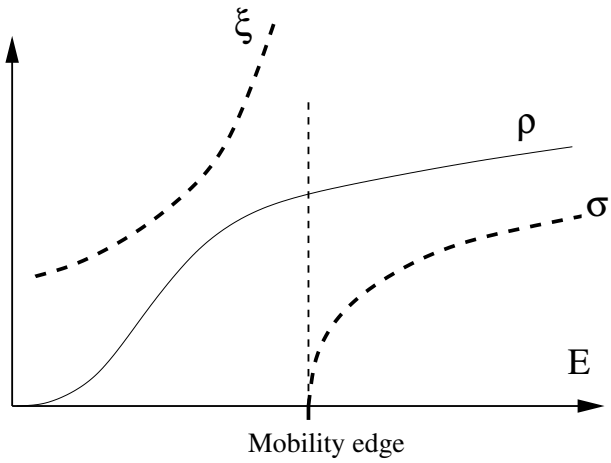


Figure 7: Behaviour in a disordered system of the density of states ρ, localisation length ξ and conductivty σ as a function of energy E. States are localised below a critical energy – the mobility edge – and extended above it.

other characteristics. That is to say, for example, we would get the same result for the final cube by combining very small, highly disordered initial cubes or larger, weakly disordered initial cubes, provided the initial value of G were the same in both cases. We express this idea by writing

$$\frac{\partial \ln g}{\partial \ln L} = \beta(g)\,.$$

The beta function $\beta(g)$ encapsulates the nature of the transition. Of course, we know only a little about it. We can, however, be confident that it is a continuous, smooth function, since it represents behaviour in a finite sample. We can also pin down its asymptotics quite easily, as follows.

Consider first the weak disorder limit. Here we expect metallic behaviour characterised by a fixed value for the conductivity σ of the material, so that $G(L) = \sigma L^{d-2}$. Hence $\beta(g) = d - 2$ at large g. Next consider strong disorder and assume states are localised. In this regime we anticipate $G \propto \exp(-L/\xi)$ which leads to $\beta(g) = \ln(g) + \text{constant}$. Connecting the two limits with a smooth monotonic curve yields the result shown in Fig. 8.

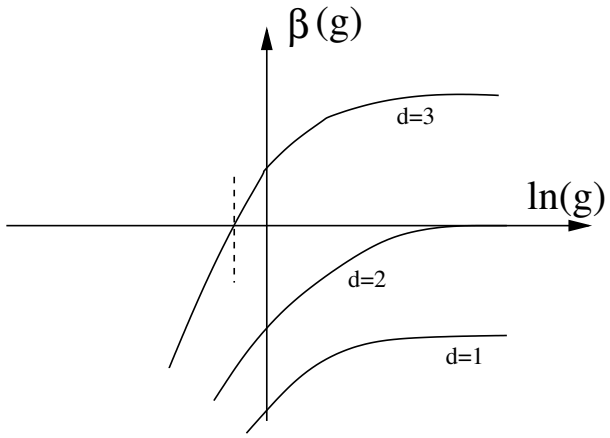


Figure 8: Dependence of the beta-function $\beta(g)$ for Anderson localisation on dimensionless conductance g, shown as a function of $\ln(g)$ in dimensions $d = 1$, 2 and 3. For $d = 3$ the position $\ln(g_c)$ of the critical point at which the beta function has a zero is marked by the vertical dashed line.

We should now examine the consequences the follow from the assumptions we have made. It is useful to separate various cases.

$d = 1$ and $d = 2$. In this case we always have $\beta(g) < 0$. As a result, $g(L)$ decreases with increasing L, eventually exponentially. That is to say, states are localised in $d = 1$ and $d = 2$ by arbitrarily weak disorder.

$d > 2$ and $g < g_c$. In this regime we again have $\beta(g) < 0$. With increasing L we find that $\beta(g)$ becomes increasingly negative, and states are localised.

$d > 2$ and $g > g_c$. In this case g increases with L and at large L reaches values for which $\beta(g) = d - 2$, so that we have an Ohmic metal with $g(L) \propto L^{d-2}$

$d > 2$ and $g = g_c$. In the case of a system with critical disorder the conductance is independent of system size – an example of scale-independence at a critical point.

12 The integer and fractional quantum Hall effects

We now switch to a discussion of the consequences of disorder and of electron-electron interactions for a two-dimensional electron gas in a strong magnetic field. This is a situation in which departures from the behaviour of an ideal, non-interacting Fermi gas without impurity scattering are essentially guaranteed, because the single particle eigenstates of the ideal problem are macroscopically degenerate. Our interest is in the ways this degeneracy may be lifted.

The experimental systems are electrons (or holes) in quasi two-dimensional semiconductor structures - either metal-oxide-semiconductor field effect transistors fabricated on silicon, or heterostructures or quantum wells fabricated from GaAs and $Al_xGa_{1-x}As$. Carrier motion in the third dimension is frozen out quantum mechanically when temperature and electron density are both small enough that only the lowest sub-band is occupied.

The quantum mechanics of electron motion in a magnetic field with flux density B is characterised by two scales. One – the cyclotron frequency $\omega_c = eB/m^*$ – is classical and material-dependent through the value of the effective mass. The other – the magnetic length $l_B = (\hbar/eB)^{1/2}$ – is quantum-mechanical and material-independent. The energy spectrum for the single-particle problem without disorder consists of a sequence of Landau levels, with energies

$$E = \left(n + \frac{1}{2}\right)\hbar\omega_c \pm \frac{1}{2}g^*\mu_B B\,.$$

For electrons in free space the orbital splitting $\hbar\omega_c$ is almost degenerate with the spin splitting $g\mu_B B$, but for electrons in a semiconductor $\hbar\omega_c$ is typically much larger than $g^*\mu_B B$, since effective masses are typically smaller than the bare mass, and the effective g-factor g^* is also reduced because of spin-orbit interactions. The number of states within each spin-split Landau level is given simply by the number N_ϕ of magnetic flux quanta passing through the area of the system. Comparing this degeneracy with the number N_e of electrons we obtain the most important parameter characterising the system: the Landau level filling factor $\nu = N_e/N_\phi$. For an electron gas with number density n we have the relation

$$n = \nu\frac{eB}{h}\,.$$

As a prelude to a discussion of experimental observations, it is useful to recall the Hall effect in an ideal system, as shown in Fig. 9. Taking the electrons to have a drift velocity $v_{\rm Drift}$, the current is $I = ev_{\rm Drift}nw = \frac{e^2}{h}\nu B v_{\rm Drift} w$ while the Hall voltage is $V_H = Bv_{\rm Drift}w$. Combining these two expressions, the Hall conductivity is $\sigma_{xy} \equiv I/V_H = e^2\nu/h$.

Experimentally, studying the resistivity tensor ρ as a function of magnetic field strength or electron density, around certain filling factors accurately quantised plateaus are observed in ρ_{xy} accompanied by vanishing ρ_{xx}. Under these conditions the forms of the resistivity tensor ρ and the conductivity tensor σ are

$$\rho = \frac{h}{\nu e^2}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \qquad \text{and} \qquad \sigma \equiv \rho^{-1} = \frac{\nu e^2}{h}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

It is striking at first sight that both ρ_{xx} and σ_{xx} should vanish: the important point is simply that the current density is perpendicular to the field, and so dissipation vanishes.

This observed behaviour is quite different from that expected for a clean, single-particle system. Instead of a Hall conductance proportional to filling factor and vanishing dissipative conductance, plateaus are seen in Hall conductance as a function of filling factor and the dissipative conductance, though small in quantum Hall plateaus, has peaks at the transtions between plateaus.

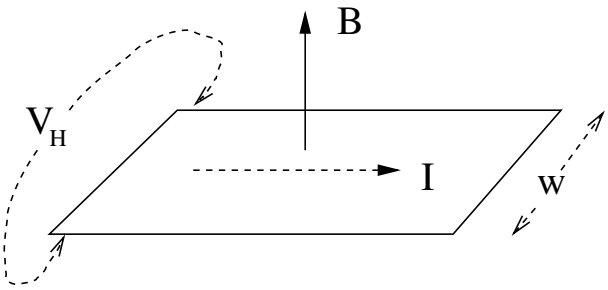


Figure 9: Schematic view of the Hall effect in a two-dimesional system: current flow I in the presence of a magnetic field B generates a Hall voltage $V_{\rm H}$ in a sample of width w.

12.1 Integer quantum Hall effect

The existence of plateaus around integer values of ν can be understood if we suppose that disorder affects the states in a Landau level as illustrated in Fig. 10, with Anderson localised states in both the low and high energy tails of the disorder-broadened Landau level, and a divergence in the localisation length near the Landau level centre. Then, if the Fermi energy lies between the centres of two Landau levels, changes in filling factor do not alter the number of extended, current-carrying states that are occupied, and so the Hall conductance is unchanged. In this situation the absence of dissipation can also be understood, since dissipative processes require excitation of an electron initially in a current-carrying state to an empty final state. As the occupied, current carrying states are buried a finite energy below the Fermi energy, such processes are suppressed at low temperature by an activation factor.

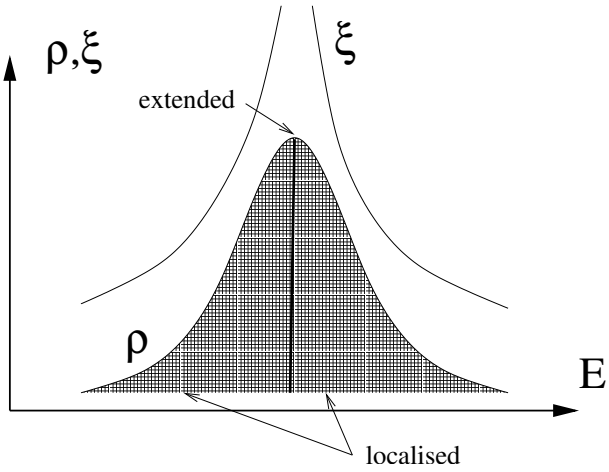


Figure 10: Disorder-broadened Landau level, showing localisation physics necessary to explain the integer quantum Hall effect. The dependence of the density of states ρ and the localisation length ξ is given as a function of energy E. Almost all states in the Landau level are localised, but the localisation length diverges at a critical point near the centre of the Landau level.

12.1.1 Exactness of quantisation

While we can understand the existence of Hall plateaus if we accept this picture for the influence of disorder on electron eigenstates, the precise quantisation of Hall conductance immediately becomes a surprise, since one might have expected the reduction in the number of extended states to be accompanied by a reduced value for the Hall conductance. Clearly, the remaining extended states must carry an extra current in a way that exactly compensates for their reduced number.

One way to understand the exactness of quantisation is from the thought-experiment illustrated in Fig 11. We consider a quantum Hall sample in the form of an annulus. In addition to the magnetic field responsible for the quantum Hall effect, which pierces the surface of the annulus, we introduce a second magnetic flux Φ, threading

through the hole at the centre of the annulus. Allowing this flux to vary as a function of time, we generate a voltage V around the circumference of the quantum Hall sample. From Faraday's law, we have

$$V = -\frac{\mathrm{d}\Phi}{\mathrm{d}t} \, .$$

Within a Hall plateau, this produces a current flow

$$I = \sigma_{xy} V$$

in the perpendicular direction, which is radial. Integrating the rates of flux change and current flow over time, a given flux difference $\Delta\Phi$ corresponds to the transport of a certain charge Q between the inner and outer edges of the annulus. Now, we expect that a change in Φ of one flux quantum (h/e) will return the interior of the quantum Hall system to its initial state, implying that an integer number of electrons have then been transported across the annulus. We have

$$\text{integer} \times e = Q = \sigma_{xy} \cdot \Delta\Phi = \sigma_{xy} \frac{h}{e}$$

and hence

$$\sigma_{xy} = \text{integer} \times \frac{e^2}{h} \, .$$

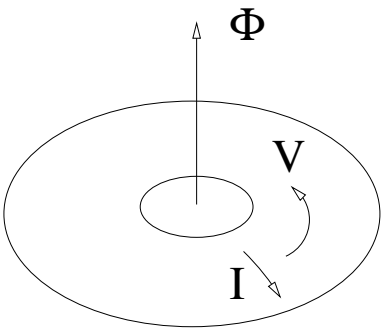


Figure 11: Geometry considered in Laughlin argument for exactness of quantisation of Hall conductance

12.1.2 Localisation

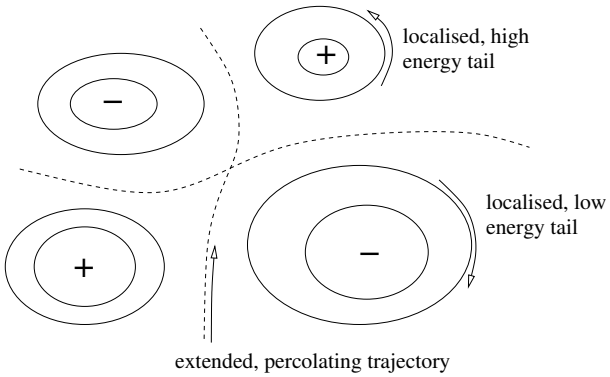


Figure 12: Trajectories for cyclotron motion guiding centres of charged particles in a smooth random potential and perpendicular magnetic field, showing how an extended trajectory arises near the Landau level centre, and localised trajectories in the low and high energy tails. Lines represent equipotentials and $+$ and $-$ signs denote maxima and minima of the potential.

Anderson localisation in the context of the quantum Hall effect can be understood to some extent by appealing to a semiclassical limit, in which the magnetic length is much smaller than the correlation length for the random potential that represents disorder. In this case there is a separation of scales: considered classically, the electron motion can be separated into a fast cyclotron orbit around an instantaneous guiding centre, and slow drift of this guiding centre. The drift is in a direction perpendicular to the local electric field and therefore along equipotentials. Quantisation of the rapid cyclotron motion generates Landau levels, and the energies of electrons within each Landau level are displaced by the energy of the equipotential along which they move. In this picture, questions about localisation reduce to ones about the nature of equipotential lines. As illustrated in Fig. 12, we expect these to exhibit a percolation transition, separating low energy closed trajectories which orbit minima in the potential from high energy trajectories which orbit maxima in the potential.

An obvious question now is how this picture of a delocalisation transition in quantum Hall systems can be related to the scaling theory described in Section 11.4. In particular, whereas we had concluded earlier that in a two-dimensional system states should be localised, we now see that this is not always the case. A scaling flow diagram that reconciles the earlier discussion with behaviour in quantum Hall systems is shown in Fig. 13. Crucially, in a magnetic field we consider the scaling flow of the two independent components of the conductivity tensor. The scaling flow diagram shows how these evolve as a system is probed on increasing lengthscales – experimentally, by lowering temperature to increase the inelastic scattering length that cuts off phase coherent localisation effects. The scaling flow diagram is periodic in the Hall conductance, measured here in units of the quantum e^2/h and contains fixed points of two types. Stable fixed points at $(\sigma_{xx}, \sigma_{xy}) = (0, n)$ with n integer describe the insulating state in zero magnetic field and also the system in quantum Hall plateaus. Unstable fixed points located at $(\sigma_{xx}, \sigma_{xy}) = (\sigma^*, n+1/2)$ describe the delocalisation transitions that separate different quantum Hall plateaus.

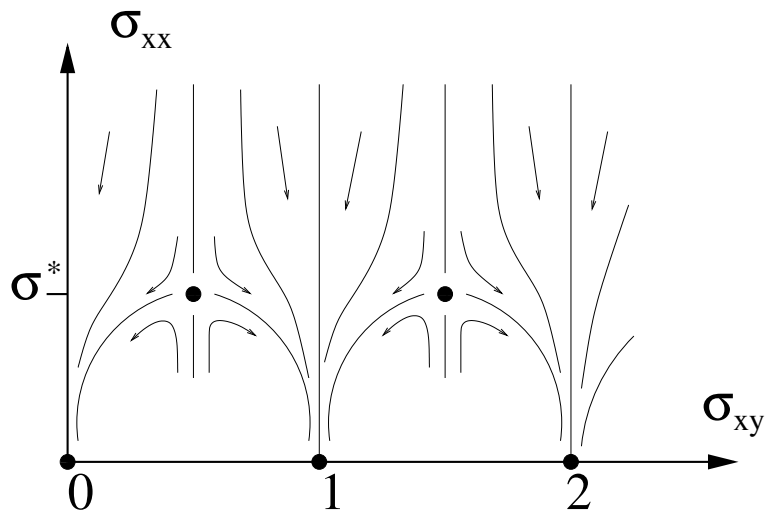


Figure 13: Scaling flow diagram for quantum Hall plateau transitions.

12.2 Fractional quantum Hall effect

In sufficiently high mobility samples, quantum Hall plateaus are observed with certain simple *fractional* values of the Hall conductance, in addition to the integer plateaus we have discussed. Specifically, one finds $\sigma_{xy} = (p/q) \cdot (e^2/h)$, with p and q integer, and q (in nearly all cases) odd. The fact that this is observed in samples with low disorder suggests it is interactions that lift the degeneracy of the partially filled Landau level, forming special correlated ground states at certain filling factors. And the fact that $\sigma_{xx} \to 0$ at low temperature (with an activated temperature dependence) suggests there is a gap for excitations from these correlated ground states. Moreover, Laughlin's argument for quantisation of Hall conductance suggests that excitations have fractional charge, since we argued that $\sigma_{xy} \times (h/e)$ should be equal to the charge transported across an annular system when one flux quantum is inserted. Assuming this charge consists of an integer number of quasiparticles, each of charge e^*, we have $(p/q)e = \text{integer} \times e^*$.

12.2.1 Single particle in a magnetic field

As a first step, it is useful to summarise some results for the quantum mechanics of a particle in a magnetic field. We take charge $-e$ and magnetic field $(0, 0, -B)$, with e and B positive. The Hamiltonian is

$$\mathcal{H} = \frac{1}{2m^*}(\pi_x^2 + \pi_y^2) \qquad \text{with} \qquad \vec{\pi} = -\mathrm{i}\hbar\vec{\nabla} + e\vec{A}$$

and

$$[\pi_x, \pi_y] = -\mathrm{i}\hbar e(\partial_x A_y - \partial_y A_x) = \frac{\mathrm{i}\hbar^2}{l_\mathrm{B}^2}\,.$$

We define raising and lowering operators

$$a^\dagger = \frac{l_\mathrm{B}}{\sqrt{2}\hbar}(\pi_x - \mathrm{i}\pi_y) \qquad \text{and} \qquad a = \frac{l_\mathrm{B}}{\sqrt{2}\hbar}(\pi_x + \mathrm{i}\pi_y) \qquad \text{with} \qquad [a, a^\dagger] = 1$$

so that

$$\mathcal{H} = \hbar\omega_\mathrm{c}(a^\dagger a + 1/2)\,.$$

Now, it is convenient to combine coordinates into a complex number, by writing $z = x + \mathrm{i}y$ and $\overline{z} = x - \mathrm{i}y$. We also use the notation

$$\partial_z = \frac{1}{2}(\partial_x - \mathrm{i}\partial_y) \qquad \text{and} \qquad \partial_{\overline{z}} = \frac{1}{2}(\partial_x + \mathrm{i}\partial_y)\,,$$

which is set up so that, for example,

$$\partial_z z = \partial_{\overline{z}}\overline{z} = 1 \qquad \text{and} \qquad \partial_z\overline{z} = \partial_{\overline{z}} z = 0\,.$$

Choosing units in which $l_\mathrm{B} = 1$ and taking the gauge $\vec{A} = (B/2)(y, -x, 0)$ we then have

$$a = -\frac{\mathrm{i}}{\sqrt{2}}(2\partial_{\overline{z}} + \frac{1}{2}z)\,.$$

Ground state wavefunctions ψ_m satisfy $a\,\psi_m = 0$, and a complete set is

$$\psi_m \propto z^m \mathrm{e}^{-|z|^2/4}$$

with $m = 0, 1, 2\ldots$. The proabability density of the m^th state is a ring with (after restoring units) radius $(2m)^{1/2}l_\mathrm{B}$ and width l_B.

12.2.2 Two particle problem

Unsymmetrised two-particle basis states from the lowest Landau level have the form

$$\psi(z_1, z_2) \propto z_1^l z_2^m \mathrm{e}^{-(|z_1|^2+|z_2|^2)/4}$$

with l, m non-negative integers. We will consider combinations of these that are eigenfunctions of relative and centre-of-mass angular momentum. They have the form

$$\psi(z_1, z_2) \propto (z_1 - z_2)^l (z_1 + z_2)^m \mathrm{e}^{-(|z_1|^2+|z_2|^2)/4}\,.$$

The value of l completely determines the typical particle separation in the state and fixes the energy given a form for the interaction potential. For this reason, and in contrast to a two-particle problem without restriction to a single Landau level, the pair is incompressible. By this we mean that the two-particle system cannot respond smoothly to an external potential that for example squeezes the particles together. Instead, with increasing external potential we expect a sequence of level crossings at which the ground state value of l changes: at each crossing there is a jump in the separation of the particles, but between jumps their separation is independent of the external potential.

Note that l is required by the statistics of the particles to be odd for fermions, or even for bosons.

12.2.3 Laughlin wavefunction

With this discussion of the two-particle problem as orientation, we can consider the many-particle wavefunction proposed by Laughlin

$$\psi(z_1 \dots z_N) = \prod_{i<j} (z_i - z_j)^q \mathrm{e}^{-\frac{1}{4}\sum_k |z_k|^2} .$$

As motivation for this form, one can begin with the intention of writing a variational state containing so-called Jastrow factors of the form $f(z_i - z_j)$. Trial wavefunctions of this form have been used successfully, for example, in the theory of superfluid ^{4}He, where an appropriate choice for the function making up the Jastrow factor enables one to build two-particle correlations into the wavefunction that minimise the energy of the system. Restriction to the lowest Landau dramatically constrains this choice of function, to the discrete set $(z_i - z_j)^q$, with q a positive integer, which must be odd for fermions or even for bosons. The value $q = 1$ describes a full Landau level. Higher values ensure that the probability density falls faster to zero as a pair of particles approach each other (as $|z_i - z_j|^{2q}$) than in a generic antisymmetric state, and for this reason are good variational choices when the interparticle interaction is repulsive. In fact, this wavefunction is an exact ground state for a certain interparticle potential (a generalised delta function), and is known from numerical studies on systems containing a small number of particles to have very high overlap with exact ground states for systems with Coulomb interactions.

In many-particle quantum mechanics, knowledge of the wavefunction is not all, since extraction of physical information from this function of many coordinates constitutes a problem as hard as the one in classical statistical mechanics of calculating physical averages from a Boltzmann factor that is a function of the coordinates of all particles in a system.

The simplest question we might ask is what filling factor is represented by the Laughlin wavefunction. Expanding, we see that the largest power of each coordinate is $z_i^{q(N-1)}$. Recalling that the single-particle wavefuction $z\,\mathrm{e}^{-|z|^2/4}$ has its probability density concentrated on a ring of radius $\propto m^{1/2}$, we see that in the Laughlin state N particles fill the area that would be occupied by $qN + 1 - q$ particles if the Landau level were completely filled. Hence we see for large N that

$$\nu = \frac{1}{q} .$$

As a next and much harder step we would like to understand correlations in the Laughlin state. A good approach to this problem is to use what is called the plasma analogy: we think of the probability density arising from this wavefunction as if it were the Boltzmann weight for a problem in classical statistical mechanics. We can read off the Hamiltonian for this classical problem, which turns out to be a Coulomb gas, or plasma. If we have good physical intuition for the statistical mechanics of plasmas, we can apply this to draw conclusions about the Laughlin state.

Some details are as follows. We define a fictitious inverse temperature β and classical Hamiltonian $\mathcal{H}_{\mathrm{cl}}$ via

$$|\psi(z_1 \dots z_n)|^2 = \mathrm{e}^{-\beta \mathcal{H}_{\mathrm{cl}}} .$$

This yields

$$\mathcal{H}_{\mathrm{cl}} = \frac{1}{2\beta} \sum_k |z_k|^2 - \frac{2q}{\beta} \sum_{i<j} \ln |z_i - z_j| .$$

To interpret this form we should recall electrostatics in two dimensions: a point charge Q at the origin gives rise at radius r to an electric field

$$E(r) = \frac{Q}{2\pi\epsilon_0 r} \qquad \text{and a potential} \qquad V(r) = -\frac{Q}{2\pi\epsilon_0} \ln r .$$

It is convenient to set the (arbitrary) inverse temperature to $\beta = 4\pi\epsilon_0/q$. Then the two-particle term in $\mathcal{H}_{\mathrm{cl}}$ becomes

$$-\frac{q^2}{2\pi\epsilon_0} \sum_{i<j} \ln |z_i - z_j| ,$$

which represents the electrostatic interaction of particles with charge q. The single particle term is

$$\frac{q}{8\pi\epsilon_0} \sum_k |z_k|^2 .$$

It would arise for particles of charge q moving in an electrostatic potential $|z|^2/(8\pi\epsilon_0)$. We can view this potential as arising from a background charge distribution, and find the density of this charges using Poisson's equation. We obtain a density

$$-\epsilon_0\nabla^2\frac{1}{8\pi\epsilon_0}|z|^2 = -\frac{1}{2\pi}\,.$$

As a first conclusion from this picture, we can recover the value of the Landau level filling factor ν. The plasma consisting of N particles each with charge q will arrange itself so as to cancel on average the background charge, by adopting a mean number density of $1/(2\pi q)$. Since this is the number density of electrons in the Laughlin state (remember that we set $l_{\rm B}$ to unity), and since its value is $1/q$ of the value for the filled Landau level, we obtain again $\nu = q^{-1}$. Going futher, we can use the plasma analogy to discuss electron-electron correlations in the Laughlin wavefunction. In fact, the plasma is known to be fluid for $q \lesssim 70$ and crystalline for $q \gtrsim 70$, and so we see that the Laughlin wavefunction represents a liquid state for the important values $q = 3, 5 \ldots$.

One can also use Laughlin's approach to write down a wavefunction including a hole, and use the plasma analogy to examine its properties. We have

$$\psi_{\rm hole}(z_1 \ldots z_N) = \prod_l (z_l - \xi) \times \psi_{\rm Laughlin}(z_1 \ldots z_N)\,.$$

The extra factor has the effect of excluding electrons from the vicinity of the point ξ where the hole is located. Within the plasma analogy we find

$$\mathcal{H}_{\rm cl} \to \mathcal{H}_{\rm cl} - \frac{q}{2\pi\epsilon_0}\sum_l \ln|z_l - \xi|\,.$$

Thus the hole translates to a particle of unit charge in the plasma, located at ξ and interacting with other particles, of charge q at z_l. Since plasmas screen, this will induce a compensatiing reduction in the plasma density around it, but since the plasma particles have charge q, a deficit of only $1/q$ plasma particles is sufficient for exact compensation. Back in the language of electrons, a deficit of $1/q$ of an electron means that the hole has charge $1/q$. This is a spectacular instance of fractionalisation, and fractional charge has been observed reasonably directly in shot noise measurements on current carried by fractional quantum Hall quasiparticles.