

SUPERCONDUCTIVITY
763645S

Erkki Thuneberg

Department of Physics
University of Oulu
2015

Practicalities

Write your name, department/group, year class, and email address on the list.

The lectures are given in English if there is sufficient demand for it, otherwise in Finnish.

The web page of the course is

<https://www oulu.fi/tf/sj/index.html>

The web page contains the lecture material (these notes), the exercises and later also the solutions to the exercises. See the web page also for possible changes in lecture and exercise times.

Time table 2015

Lectures: Wednesday 14-16, FY1103 (9.9.-2.12.)

Exercises: Thursdays 8-10, FY1103 (17.9-3.12.)

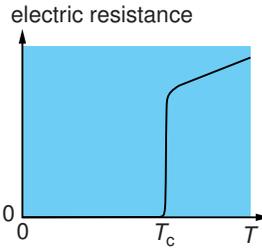
Examination: to be agreed on lecture on 18.11.

Teaching assistant: Sami Laine

Doing exercises is essential for learning. In addition, showing completed exercises will affect your final evaluation. (You can improve by one, for example, from 3 to 4.) Start calculating (at home) before the exercise time, 2 hours is too short time to start from scratch.

1. Introduction

By superconductivity we mean a phenomenon where the electrical resistivity of a material (e.g. metal) disappears below some temperature.



We are interested in superconductivity because of the following reasons:

- The theory of superconductivity is very interesting.
- Superconductivity has several technical applications.
- Superconductivity is under active study around the world (e.g. in connection with quantum computing).
- We do theoretical research related to superconductivity in Oulu, and can suggest some topics for master thesis.

Short content of the course

- short review/introduction to some fundamental results of statistical and condensed matter physics
- Thermodynamics in magnetic field
- BCS theory
- Ginzburg-Landau theory
 - Type II superconductivity
- Josephson effect

Books

- M. Tinkham, Introduction to Superconductivity (1975, 1996). Very widely used book. More experimental view, and therefore not ideal for this theory course.
- A.L. Fetter and J.D. Walecka, Quantum theory of many-particle systems (1971). Superconductivity studied in chapters 10 and 13. The problem is that most of the microscopic theory is treated using Green's functions, which are avoided in this course.
- J.B. Ketterson and S.N. Song, Superconductivity (1999). Just another book.

- A.A. Abrikosov, Fundamentals of the Theory of Metals (1988). Very extensive book. Half of the book discusses normal state metals. Derivation of BCS theory not good.
- P.G. de Gennes, Superconductivity of Metals and Alloys (1961). Old but still useful.
- K. Fossheim and A. Sudbø, Superconductivity: Physics and Applications (2004).
- Many books on solid state physics include an introduction to superconductivity, for example N. Ashcroft and D. Mermin, Solid state physics (AM).

Acknowledgement

I thank Janne Viljas for suggesting many changes to these notes.

1.1 Properties of superconductors

Occurrence of superconductivity

- several metallic elements : Al, Nb, Sn, (but not in magnetic metals and in noble metals: Cu, Au, Ag)
- many alloys, e.g. Nb-Ti
- some compounds: Nb₃Ge, MgB₂, Y-Ba-Cu-O etc.

The temperature below which superconductivity occurs is called *critical temperature*, T_c . The list gives some critical temperatures.

material	T_c (K)	$\mu_0 H_c(T=0)$ (mT)
Al	1.196	9.9
Hg	4.15	41
In	3.40	29.3
Pb	7.19	80.3
Nb	9.25	
Nb ₃ Ge	23	
MgB ₂	39	
YBa ₂ Cu ₃ O _{6+x}	98	
Tl ₂ Ca ₂ Ba ₂ Cu ₃ O ₁₀	125	

Infinite conductivity

In normal state metals the electric current \mathbf{j} is proportional to the electric field \mathbf{E} :

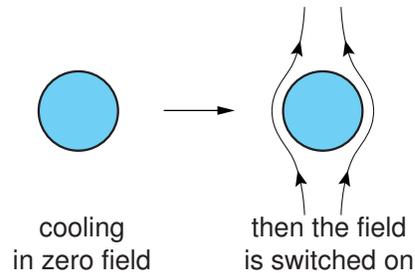
$$\mathbf{j} = \sigma \mathbf{E}. \quad (1)$$

If $\sigma \rightarrow \infty$ then $\mathbf{E} \rightarrow 0$. Maxwell's equation

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad (2)$$

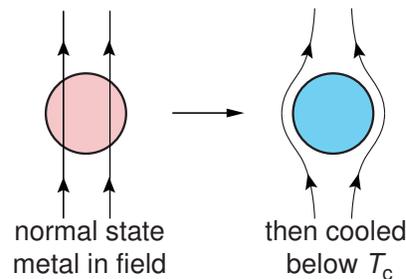
then gives that the magnetic field \mathbf{B} is constant. Let us apply this to the case that a superconductor is cooled

below T_c while $\mathbf{B} = 0$. When the field is switched on, it will not penetrate into the superconductor.



Meissner effect

A more fundamental phenomenon than infinite conductivity is seen when a normal state metal is first placed in magnetic field, and is then cooled into the superconducting state. It is observed that the magnetic field is expelled from the sample. This is called *Meissner effect*.

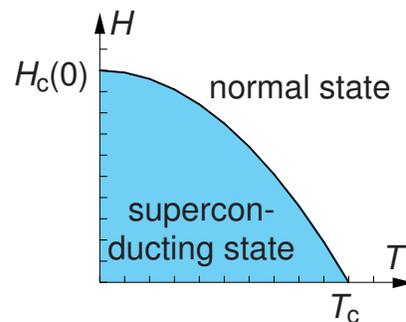


Thus the fundamental property is that the magnetic field is zero inside a superconductor. (Not only constant, as would follow from infinite conductivity.)

Critical field

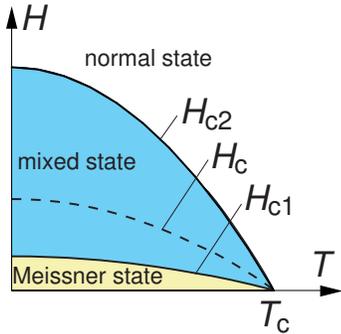
The Meissner effect is observed only if the field is not too large. Let us for simplicity consider a thin bar sample that is parallel to the field. (In this case the magnetizing field \mathbf{H} is constant.) It is observed that a transition between superconducting state and normal state takes place in *critical field* H_c , whose dependence on temperature is approximately

$$H_c(T) = H_c(0) \left[1 - \left(\frac{T}{T_c} \right)^2 \right]. \quad (3)$$



Material showing this behavior are called *type I superconductors*.

Some superconducting materials have a mixed state between Meissner and normal states. These are called *type II superconductors*.

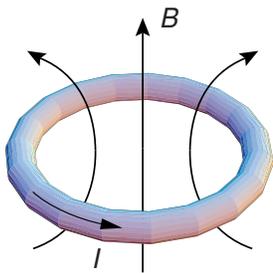


Persistent currents and flux quantization

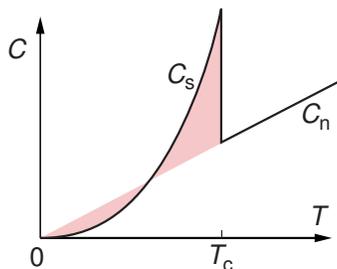
Let us place a normal state ring in perpendicular magnetic field. When it is cooled below T_c , the magnetic field is expelled from the inside of the superconductor, but a magnetic flux through the ring remains. When the external field is removed, this flux remains unchanged. Thus a persistent current I is induced in the superconducting ring that generates the magnetic field B . In addition, the magnetic flux $\Phi = \int d\mathbf{a} \cdot \mathbf{B}$ through the ring is quantized: it is an integer multiple of the flux quantum

$$\Phi_0 = \frac{h}{2|e|} = 2.07 \times 10^{-15} \text{ Wb.} \quad (4)$$

Here h is Planck's constant and e the charge of an electron. [Because $e < 0$, the absolute value is taken in (4).]



Specific heat



The transition between normal and superconducting states is also seen in thermodynamic properties. The specific heat has discontinuity but no latent heat (in zero field). This kind of change of state is called second order phase transition. In the normal state the specific heat is linear at low temperatures. In superconducting state the specific heat is exponential when $T \rightarrow 0$:

$$C_s \propto \exp\left(-\frac{\Delta_0}{k_B T}\right). \quad (5)$$

This can be understood so that there is an energy gap Δ_0 between the ground state and the lowest excited states. This gap is somewhat less than $2k_B T_c$ in most superconductors.

Isotope effect

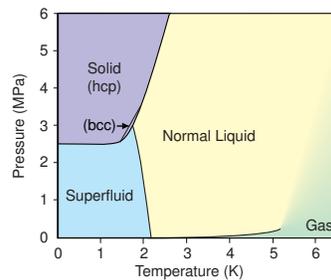
Different isotopes of the same element have difference in T_c , which depends on the ion mass M approximately as

$$T_c \propto M^{-1/2}. \quad (6)$$

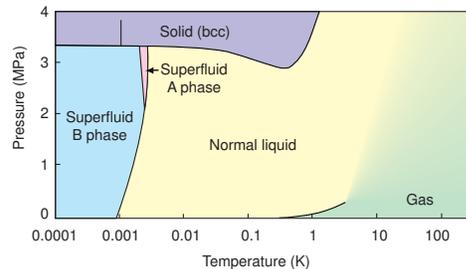
This can be used to deduce that the motion of the ions is important for superconductivity.

1.2 Superfluidity

An analog phenomenon to superconductivity is *superfluidity*. It means the a fluid can flow persistently (without viscosity), for example, in a ring shaped tube. There are two well know superfluids: ^4He



and ^3He



The superfluid state is achieved in some gases of alkali metals ^{87}Rb , ^7Li , ^{23}Na , $^1\text{H}, \dots$

1.3 History

- 1911 H. Kamerlingh Onnes finds superconductivity in mercury

- 1933 Meissner effect
- 1935 London theory
- 1950 Ginzburg-Landau theory
- 1957 Bardeen-Cooper-Schrieffer theory
- 1957 theory of type II superconductivity
- 1962 Josephson effect
- 1986 Bednorz and Müller find “high temperature superconductors”

Superfluidity

- 1938 discovery of superfluid state of ^4He
- 1972 discovery of superfluid states of ^3He
- 1995 Bose-Einstein condensation in alkali atom gases

2. Fundamental results of statistical physics

In order to understand superconductivity one has to know quantum mechanics and statistical physics. This is a short collection of some central results of statistical physics.

Let us consider a system consisting of a large number of particles. It is described by a Hamiltonian operator \hat{H} . It has eigenstates Ψ_i :

$$\hat{H}\Psi_i = E_i\Psi_i. \quad (7)$$

It should be emphasized that the system we study has on the order of 10^{23} particles, so the function $\Psi_i(\mathbf{r}_1, \mathbf{r}_2, \dots)$ has equally many arguments.

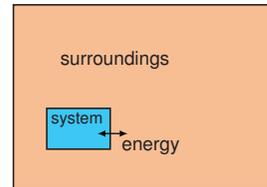
Equilibrium

An important basic result is *Gibbs distribution*: the probability ρ_i that the state Ψ_i occurs is

$$\rho_i = e^{\beta(F-E_i)}. \quad (8)$$

Here the constant β can be interpreted as inverse temperature: $\beta = 1/(k_B T)$. Here is Boltzmann’s constant $k_B = 1.38 \times 10^{-23}$ J/K, which is needed to express the temperature T in Kelvins. The constant F is determined by the condition $\sum_i \rho_i = 1$.

The Gibbs distribution can be derived under the following assumptions: 1) the system we study interacts with a much larger system called surroundings. 2) all states of the whole system (system under study + surroundings) occur with equal probability within some energy interval. The most straightforward derivation of this I have seen in the book R.P. Feynman, Statistical mechanics.



The expectation value $\langle \hat{A} \rangle$ of an arbitrary operator \hat{A} is given by

$$\langle \hat{A} \rangle = \sum_i \rho_i \langle \Psi_i | \hat{A} | \Psi_i \rangle = \sum_i \langle \Psi_i | \hat{A} e^{\beta(F-\hat{H})} | \Psi_i \rangle = \text{Tr}(\hat{A}\hat{\rho}), \quad (9)$$

where we have defined a (probability) density operator

$$\hat{\rho} = e^{\beta(F-\hat{H})}. \quad (10)$$

This is the Gibbs distribution represented in operator form. The normalization condition $\text{Tr}\hat{\rho} = 1$ gives

$$F = -\frac{1}{\beta} \ln(\text{Tr}e^{-\beta\hat{H}}) = -\frac{1}{\beta} \ln\left(\sum_i e^{-\beta E_i}\right). \quad (11)$$

We additionally define the entropy

$$S = -k_B \langle \ln \hat{\rho} \rangle \quad (12)$$

and internal energy $E = \langle \hat{H} \rangle$. Show as an exercise that

$$F = E - ST. \quad (13)$$

Let us suppose that the Hamiltonian depends on a parameter λ : $\hat{H}(\lambda)$. Differentiating the normalization condition $\text{Tr} \hat{\rho} = 1$ show as an exercise that

$$dF = -SdT + \left\langle \frac{d\hat{H}}{d\lambda} \right\rangle d\lambda. \quad (14)$$

Supposing that λ is the volume V of the system and defining the pressure

$$p = -\left\langle \frac{d\hat{H}}{dV} \right\rangle, \quad (15)$$

we get equation (14) into the form

$$dF = -SdT - pdV. \quad (16)$$

We recognize that equations (13) and (16) are familiar from thermodynamics, and they could be used to derive the first law of thermodynamics

$$dE = TdS - pdV. \quad (17)$$

Thus the quantities T , S etc. can be identified as the same quantities as defined in thermodynamics. Especially we identify $TdS = dQ$ as the heat absorbed by the system, and we can define the specific heat in constant volume

$$C_V = \left(\frac{dQ}{dT} \right)_V = T \left(\frac{\partial S}{\partial T} \right)_V = -T \left(\frac{\partial^2 F}{\partial T^2} \right)_V. \quad (18)$$

Next we show that the temperature T as defined above has properties that we expect it to have.

We study a system consisting of two parts: $\hat{H} = \hat{H}_1 + \hat{H}_2$. According to Gibbs distribution the equilibrium has

$$\hat{\rho} = e^{\beta(F-\hat{H})} = e^{\beta(F_1-\hat{H}_1)} e^{\beta(F_2-\hat{H}_2)}. \quad (19)$$

So we see that the temperatures $T = 1/(k_B\beta)$ of both subsystems are the same. This is the crucial property that we require the temperature to satisfy.

Usually there is no upper bound for the energy eigenvalues E_i . In order to the Gibbs distribution (8) to be reasonable, we must have $\beta > 0$, which implies $T \geq 0$.

We see directly from the Gibbs distribution (8) that higher temperature (smaller β) means that states with high energy have larger probability. Thus the internal energy E is a monotonically increasing function of T (assuming V is constant). If $T = 0$, only the ground state (which has the minimum E_i) is possible.

Nonequilibrium

The previous analysis can be extended to *nonequilibrium* systems as follows. We suppose that the system consists

of two macroscopic parts. (The generalization to arbitrary number of parts is trivial.) We suppose both parts are internally in equilibrium, but the parts are not in equilibrium with each other. For simplicity we assume that the parts can only exchange heat with each other (the volumes are constants). For both subsystems ($i = 1, 2$) the first law (17) gives

$$dE_i = T_i dS_i. \quad (20)$$

We define that total entropy as sum of the entropies of the subsystems,

$$S = S_1 + S_2. \quad (21)$$

The change of entropy is given by

$$dS = dS_1 + dS_2 = \frac{dE_1}{T_1} + \frac{dE_2}{T_2} = \left(\frac{1}{T_1} - \frac{1}{T_2} \right) dE_1, \quad (22)$$

because the total energy is conserved. We suppose that $T_1 > T_2$. Because T is a monotonically increasing function of energy, the time derivative of E_1 has to be negative, $dE_1/dt < 0$, in order for the equilibrium state to be stable. This case (and also the opposite case $T_1 < T_2$) implies that $dS/dt > 0$. Thus *the entropy always grows in transition from nonequilibrium to the equilibrium state.*

Let us return to study a system connected to a surroundings. The surroundings can be interpreted as an ideal heat bath, whose energy change satisfies $dE_b = TdS_b$. Here T is constant because the surroundings is much bigger than the system. Because the total system is closed, we have $dE_{\text{tot}} = dE + dE_b = 0$, and

$$\frac{dS_{\text{tot}}}{dt} = \frac{dS}{dt} + \frac{dS_b}{dt} \geq 0. \quad (23)$$

So we get $dE - TdS \leq 0$. We define the nonequilibrium F by the expression $F = E - TS$. It follows

$$\frac{dF}{dt} \leq 0 \quad (T \text{ and } V \text{ constants}). \quad (24)$$

Thus in equilibrium the free energy F has its minimum value.

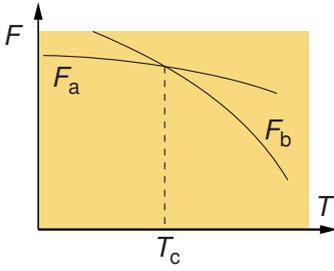
Let us restate the previous result more precisely. In nonequilibrium state $F(T, V, \lambda_1, \lambda_2, \dots)$ depends on several internal degrees of freedom λ_i . In equilibrium it is

$$F(T, V) = \min_{\lambda_1, \lambda_2, \dots} F(T, V, \lambda_1, \lambda_2, \dots). \quad (25)$$

The condition $F = E - TS = \text{minimum}$ generalizes the zero-temperature condition $E = \text{minimum}$ to finite temperature.

Example Phase equilibrium

Often the system can appear in two different phases, for example liquid and gas. Suppose that we have calculated the free energies $F_a(T, V)$ and $F_b(T, V)$ for the two phases as shown in the figure.



According to the previous result, the phase that is realized in equilibrium is the one having lower free energy. We conclude that there is a phase transition between the phases at temperature where

$$F_a = F_b. \quad (26)$$

Variable particle number

Often it is mathematically easier to study case, where the particle number is not fixed. This can be achieved by thinking the system as connected to a “particle bath”, an ideal reservoir of particles at constant energy μ , which is called the chemical potential. In Gibbs distribution (10) we can generalize

$$\hat{H} \rightarrow \hat{H} + \mu \hat{N}_b = \hat{H} - \mu \hat{N} + \text{constant} \quad (27)$$

(because $N_{\text{tot}} = N + N_b = \text{constant}$). (Here \hat{N} is the particle-number operator of the system: $\hat{N}\Psi_i = N_i\Psi_i$, where N_i is the number of particles in state Ψ_i .) In making the substitution (27) one also replaces the constant F (Helmholtz free energy) with another constant Ω (grand potential). Therefore

$$\hat{\rho} = e^{\beta(\Omega - \hat{H} + \mu \hat{N})}. \quad (28)$$

In the same way as for F , one can derive for Ω the definition (also nonequilibrium, $N = \langle \hat{N} \rangle$)

$$\Omega = E - \mu N - ST, \quad (29)$$

the equilibrium expression

$$\Omega = -\frac{1}{\beta} \ln \left[\text{Tr} e^{-\beta(\hat{H} - \mu \hat{N})} \right] = -\frac{1}{\beta} \ln \left[\sum_i e^{-\beta(E_i - \mu N_i)} \right], \quad (30)$$

the differential for equilibrium states

$$d\Omega = -SdT - pdV - Nd\mu \quad (31)$$

and the time development

$$\frac{d\Omega}{dt} \leq 0 \quad (T, V \text{ and } \mu \text{ constants}). \quad (32)$$

In addition we deduce from equation (31)

$$S = -\left(\frac{\partial \Omega}{\partial T}\right)_{V, \mu}, \quad p = -\left(\frac{\partial \Omega}{\partial V}\right)_{T, \mu}, \quad N = -\left(\frac{\partial \Omega}{\partial \mu}\right)_{T, V}. \quad (33)$$

Ideal Fermi gas

Above we formally discussed the many-body wave functions Ψ_i . The calculation of these is possible only in very special cases. One case is an ideal gas, where we assume that there are no interactions between the particles.

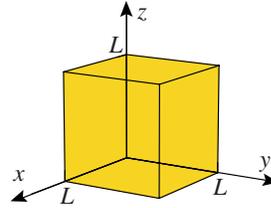
The natural choice for wave functions of a single free particle are plane wave states

$$\phi_{\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{V}} e^{i\mathbf{k}\cdot\mathbf{r}}, \quad (34)$$

where the wave vector \mathbf{k} appears as a parameter. The energy of these states is $\epsilon_k = \hbar^2 k^2 / 2m$. In order to count the states, it is most simple to require that the wave functions are periodic in a cube of volume $V = L^3$, which allows the wave vectors \mathbf{k} (n_x, n_y and n_z integers)

$$k_x = \frac{2\pi n_x}{L}, \quad k_y = \frac{2\pi n_y}{L}, \quad k_z = \frac{2\pi n_z}{L}. \quad (35)$$

We suppose that the volume V is very large. Then we can take the limit $V \rightarrow \infty$ in quantities that do not essentially depend on V .



In addition to the location \mathbf{r} , the wave function of a fermion depends on spin index σ , which describes the component of spin angular momentum on some chosen z axis. For spin- $\frac{1}{2}$ fermions this can have two values. These can be denoted by $\sigma = \pm\frac{1}{2}$, or alternatively by \uparrow and \downarrow . For free particles we can thus choose “spin-up levels”

$$\phi_{\mathbf{k}\uparrow}(\mathbf{r}, \sigma) = \begin{cases} \frac{1}{\sqrt{V}} e^{i\mathbf{k}\cdot\mathbf{r}} & \text{if } \sigma = \frac{1}{2} \\ 0 & \text{if } \sigma = -\frac{1}{2} \end{cases}. \quad (36)$$

and “spin-down levels”

$$\phi_{\mathbf{k}\downarrow}(\mathbf{r}, \sigma) = \begin{cases} 0 & \text{if } \sigma = \frac{1}{2} \\ \frac{1}{\sqrt{V}} e^{i\mathbf{k}\cdot\mathbf{r}} & \text{if } \sigma = -\frac{1}{2} \end{cases}. \quad (37)$$

Note that (following book AM) we call the single-particle $\phi_{\mathbf{k}\uparrow}$ and $\phi_{\mathbf{k}\downarrow}$ “levels” in order to clearly distinguish them from the “states” Ψ_i of the many body system.

There are several ways to present many-body states Ψ_i , as will be discussed later. One useful way to think of these states is first to list all levels (36)-(37) in some arbitrary order, for example

$$\phi_{\mathbf{0}\uparrow}, \phi_{\mathbf{0}\downarrow}, \phi_{\mathbf{k}_1\uparrow}, \phi_{\mathbf{k}_1\downarrow}, \phi_{\mathbf{k}_2\uparrow}, \phi_{\mathbf{k}_2\downarrow}, \dots \quad (38)$$

Then the basis states of the many-body space can be expressed by telling how many particles is in any of the

levels,

$$|\Psi_i\rangle = |n_1, n_2, n_3, \dots, n_\infty\rangle. \quad (39)$$

Here n_α is the number of particles in the α 'th level (38). [In practice the writing of the state (39) is difficult because there is an infinite number of levels and thus the great majority of the numbers n_α are zeros.] The energy of the many-body state (39) is $E = \sum_\alpha n_\alpha \epsilon_\alpha$. Fermions obey the Pauli exclusion principle and thus all occupations n_α are either 0 or 1.

In order to determine the thermal equilibrium state, it is easiest to use the formulas for variable particle number. Starting from equation (30) we get

$$\begin{aligned} e^{-\beta\Omega} &= \text{Tr} e^{-\beta(\hat{H}-\mu\hat{N})} \\ &= \sum_{n_1} \sum_{n_2} \dots \langle n_1, n_2, \dots | e^{-\beta(\hat{H}-\mu\hat{N})} | n_1, n_2, \dots \rangle \\ &= \sum_{n_1} \sum_{n_2} \dots e^{-\beta(\epsilon_1-\mu)n_1} e^{-\beta(\epsilon_2-\mu)n_2} \dots \\ &= \sum_{n_1} e^{-\beta(\epsilon_1-\mu)n_1} \sum_{n_2} e^{-\beta(\epsilon_2-\mu)n_2} \dots \\ &= \prod_\alpha \sum_{n_\alpha} e^{-\beta(\epsilon_\alpha-\mu)n_\alpha} \\ &= \prod_\alpha \left[1 + e^{-\beta(\epsilon_\alpha-\mu)} \right]. \end{aligned} \quad (40)$$

Thus

$$\Omega = -\frac{1}{\beta} \ln \prod_\alpha \left[1 + e^{-\beta(\epsilon_\alpha-\mu)} \right] = -\frac{1}{\beta} \sum_\alpha \ln \left[1 + e^{-\beta(\epsilon_\alpha-\mu)} \right]. \quad (41)$$

From this we can calculate all thermodynamic quantities. Especially the particle number (31) is

$$\begin{aligned} N &= -\left(\frac{\partial \Omega}{\partial \mu} \right)_{T,V} = \frac{1}{\beta} \sum_\alpha \frac{e^{-\beta(\epsilon_\alpha-\mu)} \beta}{1 + e^{-\beta(\epsilon_\alpha-\mu)}} \\ &= \sum_\alpha \frac{1}{e^{\beta(\epsilon_\alpha-\mu)} + 1}. \end{aligned} \quad (42)$$

Here we see that the average occupation probability of each level depends on its energy ϵ and is

$$f(\epsilon) = \frac{1}{e^{\beta(\epsilon-\mu)} + 1}. \quad (43)$$

This is the familiar Fermi-Dirac distribution.

In a similar fashion we can derive for an ideal Bose gas (possible occupation numbers $n_\alpha = 0, 1, 2, \dots, \infty$) the Bose-Einstein distribution

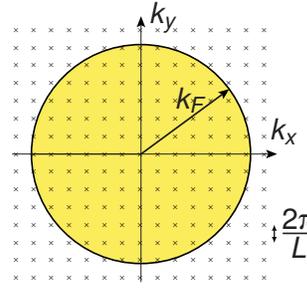
$$f(\epsilon) = \frac{1}{e^{\beta(\epsilon-\mu)} - 1}. \quad (44)$$

The purpose above was to show that ideal Bose and Fermi distributions can be derived from the more general Gibbs distribution, which can be applied to arbitrary interacting systems as well.

Let us remind about the main features of Fermi distribution. When the temperature $T \rightarrow 0$, the occupation becomes a step function

$$f(\epsilon) = \begin{cases} 1 & \text{for } \epsilon < \mu \\ 0 & \text{for } \epsilon > \mu, \end{cases} \quad (45)$$

where all levels below the chemical potential μ are filled. The kinetic energy at highest filled level is called *Fermi energy* ϵ_F , and expressed in temperature units it is called *Fermi temperature* T_F : $\mu(T=0) = \epsilon_F = k_B T_F$. We also define the *Fermi wave vector* k_F and the *Fermi momentum* $p_F = \hbar k_F$ corresponding to the Fermi energy, $\hbar^2 k_F^2 / 2m = \epsilon_F$. In momentum space all levels inside ($k < k_F$) of the Fermi surface ($k = k_F$) are occupied, and the ones outside are empty.



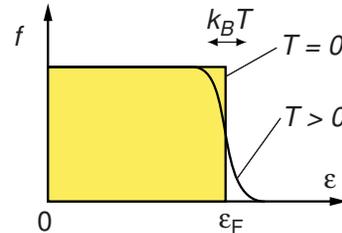
Equation (42) gives at zero temperature

$$N = 2 \sum_{k < k_F} 1 = 2 \frac{\frac{4}{3} \pi k_F^3}{(2\pi/L)^3},$$

where the factor 2 comes from spin. From this we get a relation between the Fermi wave vector and the particle density,

$$\frac{N}{V} = \frac{k_F^3}{3\pi^2}. \quad (46)$$

When $T > 0$, the occupation $f(\epsilon)$ gets rounded so that the change from $f \approx 1$ to $f \approx 0$ takes place in the energy interval $\approx k_B T$.



3. Thermodynamics of superconductors

When we study the thermodynamics of a magnetic material, we should consider the field as an additional variable. The equation (14) can be written (supposing the volume V constant)

$$dF = -SdT + V\mathbf{H} \cdot d\mathbf{B}. \quad (47)$$

In the normal way one gets from this

$$S = -\left(\frac{\partial F}{\partial T}\right)_B, \quad \mathbf{H} = \frac{1}{V}\left(\frac{\partial F}{\partial \mathbf{B}}\right)_T, \quad (48)$$

where the vector relation should be understood as $H_x = \partial F / V \partial B_x$ etc.

Here is more detailed justification of equation (47). The Maxwell equations in the presence of material are (Appendix A)

$$\nabla \cdot \mathbf{D} = \rho_f, \quad (49)$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \quad (50)$$

$$\nabla \cdot \mathbf{B} = 0, \quad (51)$$

$$\nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t} + \mathbf{j}_f. \quad (52)$$

Together with material relations, e.g.

$$\mathbf{D}(\mathbf{E}) \text{ ja } \mathbf{H}(\mathbf{B}), \quad (53)$$

they form a complete set of equations. From these one can derive the conservation law

$$\begin{aligned} \int dV (\mathbf{E} \cdot d\mathbf{D} + \mathbf{H} \cdot d\mathbf{B} + \mathbf{E} \cdot \mathbf{j}_f dt) \\ = -dt \oint d\mathbf{a} \cdot (\mathbf{E} \times \mathbf{H}). \end{aligned} \quad (54)$$

The left hand side is energy differential integrated over the system and the right hand side is the flux of electromagnetic energy through the surface of the system. When $\mathbf{D} = 0$ and $\mathbf{j}_f = 0$, we get the magnetic energy change of a homogeneous system as given by equation (47). ■

In order to study the sample in a given external field \mathbf{H} , we have to add to F an energy term (“bath”) that gives rise to this field. The needed term can be justified as Legendre transformation from variable \mathbf{B} to \mathbf{H} . So we get as the quantity to be minimized

$$G = F - V\mathbf{H} \cdot \mathbf{B}. \quad (55)$$

and its differential in equilibrium state

$$dG = -SdT - V\mathbf{B} \cdot d\mathbf{H}. \quad (56)$$

More physically, the additional term in equation (55) can be understood as the energy of a current source. The

current source drives current to a coil that generates the magnetizing field \mathbf{H} . (Both the coil and the current source are idealized as dissipationless.) When \mathbf{B} in the sample changes, it induces a voltage in the coil and changes the energy of the current source just by the change of the additional term. (more details as an exercise)

Let us apply this to a long cylindrical superconducting sample in parallel magnetic field. When H increases from zero, equation (56) gives

$$G(T, H) - G(T, 0) = -V \int_0^H dH' B(H'). \quad (57)$$

In the normal state of a superconductor, the magnetization is very weak. Thus to a good approximation $\mathbf{B} = \mu_0 \mathbf{H}$ (Appendix A). We get

$$G_n(T, H) - G_n(T, 0) = -\frac{1}{2}V\mu_0 H^2. \quad (58)$$

In the superconducting state $\mathbf{B} = 0$, so we get

$$G_s(T, H) = G_s(T, 0). \quad (59)$$

Because these two states are in equilibrium at the critical field H_c , their potentials must be the same (26):

$$G_n(T, H_c) = G_s(T, H_c). \quad (60)$$

From above we deduce

$$G_s(T, 0) = G_n(T, 0) - \frac{1}{2}V\mu_0 H_c^2. \quad (61)$$

So the energy of the superconducting state is lower than that of the normal state by the *condensation energy* $\frac{1}{2}V\mu_0 H_c^2$.

Rearranging we get from formulas (58), (59) and (61)

$$G_s(T, H) - G_n(T, H) = \frac{1}{2}V\mu_0(H^2 - H_c^2). \quad (62)$$

According to equation (56) we see that the corresponding difference in entropy is obtained by taking the derivative

$$S_s(T, H) - S_n(T, H) = V\mu_0 H_c(T) \frac{dH_c(T)}{dT}. \quad (63)$$

Because $H_c(T)$ decreases as the temperature increases (equation 3), we see that the entropy in the superconducting state is lower than in the normal state. Note that the entropy difference (63) is independent of the field H . From equations (63) we can also deduce that the latent heat in the transition between normal and superconducting states $T(S_s - S_n)$ vanishes at $T = 0$ and at $T = T_c$, but not at intermediate temperatures.

The specific heat in constant field can be calculated using

$$C_H = T \left(\frac{\partial S}{\partial T} \right)_H \quad (64)$$

From this we get the discontinuity of the specific heat

$$\begin{aligned} C_s(T, H) - C_n(T, H) \\ = TV\mu_0 \left[\left(\frac{dH_c}{dT} \right)^2 + H_c \frac{d^2 H_c}{dT^2} \right]. \end{aligned} \quad (65)$$

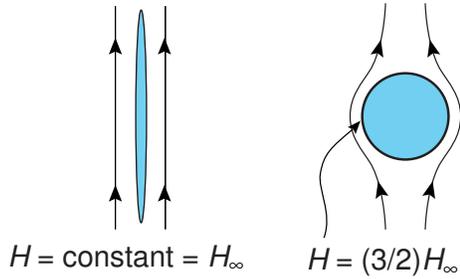
In the special case $T = T_c$ this reduces to

$$C_s(T, H) - C_n(T, H) = \mu_0 VT_c \left(\frac{dH_c}{dT} \right)_{T_c}^2. \quad (66)$$

In summary, the thermodynamics sets some conditions between different measurable quantities.

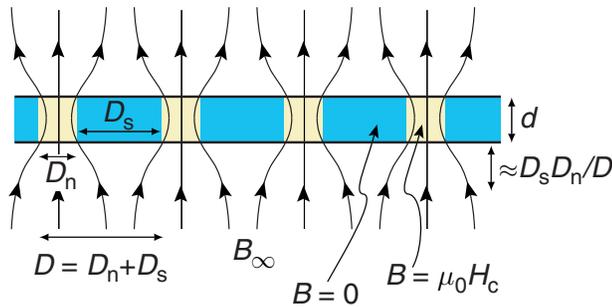
3.1 Intermediate state

The discussion above is valid for a thin sample that in parallel field, where the magnetizing field $\mathbf{H} \approx \text{constant}$. [Show this starting from equation $\nabla \times \mathbf{H} = 0$ (52).] In other cases $\mathbf{H}(\mathbf{r}) \neq \text{constant}$. For example, for a sphere we get that the field at polar angle $\theta = \pi/2$ is $H = \frac{3}{2}H(r = \infty)$.



If $H(r = \infty) > \frac{2}{3}H_c$, the critical field is exceeded on the equator. On the other hand, the whole sphere cannot go into the normal state below $H(r = \infty) = H_c$. Between these fields one gets so called intermediate state (of type I superconductor).

Let us consider superconducting slab in perpendicular magnetic field. We estimate orders of magnitude assuming the intermediate state consists of alternating layers of normal and superconducting states (thicknesses D_s and D_n).



It follows from Maxwell equations (51) that the average flux through the slab is the same as the field B_∞ far from the slab. This goes through the normal state part only. In

order the normal and superconducting states to be in equilibrium, the normal state must have $B = \mu_0 H_c$. This gives

$$B_\infty(D_s + D_n) = \mu_0 H_c D_n. \quad (67)$$

It is obvious that there is some extra energy associated with the interface between normal and superconducting states, which is proportional to the area A of the interface. We write this energy in the form

$$F_{\text{surface}} = \sigma A, \quad \sigma = \frac{\mu_0}{2} H_c^2 \delta, \quad (68)$$

where δ is some microscopic quantity of dimension length. The surface energy per unit area of the slab is

$$f_{\text{surface}} = \frac{2d\sigma}{D} = \frac{\mu_0}{2} H_c^2 \frac{2d\delta}{D}, \quad (69)$$

where $D = D_n + D_s$. The minimization of this would lead to minimizing the interface by $D \rightarrow \infty$. This tendency is opposed by the fact that then the field outside the slab would be very inhomogeneous. In vacuum $\mathbf{H} = \mathbf{B}/\mu_0$ and the energy density is $\frac{1}{2\mu_0} B^2$. We estimate the order of magnitude of the additional energy caused by the inhomogeneity of the field by

$$f_{\text{inhomog}} = \left(\frac{\mu_0}{2} H_c^2 \frac{D_n}{D} - \frac{1}{2\mu_0} B_\infty^2 \right) \frac{2D_s D_n}{D} \quad (70)$$

The expression in parenthesis gives the field energy near the slab (weighted by the area factor D_n/D), from which the energy of homogeneous field is subtracted. The factor $2D_s D_n/D \approx 2 \min(D_s, D_n)$ estimated the thickness where the inhomogeneity is important. The total energy is

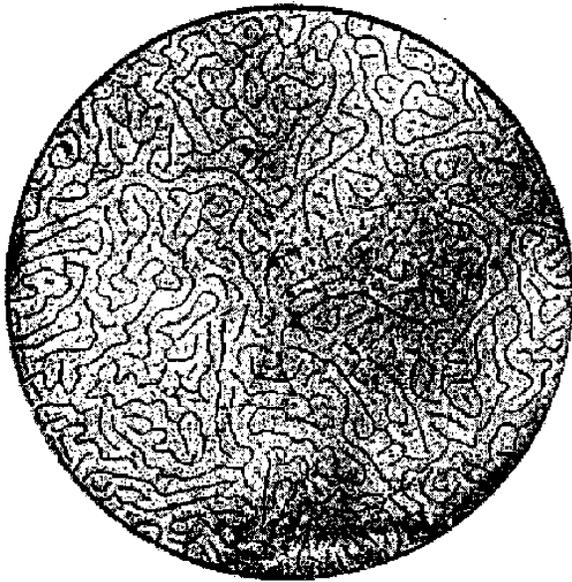
$$f = \mu_0 H_c^2 \left[\frac{d\delta}{D} + D \left(\frac{B_\infty}{\mu_0 H_c} \right)^2 \left(1 - \frac{B_\infty}{\mu_0 H_c} \right)^2 \right] \quad (71)$$

Minimizing with respect to D we get

$$D = \frac{\sqrt{d\delta}}{\left(\frac{B_\infty}{\mu_0 H_c} \right) \left(1 - \frac{B_\infty}{\mu_0 H_c} \right)} \quad (72)$$

The essential thing is the numerator, which gives the geometric mean of a microscopic and macroscopic length, δ and d .

Experiments show alternating normal and superconducting layers that can form quite complicated structures.



Intermediate state of indium. The superconducting regions are made visible with niobium powder (black) as the superconducting powder particles tend to concentrate in regions of low magnetic field. The applied field is close to the critical field ($H/H_c = 0.931$). F. Haenssler and L. Rinderer, *Helv. Phys. Acta* **40**, 659 (1967).

The intermediate state described above occurs in type I superconductors. Type II superconductors show essentially different behavior in magnetic field, as will be seen later.

4. Microscopic theory

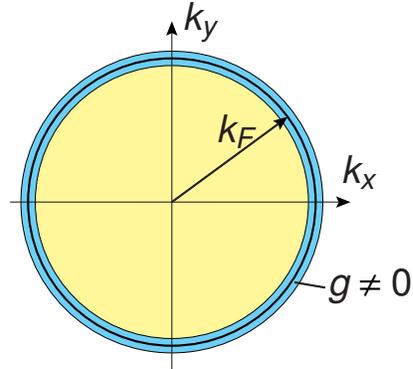
4.1 Normal state metal

The properties of metals are more thoroughly studied in the course on condensed matter physics. Here we only consider a very simplified model of a metal, which still is sufficient as a starting point for understanding basic properties of superconductivity.

We assume that there are *conduction electrons*, which can move freely like ideal gas through the metal. We assume that the rest of the electrons called *valence electrons* are bound to the atomic nuclei so that they do not contribute to the electric conductivity. The kinetic energy $\epsilon_{\mathbf{k}}$ of the conduction electrons depends on the wave vector \mathbf{k} . In the simplest case this dependence is of the form

$$\epsilon_{\mathbf{k}} = \frac{\hbar^2 k^2}{2m}. \quad (73)$$

Here m is the effective mass, which can be different from the free electron mass. We can estimate that the Fermi energy ϵ_F is on the order of a few eV corresponding to $T_F \approx 10^5$ K. This should be compared with the typical temperatures $T \sim T_c \sim 10$ K that occur in superconductivity. The theory thus has two very different energy scales, $T_c/T_F \sim 10^{-4}$. We will often make use of the relation $T \ll T_F$.



The figure illustrates the neighborhood of the Fermi surface. (A cut $k_z = 0$ of the momentum space is shown.)

Example. One often encounters integration over \mathbf{k} space of a function $g(\mathbf{k})$. Show as an exercise that

$$\frac{1}{L^3} \sum_{\mathbf{k}} g(\mathbf{k}) = \int \frac{d^2\Omega}{4\pi} \int d\epsilon N(\epsilon) g(\mathbf{k}). \quad (74)$$

This contains integration over the solid angle

$$\int d^2\Omega = \int_0^{2\pi} d\phi \int_0^\pi d\theta \sin \theta, \quad (75)$$

where θ and ϕ are the polar and azimuthal angles of \mathbf{k} . The second integration is over energy, $\int d\epsilon$. The factor $N(\epsilon)$ is the density of levels per unit energy (and per

volume). If $g(\mathbf{k})$ is different from zero only near the Fermi surface (figure), we can further approximate

$$\frac{1}{L^3} \sum_{\mathbf{k}} g(\mathbf{k}) = N(0) \int \frac{d^2\Omega}{4\pi} \int d\epsilon g(\mathbf{k}), \quad (76)$$

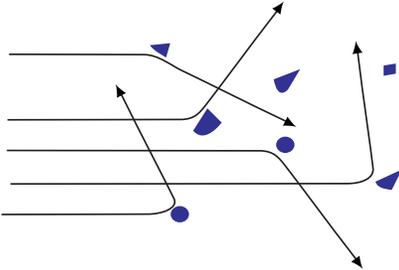
where

$$N(0) = \frac{mk_F}{2\pi^2\hbar^2} \quad (77)$$

is the density of levels at the Fermi surface. [The notation $N(0)$ instead of $N(\epsilon_F)$ comes from the fact that often it is most convenient to count the energy starting from the Fermi level. In condensed matter physics the zero value of the energy can usually be chosen freely, and the Fermi level is often a better choice than the bottom of the conduction band.]

In the free electron model it is easy to calculate the specific heat. The result is that C is linearly proportional to the temperature when $T \ll T_F$. This explains the specific heat of the normal metal mentioned in the introduction. This will be calculated in detail later on in this course.

How can we understand the electrical resistivity of a normal metal? The reason for electrical resistance is that the conduction electrons collide with lattice vibrations and impurities that occur in the metal. If the ideal gas has net momentum in the beginning, it will decay gradually as individual electrons collide and are scattered in random directions.



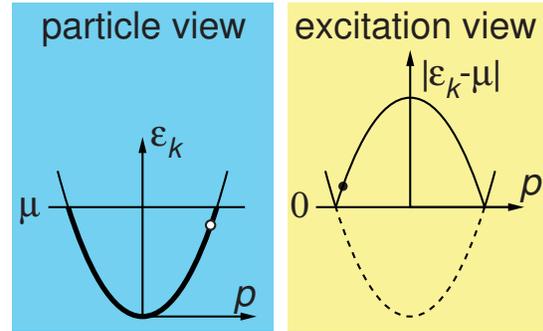
Excitation picture

Previously we discussed that the ground state (= lowest energy state) of a Fermi gas is such that all energy levels inside the Fermi surface are filled and the levels outside are empty. The simplest excited states are of two different types.

- one extra electron is in a level with \mathbf{k} whose $\epsilon_k > \mu$. This excitation has energy $\xi_k \equiv \epsilon_k - \mu$ and momentum $\mathbf{p} = \hbar\mathbf{k}$.
- one electron is missing from a level with \mathbf{k} whose $\epsilon_k < \mu$. This is called a *hole*. Its excitation energy is $\mu - \epsilon_k > 0$ and momentum $\mathbf{p} = -\hbar\mathbf{k}$.

In both cases the excitation energies are positive. The figure illustrates the same hole type excitation seen both

as a missing particle and as an excitation. (The momentum space is simplified as one-dimensional in the figure.)



If the particle number is fixed, the particle and hole type excitations must always appear as pairs. That is, the number of particle type excitations has to be equal to the number of hole type excitations.

We note that also the spin of the hole type excitation has to be opposite to the spin of the missing electron. If an electron at \mathbf{k} ($\epsilon_k < \mu$) and spin up (\uparrow) is missing, it corresponds to an excitation of momentum $\mathbf{p} = -\hbar\mathbf{k}$ and spin down (\downarrow).

4.2 Ideas about superconductivity

Materials can be studied on two very different scales:

- atomic scale. Quantum mechanics is essential. There is no friction.
- macroscopic scale. Material is described by laws of classical physics (e.g. theory of elasticity). Friction is essential.

The phenomena on atomic scale are not usually visible on the macroscopic scale. The reason for this is that macroscopic bodies consist of enormously large number of particles ($\sim 10^{23}$). The different particles are in general in different quantum levels, and only the average of them is visible. Friction is needed because it is not possible to take into account the motion of individual particles.

Idea: Superconductivity is an exception to the rule above: it is a quantum phenomenon that is still preserved on a macroscopic scale.

How is this possible? As a simple example, let us consider ideal Bose gas. At zero temperature its distribution function (44) reduces to the form

$$f(\epsilon) = \begin{cases} N & \text{lowest level } (\epsilon_i = \epsilon_0) \\ 0 & \text{other levels } (\epsilon_i > \epsilon_0). \end{cases} \quad (78)$$

Also at finite temperatures below so-called Bose condensation temperature, the occupation of the lowest level $N_0/N > 0$, whereas for all other levels $f_i/N \rightarrow 0$ when $N, V \rightarrow \infty$. It is expressed by saying that the wave

function of the lowest level becomes *macroscopic wave function* because a macroscopic number of particles is in the same level.

It can be said that *the superfluid phases of alkali atom gases and ^4He liquid are based on Bose condensation*. However, the ideal Bose gas is a too simplified model for them because more detailed study shows that the interactions between particles have an essential role in the superfluid state. In this course we will not study this interesting problem more.

Electrons are fermions. One can put only one fermion into a single level. Thus the discussion above as such cannot explain the superconductivity of metals.

Let us study the wave functions of spin- $\frac{1}{2}$ fermions. The wave function $\phi(\mathbf{r}\sigma)$ of a level [e.g. (36) or (37)] depends on the location \mathbf{r} and on the spin index $\sigma = \pm\frac{1}{2}$. In general form the Pauli exclusion principle says that the wave function has to be antisymmetric in any exchange of two electron coordinates. For a two-electron wave function $\Psi(\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2)$ this requirement is

$$\Psi(\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2) = -\Psi(\mathbf{r}_2\sigma_2, \mathbf{r}_1\sigma_1). \quad (79)$$

Starting from an arbitrary function $\psi_0(\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2)$, one can always by antisymmetrizing construct a function that satisfies this condition:

$$\Psi(\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2) = \psi_0(\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2) - \psi_0(\mathbf{r}_2\sigma_2, \mathbf{r}_1\sigma_1). \quad (80)$$

Thus the pair state formed from two levels ϕ_1 and ϕ_2 has the wave function

$$\Psi(\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_1) = \phi_1(\mathbf{r}_1\sigma_1)\phi_2(\mathbf{r}_2\sigma_2) - \phi_1(\mathbf{r}_2\sigma_2)\phi_2(\mathbf{r}_1\sigma_1), \quad (81)$$

so called Slater determinant. We see immediately that this vanishes, $\Psi \equiv 0$, if the levels are the same: $\phi_1 = \phi_2$. Thus two fermions cannot be placed into the same level.

In case of many particles, the antisymmetry is required in any pairwise exchange of the coordinates of two particles

$$\begin{aligned} \Psi(\dots, \mathbf{r}_i\sigma_i, \mathbf{r}_{i+1}\sigma_{i+1}, \dots, \mathbf{r}_k\sigma_k, \dots) \\ = -\Psi(\dots, \mathbf{r}_k\sigma_k, \mathbf{r}_{i+1}\sigma_{i+1}, \dots, \mathbf{r}_i\sigma_i, \dots) \end{aligned} \quad (82)$$

Idea: fermions form *pairs*.

$$\begin{aligned} \psi_0(\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2, \mathbf{r}_3\sigma_3, \dots) = \phi(\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2) \times \\ \phi(\mathbf{r}_3\sigma_3, \mathbf{r}_4\sigma_4)\phi(\mathbf{r}_5\sigma_5, \mathbf{r}_6\sigma_6) \dots, \end{aligned} \quad (83)$$

All *pair states are the same!* This function does *not* vanish in antisymmetrization as long as the pair function is antisymmetric,

$$\phi(\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2) = -\phi(\mathbf{r}_2\sigma_2, \mathbf{r}_1\sigma_1). \quad (84)$$

In exchanging pairs one gets the factor $(-1)^2 = 1$, similarly to bosons.

We see that superconductivity could arise from macroscopic occupation of a pair state. Before we can accept this claim, we must be able to answer the following questions.

- Is there a force that binds the pairs?
- Is it sufficiently strong?
- Does the pair state have a lower energy than the normal state?

4.3 Cooper problem

Let us study two fermions that interact with each other but not with other fermions at temperature $T = 0$. We ignore the spin for a while. The Schrödinger equation is

$$\left[-\frac{\hbar^2}{2m}\nabla_1^2 - \frac{\hbar^2}{2m}\nabla_2^2 + V(\mathbf{r}_1, \mathbf{r}_2) \right] \phi(\mathbf{r}_1, \mathbf{r}_2) = E\phi(\mathbf{r}_1, \mathbf{r}_2). \quad (85)$$

It is likely that in the lowest energy state, the center of the mass of the pair is at rest. Therefore we write the wave function as dependent only on the difference $\mathbf{r}_1 - \mathbf{r}_2$,

$$\phi(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{L^3} \sum_{\mathbf{k}} \chi(\mathbf{k}) e^{i\mathbf{k}\cdot(\mathbf{r}_1 - \mathbf{r}_2)}. \quad (86)$$

We substitute this into the Schrödinger equation (85). We multiply by $e^{-i\mathbf{k}'\cdot(\mathbf{r}_1 - \mathbf{r}_2)}$, integrate over \mathbf{r}_1 and \mathbf{r}_2 and get

$$(2\epsilon_{\mathbf{k}'} - E) \chi(\mathbf{k}') = - \sum_{\mathbf{k}} \langle \mathbf{k}', -\mathbf{k}' | V | \mathbf{k}, -\mathbf{k} \rangle \chi(\mathbf{k}). \quad (87)$$

Here we have used the notation

$$\begin{aligned} \langle \mathbf{k}'_1, \mathbf{k}'_2 | V | \mathbf{k}_1, \mathbf{k}_2 \rangle = \frac{1}{L^6} \int d^3r_1 \int d^3r_2 \\ \times e^{-i\mathbf{k}'_1\cdot\mathbf{r}_1} e^{-i\mathbf{k}'_2\cdot\mathbf{r}_2} V(\mathbf{r}_1, \mathbf{r}_2) e^{i\mathbf{k}_1\cdot\mathbf{r}_1} e^{i\mathbf{k}_2\cdot\mathbf{r}_2}, \end{aligned} \quad (88)$$

which looks slightly clumsy here but is useful later.

The presence of other electrons is taken into account only through the Pauli exclusion principle. They restrict the pair function so that

$$\chi(\mathbf{k}) = 0 \quad \text{for } k < k_F. \quad (89)$$

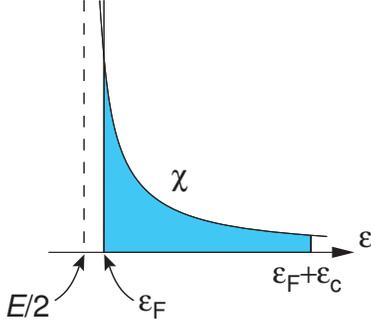
We proceed from equation (87) by assuming that V is a constant in a thin layer on both sides of the Fermi surface, and zero elsewhere,

$$\begin{aligned} \langle \mathbf{k}', -\mathbf{k}' | V | \mathbf{k}, -\mathbf{k} \rangle \\ = \begin{cases} -g/L^3 & \text{for } |\epsilon_{\mathbf{k}'} - \epsilon_F| < \epsilon_c \text{ and } |\epsilon_{\mathbf{k}} - \epsilon_F| < \epsilon_c, \\ 0 & \text{otherwise,} \end{cases} \end{aligned} \quad (90)$$

where we assume that $g > 0$ and $\epsilon_c \ll \epsilon_F$. [The cutting of the interaction (90) is slightly problematic since according to (88) the interaction $\langle \mathbf{k}', -\mathbf{k}' | V | \mathbf{k}, -\mathbf{k} \rangle$ depends only on

$\mathbf{k}' - \mathbf{k}$. Later we will see that indeed the interaction has a more complicated form. The cutting also will not have a serious effect since $\chi(\mathbf{k})$ will be small when $\epsilon_k \approx \epsilon_F + \epsilon_c$.] We see that inside of the layer the right hand side of the equation (87) is independent of \mathbf{k}' , and we mark it by I . From the left hand side we get easily

$$\chi(\mathbf{k}) = \frac{I}{2\epsilon_k - E} \text{ for } k > k_F. \quad (91)$$



Substituting back on the right hand side and cancelling the common factor I , we get

$$\frac{1}{g} = \frac{1}{L^3} \sum_{\epsilon_F < \epsilon_k < \epsilon_F + \epsilon_c} \frac{1}{2\epsilon_k - E}. \quad (92)$$

Applying the integration formula (76) we get

$$\frac{1}{N(0)g} = \int_{\epsilon_F}^{\epsilon_F + \epsilon_c} d\epsilon \frac{1}{2\epsilon - E} = \frac{1}{2} \ln \frac{2\epsilon_F - E + 2\epsilon_c}{2\epsilon_F - E}. \quad (93)$$

We solve

$$\begin{aligned} E &= 2\epsilon_F + \frac{2\epsilon_c}{1 - e^{2/gN(0)}} \\ &= 2\epsilon_F - 2\epsilon_c e^{-2/gN(0)}, \end{aligned} \quad (94)$$

where we have made *weak coupling approximation* $gN(0) \ll 1$ and used $g > 0$ on the second line. We see that the energy is lower than the energy $2\epsilon_F$ of a noninteracting pair when $g > 0$ (90).

It is interesting to note that this *bound state is formed even for arbitrarily weak attractive interaction*. [When there are no other particles, the attractive interaction for forming a bound state must exceed some threshold value. This corresponds to the case $k_F = N(0) = 0$ (77).]

The wave function (91) of the pair is independent of the direction \mathbf{k} of the wave vector. [Then, according to equation (86), the wave function in the \mathbf{r} space is independent of the direction of $\mathbf{r} - \mathbf{r}'$.] Thus the pair is formed in s-wave state (not p, d, f ...). The s state is symmetric in the exchange of coordinates \mathbf{r} and \mathbf{r}' . In order the total wave function to be antisymmetric, the spin state must be antisymmetric, and thus singlet:

$$\phi^{\text{tot}}(\mathbf{r}\sigma, \mathbf{r}'\sigma') = \phi(\mathbf{r} - \mathbf{r}') \frac{1}{\sqrt{2}} [\delta_{\sigma, \frac{1}{2}} \delta_{\sigma', -\frac{1}{2}} - \delta_{\sigma, -\frac{1}{2}} \delta_{\sigma', \frac{1}{2}}]. \quad (95)$$

As a function of the magnitude of the wave vector k , the wave function has its maximum at the Fermi surface, and decreases when k moves away from it. It is thus likely that the cut off of the interaction (90) is not essential.

We see that an attractive interaction makes a Fermi gas unstable against formation of pairs at $T = 0$.

4.4 Attractive interaction

Between electrons there is the repulsive Coulomb interaction

$$V(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \frac{e^2}{r}. \quad (96)$$

What could cause an attraction so that pairs could be formed?

Warning: in the following we only attempt to make it plausible that an attractive interaction might exist.

The first observation is that a metal has a great number of conduction electrons and ions. (By an ion we mean the nucleus and the valence electrons of the atom.) The whole system is charge neutral because of the ions. When we look at the interaction between electrons, we also should take into account also other electrons. These tend to go away from a negative charge, and thus the electron is surrounded by a net positive charge from the ions. This is called *screening* of the electron charge. In the simplest case this leads to a potential (λ is constant)

$$V(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \frac{e^2 \exp(-\lambda r)}{r}, \quad (97)$$

where the potential decays exponentially at large distances. A more accurate calculation gives also an oscillating component to the potential. However, the attraction caused by this overscreening is too weak to cause formation of pairs in usual superconductors.

It is found experimentally that the superfluid state depends on the mass of the ion (isotope effect). The ion mass can appear because the ion lattice is not at rest but oscillates. In the following we study what effect the ion motion has on the interaction between electrons.

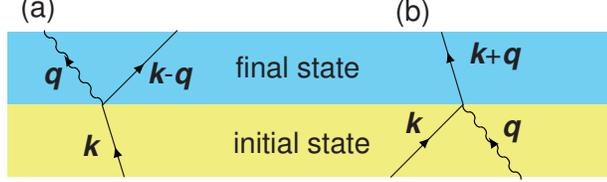
Lattice vibrations are studied in more detail in the course of condensed matter physics. Here we only state that the eigenstates of the lattice are vibrations that have a well defined wave vector \mathbf{k} and a frequency $\omega(\mathbf{k})$, which depends on the wave vector. According to quantum mechanics, the vibrations are quantized, i.e. they consists of *phonons* that have energy $\hbar\omega(\mathbf{k})$ and momentum $\hbar\mathbf{k}$. This is analogous to the photons, which are oscillation quanta of the electromagnetic field.

The coupling of lattice vibrations with the electrons is described as follows. An electron with momentum $\hbar\mathbf{k}$ emits a phonon with momentum $\hbar\mathbf{q}$, and the momentum remaining for the electron is $\hbar\mathbf{k}'$. This process has the

matrix element

$$\langle \psi_m^{(0)} | H_1 | \psi_n^{(0)} \rangle \propto \int dV e^{i\mathbf{k}' \cdot \mathbf{r}} e^{-i\mathbf{q} \cdot \mathbf{r}} e^{-i\mathbf{k} \cdot \mathbf{r}}. \quad (98)$$

It follows that the momentum is conserved, $\mathbf{k}' = \mathbf{k} - \mathbf{q}$. The process is often depicted by the graphs (a) in the figure.



Correspondingly, there must exist a process (b), where the electron absorbs a phonon.

We study the effect of lattice vibrations on electrons using perturbation theory. We recall time independent perturbation theory (Quantum mechanics I, nondegenerate case), which gives

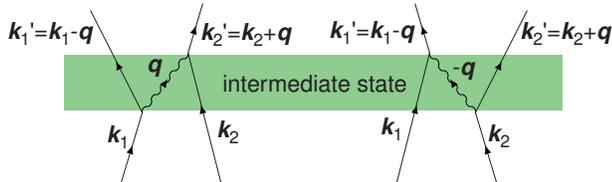
$$E_n = E_n^{(0)} + \langle \psi_n^{(0)} | H_1 | \psi_n^{(0)} \rangle + \sum_{\alpha \neq n} \frac{|\langle \psi_\alpha^{(0)} | H_1 | \psi_n^{(0)} \rangle|^2}{E_n^{(0)} - E_\alpha^{(0)}} + O(H_1^3). \quad (99)$$

In a similar way but with slightly more work (Landau-Lifshitz, Quantum mechanics) one can show that the effective interaction between two degenerate states ψ_m and ψ_n ($E_n^{(0)} = E_m^{(0)}$) is

$$\begin{aligned} \langle \psi_m | V^{\text{eff}} | \psi_n \rangle &= \langle \psi_m^{(0)} | H_1 | \psi_n^{(0)} \rangle \\ &+ \sum_{\alpha} \frac{\langle \psi_m^{(0)} | H_1 | \psi_\alpha^{(0)} \rangle \langle \psi_\alpha^{(0)} | H_1 | \psi_n^{(0)} \rangle}{E_m^{(0)} - E_\alpha^{(0)}} \\ &+ O(H_1^3), \end{aligned} \quad (100)$$

where α goes through all nondegenerate states.

We estimate the effect of lattice vibrations on the interaction between two electrons by calculating the effective matrix element (100) in the case where the initial state ψ_n has two electrons at wave vectors \mathbf{k}_1 and \mathbf{k}_2 and the final state ψ_m at wave vectors \mathbf{k}'_1 and \mathbf{k}'_2 . The first order term vanishes. The intermediate state in second order $\psi_\alpha^{(0)}$ has two alternatives



Assuming the matrix elements are constants, we get the corresponding terms

$$\begin{aligned} &\langle \mathbf{k}_1 - \mathbf{q}, \mathbf{k}_2 + \mathbf{q} | V^{\text{eff}} | \mathbf{k}_1, \mathbf{k}_2 \rangle \\ &= \sum_{\mathbf{q}} \frac{|V_{\mathbf{q}}|^2}{\epsilon(\mathbf{k}_1) - \epsilon(\mathbf{k}_1 - \mathbf{q}) - \hbar\omega(\mathbf{q})} \end{aligned}$$

$$+ \sum_{\mathbf{q}} \frac{|V_{\mathbf{q}}|^2}{\epsilon(\mathbf{k}_2) - \epsilon(\mathbf{k}_2 + \mathbf{q}) - \hbar\omega(-\mathbf{q})} + O(H_1^3).$$

Now we use energy conservation

$\epsilon(\mathbf{k}_1) + \epsilon(\mathbf{k}_2) = \epsilon(\mathbf{k}'_1) + \epsilon(\mathbf{k}'_2)$ and because $\omega(-\mathbf{q}) = \omega(\mathbf{q})$, we get

$$\begin{aligned} &\langle \mathbf{k}_1 - \mathbf{q}, \mathbf{k}_2 + \mathbf{q} | V^{\text{eff}} | \mathbf{k}_1, \mathbf{k}_2 \rangle \\ &= - \sum_{\mathbf{q}} \frac{2\hbar\omega(\mathbf{q})|V_{\mathbf{q}}|^2}{[\hbar\omega(\mathbf{q})]^2 - [\epsilon(\mathbf{k}_1) - \epsilon(\mathbf{k}'_1)]^2} \\ &+ O(H_1^3) \end{aligned} \quad (101)$$

In the Cooper problem the total momentum of the pair was assumed zero and therefore the used potential $\langle \mathbf{k}', -\mathbf{k}' | V | \mathbf{k}, -\mathbf{k} \rangle$ is obtained as the special case $\mathbf{k}_2 = -\mathbf{k}_1$.

We see that the interaction (101) is attractive when $\hbar\omega(\mathbf{q}) > |\epsilon(\mathbf{k}_1) - \epsilon(\mathbf{k}'_1)|$. The former quantity is typically on the order of the Debye temperature ~ 100 K, which means that this inequality possibly is satisfied. Ultimately, the formation of pairs depends if this attractive force is sufficiently strong that it wins the repulsive Coulomb force. With some more work one could show that these two forces have the same order of magnitude. Therefore, the formation of Cooper pairs and thus superconductivity depends on the detailed structure of each metal.

4.5 Creation and annihilation operators

This part is well presented in FW pages 3-19.

Above we already studied many-body states a bit. Before we continue, it is good to introduce a new notation, which is commonly used to describe many-body systems. This involves the introduction of the so-called *creation and annihilation operators*. These are also referred to as the operators of *second quantization* (we talk about the names later).

The basic idea, which appears also more generally, is the following. Originally we have some notation, where there are extra (unphysical) degrees of freedom. Then it makes sense to take into use a notation which automatically excludes the unphysical degrees of freedom.

Examples:

- moving from coordinates x, y and z to a vector \mathbf{r} , which is not dependent on the choice of the coordinate system
- going over from time t and position \mathbf{r} to the four-space x^α in relativity theory

The case under study here is a many-body system. As a starting point we have a complete set of levels $\phi_i(\mathbf{r}, \sigma)$ [e.g. (34)], $i = 1, \dots, \infty$. From these we form simple

n -particle states as a the product

$$\Psi(\mathbf{r}_1\sigma_1, \dots, \mathbf{r}_n\sigma_n) = \phi_{i_1}(\mathbf{r}_1\sigma_1)\phi_{i_2}(\mathbf{r}_2\sigma_2) \dots \phi_{i_n}(\mathbf{r}_n\sigma_n) \quad (102)$$

Based on experimental evidence it has been judged that all such states are not found in nature, but that physical states must additionally satisfy a symmetry with regard to exchange of indistinguishable particles:

$$\begin{aligned} & \Psi(\dots, \mathbf{r}_i\sigma_i, \mathbf{r}_{i+1}\sigma_{i+1}, \dots, \mathbf{r}_k\sigma_k, \dots) \\ & = \pm \Psi(\dots, \mathbf{r}_k\sigma_k, \mathbf{r}_{i+1}\sigma_{i+1}, \dots, \mathbf{r}_i\sigma_i, \dots), \end{aligned} \quad (103)$$

where $+$ is for bosons and $-$ is for fermions. We can make states of the type (102) physical by symmetrizing (bosons) or antisymmetrizing (fermions) them. A general many-body state is a linear combination of (anti)symmetrized states.

Writing the wave function in terms of coordinates is unsatisfying because it does not directly forbid writing down un-(anti)symmetrized functions. We can deal with this as already discussed earlier (39). Let us write all levels (34) in some arbitrary order $\phi_1, \phi_2, \dots, \phi_\infty$. Then we express the (noninteracting) many-body state by writing how many particles there are in each level:

$$|n_1, n_2, \dots, n_\infty\rangle, \quad (104)$$

($n_i = 0$ or 1 for fermions; $n_i = 0, 1, \dots, \infty$ for bosons). The basis states obtained in this way can be used to express an arbitrary (interacting) many-body state in the form

$$|\Psi(t)\rangle = \sum_{n_1, n_2, \dots, n_\infty} c_{n_1, n_2, \dots, n_\infty}(t) |n_1, n_2, \dots, n_\infty\rangle. \quad (105)$$

Bosons

Although the (anti)symmetrization requirement (103) was automatically satisfied with the notation (104), dealing with these states is still difficult due to the long list of numbers. To get rid of this, let us first study purely mathematically the operators \check{b}_k and \check{b}_k^\dagger , where the latter is the Hermitian conjugate of the former with respect to some inner-product space. Let us require these operators to satisfy the following commutators ($[\check{A}, \check{B}] \equiv \check{A}\check{B} - \check{B}\check{A}$)

$$\begin{aligned} [\check{b}_k, \check{b}_{k'}^\dagger] &= \delta_{k, k'}, \\ [\check{b}_k, \check{b}_{k'}] &= [\check{b}_k^\dagger, \check{b}_{k'}^\dagger] = 0 \end{aligned} \quad (106)$$

We claim that these conditions (alone) are yield the following relations

$$\begin{aligned} \check{b}_k^\dagger \check{b}_k | \dots, n_k, \dots \rangle &= n_k | \dots, n_k, \dots \rangle, \quad n_k = 0, 1, 2, \dots, \infty \\ \check{b}_k | \dots, n_k, \dots \rangle &= \sqrt{n_k} | \dots, n_k - 1, \dots \rangle \\ \check{b}_k^\dagger | \dots, n_k, \dots \rangle &= \sqrt{n_k + 1} | \dots, n_k + 1, \dots \rangle, \end{aligned} \quad (107)$$

where the states are of the form (104). Interpretation: $\check{b}_k^\dagger \check{b}_k$ is a particle-number operator, \check{b}_k is an annihilation operator, and \check{b}_k^\dagger a creation operator for bosons.

In the literature no hat seems to be written on the \check{b}_k operators. Here it has been added so that all operators of second quantization would be denoted uniformly.

Proof. (read at home) We study the operator

$$\check{n} = \check{b}^\dagger \check{b}. \quad (108)$$

It is Hermitian (verify). The eigenvalues of a Hermitian operator are real valued. Let us label the eigenstates of the operator \check{n} by using the eigenvalue n (a real number)

$$\check{n}|n\rangle = n|n\rangle. \quad (109)$$

We assume the eigenstates are normalized, $\langle n|n\rangle = 1$. We see that the eigenvalue n cannot be negative,

$$\begin{aligned} n &= \langle n|\check{n}|n\rangle = \langle n|\check{b}^\dagger \check{b}|n\rangle = \sum_m \langle n|\check{b}^\dagger|m\rangle \langle m|\check{b}|n\rangle \\ &= \sum_m |\langle m|\check{b}|n\rangle|^2 \geq 0. \end{aligned} \quad (110)$$

We easily calculate

$$[\check{n}, \check{b}] = -\check{b}. \quad (111)$$

This implies

$$\check{n}(\check{b}|n\rangle) = \check{b}\check{n}|n\rangle - \check{b}|n\rangle = (n-1)(\check{b}|n\rangle), \quad (112)$$

and we see that $\check{b}|n\rangle$ either is a state corresponding to the eigenvalue $n-1$, i.e. $\check{b}|n\rangle = c|n-1\rangle$, or else $\check{b}|n\rangle = 0$. The latter alternative implies $\check{n}|n\rangle = 0$, so it is possible only if $n=0$. In the former case normalization gives

$$|c|^2 = \langle n|\check{b}^\dagger \check{b}|n\rangle = n, \quad (113)$$

and therefore we fix

$$\check{b}|n\rangle = \sqrt{n}|n-1\rangle. \quad (114)$$

If one operates sufficiently many times with \check{b} , one should arrive at negative eigenvalues, which is in contradiction with equation (110). The way out of this is that n is an integer, so that $\check{b}|0\rangle = 0$, and the process (114) ends.

(Note the essential difference between the $n=0$ eigenstate $|0\rangle$ and the zero of the linear space 0.) Correspondingly one can deduce results for \check{b}^\dagger . Inserting the indices k we get all formulas (107). ■

The many-body Schrödinger equation in coordinate representation is (we forget spin)

$$i\hbar \frac{\partial \Psi}{\partial t}(\mathbf{r}_1, \dots, \mathbf{r}_n, t) = H\Psi(\mathbf{r}_1, \dots, \mathbf{r}_n, t), \quad (115)$$

where

$$\begin{aligned} H &= \sum_i T(\mathbf{r}_i) + \frac{1}{2} \sum_i \sum_{\substack{j \\ j \neq i}} V(\mathbf{r}_i, \mathbf{r}_j) \\ &= - \sum_i \frac{\hbar^2}{2m} \nabla_i^2 + \frac{1}{2} \sum_i \sum_{\substack{j \\ j \neq i}} V(\mathbf{r}_i, \mathbf{r}_j). \end{aligned} \quad (116)$$

How to write this for states (105)? That is, when

$$i\hbar \frac{\partial}{\partial t} |\Psi(t)\rangle = \check{H} |\Psi(t)\rangle, \quad (117)$$

what is \check{H} ?

The answer:

$$\check{H} = \sum_{i,j} \check{b}_i^\dagger \langle i|T|j\rangle \check{b}_j + \frac{1}{2} \sum_{ijkl} \check{b}_i^\dagger \check{b}_j^\dagger \langle i,j|V|k,l\rangle \check{b}_l \check{b}_k. \quad (118)$$

Here $\langle i|T|j\rangle$ and $\langle i,j|U|k,l\rangle$ are familiar matrix elements calculated for levels

$$\langle i|T|j\rangle = \int d^3r \phi_i^*(\mathbf{r}) \left(-\frac{\hbar^2}{2m} \nabla^2 \right) \phi_j(\mathbf{r}) \quad (119)$$

$$\langle i,j|V|k,l\rangle \quad (120)$$

$$= \int d^3r_1 \int d^3r_2 \phi_i^*(\mathbf{r}_1) \phi_j^*(\mathbf{r}_2) V(\mathbf{r}_1, \mathbf{r}_2) \phi_k(\mathbf{r}_1) \phi_l(\mathbf{r}_2).$$

Note in particular that $\langle i,j|V|k,l\rangle$ is calculated so that one electron scatters from level k to level i and the other from level l to level j . Also note that (88) used above is a special case of the formula (120), where plane waves $\phi_{\mathbf{k}}(\mathbf{r}) = (1/L^{3/2})e^{i\mathbf{k}\cdot\mathbf{r}}$ (34) have been used. Using these levels the kinetic energy matrix element is

$$\begin{aligned} \langle \mathbf{k}'|T|\mathbf{k}\rangle &= \frac{1}{L^3} \int d^3r e^{-i\mathbf{k}'\cdot\mathbf{r}} \left(-\frac{\hbar^2}{2m} \nabla^2 \right) e^{i\mathbf{k}\cdot\mathbf{r}} \\ &= \delta_{\mathbf{k}',\mathbf{k}} \frac{\hbar^2 k^2}{2m}. \end{aligned} \quad (121)$$

Instead of deriving (118) generally (see e.g. FW) we only note that is sensible in certain cases. 1) Normally $\langle i|T|j\rangle$ diagonal $\Rightarrow \check{H}_{\text{kin}} = \sum_i \check{n}_i \langle i|T|i\rangle$. 2) Interaction term \check{V} does not count interaction of a particle with itself. 3) In a two-particle state $\langle \check{V} \rangle$ gives both the direct interaction, and a term where the particles have been exchanged with each other. More in the exercises.

Fermions

The development in the case of fermions differs in that we require the anticommutators $\{\check{A}, \check{B}\} \equiv \check{A}\check{B} + \check{B}\check{A}$ and change the letter just to be sure

$$\begin{aligned} \{\check{a}_k, \check{a}_{k'}^\dagger\} &= \delta_{k,k'}, \\ \{\check{a}_k, \check{a}_{k'}\} &= \{\check{a}_k^\dagger, \check{a}_{k'}^\dagger\} = 0. \end{aligned} \quad (122)$$

It is now claimed that these conditions alone yield the following relations (exercise)

$$\begin{aligned} \check{a}_k^\dagger \check{a}_k | \dots, n_k, \dots \rangle &= n_k | \dots, n_k, \dots \rangle, \quad n_k = 0, 1 \\ \check{a}_k | \dots, 0_k, \dots \rangle &= 0 \\ \check{a}_k | \dots, 1_k, \dots \rangle &= | \dots, 0_k, \dots \rangle \\ \check{a}_k^\dagger | \dots, 0_k, \dots \rangle &= | \dots, 1_k, \dots \rangle \\ \check{a}_k^\dagger | \dots, 1_k, \dots \rangle &= 0, \end{aligned} \quad (123)$$

where 0_k and 1_k mean numbers 0 and 1 in the argument k of state (104). Interpretation: $\check{a}_k^\dagger \check{a}_k$ is a particle-number operator, \check{a}_k an annihilation operator, and \check{a}_k^\dagger a creation operator for fermions.

The Hamilton operator in second quantization

$$\check{H} = \sum_{i,j} \check{a}_i^\dagger \langle i|T|j\rangle \check{a}_j + \frac{1}{2} \sum_{ijkl} \check{a}_i^\dagger \check{a}_j^\dagger \langle i,j|V|k,l\rangle \check{a}_l \check{a}_k. \quad (124)$$

is exactly of the same form as for bosons. Now in particular the order of the operators $\check{a}_l \check{a}_k$ is essential, because an opposite order would change the sign. Exercises.

The interaction potential is assumed to depend only on the distances between particles: $V(\mathbf{r}, \mathbf{r}') = V(\mathbf{r} - \mathbf{r}')$. In general the interaction potential could also depend on the spins of the particles. However, let us study only spin-independent interactions, i.e. the interaction cannot change the spins of the particles, and the value of the potential does not depend on the spins.

Exercise: calculate the matrix elements (119) and (120) using the plane wave levels (36)-(37) and then derive the second-quantized Hamilton operator

$$\begin{aligned} \check{H} &= \sum_{\mathbf{k},\sigma} \epsilon_{\mathbf{k}} \check{a}_{\mathbf{k}\sigma}^\dagger \check{a}_{\mathbf{k}\sigma} + \frac{1}{2L^3} \sum_{\mathbf{k}_1,\sigma} \sum_{\mathbf{k}_2,\lambda} \sum_{\mathbf{k}_3} \sum_{\mathbf{k}_4} V(\mathbf{k}_3 - \mathbf{k}_1) \\ &\times \delta_{\mathbf{k}_1+\mathbf{k}_2,\mathbf{k}_3+\mathbf{k}_4} \check{a}_{\mathbf{k}_3\sigma}^\dagger \check{a}_{\mathbf{k}_4\lambda}^\dagger \check{a}_{\mathbf{k}_2\lambda} \check{a}_{\mathbf{k}_1\sigma}, \end{aligned} \quad (125)$$

where spin indices σ and λ can have the values \uparrow and \downarrow . Note that due to the Kronecker delta the momentum is conserved also in interactions. Let us also remind that the Fourier transformation is defined as

$$V(\mathbf{k}) = \int d^3r V(\mathbf{r}) e^{-i\mathbf{k}\cdot\mathbf{r}}, \quad (126)$$

so that its inverse transformation is

$$V(\mathbf{r}) = \frac{1}{L^3} \sum_{\mathbf{k}} V(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}}. \quad (127)$$

As a special case let us inspect an interaction that can be described with a delta function

$$V(\mathbf{r}, \mathbf{r}') = -g\delta(\mathbf{r} - \mathbf{r}'). \quad (128)$$

For this we get (125) in the form

$$\begin{aligned} \check{H} &= \sum_{\mathbf{k},\sigma} \epsilon_{\mathbf{k}} \check{a}_{\mathbf{k}\sigma}^\dagger \check{a}_{\mathbf{k}\sigma} - \frac{g}{2L^3} \sum_{\mathbf{k}_1,\sigma} \sum_{\mathbf{k}_2,\lambda} \sum_{\mathbf{k}_3} \sum_{\mathbf{k}_4} \\ &\times \delta_{\mathbf{k}_1+\mathbf{k}_2,\mathbf{k}_3+\mathbf{k}_4} \check{a}_{\mathbf{k}_3\sigma}^\dagger \check{a}_{\mathbf{k}_4\lambda}^\dagger \check{a}_{\mathbf{k}_2\lambda} \check{a}_{\mathbf{k}_1\sigma}. \end{aligned} \quad (129)$$

With a change of variables $\mathbf{k}_1 \leftrightarrow \mathbf{k}_2$ and using (122), we can see that a nonzero interaction is obtained only in the case that the spins are opposite:

$$\begin{aligned} \check{H} &= \sum_{\mathbf{k},\sigma} \epsilon_{\mathbf{k}} \check{a}_{\mathbf{k}\sigma}^\dagger \check{a}_{\mathbf{k}\sigma} - \frac{g}{L^3} \sum_{\mathbf{k}_1} \sum_{\mathbf{k}_2} \sum_{\mathbf{k}_3} \sum_{\mathbf{k}_4} \\ &\times \delta_{\mathbf{k}_1+\mathbf{k}_2,\mathbf{k}_3+\mathbf{k}_4} \check{a}_{\mathbf{k}_3\uparrow}^\dagger \check{a}_{\mathbf{k}_4\downarrow}^\dagger \check{a}_{\mathbf{k}_2\downarrow} \check{a}_{\mathbf{k}_1\uparrow}. \end{aligned} \quad (130)$$

This can be understood so that two fermions whose spins are the same cannot coexist in the same place, so they cannot feel a delta-function interaction.

This concludes the introduction to second quantization. Here we still give a short account of the terminology. When we take as a starting point the many-body wave equation (103), (115), and (116), then the second quantization is just new way of denoting it. An alternative approach, used in quantum field theory and from which the name second quantization comes from, is that one makes the switch from single-particle quantum mechanics to many-body theory by postulating the commutators (106) [or for fermions the anticommutators (122)] and only after that deduces (if needed) the formulas (103), (115), and (116).

4.6 Noninteracting system

We study fermions that do not interact with each other. It is easiest to use grand canonical ensemble. The effective Hamiltonian operator (27) is

$$\check{K} \equiv \check{H} - \mu\check{N} = \sum_{\mathbf{k},\sigma} \xi_{\mathbf{k}} \check{a}_{\mathbf{k}\sigma}^\dagger \check{a}_{\mathbf{k}\sigma}, \quad (131)$$

where $\xi_{\mathbf{k}} = \epsilon_{\mathbf{k}} - \mu$. Its ground state (45) and the elementary excitations are described above. We see that in the notation of second quantization the ground state $|\mathbf{0}\rangle$ can be written in the form

$$|\mathbf{0}\rangle = \prod_{|\mathbf{k}| < k_F} \check{a}_{\mathbf{k}\uparrow}^\dagger \check{a}_{\mathbf{k}\downarrow}^\dagger |\text{vac}\rangle, \quad (132)$$

where $\prod_{|\mathbf{k}| < k_F}$ is a product over all wave vectors having $k < k_F$ and $|\text{vac}\rangle$ is the vacuum state $|0, 0, 0, \dots\rangle$.

In order to get clear connection to the excitation picture, we define new operators

$$\check{a}_{\mathbf{k}\sigma} = \begin{cases} \check{\gamma}_{-\mathbf{k}, -\sigma}^\dagger & \text{when } \xi_{\mathbf{k}} < 0 \\ \check{\gamma}_{\mathbf{k}\sigma} & \text{when } \xi_{\mathbf{k}} > 0. \end{cases} \quad (133)$$

Here $-\mathbf{k}$ denotes the opposite momentum to \mathbf{k} and correspondingly $-\sigma$ the opposite spin to σ . The operators $\check{\gamma}_{\mathbf{k}\sigma}$ ja $\check{\gamma}_{\mathbf{k}\sigma}^\dagger$ satisfy the same commutation relations (122) as $\check{a}_{\mathbf{k}\sigma}$ ja $\check{a}_{\mathbf{k}\sigma}^\dagger$. Therefore the transformation (133) is called canonical. The operators $\check{\gamma}_{\mathbf{k}\sigma}$ ja $\check{\gamma}_{\mathbf{k}\sigma}^\dagger$ can be interpreted as annihilation and creation operators of an *excitation*. Inside the Fermi sphere we use

$$\check{a}_{\mathbf{k}\sigma}^\dagger \check{a}_{\mathbf{k}\sigma} = 1 - \check{a}_{\mathbf{k}\sigma} \check{a}_{\mathbf{k}\sigma}^\dagger \quad (134)$$

and using the notation (133) we get

$$\check{K} = \sum_{\mathbf{k},\sigma} |\xi_{\mathbf{k}}| \check{\gamma}_{\mathbf{k}\sigma}^\dagger \check{\gamma}_{\mathbf{k}\sigma} + \Omega_0, \quad (135)$$

where the last term is a constant (the grand potential of the ground state). The ground state $|\mathbf{0}\rangle$ satisfies

$$\check{\gamma}_{\mathbf{k}\sigma} |\mathbf{0}\rangle = 0 \quad \forall \mathbf{k}, \sigma. \quad (136)$$

The number operator of the excitation $\mathbf{k}\sigma$ is $\check{\gamma}_{\mathbf{k}\sigma}^\dagger \check{\gamma}_{\mathbf{k}\sigma}$, and all excitation energies are positive.

4.7 Many body problem

We study an interacting system. We write the Hamiltonian presented above (124) using a shorter notation

$$\check{K} \equiv \check{H} - \mu\check{N} = \sum_{i,j} \xi_{ij} \check{a}_i^\dagger \check{a}_j + \frac{1}{2} \sum_{ijkl} V_{ijkl} \check{a}_i^\dagger \check{a}_j^\dagger \check{a}_l \check{a}_k. \quad (137)$$

where $\xi_{ij} = \langle i|T|j\rangle - \mu\delta_{ij}$ and $V_{ijkl} = \langle i,j|V|k,l\rangle$.

The essential problem is the interaction term in (137), which is fourth order in \check{a} . This is the *many body problem*. The noninteracting case (=ideal gas) can be calculated exactly, but there is no general method to solve exactly the problem of many interacting particles. Instead, there exists numerous approximation methods that can be applied in different cases.

It turns out that in case of superconductivity, there exist a quite good approximation method (known as quasiclassical theory of Fermi liquids). There the many body problem is solved using the fact that the ratio T_c/T_F is small. In general form this theory is very complicated, and therefore it is not presented in this course. Although we have to refrain from firm justification, we attempt to make the main results understandable in the following.

A quite general approximation method in many-body systems is *Hartree-Fock approximation*. The main idea is the following. Because the difficulties in the Hamiltonian (137) arise in the fourth order term, one approximates it with a second order term

$$\check{V} \approx \frac{1}{2} \sum_{ijkl} V_{ijkl} \check{a}_i^\dagger \check{a}_j^\dagger \check{a}_l \check{a}_k \approx \sum_{ij} A_{ij} \check{a}_i^\dagger \check{a}_j \quad (138)$$

Because this term has the same form as the noninteracting system, its solution is easy (at least relatively). Now one must determine the coefficients A_{ij} . In Hartree-Fock approximation this is done by replacing the removed operators by their expectation values. Thus

$$\check{V} \approx \check{V}_{\text{HF}} = \frac{1}{2} \sum_{ijkl} V_{ijkl} (\langle \check{a}_i^\dagger \check{a}_k \rangle \check{a}_j^\dagger \check{a}_l + \check{a}_i^\dagger \check{a}_k \langle \check{a}_j^\dagger \check{a}_l \rangle - \langle \check{a}_i^\dagger \check{a}_l \rangle \check{a}_j^\dagger \check{a}_k - \check{a}_i^\dagger \check{a}_l \langle \check{a}_j^\dagger \check{a}_k \rangle) + \text{constant}. \quad (139)$$

The first term can be understood so that a particle that scatters from level l to level j feels an interaction that is averaged over all states of other particles. The second term is of the same type but has different indices. The third and fourth terms are caused by the fact that it is impossible to distinguish the particles, but one must allow them to interchange (so called exchange interaction). [We have neglected in equation (139) correction terms that arise from the fact that a particle cannot interact with itself.]

The expectation value is calculated as above $\langle \dots \rangle = \text{Tr}(\dots \check{\rho})$ (9), but in order to get a closed theory, one must use the same approximation for \check{K} in the density matrix:

$$\langle \dots \rangle = \frac{\text{Tr}(\dots e^{-\beta \check{K}_{\text{HF}}})}{\text{Tr} e^{-\beta \check{K}_{\text{HF}}}}. \quad (140)$$

This is known as the self-consistency equation. It must be solved together with (139) because they depend on each other.

The Hartree-Fock approximation is widely used in calculation of the electronic states of atoms and molecules. For conduction electrons it is clearly insufficient already in the normal state (because it does not take into account the screening of the Coulomb potential). Also, it does not lead to superconductivity.

4.8 Superconducting state

We make so called *anomalous* Hartree-Fock approximation, where we also take terms of the type $\check{a}_i^\dagger \check{a}_j^\dagger \langle \check{a}_l \check{a}_k \rangle$. This is not included in the standard Hartree-Fock approximation because the expectation values of the operators, which change the particle number, vanish. It turns out, however, that just these terms are essential for superconductivity.

The anomalous HF approximation can also be justified by first writing exactly

$$\check{a}_i \check{a}_j = \langle \check{a}_i \check{a}_j \rangle + (\check{a}_i \check{a}_j - \langle \check{a}_i \check{a}_j \rangle) \quad (141)$$

and correspondingly for the Hermitian conjugate operator. Now we assume that the expectation value is a good approximation, i.e. the term in the parenthesis on the right hand side is small. Its order of magnitude is denoted by ϵ . We substitute in the Hamiltonian (137) and drop terms that are proportional to ϵ^2 , and get

$$\check{V}_{\text{anom}} = \frac{1}{2} \sum_{ijkl} V_{ijkl} (\langle \check{a}_i^\dagger \check{a}_j^\dagger \rangle \check{a}_l \check{a}_k + \check{a}_i^\dagger \check{a}_j^\dagger \langle \check{a}_l \check{a}_k \rangle - \langle \check{a}_i^\dagger \check{a}_j^\dagger \rangle \langle \check{a}_l \check{a}_k \rangle). \quad (142)$$

This was the most essential approximation. In addition, we make the following simplifying approximations.

- the interaction is approximated by a contact interaction (128).
- the normal Hartree-Fock terms (139) are dropped assuming that they shift equally the energies of the normal and superconducting states.
- we assume that momentum is conserved in expectation values (although the particle number is not). One can show that this limits the study to the case of stationary (not flowing) state of the superconductor.

With these assumptions we get the Hamiltonian

$$\begin{aligned} \check{K}_{\text{eff}} = & \sum_{\mathbf{k}, \sigma} \xi_{\mathbf{k}} \check{a}_{\mathbf{k}\sigma}^\dagger \check{a}_{\mathbf{k}\sigma} - \frac{g}{L^3} \sum_{\mathbf{k}} \sum_{\mathbf{k}'} \\ & \times (\check{a}_{\mathbf{k}\uparrow}^\dagger \check{a}_{-\mathbf{k}\downarrow}^\dagger \langle \check{a}_{-\mathbf{k}'\downarrow} \check{a}_{\mathbf{k}'\uparrow} \rangle + \langle \check{a}_{\mathbf{k}\uparrow}^\dagger \check{a}_{-\mathbf{k}\downarrow}^\dagger \rangle \check{a}_{-\mathbf{k}'\downarrow} \check{a}_{\mathbf{k}'\uparrow} \\ & - \langle \check{a}_{\mathbf{k}\uparrow}^\dagger \check{a}_{-\mathbf{k}\downarrow}^\dagger \rangle \langle \check{a}_{-\mathbf{k}'\downarrow} \check{a}_{\mathbf{k}'\uparrow} \rangle). \end{aligned} \quad (143)$$

We write this into the form

$$\begin{aligned} \check{K}_{\text{eff}} = & \sum_{\mathbf{k}, \sigma} \xi_{\mathbf{k}} \check{a}_{\mathbf{k}\sigma}^\dagger \check{a}_{\mathbf{k}\sigma} - \sum_{\mathbf{k}} (\Delta \check{a}_{\mathbf{k}\uparrow}^\dagger \check{a}_{-\mathbf{k}\downarrow}^\dagger \\ & + \Delta^* \check{a}_{-\mathbf{k}\downarrow} \check{a}_{\mathbf{k}\uparrow}) + C. \end{aligned} \quad (144)$$

Here we have defined

$$\Delta = \frac{g}{L^3} \sum_{\mathbf{k}} \langle \check{a}_{-\mathbf{k}\downarrow} \check{a}_{\mathbf{k}\uparrow} \rangle. \quad (145)$$

For its complex conjugate we get

$$\Delta^* = \frac{g}{L^3} \sum_{\mathbf{k}} \langle \check{a}_{\mathbf{k}\uparrow}^\dagger \check{a}_{-\mathbf{k}\downarrow}^\dagger \rangle. \quad (146)$$

The Hamiltonian (144) can also be written in the form

$$\begin{aligned} \check{K}_{\text{eff}} = & \sum_{\mathbf{k}} \begin{pmatrix} \check{a}_{\mathbf{k}\uparrow}^\dagger & \check{a}_{-\mathbf{k}\downarrow} \end{pmatrix} \begin{pmatrix} \xi_{\mathbf{k}} & -\Delta \\ -\Delta^* & -\xi_{\mathbf{k}} \end{pmatrix} \begin{pmatrix} \check{a}_{\mathbf{k}\uparrow} \\ \check{a}_{-\mathbf{k}\downarrow} \end{pmatrix} \\ & + C_2. \end{aligned} \quad (147)$$

The Hamiltonian (144) is second order in \check{a} , as was desired. However, it is not yet of the same form as for ideal gas (131) because in addition to diagonal terms $\propto \check{a}_{\mathbf{k}\sigma}^\dagger \check{a}_{\mathbf{k}\sigma}$ it contains nondiagonal terms. In order to achieve a diagonal form we make *Bogoliubov transformation*. We introduce the operators $\check{\gamma}_{\mathbf{k}\uparrow}$ ja $\check{\gamma}_{\mathbf{k}\downarrow}$ by defining

$$\begin{aligned} \check{a}_{\mathbf{k}\uparrow} &= u_{\mathbf{k}}^* \check{\gamma}_{\mathbf{k}\uparrow} + v_{\mathbf{k}} \check{\gamma}_{-\mathbf{k}\downarrow}^\dagger \\ \check{a}_{\mathbf{k}\downarrow} &= u_{\mathbf{k}}^* \check{\gamma}_{\mathbf{k}\downarrow} - v_{\mathbf{k}} \check{\gamma}_{-\mathbf{k}\uparrow}^\dagger. \end{aligned} \quad (148)$$

Here $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ are generally complex valued coefficients. For the Hermitian conjugates we get

$$\begin{aligned} \check{a}_{\mathbf{k}\uparrow}^\dagger &= u_{\mathbf{k}} \check{\gamma}_{\mathbf{k}\uparrow}^\dagger + v_{\mathbf{k}}^* \check{\gamma}_{-\mathbf{k}\downarrow} \\ \check{a}_{\mathbf{k}\downarrow}^\dagger &= u_{\mathbf{k}} \check{\gamma}_{\mathbf{k}\downarrow}^\dagger - v_{\mathbf{k}}^* \check{\gamma}_{-\mathbf{k}\uparrow}. \end{aligned} \quad (149)$$

We wish to make the transformation *canonical*, which means that the new operators should satisfy the same anticommutation relations as the original operators (122). We calculate

$$\begin{aligned} \{\check{a}_{\mathbf{k}\uparrow}, \check{a}_{\mathbf{k}\uparrow}^\dagger\} &= \\ |u_{\mathbf{k}}|^2 \{\check{\gamma}_{\mathbf{k}\uparrow}, \check{\gamma}_{\mathbf{k}\uparrow}^\dagger\} &+ |v_{\mathbf{k}}|^2 \{\check{\gamma}_{-\mathbf{k}\downarrow}, \check{\gamma}_{-\mathbf{k}\downarrow}^\dagger\} \\ + u_{\mathbf{k}}^* v_{\mathbf{k}} \{\check{\gamma}_{\mathbf{k}\uparrow}, \check{\gamma}_{-\mathbf{k}\downarrow}\} &+ u_{\mathbf{k}} v_{\mathbf{k}} \{\check{\gamma}_{-\mathbf{k}\downarrow}, \check{\gamma}_{\mathbf{k}\uparrow}^\dagger\} = 1. \end{aligned}$$

This succeeds if

$$|u_k|^2 + |v_k|^2 = 1. \quad (150)$$

Exercise: Show that the inverse of the transformation (148) is

$$\begin{aligned} \tilde{\gamma}_{\mathbf{k}\uparrow} &= u_k \check{a}_{\mathbf{k}\uparrow} - v_k \check{a}_{-\mathbf{k}\downarrow}^\dagger \\ \tilde{\gamma}_{\mathbf{k}\downarrow} &= u_k \check{a}_{\mathbf{k}\downarrow} + v_k \check{a}_{-\mathbf{k}\uparrow}^\dagger. \end{aligned} \quad (151)$$

Show that the $\tilde{\gamma}$ operators satisfy all the same anticommutation rules (122) as the \check{a} operators.

Now we substitute the Bogoliubov transformation to the Hamiltonian (144). We see that the inconvenient terms drop out if one chooses

$$2\xi_k u_k v_k - \Delta u_k^2 + \Delta^* v_k^2 = 0, \quad (152)$$

and the Hamiltonian gets the form

$$\begin{aligned} \check{K}_{\text{eff}} &= \sum_{\mathbf{k}} \{ \xi_k [|u_k|^2 (\tilde{\gamma}_{\mathbf{k}\uparrow}^\dagger \tilde{\gamma}_{\mathbf{k}\uparrow} + \tilde{\gamma}_{\mathbf{k}\downarrow}^\dagger \tilde{\gamma}_{\mathbf{k}\downarrow}) \\ &+ |v_k|^2 (\tilde{\gamma}_{\mathbf{k}\uparrow}^\dagger \tilde{\gamma}_{\mathbf{k}\uparrow} + \tilde{\gamma}_{\mathbf{k}\downarrow}^\dagger \tilde{\gamma}_{\mathbf{k}\downarrow})] + (\Delta^* u_k^* v_k + \Delta u_k v_k^*) \\ &\times (\tilde{\gamma}_{\mathbf{k}\uparrow}^\dagger \tilde{\gamma}_{\mathbf{k}\uparrow} - \tilde{\gamma}_{\mathbf{k}\downarrow}^\dagger \tilde{\gamma}_{\mathbf{k}\downarrow}) \} + C. \end{aligned} \quad (153)$$

Now we suppose, for simplicity, that Δ is real valued. (We shall return to the general case later.) The solution of equations (150) and (152) is

$$u_k = \sqrt{\frac{1}{2} \left(1 + \frac{\xi_k}{E_k} \right)}, \quad v_k = \sqrt{\frac{1}{2} \left(1 - \frac{\xi_k}{E_k} \right)}, \quad (154)$$

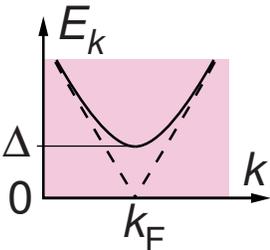
where

$$E_k = \sqrt{\xi_k^2 + \Delta^2}. \quad (155)$$

Substituting these into (153) gives the Hamiltonian in the desired form

$$\check{K}_{\text{eff}} = \sum_{\mathbf{k}, \sigma} E_k \tilde{\gamma}_{\mathbf{k}\sigma}^\dagger \tilde{\gamma}_{\mathbf{k}\sigma} + \Omega_0. \quad (156)$$

(We return to the constant Ω_0 later.) Because $E_k > 0$, it can be interpreted, similarly as in the normal state (135-136), as the excitation energy and $\tilde{\gamma}_{\mathbf{k}\sigma}^\dagger \tilde{\gamma}_{\mathbf{k}\sigma}$ as the number operator of excitations. The excitation energy (155) in the neighborhood of the Fermi surface is depicted in the figure. In the superconducting state the particle and hole type excitations change to each other smoothly as k changes. The excitations have a minimum energy Δ , which is known as energy gap. The excitation spectrum reduces to the one in normal state when $\Delta = 0$.



Gap equation

In order to fix the theory, we should determine Δ (145). Similarly as in the Hartree-Fock approximation we have to calculate the expectation value using the effective Hamiltonian (156). We get

$$\Delta = \frac{g}{L^3} \sum_{\mathbf{k}} \langle \check{a}_{-\mathbf{k}\downarrow} \check{a}_{\mathbf{k}\uparrow} \rangle = \frac{g}{L^3} \sum_{\mathbf{k}} \frac{\text{Tr}(\check{a}_{-\mathbf{k}\downarrow} \check{a}_{\mathbf{k}\uparrow} e^{-\beta \check{K}_{\text{eff}}})}{\text{Tr} e^{-\beta \check{K}_{\text{eff}}}}. \quad (157)$$

Using the transformation formulas (148) and expressing the trace (Tr) in eigenstates of the excitation number operators we get

$$\Delta = \frac{g}{L^3} \sum_{\mathbf{k}} u_k v_k [1 - 2n(E_k)], \quad (158)$$

where the Fermi distribution

$$n(\epsilon) = \frac{1}{e^{\beta\epsilon} + 1}. \quad (159)$$

Using still the expressions for u_k and v_k (154) and rearranging one gets the consistency equation in the form

$$\Delta = \frac{g}{2L^3} \sum_{\mathbf{k}} \frac{\Delta}{E_k} \tanh \frac{E_k}{2k_B T}. \quad (160)$$

This has at least the trivial solution $\Delta = 0$, which corresponds to the normal state. Any other solutions should satisfy

$$1 = \frac{g}{2L^3} \sum_{\mathbf{k}} \frac{1}{E_k} \tanh \frac{E_k}{2k_B T}. \quad (161)$$

We see that because of the factor $1/E_k$ the summand is largest at the Fermi surface. Indeed the contact interaction (128), whose Fourier transform is a constant, is too idealized, and at large momenta the Fourier transform should drop to zero. Similarly as in the Cooper problem, we cut off the summation at energy ϵ_c . Using the summation formula (76) we get

$$1 = \frac{gN(0)}{2} \int_{-\epsilon_c}^{\epsilon_c} d\xi \frac{1}{\sqrt{\xi^2 + \Delta^2}} \tanh \frac{\sqrt{\xi^2 + \Delta^2}}{2k_B T}. \quad (162)$$

Noticing the symmetry of the integrand we get it to the form

$$\frac{1}{gN(0)} = \int_0^{\epsilon_c} d\xi \frac{1}{\sqrt{\xi^2 + \Delta^2}} \tanh \frac{\sqrt{\xi^2 + \Delta^2}}{2k_B T}. \quad (163)$$

This *gap equation* determines the energy gap as a function of temperature, $\Delta(T)$. Generally it should be solved numerically. We study two limiting cases.

1) $T = 0$. One gets

$$\begin{aligned} \frac{1}{gN(0)} &= \int_0^{\epsilon_c} d\xi \frac{1}{\sqrt{\xi^2 + \Delta^2}} \\ &= \ln \frac{2\epsilon_c}{\Delta}, \end{aligned} \quad (164)$$

where the integral can be calculated with Mathematica, and the result is valid in the limit $\epsilon_c \gg \Delta$. We get

$$\Delta(T=0) = 2\epsilon_c e^{-\frac{1}{gN(0)}}. \quad (165)$$

2) $T = T_c$. Here the superconducting state vanishes, $\Delta \rightarrow 0$. When $\epsilon_c \gg k_B T_c$ we get

$$\begin{aligned} \frac{1}{gN(0)} &= \int_0^{\epsilon_c} d\xi \frac{1}{\xi} \tanh \frac{\xi}{2k_B T_c} \\ &= \ln \frac{2\epsilon_c e^\gamma}{\pi k_B T_c}, \end{aligned} \quad (166)$$

where the Euler constant $\gamma = 0.5772$, $e^\gamma = 1.781$. We get

$$k_B T_c = \frac{2e^\gamma}{\pi} \epsilon_c e^{-\frac{1}{gN(0)}}. \quad (167)$$

The result (167) in principle predicts the transition temperature of the superconductor. It contains the parameters $N(0)$, ϵ_c and g . Out of these $N(0)$ (77) can be measured independently and ϵ_c is estimated to be on same order of magnitude as the Debye temperature (~ 100 K). The greatest uncertainty appears in the constant g . Because T_c depends exponentially on g , it makes the calculation of T_c very uncertain. Therefore the equation (167) as such is not as remarkable as one could initially think.

Considerably more reliable result is obtained if we eliminate $\epsilon_c e^{-1/gN(0)}$ from equations (165) and (167):

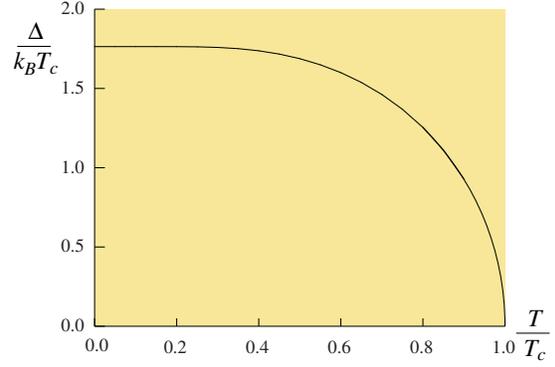
$$\Delta(T=0) = \pi e^{-\gamma} k_B T_c = 1.764 k_B T_c. \quad (168)$$

We see that the energy gap and the transition temperature are of the same order of magnitude, and also get a precise factor of proportionality between the two.

More generally, g and ϵ_c can be eliminated from the gap equation in the *weak coupling limit* $\epsilon_c \gg k_B T_c$. This can be accomplished by subtracting from the gap equation (163) its T_c condition (166). Utilizing the limit $\epsilon_c \gg k_B T_c$ the resulting equation can be written in the form (exercise)

$$\begin{aligned} &\int_0^\infty \left[\frac{\tanh(\xi/2k_B T)}{\xi} - \frac{\tanh(\sqrt{\xi^2 + \Delta^2}/2k_B T)}{\sqrt{\xi^2 + \Delta^2}} \right] d\xi \\ &= \ln \frac{T_c}{T}. \end{aligned} \quad (169)$$

We see that $\Delta(T)/k_B T_c$ is a universal (= independent of the material) function of T/T_c . The only material dependent parameter is thus T_c .



In the neighborhood of T_c one can derive ($\zeta(3) = 1.202$)

$$\Delta(T) \approx k_B T_c \pi \sqrt{\frac{8}{7\zeta(3)}} \sqrt{1 - \frac{T}{T_c}}. \quad (170)$$

Let us study how well the previous assumptions are satisfied. The table shows experimental values except that $gN(0)$ is calculated from equation (167). We see that for elemental metals $\epsilon_c \gg k_B T_c$ holds reasonably. Equation (168) is rather well satisfied. The greatest deviation appears in Lead. The deviations can be more or less understood using strong coupling theory, which takes more accurately into account e.g. the energy dependence of the phonon mediated interaction.

	T_c (K)	T_D (K)	$gN(0)$	$\frac{\Delta(0)}{k_B T_c}$
BCS				1.764
Cd	0.56	164	0.18	1.6
Al	1.2	375	0.18	1.3-2.1
Sn	3.75	195	0.25	1.6
Pb	7.22	96	0.39	2.2

The quasiclassical theory of Fermi liquids is based on the following idea. There are a few microscopic parameters like T_c , effective mass etc., whose values cannot be calculated in this theory, but their values can be determined experimentally. Once the parameters are known, the quasiclassical theory can be used to calculate several properties of the superconducting state, e.g. the gap function.

BCS ground state

In their original work Bardeen, Cooper and Schrieffer presented the ground state of the superconductor

$$|\psi_0\rangle = \prod_{\mathbf{k}} (u_{\mathbf{k}} + v_{\mathbf{k}} \tilde{a}_{\mathbf{k}\uparrow}^\dagger \tilde{a}_{-\mathbf{k}\downarrow}^\dagger) |\text{vac}\rangle, \quad (171)$$

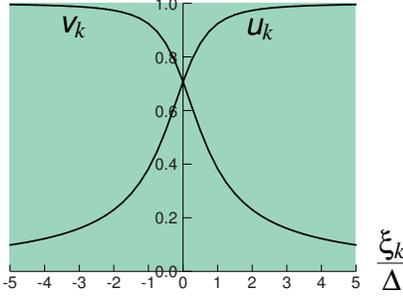
where $\prod_{\mathbf{k}}$ is the product over all wave vectors and $|\text{vac}\rangle$ is the vacuum state. As the first step show that expression (171) does not depend on the order in which the elements of the product are written. This follows from the commutator

$$[u_{\mathbf{k}} + v_{\mathbf{k}} \tilde{a}_{\mathbf{k}\uparrow}^\dagger \tilde{a}_{-\mathbf{k}\downarrow}^\dagger, u_{\mathbf{q}} + v_{\mathbf{q}} \tilde{a}_{\mathbf{q}\uparrow}^\dagger \tilde{a}_{-\mathbf{q}\downarrow}^\dagger] = 0. \quad (172)$$

Next show as an exercise that the BCS state (171) is consistent with the theory above (156) by showing that $\gamma_{\mathbf{k}\sigma}$ (151) annihilates this state,

$$\tilde{\gamma}_{\mathbf{k}\sigma}|\psi_0\rangle = 0. \quad (173)$$

The functions u_k and v_k (154) near the Fermi surface are shown in the figure.



The state (171) shows that the occupations of the levels $\mathbf{k}\uparrow$ and $-\mathbf{k}\downarrow$ are correlated. If a particle is present in level $\mathbf{k}\uparrow$, it is also present in level $-\mathbf{k}\downarrow$, and if not, also not in the other level. For ideal gas ground state (132) this is satisfied trivially because the occupation number is always either 0 or 1, depending on k . The BCS state differs from the normal state at those k for which $u_k v_k \neq 0$. This takes place around the Fermi surface in the energy range of a few Δ , or, equivalently, of a few $k_B T_c$.

One can also show that the state (171) is the one consisting of pairs, as was expected above using the wave function presentation (83) (see de Gennes).

The size of a pair

In order to see the general picture, it is important to study the size of one pair. This can be justified in more detail, for example by studying the wave function obtained from the Cooper problem (see Ketterson-Song), but the following gives the same result more directly. The superconducting state differs from the normal one in the energy range $\delta\xi \sim k_B T_c$. By calculating

$$\delta\xi = \delta\epsilon = \delta\left(\frac{p^2}{2m}\right) = \left(\frac{p}{m}\right)\delta p \approx v_F \delta p \quad (174)$$

we get that in momentum this corresponds to a shell of thickness $\delta p = \hbar\delta k \sim k_B T_c / v_F$ around the Fermi surface. It is a general property of wave motion that the minimum size δx for a localized wave packet satisfies $\delta x \delta k \sim 1$ when the available wave vectors are in range δk . Applying this to the superconducting state we get that the minimum size of a Cooper pair has the magnitude

$$\xi_0 = \frac{\hbar v_F}{2\pi k_B T_c}, \quad (175)$$

which is called *coherence length*. (Because it is an order of magnitude estimate, the factor 2π can be added without

justification.) The table shows calculated values for a few metals.

	ξ_0 (nm)
Al	1500
Sn	480
Pb	160
Nb	14

The essential result is that ξ_0 is much larger than the atomic scale ~ 0.1 nm. The pairs are so huge that within each pair there are $\sim 10^{10}$ conduction electrons! The situation is quite opposite, for example, to O atoms forming O_2 molecules in oxygen gas.

Because the pairs are strongly overlapping, the identity of electrons and the Pauli exclusion principle are quite essential in the theory of superconductivity. For example, one cannot answer the question which two electrons form the pair. Therefore the BCS ground state (171), which only indicates pair correlations, is much more useful way of presenting than the wave function presentation (83) where the effect of antisymmetrization is essential.

The size of a pair compared to the atomic scale $a \sim k_F^{-1}$ is equivalent with the previously mentioned fact that $k_B T_c \ll \epsilon_F$:

$$\frac{\xi_0}{a} \sim \frac{\epsilon_F}{k_B T_c} \gg 1. \quad (176)$$

Superconductivity is a relatively low energy phenomenon and therefore the associated length scale is large.

Particle number and phase

We attempt to understand that the number of particles in the BCS state is not fixed.

Exercise: suppose that Δ defined in equation (145) is not real valued but

$$\Delta = e^{i\phi} |\Delta|, \quad (177)$$

where ϕ is a real *phase*. Show that all the preceding results are valid also in this case when equations (154) and (155) are replaced by

$$u_k^2 = \frac{1}{2} \left(1 + \frac{\xi_k}{E_k}\right), \quad v_k^2 = \frac{1}{2} e^{2i\phi} \left(1 - \frac{\xi_k}{E_k}\right), \quad (178)$$

$$E_k = \sqrt{\xi_k^2 + |\Delta|^2}. \quad (179)$$

and in equations (162)-(170) one replaces $\Delta \rightarrow |\Delta|$. ■

We conclude that the BCS state (171) is degenerate, i.e. corresponding to each value of ϕ there is a state with the same energy, where the original v_k (154) is multiplied by a phase factor $\exp(i\phi)$:

$$|\psi_\phi\rangle = \prod_{\mathbf{k}} (u_k + v_k e^{i\phi} \tilde{a}_{\mathbf{k}\uparrow}^\dagger \tilde{a}_{-\mathbf{k}\downarrow}^\dagger) |\text{vac}\rangle. \quad (180)$$

Let us consider more generally the relationship between phase and pair number. We define the pair number

operator \tilde{n} . Its eigenstates are $|n\rangle$,

$$\tilde{n}|n\rangle = n|n\rangle. \quad (181)$$

[For mathematical completeness we consider the integer n in the range $(-\infty, +\infty)$ but in practice a much smaller range is sufficient.] In addition we define the operator

$$\tilde{T} = \sum_n |n-1\rangle\langle n|. \quad (182)$$

Operated on an arbitrary pair state it removes one pair out of it. We look for the eigenstates of \tilde{T} :

$$\tilde{T}|t\rangle = t|t\rangle. \quad (183)$$

In order to achieve this we write

$$|t\rangle = \sum_n f_n |n\rangle. \quad (184)$$

Now the condition (183) gives $f_{n+1} = t f_n$, whose solution is $f_n = C \exp(in\phi)$ and the eigenvalue $t = \exp(i\phi)$. Thus corresponding to each eigenstate of the operator \tilde{T} there is a corresponding value of ϕ . We use this value to label the eigenstates and therefore write instead of (183)

$$\tilde{T}|\phi\rangle = e^{i\phi}|\phi\rangle. \quad (185)$$

This eigenstate of the phase can be presented using pair number eigenstates

$$|\phi\rangle = \frac{1}{\sqrt{2\pi}} \sum_n e^{in\phi} |n\rangle. \quad (186)$$

A calculation gives the inverse relation

$$|n\rangle = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} d\phi e^{-in\phi} |\phi\rangle. \quad (187)$$

We see that the BCS state (180) has the form

$$|\psi_\phi\rangle = \sum_n e^{in\phi} f_n |n\rangle. \quad (188)$$

Here f_n is real and it differs from zero in a range of n 's of width $\Delta n \gg 1$. Here the phase is rather well defined but the pair number is uncertain. If one wants a state with fixed number of pairs, it can be formed as linear combination of phase eigenstates, as in equation (187). Then the phase is fully uncertain. More generally one can construct an uncertainty relation

$$\Delta n \Delta \phi \gtrsim 1. \quad (189)$$

Justification: We consider an arbitrary state $|\psi\rangle$. Its representation in the pair number eigenbasis is $\psi_n = \langle n|\psi\rangle$ and in phase eigenbasis $\psi(\phi) = \langle \phi|\psi\rangle$. From equation (187) we get

$$\psi_n = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} d\phi e^{in\phi} \psi(\phi). \quad (190)$$

This shows that these representations are constructed from each other by Fourier transformation.

The relation (190) is similar as between the momentum and coordinate representations

$$\psi_p = \frac{1}{\sqrt{L}} \int_{-L/2}^{L/2} dx e^{-ipx/\hbar} \psi(x) \quad (191)$$

for a particle in a one-dimensional box of width L . Based on this one can derive the Heisenberg uncertainty relation $\Delta p \Delta x \geq \hbar/2$ in the limit $L \rightarrow \infty$. The case of (190) is slightly different because ϕ is a periodic variable and thus n is discrete. Anyway, (189) can be derived principally in the same way. Here we do it only approximately. Let $\psi(\phi)$ be different from zero when $|\phi| < \Delta\phi/2$. In order for ψ_n to change from its maximum value, n has to change so that the exponent in equation (190) changes essentially, and thus $\Delta n \Delta\phi \sim 1$. ■

As a conclusion we realize that the anomalous Hartree-Fock approximation leads to BCS states where the phase is well defined but the pair number is uncertain. As a linear combination of such states it is possible to form states of fixed number of pairs.

4.9 Thermodynamics

We follow the principle discussed in the beginning of the course that we first calculate the thermodynamic potential, and from it we get all thermodynamic quantities as derivatives.

To simplify the notation, we assume Δ real. We first calculate the constant appearing in the diagonalized \tilde{K} (156) (exercise), and get

$$\Omega_0 = 2 \sum_{\mathbf{k}} (\xi_{\mathbf{k}} v_{\mathbf{k}}^2 - \Delta u_{\mathbf{k}} v_{\mathbf{k}}) + \frac{L^3}{g} \Delta^2. \quad (192)$$

Because \tilde{K} (156) is diagonal, the grand potential (30) can be calculated in the same way as for ideal Fermi gas and we get

$$\Omega = \Omega_0 - \frac{2}{\beta} \sum_{\mathbf{k}} \ln(1 + e^{-\beta E_{\mathbf{k}}}). \quad (193)$$

We see that at zero temperature $\Omega(T=0) = \Omega_0$.

The energy functional (193) has a couple of interesting properties.

- 1) If one minimizes Ω_0 with respect to $v_{\mathbf{k}}$ [taking into account (150), $\Delta = \text{constant}$] one gets the same condition (152) that was derived above in another way.
- 2) In the reminder of statistical physics we stated that Ω has a minimum with respect to internal degrees of freedom. In particular, this should apply to Δ . Verify as an exercise that the condition

$$\frac{\partial \Omega(T, V, \mu, \Delta)}{\partial \Delta} = 0 \quad (194)$$

is equivalent with the gap equation (158).

Based on the energy functional (193) one can calculate all thermodynamic quantities. Here we calculate the difference in energy between the superconducting and the normal state at zero temperature:

$$\begin{aligned} \Omega_0 - \Omega_0(\Delta = 0) &= 2 \sum_{\mathbf{k}} (\xi_k v_k^2 - \Delta u_k v_k) + \frac{L^3}{g} \Delta^2 \\ &\quad - 2 \sum_{\mathbf{k}} \xi_k \frac{1}{2} \left(1 - \frac{\xi_k}{|\xi_k|}\right). \end{aligned} \quad (195)$$

In calculating the integrals one can proceed as follows. Because of the gap equation (158) the term $L^3 \Delta^2 / g$ cancels half of the term $-2 \sum_{\mathbf{k}} \Delta u_k v_k$. Using expressions (154) and (155) we get the summation in the form

$$\Omega_0 - \Omega_0(\Delta = 0) = -\frac{1}{2} \sum_{\mathbf{k}} \frac{\Delta^4}{E_k (E_k + |\xi_k|)^2}. \quad (196)$$

We see that this converges well for large ξ_k . Therefore we can use formula (76). We get

$$\Omega_0 - \Omega_0(\Delta = 0) = -\frac{1}{2} L^3 N(0) \Delta^2. \quad (197)$$

This is the condensation energy of the superconducting state, whose existence we deduced above by studying thermodynamics in magnetic field (61). We see that the energy of the superconducting state is lower than that of the normal state. The energy difference (197) can be roughly understood that the energy of those levels, which are around the Fermi surface in a shell of energy Δ , is reduced by Δ .

The entropy can be calculated from the formula (33) and the specific heat from formula (18). In calculating the entropy one should in principle differentiate with respect to all temperature dependent parameters, but because of relation (194) the temperature dependence of Δ does not contribute to the final result. Show as exercise that the specific heat is given by

$$C = \frac{L^3 N(0)}{2k_B T^2} \int_{-\infty}^{\infty} d\xi \frac{1}{\cosh^2 \frac{\sqrt{\xi^2 + \Delta^2}}{2k_B T}} \left(\xi^2 + \Delta^2 - T \Delta \frac{d\Delta}{dT} \right). \quad (198)$$

Using numerical calculation one could show that this gives similar curve as plotted on page 3.

Exercise. Show from formula (198) that the specific heat of the normal state is given by

$$C = \frac{2\pi^2}{3} L^3 N(0) k_B^2 T, \quad (199)$$

which is linear in T . From this one can determine $N(0)$ experimentally.

Note. Because we used grand canonical ensemble (28) the specific heat (198) is calculated at constant μ . In the courses of condensed matter physics and statistical

physics it is shown that in the case studied ($T \ll T_F$) the specific heat at constant volume is the same.

4.10 Inhomogeneous superconductor

Previously in equation (144) we assumed that only those expectation values conserving the momentum were nonzero. We start to generalize the calculation to the case that this assumption is not made. The effective Hamiltonian can be written in the form

$$\begin{aligned} \tilde{K}_{\text{eff}} &= \sum_{\mathbf{k}, \sigma} \xi_k \tilde{a}_{\mathbf{k}\sigma}^\dagger \tilde{a}_{\mathbf{k}\sigma} - \frac{1}{L^3} \sum_{\mathbf{k}'} \sum_{\mathbf{q}} \\ &\times (\tilde{a}_{\frac{1}{2}\mathbf{q}+\mathbf{k}'\uparrow}^\dagger \tilde{a}_{\frac{1}{2}\mathbf{q}-\mathbf{k}'\downarrow}^\dagger \Delta(\mathbf{q}) + \Delta^*(\mathbf{q}) \tilde{a}_{\frac{1}{2}\mathbf{q}-\mathbf{k}'\downarrow} \tilde{a}_{\frac{1}{2}\mathbf{q}+\mathbf{k}'\uparrow}) \\ &\quad + C, \end{aligned} \quad (200)$$

where

$$\Delta(\mathbf{q}) = g \sum_{\mathbf{k}} \langle \tilde{a}_{\frac{1}{2}\mathbf{q}-\mathbf{k}\downarrow} \tilde{a}_{\frac{1}{2}\mathbf{q}+\mathbf{k}\uparrow} \rangle. \quad (201)$$

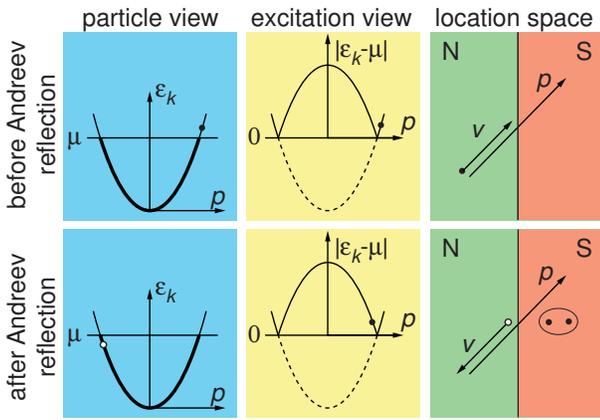
We define the inverse Fourier transform of $\Delta(\mathbf{q})$ in the usual way

$$\begin{aligned} \Delta(\mathbf{r}) &= \frac{1}{L^3} \sum_{\mathbf{q}} e^{i\mathbf{q}\cdot\mathbf{r}} \Delta(\mathbf{q}) \\ &= \frac{g}{L^3} \sum_{\mathbf{k}} \sum_{\mathbf{q}} e^{i\mathbf{q}\cdot\mathbf{r}} \langle \tilde{a}_{\frac{1}{2}\mathbf{q}-\mathbf{k}\downarrow} \tilde{a}_{\frac{1}{2}\mathbf{q}+\mathbf{k}\uparrow} \rangle. \end{aligned} \quad (202)$$

We realize that $\hbar\mathbf{q}$ the total momentum of a pair. Thus we can deduce that $\Delta(\mathbf{r})$ can be interpreted as the wave function describing the center of mass of the pair. The function $\Delta(\mathbf{r})$ is called *order parameter*. If the momentum in the expectation value (202) is conserved, $\Delta(\mathbf{r})$ reduces to the same constant as above (145).

We realize that this generalization is needed in the case where the superfluid is in motion or in other inhomogeneous state. We will not continue the analysis in general form any further, but we will later use another approach to study this very important case.

An interesting process taking place in inhomogeneous superconductors is *Andreev reflection*. Consider the interface of a normal state (N) and a superconducting state (S) of a metal. Consider a particle type excitation on the normal side approaching the interface. Assume that the excitation energy is smaller than the energy gap on the superconducting side, $\epsilon_k - \mu < \Delta$. In this case the excitation cannot continue to the superconducting side since there are no states at the same energy. What can happen is that the excitation forms a Cooper pair in the superconductor. But a Cooper pair needs two electrons. This means that a hole-type excitation is created at the interface and its velocity \mathbf{v} is opposite to that of the initial particle excitation. The extra Cooper pair is depicted by a circled pair of dots in the figure.



superconductors the pairs are found to form in the d wave state $v_{\mathbf{k}} \propto k_x^2 - k_y^2$. The superfluid states of ^3He are studied a lot because there one can avoid such complications occurring in metals such as the ion lattice (which causes non spherically symmetric Fermi surface) and impurities.

Energy and momentum are conserved in the process. The momentum change of the excitation is small, $\Delta p \ll p_F$, in contrast to usual reflection (from a potential barrier or impurities), where the momentum change is on the order of p_F . Note that the reflection direction differs from specular (=mirror like) reflection, where only the normal component of the velocity changes sign. To emphasize this, Andreev reflection is called retroreflection. Andreev reflection appears predominantly when the contact between the metals is good. The opposite case of a weakly transmitting contact is discussed later in Sec. 6.

4.11 Superfluid ^3He

As an exercise we found out that ^3He atom is a fermion and its Fermi temperature $T_F \sim 1$ K. It is found experimentally that liquid ^3He has a transition to superfluid state at temperatures $T_c = 1 \dots 2.5$ mK. This superfluid state can to a large extent be understood similarly as superconductivity above.

An essential difference is that in ^3He the pairs form in a p wave state instead of the s wave state found in most superconductors. This means that v_k has to be replaced by $v_{\mathbf{k}}$, which also depends on the direction of \mathbf{k} according to a combination of spherical harmonic functions $Y_1^m(\theta, \phi)$ [thus $\ell = 1$ ja $m = 0, \pm 1$]. Simultaneously the spin state has to be written more generally. The BCS state (171) thus has the form

$$|\psi_0\rangle = \prod_{\mathbf{k}} \left[u_{\mathbf{k}} + \sum_{\sigma, \sigma'} v_{\mathbf{k}\sigma\sigma'} \tilde{a}_{\mathbf{k}\sigma}^\dagger \tilde{a}_{-\mathbf{k}\sigma'}^\dagger \right] |\text{vac}\rangle. \quad (203)$$

Because $v_{-\mathbf{k}\sigma\sigma'} = -v_{\mathbf{k}\sigma\sigma'}$ we see that only the spin symmetric part $v_{\mathbf{k}\sigma\sigma'} + v_{\mathbf{k}\sigma'\sigma}$ gives something nonzero (exercise). Thus the spin state is triplet $[\uparrow\uparrow, \downarrow\downarrow$ or $(\uparrow\downarrow + \downarrow\uparrow)/\sqrt{2}]$ instead of the singlet in superconductors (95). It follows that the order parameter of ^3He is a 3×3 matrix, where the indices refer to three p wave states and to three spin triplet states. j

Also some superconductors show properties from which one can infer an order parameter consisting of more than one component (e.g. UPt_3). In high temperature

5. Ginzburg-Landau theory

5.1 Introduction

This part is well presented in FW pages 430-439

Similarly as Ginzburg and Landau (GL), we derive the GL theory phenomenologically. We will discuss later, how it can be derived from microscopic theory.

Ginzburg and Landau presumed that the superconducting state is described by a complex-valued *order parameter* Ψ . This parameter is assumed to be different from zero only in the superconducting state. Further it is assumed that Ψ is small near the transition temperature. We suppose that near the transition temperature the free energy density $f = F/L^3$ can be written as Taylor series in Ψ and Ψ^* ,

$$f = f_0 + \alpha|\Psi|^2 + \frac{1}{2}\beta|\Psi|^4 + \dots \quad (204)$$

The terms appearing here are restricted by the fact that F_s has to be real valued for arbitrary complex valued Ψ . Therefore, the term $c\Psi$ cannot appear. Instead, Ψ has to appear in product with Ψ^* : $\Psi^*\Psi = |\Psi|^2$. Also the term $c\text{Re}\Psi$ is not accepted. The reason is that we require F_s to remain unchanged in the transformation $\Psi \rightarrow e^{i\phi}\Psi$, where ϕ is a real-valued constant.

The expansion (204) is incomplete because nothing in it prevents a spatial dependence $\Psi(\mathbf{r})$. Such a spatial dependence can be limited by adding a term $|\nabla\Psi|^2$ that increases the energy of inhomogeneous states. However, also this is unsatisfactory in the case of a nonzero magnetic field. The magnetic field \mathbf{B} can be described with a vector potential \mathbf{A} :

$$\mathbf{B} = \nabla \times \mathbf{A}. \quad (205)$$

From the course of analytical mechanics we know that the real momentum $m\mathbf{v} = \mathbf{p} - q\mathbf{A}$. Here \mathbf{p} is a canonical momentum, which in quantum mechanics is replaced by the operator $\frac{\hbar}{i}\nabla$. Analogously to this, GL chose the additional energy term to be

$$\gamma \left| \left(\frac{\hbar}{i}\nabla - q\mathbf{A} \right) \Psi \right|^2, \quad (206)$$

where q is some charge. Let us take into account also the energy density of the magnetic field in the sample

$$\frac{1}{2\mu_0} B^2. \quad (207)$$

In this way we obtain the total energy in Ginzburg-Landau theory as

$$\begin{aligned} F &= F_0 + \int d^3r f, \\ f &= \alpha|\Psi|^2 + \frac{1}{2}\beta|\Psi|^4 + \gamma \left| \left(\frac{\hbar}{i}\nabla - q\mathbf{A} \right) \Psi \right|^2 \\ &\quad + \frac{1}{2\mu_0} B^2. \end{aligned} \quad (208)$$

Often one wants to study a system in a given external magnetic field. Then, instead of F , one should minimize G (55), in this case

$$G = F - \int d^3r \mathbf{H} \cdot \mathbf{B}. \quad (209)$$

Let us still write the G in GL theory in its full form

$$\begin{aligned} G &= F_0 + \int d^3r g, \\ g &= \alpha|\Psi|^2 + \frac{1}{2}\beta|\Psi|^4 + \gamma \left| \left(\frac{\hbar}{i}\nabla - q\mathbf{A} \right) \Psi \right|^2 \\ &\quad + \frac{1}{2\mu_0} B^2 - \mathbf{B} \cdot \mathbf{H}. \end{aligned} \quad (210)$$

GL differential equations

Earlier we derived the result that in equilibrium the free energy must be minimized. In a given external field one must thus minimize G (210) both with respect to Ψ and to \mathbf{A} . {In minimizing with respect to Ψ the independent variables [e.g. ($\text{Re}\Psi, \text{Im}\Psi$) or ($|\Psi|, \arg\Psi$)] can be chosen arbitrarily. The shortest calculation follows by treating Ψ and Ψ^* as independent variables.} Let us leave the minimization as an exercise. As a result we obtain the GL differential equations

$$\gamma \left(\frac{\hbar}{i}\nabla - q\mathbf{A} \right)^2 \Psi + \alpha\Psi + \beta|\Psi|^2\Psi = 0, \quad (211)$$

$$\begin{aligned} \frac{1}{\mu_0} \nabla \times \mathbf{B} &= \frac{q\hbar\gamma}{i} (\Psi^* \nabla \Psi - \Psi \nabla \Psi^*) \\ &\quad - 2q^2\gamma|\Psi|^2 \mathbf{A}. \end{aligned} \quad (212)$$

The surface terms arising from integration by parts must also vanish. From this we get the boundary conditions at the surface of a superconductor

$$\hat{\mathbf{n}} \cdot \left(\frac{\hbar}{i}\nabla - q\mathbf{A} \right) \Psi = 0, \quad (213)$$

$$\hat{\mathbf{n}} \times (\mathbf{B} - \mu_0\mathbf{H}) = 0. \quad (214)$$

[It is noted in passing that the transformation (209) is essential only for the surface terms.]

It is noted that based on the Maxwell equation

$$\nabla \times \mathbf{B} = \epsilon_0\mu_0 \frac{\partial \mathbf{E}}{\partial t} + \mu_0 \mathbf{j} \quad (215)$$

we identify the quantity appearing in (212) as an electric current density

$$\mathbf{j} = \frac{q\hbar\gamma}{i} (\Psi^* \nabla \Psi - \Psi \nabla \Psi^*) - 2q^2\gamma|\Psi|^2 \mathbf{A}. \quad (216)$$

Exercise: show that the equations guarantee current conservation

$$\nabla \cdot \mathbf{j} = 0, \quad \hat{\mathbf{n}} \cdot \mathbf{j} = 0. \quad (217)$$

5.2 Special cases

The GL equations (211) and (212) constitute a coupled set of differential equations, whose solution gives $\Psi(\mathbf{r})$ and $\mathbf{A}(\mathbf{r})$. In the general case this is very complicated. Let us start by considering simple special cases.

1) Homogeneous superconductor, $\mathbf{H} = \mathbf{A} = 0$. Equation (211) gives as possible solutions

$$\Psi = 0, \quad (218)$$

$$|\Psi|^2 = -\frac{\alpha}{\beta}. \quad (219)$$

The former solution describes normal state. The latter, superconducting state, is possible only if $\alpha/\beta < 0$. In order for F (208) to be sensible (minimum energy must be achieved with a finite Ψ) we must always have $\beta > 0$. The condition for the latter state is therefore $\alpha < 0$. The energies corresponding to the states (218) and (219) are found by inserting into the functional (208):

$$F = F_0, \quad (220)$$

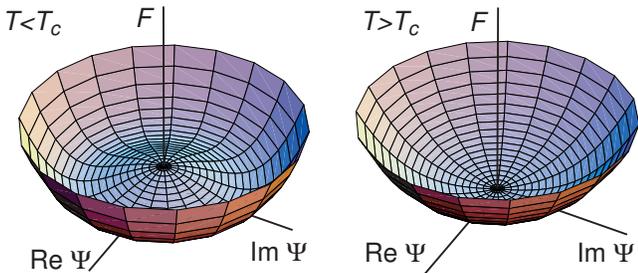
$$F = F_0 - V \frac{\alpha^2}{2\beta}. \quad (221)$$

We thus see that if $\alpha > 0$, only the normal state is possible, while in the case $\alpha < 0$ the superconducting state has the lowest energy. The transition temperature T_c thus corresponds to the point where $\alpha = 0$. In GL theory it is assumed that the temperature dependence of the coefficient α is linear

$$\alpha(T) = \alpha' \left(\frac{T}{T_c} - 1 \right), \quad (222)$$

and the other coefficients (β , γ , q) are temperature-independent.

The dependence of the free energy on the order parameter can be illustrated with the following pictures.



It is observed that in the normal state $\Psi = 0$ is completely determined, but in the superconducting state only the absolute value $|\Psi|$ of the order parameter is fixed while the phase $\arg \Psi$ is arbitrary.

2) Changing $|\Psi|$, $\mathbf{A} = 0$. From the boundary condition (213) it follows that $\hat{\mathbf{n}} \cdot \nabla \Psi = 0$. Thus a position-independent $|\Psi|$ (219) is a valid solution everywhere in the superconductor, also close to boundaries. Despite this we consider a case where Ψ

deviates from its equilibrium value. From equation (211) we find

$$\hbar^2 \gamma \nabla^2 \Psi - \alpha \Psi - \beta |\Psi|^2 \Psi = 0. \quad (223)$$

Assuming Ψ to be real and writing $\Psi = \sqrt{|\alpha|/\beta} f$ we put this in the form

$$\xi_{GL}^2 \nabla^2 f + f - f^3 = 0, \quad (224)$$

where we have defined the *GL coherence length*

$$\xi_{GL} = \sqrt{\frac{\hbar^2 \gamma}{|\alpha|}}. \quad (225)$$

We see that ξ_{GL} determines that length scale on which Ψ can vary essentially. As an example we give the solution of equation (224) in the case of a one-dimensional dependence:

$$f(x) = \tanh \frac{x}{\sqrt{2} \xi_{GL}}. \quad (226)$$

3) Let us investigate the case

$$\Psi(\mathbf{r}) = e^{i\phi(\mathbf{r})} |\Psi| \quad (227)$$

where $|\Psi|^2 \approx |\alpha|/\beta$ is constant. By inserting into the expression of current (216) we find

$$\mathbf{j} = 2q\gamma |\Psi|^2 (\hbar \nabla \phi - q \mathbf{A}). \quad (228)$$

By taking the rotor of this we have the *London equation*

$$\nabla \times \mathbf{j} = -2q^2 \gamma |\Psi|^2 \mathbf{B}. \quad (229)$$

By using the Maxwell equations (215) and (51) this yields

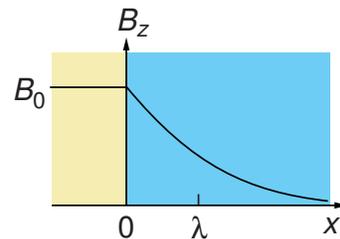
$$\begin{aligned} \mathbf{B} &= -\frac{\nabla \times \mathbf{j}}{2q^2 \gamma |\Psi|^2} = -\frac{\nabla \times (\nabla \times \mathbf{B})}{2\mu_0 q^2 \gamma |\Psi|^2} \\ &= \frac{\nabla^2 \mathbf{B}}{2\mu_0 q^2 \gamma |\Psi|^2}. \end{aligned} \quad (230)$$

Thus we have

$$\mathbf{B} = \lambda^2 \nabla^2 \mathbf{B}, \quad (231)$$

where

$$\lambda = \sqrt{\frac{\beta}{2\mu_0 q^2 \gamma |\alpha|}}. \quad (232)$$



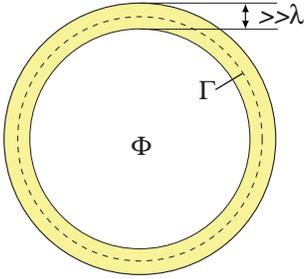
Now we study a superconducting half-space $x > 0$. Let us assume that outside the superconductor ($x < 0$) there is a

field $\mathbf{B}_0 = B_0 \hat{\mathbf{z}}$ parallel to the surface of the superconductor. The solution of equation (231) inside the superconductor is

$$B_z(x) = B_0 e^{-x/\lambda}. \quad (233)$$

Thus we explain the Meissner effect: the magnetic field does not penetrate into the superconductor, apart from a layer with thickness on the order of λ . In this layer a current is flowing (calculate it), which cancels the external field inside the superconductor. The result thus also implies the existence of a dissipationless current. Below some experimental values for the penetration dept are given.

	$\lambda(T \ll T_c)$ (nm)
Al	49
Sn	51
Pb	39



4) We look at the superconducting state in a ring (loop, torus), whose cross-section is considerably larger than the penetration depth. Then inside the ring $\mathbf{j} = 0$. From (228) we find that for a path Γ going around inside the ring we have

$$0 = \oint d\mathbf{l} \cdot (\hbar \nabla \phi - q \mathbf{A}) = \hbar 2\pi N - q \int d\mathbf{a} \cdot \nabla \times \mathbf{A}. \quad (234)$$

Here N is an integer, which follows from that the fact that a unique single-valued Ψ (227) only allows for ϕ to change by a multiple of 2π when going around Γ . Thus for the magnetic flux threading the loop we find

$$\Phi = \int d\mathbf{a} \cdot \mathbf{B} = N \frac{2\pi \hbar}{q}. \quad (235)$$

It has been experimentally observed that the flux is quantized according to this formula. From the magnitude from the observed flux quantum, $\Phi_0 = \frac{h}{2|e|}$ (4), we deduce that q is twice the charge e of an electron (the sign of the charge cannot be deduced from this).

5) Above we have defined two length: the GL coherence length ξ_{GL} (225) and the penetration length λ (232). Both have the temperature dependence

$$\lambda(T), \xi_{GL}(T) \propto \frac{1}{\sqrt{|\alpha|}} \propto \frac{1}{\sqrt{1 - T/T_c}}, \quad (236)$$

so that they diverge when $T \rightarrow T_c$. The ratio of the

lengths is called the *GL parameter*

$$\kappa = \frac{\lambda(T)}{\xi_{GL}(T)} = \sqrt{\frac{\beta}{2\mu_0 q^2 \hbar^2 \gamma^2}}. \quad (237)$$

It is a temperature-independent constant. By writing the GL equations in a dimensionless form we observe that this is the only dimensionless parameter in the theory.

6) The equilibrium between normal and superconducting states in an external field was studied already in the beginning of the course, but it is instructive to see the same by starting from the GL functional (210). In the superconducting state we obtain from the terms $\alpha|\Psi|^2 + \frac{1}{2}\beta|\Psi|^4$ a negative contribution that was calculated above (221). This is independent of the field H , because $B \equiv 0$. In the normal state only the terms $\frac{1}{2\mu_0}B^2 - \mathbf{B} \cdot \mathbf{H}$ are nonzero. By minimizing G with respect to \mathbf{B} we find

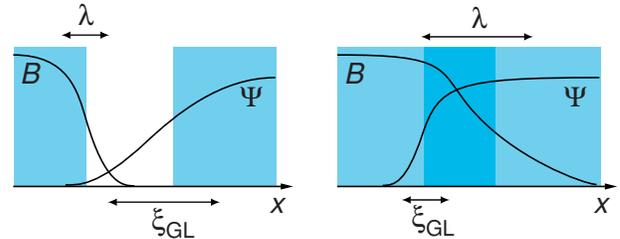
$$G = F_0 - \frac{1}{2}V\mu_0 H^2, \quad \mathbf{B} = \mu_0 \mathbf{H}. \quad (238)$$

In order for the energies to be equal when $H = H_c$ we have

$$\frac{1}{2}\mu_0 H_c^2 = \frac{\alpha^2}{2\beta}. \quad (239)$$

From this and the linearity (222) of $\alpha(T)$ it is concluded that $H_c(T)$ is linear close to T_c . This is consistent with the experimental observation (3).

7) Next we investigate the interface between normal and superconducting states. This requires the field H to be of the critical magnitude H_c , because the interface can only be stable if the two phases are in equilibrium. The structure of the interface can be solved exactly from the GL equations, be here we are satisfied by a qualitative analysis. In the previous point we identified the essential terms of normal and superconducting states in the functional (210). Let us see how these are involved in the interface.



In the figure a light shading roughly describes the regions where the energy is lowered from F_0 by the amount (239).

The situation in an S-N interface depends essentially on the ratio of the penetration depth and the GL coherence length. If $\lambda \ll \xi_{GL}$ a region of thickness $\approx \xi_{GL}$ is formed, where neither of the negative contributions is reached. This means an interface energy $\sigma \approx \frac{1}{2}\xi_{GL}\mu_0 H_c^2$ [compare to equation (68)]. In the opposite case $\lambda \gg \xi_{GL}$, both negative contributions are present within a thickness $\approx \lambda$

and we find a negative interface energy $\sigma \approx -\frac{1}{2}\lambda\mu_0 H_c^2$. This latter case leads to completely new types of properties. The description of the intermediate state given in the beginning of the course is clearly not valid in this case.

A superconductor where the interface energy is negative is called a type II superconductor, as opposed to the type I superconductor that has a positive interface energy. By solving the GL equations we find that the limit between the two cases goes at the value $\kappa = 1/\sqrt{2}$ of the GL parameter. Thus for a type I superconductor $\kappa < 1/\sqrt{2}$ and for a type II superconductor $\kappa > 1/\sqrt{2}$.

5.3 Derivation from microscopic theory

The Ginzburg-Landau theory can be derived starting from microscopic theory. The correspondence to microscopic theory is achieved when one identifies

$$\Psi(\mathbf{r}) = \Delta(\mathbf{r}), \quad (240)$$

where Δ is defined by equation (202). [Note that relation (240) can contain an arbitrary constant factor of proportionality, and it is often also used.]

We state the conditions under which the general theory reduces to the GL theory:

- 1) The temperature is near the transition temperature, $T_c - T \ll T_c \Leftrightarrow \Delta \ll k_B T_c$.
- 2) The order parameter is not changing too steeply, $|\nabla\Delta| \ll \Delta/\xi_0$.

The Ginzburg-Landau theory can be understood as Taylor expansion in both Δ and $\nabla\Delta$ where one keeps only the lowest order terms.

From microscopic theory one can derive the following expressions for the parameters of the GL theory

$$\alpha = N(0) \frac{T - T_c}{T_c} \quad (241)$$

$$\beta = \frac{7\zeta(3)N(0)}{8(\pi k_B T_c)^2} \quad (242)$$

$$\gamma = \frac{7\zeta(3)N(0)}{12\hbar^2} \xi_0^2 \quad (243)$$

$$q = 2e, \quad (244)$$

where $\zeta(3) = 1.202$, ξ_0 is defined in equation (175) and e is the electron charge ($e < 0$). The last relation can be easily understood: because Ψ the wave function of a pair, its kinetic energy contains the pair charge $q = 2e$.

The two first equations [(241) and (242)] can be derived directly from the energy functional (193). The calculation is somewhat complicated though and therefore is not done here. The two other equations [(243) and (244)] can be derived from the inhomogeneous state theory mentioned above (200) when also the vector potential is

included in the kinetic energy (119):

$$-\frac{\hbar^2}{2m} \nabla^2 \rightarrow \frac{1}{2m} \left(\frac{\hbar}{i} \nabla - e\mathbf{A} \right)^2. \quad (245)$$

From coefficients (241) and (243) we get for the GL coherence length (225)

$$\begin{aligned} \xi_{GL}(T) &= \sqrt{\frac{\hbar^2 \gamma}{|\alpha|}} = \xi_0 \sqrt{\frac{7\zeta(3)}{12}} \frac{1}{\sqrt{1 - T/T_c}} \\ &= 0.837 \xi_0 \frac{1}{\sqrt{1 - T/T_c}}. \end{aligned} \quad (246)$$

The coherence lengths are thus of same order of magnitude except the case $T \rightarrow T_c$, where $\xi_{GL}(T) \rightarrow \infty$.

It turns out that when one goes very near T_c , the GL theory is not valid any more. This is caused by *critical fluctuations*, that are common to all second order phase transitions. These are discussed more in the course of statistical physics. In ordinary superconductors the temperature region where critical fluctuations are important is vanishingly small.

Effect of impurities

Metals always have impurities. It was discussed above that these scatter the conduction electrons and thus cause the electrical resistance in the normal state. What happens to the superconducting state when impurities are present? Are the pairs broken?

The effect of impurities can be studied theoretically by adding an external potential $U(\mathbf{r})$ besides the kinetic energy (119),

$$-\frac{\hbar^2}{2m} \nabla^2 \rightarrow -\frac{\hbar^2}{2m} \nabla^2 + U(\mathbf{r}). \quad (247)$$

In the simplest case one can study a delta-function potential $U(\mathbf{r}) = u\delta(\mathbf{r} - \mathbf{r}_0)$, where \mathbf{r}_0 is the location of the impurity. [This scatters electrons from one plane wave level (34) to another but takes up no volume.] The interesting result of this calculation is that out of the GL coefficients (241)-(244) only γ (243) changes. The fact that T_c is unchanged could be understood by saying that although electrons scatter from one level to another, they always can find a new partner to form a pair.

The previous does not hold in “unusual” superfluid states where $v_{\mathbf{k}}$ depends on direction, i.e. in high temperature superconductors or in ^3He , and pairs are broken there.

What happens to γ ? Let us study the case where the density of impurities is so large that the mean free path ℓ of a particle between scattering events is much smaller than ξ_0 . This case $\ell \ll \xi_0$ is called the “dirty limit”. The electron propagates randomly as its direction changes after an average flight by ℓ . A simple calculation gives

that if the total distance travelled by the the particle is ξ_0 , it is at the distance $R \approx \sqrt{\ell \xi_0}$ from its starting point:

$$R^2 = \left(\sum_i \Delta \mathbf{x}_i \right)^2 = \sum_i (\Delta \mathbf{x}_i)^2 + \sum_i \sum_{\substack{j \\ j \neq i}} \Delta \mathbf{x}_i \cdot \Delta \mathbf{x}_j$$

$$\approx \sum_i (\Delta \mathbf{x}_i)^2 \approx \frac{\xi_0}{\ell} \ell^2 = \xi_0 \ell. \quad (248)$$

This is because on the average $\Delta \mathbf{x}_i \cdot \Delta \mathbf{x}_j = 0$ when $j \neq i$, and ξ_0/ℓ is the number of terms in the last summation

Supposing now that the estimate about the pair size (175) concerns the total path length, we arrive at the result that in the dirty limit the pair size is reduced to $\sim \sqrt{\ell \xi_0} \ll \xi_0$. This could be described by saying that in dirty case the members of the pair lose each other slower than in the pure case.

For the GL parameter γ instead of (243) we get

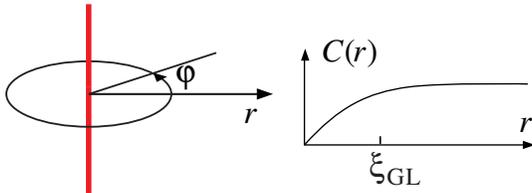
$$\gamma \sim \frac{N(0)}{\hbar^2} \ell \xi_0, \quad (249)$$

i.e. a gradient of the order parameter does not increase the energy as much as in the case of a pure superconductor. It follows from equation (237) that κ increases with increasing impurity. Pure elemental metals are almost exclusively of type I, but they change to type II with increasing impurity.

5.4 Type II superconductivity

Previously we studied the intermediate state of a type I superconductor. In type II superconductor the interface energy is negative. It follows that the magnetic field penetrates into the sample in as small units as possible in order to maximize the amount of the interface. Because of flux quantization (235) we deduce that the smallest unit is one flux quantum Φ_0 . We sketch the corresponding solution of the GL equations. In cylindrical coordinates

$$\Psi(r, \varphi, z) = C(r) e^{i\varphi}. \quad (250)$$



Here the phase ϕ of the order parameter is the same as the azimuthal angle φ of the cylindrical coordinates. Because the order parameter is independent of z , we consider it in the x - y plane. Ψ has to be continuous everywhere. It has a zero at $r = 0$, where it is analytic in spite of the singularity of the coordinate system, $\Psi(x, y, z) = a(x + iy) + O(r^2)$.

The dependence $e^{i\varphi}$ (250) on the phase causes a current (228) that circulates around the z axis. We suppose that

also the vector potential \mathbf{A} is in the direction of the azimuthal angle, $\mathbf{A} = A(r)\hat{\varphi}$. The current

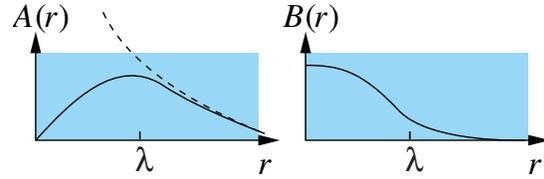
$$\mathbf{j} = 2q\gamma C^2(r) \left[\frac{\hbar}{r} - qA(r) \right] \hat{\varphi}. \quad (251)$$

At large r the order parameter approaches its equilibrium value (219). There the current (251) must vanish (exponentially). Thus

$$\mathbf{A}(r) = \frac{\hbar}{q} \frac{\hat{\varphi}}{r} \quad (r \gg \lambda). \quad (252)$$

Requiring that $\mathbf{A}(r)$ is regular at origin, we can guess its shape. Finally we calculate

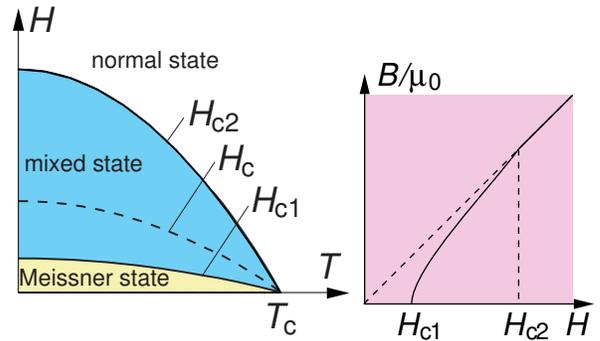
$$\mathbf{B} = \nabla \times \mathbf{A} = \frac{\hat{z}}{r} \frac{d(rA)}{dr}. \quad (253)$$



The accurate forms of the functions are obtained by solving the GL equations, which generally is possible only numerically.

The solution of the type (250) is called a *quantized vortex* or *vortex* or *flux line*. We see from equation (252) that the magnetic flux associated to a vortex is precisely one flux quantum Φ_0 (4).

For a type II superconductor one gets the following phase diagram.

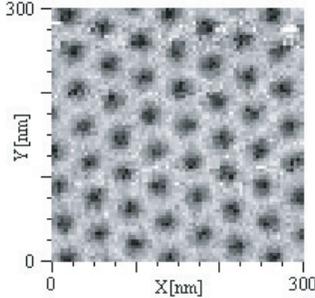


Between the critical fields H_{c1} and H_{c2} the magnetic field partly penetrates to the sample. The density of vortices is $n = B/\Phi_0$. The solution for one vortex describes the situation at fields near H_{c1} , where the vortices are far apart from each other. In increasing field also the density of vortices increases until at H_{c2} they are so dense that there is no space for superconductivity in between. Because the vortex core size is approximately ξ_{GL} , we estimate from this that $B_{c2} \sim \Phi_0/\xi_{GL}^2$. An accurate calculation with GL theory gives

$$B_{c2} = \mu_0 H_{c2} = \frac{\hbar}{2|e|\xi_{GL}^2} = \frac{\Phi_0}{2\pi\xi_{GL}^2}. \quad (254)$$

Near H_{c2} the order parameter goes continuously to zero. In this case the third order term $\Psi|\Psi|^2$ in the GL equation (211) can be dropped, and the remaining equation is the same as the Schrödinger equation for a charged particle in constant magnetic field \mathbf{B} . We leave the mathematics of this problem to the condensed matter course.

In the equilibrium state the vortices fill the sample as uniformly as possible. This leads to a lattice that is hexagonal.



Spectroscopic image of the vortex lattice in NbS₂ at 4.2 Kelvin and 1 Tesla. Dark corresponds to the normal vortex cores, and bright to the superconducting regions. The vortex lattice imaging by scanning tunneling spectroscopy relies on spatial variations of the density of levels in the mixed state. Indeed, the local density of levels is different at the center of vortex cores compared to the surrounding superconducting regions. Plotting these differences as a function of position yields a spectroscopic real space image of the Abrikosov vortex lattice. (figure from http://dpmc.unige.ch/gr_fischer/)

Check as an exercise that the dimensions of the figure and the given field are consistent.

Force on a flux line

A Lorentz force

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}). \quad (255)$$

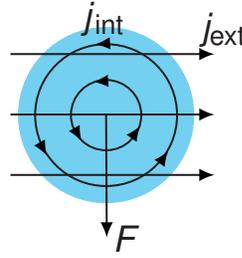
acts on a charged particle. This can be generalized to continuously distributed matter as

$$\mathbf{F} = \int d^3r (\rho \mathbf{E} + \mathbf{j} \times \mathbf{B}), \quad (256)$$

where ρ is the charge density and \mathbf{j} the electric current density. We apply this to a single flux line in an applied flow field $\mathbf{j}_{\text{ext}}(\mathbf{r})$. Because of charge neutrality $\rho = 0$. Supposing that $\mathbf{j}_{\text{ext}}(\mathbf{r})$ is approximately constant in the cross section of the vortex (area $\approx \lambda^2$), we can calculate the integral in transverse plane and get the force acting on the flux line

$$\mathbf{F} = \Phi_0 \int \mathbf{j}_{\text{ext}} \times d\mathbf{l}, \quad (257)$$

where $d\mathbf{l}$ is the line element of the flux line. This force drives the vortex in direction that is perpendicular to the applied current.



Notice that the direction of the force is such that it tends to decrease the region where the total flow velocity is largest (compare to the Magnus force).

One application of superconductivity is to build strong magnets. Because H_c is relatively small (see the table on page 2), superconductors of type II are used. A large part of applications use Nb-Ti alloy, where 45 weight per cent is titanium. Because the alloy is disordered, the mean free path is very short and therefore H_{c2} is high, $B_{c2} \approx 10$ T at $T = 4$ K. Notice that if the alloy would be ordered, the mean free path could in principle be as long as in a pure element (see condensed matter course for justification).

The motion of flux lines leads to dissipation, which means that the “superconductor” is not conducting without resistance. This can be prevented by grain boundaries, precipitates or other impurities that trap flux lines. Especially Nb-Ti has titanium precipitates which are not superconducting. In these regions the flux lines have lower energy and thus are trapped there.

Rotation of a superfluid

Many of the results described above are also valid in ⁴He and ³He superfluids. An essential difference is that they are uncharged (effectively $q = 0$) so that instead of electric current (228) one gets mass current

$$\mathbf{j}_{\text{mass}} = 2m\gamma|\Psi|^2\hbar\nabla\phi. \quad (258)$$

Here $m = m_4$ for ⁴He and $m = 2m_3$ for ³He, where m_4 and m_3 are the corresponding masses of the atoms (why so?). When $|\Psi|^2$ is constant one can define *superfluid velocity* \mathbf{v}_s : because $\hbar\nabla\phi$ is momentum,

$$\mathbf{v}_s = \frac{\hbar}{m}\nabla\phi. \quad (259)$$

It follows that

$$\nabla \times \mathbf{v}_s = 0. \quad (260)$$

We compare this to uniform rotation where

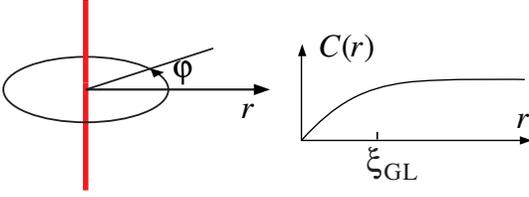
$$\mathbf{v} = \boldsymbol{\Omega} \times \mathbf{r} \Rightarrow \nabla \times \mathbf{v} = 2\boldsymbol{\Omega}. \quad (261)$$

We thus get the interesting result that superfluid cannot rotate in such a way that $|\Psi|^2$ is constant. This is called the rotation paradox.

The solution of the rotation paradox is that quantized vortices are formed. Because $q = 0$, we get for a single vortex in the simplest case

$$\Psi(r, \varphi, z) = C(r)e^{i\varphi}, \quad (262)$$

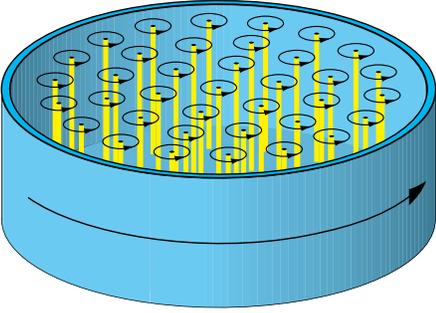
$$\mathbf{j}_{\text{mass}} = 2m\gamma C^2(r) \frac{\hbar}{r} \hat{\phi}. \quad (263)$$



The circulation of the superfluid velocity around a single vortex is

$$\oint d\mathbf{l} \cdot \mathbf{v}_s = \frac{\hbar}{m} \oint d\mathbf{l} \cdot \nabla\phi = \frac{h}{m}.$$

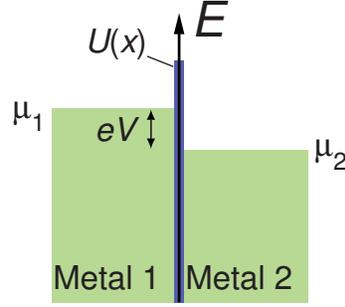
Vortices are formed in a superfluid when the container is rotated.



6. Josephson effect

6.1 Tunneling

Let us consider two metals that separated from each other by a thin insulating layer. Although the potential in the insulator is higher than the energy of electrons, according to quantum mechanics the electrons can tunnel through the potential barrier. In equilibrium, the chemical potentials of the metal are equal (why?) We study the case when a current source is connected between the metals, and it causes a voltage V across the junction: $\mu_1 - \mu_2 = eV$.



(Note. The figures are often drawn as if $eV > 0$. If one wants to take into account that $e < 0$, it is easiest to think that also $V < 0$.)

For the tunneling current we get the expression

$$\begin{aligned} J = c \int dE \tau(E) & \\ \times \{ N_1(E) f(E) N_2(E + eV) [1 - f(E + eV)] & \\ - N_1(E) [1 - f(E)] N_2(E + eV) f(E + eV) \}. & \end{aligned} \quad (264)$$

Here c is a constant, $\tau(E)$ is the tunneling probability and N_1 and N_2 are the densities of level in the metals. The first term in equation (264) describes tunneling from left to right and the second from right to left. The Fermi functions f (43) in both terms take into account that only those cases are counted where the state on the starting side is initially occupied, and that the electron arrives on the other side to a state that initially was empty. From equation (264) we get by direct calculation

$$\begin{aligned} J = c \int dE \tau(E) N_1(E) N_2(E + eV) & \\ \times [f(E) - f(E + eV)]. & \end{aligned} \quad (265)$$

Notice that the factor containing the Fermi functions is different from zero only in an energy interval of $\sim eV + k_B T$.

We study tunneling between two normal state metals. There the density of states is approximately constant and we get

$$J \approx c N_1(0) N_2(0) \tau(0) \int dE$$

$$\begin{aligned} & \times [f(E) - f(E + eV)] \\ & = cN_1(0)N_2(0)\tau(0)eV, \end{aligned} \quad (266)$$

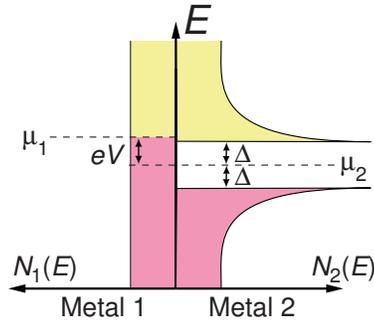
where we also assumed that τ depends only weakly on energy. The result is linear in voltage V . Thus the factor multiplying it can be identified as conductance (inverse of resistance) $G = 1/R$, and

$$J = V/R. \quad (267)$$

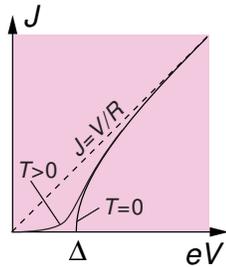
A more complicated result for the current is obtained between a superconducting and normal state metal. The density of levels in the superconducting state is

$$N_s(E) = \begin{cases} \frac{N(0)E}{\sqrt{E^2 - \Delta^2}} & \text{when } |E| > \Delta \\ 0 & \text{when } |E| < \Delta, \end{cases} \quad (268)$$

which is easily obtained from the dispersion relation (155) taking into account that the levels are uniformly distributed in ξ_k (exercise).

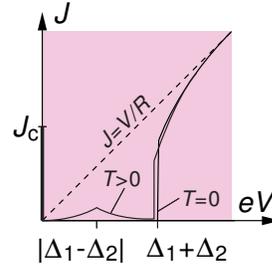


At zero temperature the current is different from zero only if $|eV| > \Delta$. One can deduce that the current-voltage relationship is qualitatively of the type shown in the figure. (More accurately one should take into account that the tunneling objects (electrons) are not the same as the excitations on the superconductor side, but this does not change the the result, see Tinkham).



This kind of dependence of the current on the energy gap is applied, among other things, in measuring the figure on page 30.

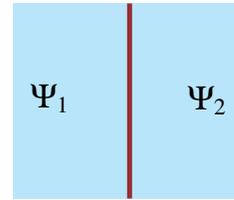
Between two superconductors one gets even more complicated current-voltage relationships depending on the magnitudes of the energy gaps.



The most interesting feature is the current that is obtained at precisely zero voltage.

6.2 Josephson effect

We consider two superconductors that are weakly coupled to each other. We assume in the beginning that $\mathbf{B} = \mathbf{A} = 0$.



Analogously to the phenomenological derivation of the GL theory, we form an expression for the energy associated with such a junction. Let Ψ_1 and Ψ_2 denote the order parameters on the left and right hand sides. We require 1) reality, 2) independence of a constant phase factor $\exp(i\phi)$. 3) independence on exchange of the two sides, and 4) take only the leading order terms. This way we get the *Josephson energy*

$$F_J = -a(\Psi_1^* \Psi_2 + \Psi_1 \Psi_2^*) = -2a \operatorname{Re}(\Psi_1^* \Psi_2). \quad (269)$$

We substitute

$$\Psi_1 = C e^{i\phi_1}, \quad \Psi_2 = C e^{i\phi_2}, \quad (270)$$

and get

$$F_J = -E_J \cos \Delta\phi. \quad (271)$$

We have defined the *phase difference*

$$\Delta\phi = \phi_2 - \phi_1 \quad (272)$$

and $E_J = 2aC^2$.

From the Josephson energy (271) we get the electric current through the junction

$$J = J_c \sin \Delta\phi, \quad (273)$$

where $J_c = (q/\hbar)E_J = (2e/\hbar)E_J$.

Justification of (273): We substitute (227) in to the GL energy (208). We make variation of it with respect ϕ on both sides (1-dimensional model is sufficient). The variation gives surface terms, and these counted together with the Josephson energy (271) should vanish, which gives relation (273). ■

Another important relation is the following, which gives the time derivative of the phase,

$$\frac{d\phi}{dt} = -\frac{2\mu}{\hbar}. \quad (274)$$

Justification of (274): the order parameter Ψ of the GL theory was interpreted as the wave function describing a Cooper pair. In equilibrium the Cooper pairs are in equilibrium with electrons so that the energy of a Cooper pair is 2μ , twice the electron chemical potential. The time dependence of an energy eigenstate in quantum mechanics comes from the factor $\exp(-iEt/\hbar)$, which for the order parameter $\Psi = e^{i\phi}|\Psi|$ is $\exp(-i2\mu t/\hbar)$, and thus one gets (274). ■

We apply (274) to a Josephson junction. For the phase difference (272) we get

$$\frac{d\Delta\phi}{dt} = \frac{2eV}{\hbar}, \quad (275)$$

since the difference in the chemical potentials is related to the voltage V by $\Delta\mu = \mu_2 - \mu_1 = -eV$.

The equations (273) and (275) are known as Josephson equations. The first gives that in equilibrium ($V = 0$) a constant current flows through that depends sinusoidally on the phase difference $\Delta\phi \equiv \phi_2 - \phi_1$. This is known as *dc Josephson effect*.

If the voltage V is constant, one gets from equation (275) that the phase grows linearly in time,

$$\Delta\phi = \frac{2eV}{\hbar}t. \quad (276)$$

Substituting this in to equation (273) one gets alternating current at angular frequency

$$\omega = \frac{2e}{\hbar}V. \quad (277)$$

This is known as *ac Josephson effect*.

At voltage 0.1 mV (which is typical in the figure on page 32) the equation (277) gives the frequency $\nu = \omega/2\pi = 48$ GHz.

Using Josephson junctions it is possible to make sensitive measuring devices. For example, equation (277) makes possible a voltage standard, as the frequency can accurately be measured.

The current source driving the junction makes in time dt the work $VJdt$. According to energy conservation we must have

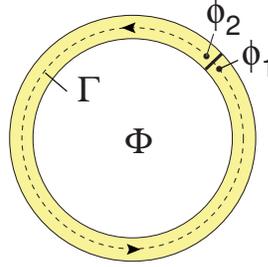
$$\frac{dF_J}{dt} = VJ. \quad (278)$$

We see that using this relation we can derive the third relation based on any pair of the relations (271), (273) and (275).

If the vector potential \mathbf{A} is different from zero, the phase difference (272) used above should be redefined as follows.

Analogously to equation (228) one defines a gauge invariant phase difference

$$\Delta\phi = \phi_2 - \phi_1 - \frac{2e}{\hbar} \int_1^2 \mathbf{dl} \cdot \mathbf{A}. \quad (279)$$



Let us consider a ring that contains one Josephson junction. The total flux is

$$\Phi = \int d\mathbf{a} \cdot \mathbf{B} = \oint \mathbf{dl} \cdot \mathbf{A} = \int_1^2 \mathbf{dl} \cdot \mathbf{A} + \int_2^1 \mathbf{dl} \cdot \mathbf{A} \quad (280)$$

where the last form has two terms, the former across the junction and the latter over the rest of the ring. Using (228) and $j = 0$ the latter contribution is

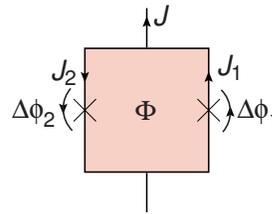
$$\int_2^1 \mathbf{dl} \cdot \mathbf{A} = \frac{\hbar}{2e} \int_2^1 \mathbf{dl} \cdot \nabla\phi = \frac{\hbar}{2e}(\phi_1 - \phi_2 + 2\pi N). \quad (281)$$

Substituting in (280) we get the phase difference across the junction

$$\Delta\phi = \frac{2\pi\Phi}{\Phi_0} + 2\pi N. \quad (282)$$

If there are more junctions, this generalizes to

$$\sum_j \Delta\phi_j = \frac{2\pi\Phi}{\Phi_0} + 2\pi N. \quad (283)$$



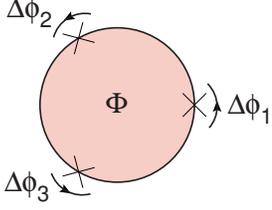
Consider the circuit that has thick ($\gg \lambda$) superconducting wires and two Josephson junctions. This is described by equations

$$\begin{aligned} \Delta\phi_1 + \Delta\phi_2 &= \frac{2\pi\Phi}{\Phi_0} + 2\pi N \\ J &= J_{c1} \sin(\Delta\phi_1) - J_{c2} \sin(\Delta\phi_2). \end{aligned} \quad (284)$$

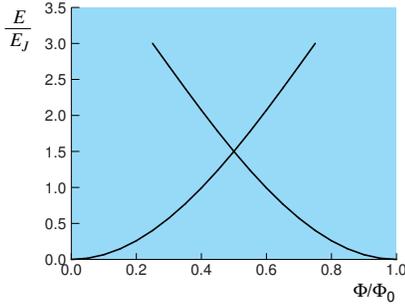
For simplicity assume $J_{c1} = J_{c2}$. By calculation one sees that the two junctions behave as if a single junction (273), whose critical current J_c depends on the flux

$$J_c = 2J_{c1} \left| \cos \frac{\pi\Phi}{\Phi_0} \right|. \quad (285)$$

Changing the field thus changes the critical current. This is used in very sensitive measuring devices of magnetic field. The device based on the circuit is called dc SQUID (Superconducting quantum interference device).



We can study a ring with three Josephson junctions. At certain fields $\Phi \approx \frac{1}{2}\Phi_0$ this has two possible energy states that correspond to currents flowing in opposite directions.



6.3 Alternative treatment of the Josephson effect

We showed above that the BCS state (171) can be represented as a superposition of states $|n\rangle$, where the pair number n is fixed,

$$|\psi\rangle = \sum_n e^{in\phi} f_n |n\rangle. \quad (286)$$

Here we have written the phase ϕ (177) explicitly and therefore f_n is real valued.

We consider two superconductors L and R, which are weakly coupled to each other. The coupling can simplest be described by adding to the Hamiltonian the term

$$\begin{aligned} \check{H}_T = & -\frac{E_J}{2} \sum_m \sum_n \left(|m-1, n+1\rangle \langle m, n| \right. \\ & \left. + |m+1, n-1\rangle \langle m, n| \right). \end{aligned} \quad (287)$$

Here $|m, n\rangle$ is a state where the superconductor L has m pairs and superconductor R has n pairs. The first term describes process where one pair jumps from L to R, and the second term describes the opposite process. E_J is a constant describing the strength of the coupling.

The states of uncoupled superconductors are described by the product of two states of the type (286):

$$|\psi\rangle = \sum_k \sum_l e^{ik\phi_L} e^{il\phi_R} f_k f_l |k, l\rangle. \quad (288)$$

If the coupling is weak, it can be treated as a small perturbation. The first order correction to the energy is obtained from the expectation value

$$E_1 = \langle \psi | \check{H}_T | \psi \rangle. \quad (289)$$

Substituting the expressions (287) and (288) we get

$$\begin{aligned} E_1 &= -\frac{E_J}{2} \sum_m \sum_n (e^{i(\phi_L - \phi_R)} f_{m-1} f_{n+1} \\ &\quad + e^{-i(\phi_L - \phi_R)} f_{m+1} f_{n-1}) f_m f_n \\ &\approx -E_J \cos(\phi_R - \phi_L) \end{aligned} \quad (290)$$

supposing $\sum_m f_{m-1} f_m \approx \sum_m f_m^2 = 1$. The result is the same as obtained above (271) in a different way.

Based on the same assumptions, we can also calculate the current

$$J = 2e \frac{d}{dt} \langle \psi | \check{n}_R | \psi \rangle, \quad (291)$$

where

$$\check{n}_R = \sum_m \sum_n n |m, n\rangle \langle m, n| \quad (292)$$

is the pair number operator on the side R.

The time derivative in equation (291) can easily be calculated using time dependent Schrödinger equation. More directly the same result is obtained in the Heisenberg picture, where the states are time independent but the operators obey the equation of motion

$$\frac{d\check{n}_R}{dt} = \frac{i}{\hbar} [\check{H}, \check{n}_R] = \frac{i}{\hbar} [\check{H}_T, \check{n}_R]. \quad (293)$$

Here the latter equality follows because the only term in the Hamiltonian that does not commute with \check{n}_R is the tunneling term (287). Substituting (287) we get

$$\begin{aligned} \frac{d\check{n}_R}{dt} = & -\frac{iE_J}{2\hbar} \sum_m \sum_n \left(-|m-1, n+1\rangle \langle m, n| \right. \\ & \left. + |m+1, n-1\rangle \langle m, n| \right). \end{aligned} \quad (294)$$

The expectation value of this can be calculated similarly as above, and we get

$$J = \frac{2e}{\hbar} E_J \sin(\phi_R - \phi_L). \quad (295)$$

Also this result is the same as above (273).

We apply the time dependent Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} |\psi\rangle = \check{H} |\psi\rangle \quad (296)$$

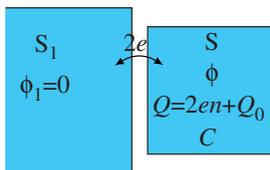
to the state (186). By definition of the chemical potential μ , the energies of states with different number of particles differ by μ , i.e., $\check{H} = 2\mu\check{N} + \text{constant}$. We get

$$\frac{d\phi}{dt} = -\frac{2\mu}{\hbar}, \quad (297)$$

which also was derived above in (274).

6.4 Macroscopic quantum mechanics

Previously we considered the order parameter and its phase as classical quantities, which have definite values. The precision of the phase, however, is limited by the uncertainty relation (189). Let us study when this restriction becomes essential.



We consider a small superconducting body S, which we call island. It is connected via a Josephson junction to another larger superconductor S₁. (Without losing generality we can choose the phase of S₁ to vanish, and the phase of S is marked by φ.) The tunneling of one pair changes the electric charge Q of the island by 2e. This leads to a change of the electrostatic potential Q²/2C, where C is the capacitance of the island. The capacitance can be estimated using the plate-capacitor formula C = εA/d, where A is the area of the Josephson junction, ε the permeability of the insulator and d the thickness of the insulating layer. The order of magnitude of the charging energy is given by E_c ≡ e²/2C.

The phase is well defined if the charging energy is small compared to the Josephson energy, E_c ≪ E_J. In the opposite case, when E_c has same order of magnitude as E_J (or is larger), the phase is a quantum mechanical variable in a sense that it should be described by a wave function ψ(φ). This is known as *macroscopic quantum mechanics*, as φ describes a large number of electrons but still behaves according to quantum laws.

Other conditions for the observation of quantum behavior of phase is low temperature, and that the charge cannot escape too fast by other means. In practice this means very small junctions of area A < 10⁻¹² m².

In order to get a quantitative theory, we construct a Hamiltonian. The charging energy is 4E_c(n + Q₀/2e)². This is the capacitive energy Q²/2C caused by charge Q = 2en + Q₀, where 2en describes the charge dependent on the number of Cooper pairs and Q₀ is an independent contribution that is called the background charge. Although n gets only integer values, Q₀ can be an arbitrary real number, since it is changed, for example, by the motion of an impurity ion in the insulator between the capacitor plates. The charging energy is diagonal when using the pair number eigenstates |n⟩. The tunneling part was postulated above in (287), but here we concentrate only on the island. (We suppose the other superconductor so large that the changes in its particle

number are unimportant.) Thus we get the Hamiltonian

$$\check{H} = \sum_n \left[4E_c |n\rangle \left(n + \frac{Q_0}{2e} \right)^2 \langle n| - \frac{E_J}{2} (|n+1\rangle\langle n| + |n-1\rangle\langle n|) \right]. \quad (298)$$

It is interesting to express this Hamiltonian using the phase eigenstates. Using the transformation formulas (186) and (187) we get the form

$$\check{H} = \int_{-\pi}^{\pi} d\phi |\phi\rangle \left[4E_c \left(i \frac{\partial}{\partial \phi} + \frac{Q_0}{2e} \right)^2 - E_J \cos \phi \right] \langle \phi|. \quad (299)$$

Verify this as an exercise. We see that in φ representation the operator giving n is

$$n_{\text{op}} = i \frac{\partial}{\partial \phi}. \quad (300)$$

This and the phase have commutation relation [φ, n_{op}] = -i, as generally applies to quantities that are obtained by Fourier transform from each other. [Depending on definitions the sign of the commutator is ±i. Here the sign of the exponent in equation (186) is chosen such that it leads to the same definition of φ as usually used in the literature of superconductivity.]

Because the Hamiltonian (299) is diagonal in the phase eigenstates, it is simpler to write it as

$$H = 4E_c \left(i \frac{\partial}{\partial \phi} + \frac{Q_0}{2e} \right)^2 - E_J \cos \phi. \quad (301)$$

This Hamiltonian determines the form of the phase wave function ψ(φ). If E_c ≪ E_J, can ψ(φ) be almost a delta function, and the phase is well defined. In the opposite case ψ(φ) is distributed and the phase uncertain.

Finally we mention that the macroscopic quantum mechanics is actively studied at present. One of the motivations is to make a qubit, the bit of a quantum computer. For example, the quantum tunneling between two macroscopic states has been observed experimentally in the ring with three Josephson junctions described above [van der Wal et al, Science **290**, 773 (2000).].

Alternative derivation

Here we give alternative derivation for the quantum properties of phase. If you are satisfied with the above, you can skip this.

Tunnel junctions always have also electric capacitance C. Its energy is

$$E_Q = \frac{1}{2} CV^2 = \frac{1}{2} C \left(\frac{\hbar}{2e} \right)^2 \dot{\phi}^2 \quad (302)$$

where we have used formula (276). This can be thought as some kinetic energy and the corresponding potential energy is then the Josephson energy (271). Out of these we can form the Lagrange function

$$L = \frac{1}{2}C \left(\frac{\hbar}{2e} \right)^2 \dot{\phi}^2 + E_J \cos \phi. \quad (303)$$

We calculate the canonical momentum

$$p = \frac{\partial L}{\partial \dot{\phi}} = C \left(\frac{\hbar}{2e} \right)^2 \dot{\phi} = \frac{\hbar}{2e} CV = \frac{\hbar}{2e} Q. \quad (304)$$

The Hamiltonian is

$$H = \frac{Q^2}{2C} - E_J \cos \phi. \quad (305)$$

The commutation rule for canonical variables x ja p gives

$$[\check{p}, \check{x}] = -i\hbar, \quad (306)$$

and based on analogy we have

$$[\check{Q}, \check{\phi}] = -2ei. \quad (307)$$

Using $\check{Q} = 2e\check{n} + Q_0$ we see that we get the same formulas as before (300) and (301) except a sign difference in (300) and the Q_0 contribution in (301). (The sign difference could be avoided by reversing the definition of the phase difference over the junction, and therefore should not be a fundamental problem. The Q_0 contribution can be added as discussed in the course Quantum optics in electric circuits.)

We note that the classical problem considered here (303) is the same as for a simple pendulum, and the quantum mechanical problem (301) is the same as for an electron in a 1-dimensional crystal with sinusoidal periodic potential.

7. Conclusion

The course had two main topics: the microscopic theory (BCS) and the macroscopic theories (thermodynamics, GL, Josephson effect). The purpose was to give basic understanding of both. In addition we considered several more detailed questions. In conclusion one should read again the introduction to see if the topics mentioned there were sufficiently understood.

Appendix

A. Maxwell's equations

Maxwell's equations are

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}, \quad (308)$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \quad (309)$$

$$\nabla \cdot \mathbf{B} = 0, \quad (310)$$

$$\nabla \times \mathbf{B} = \epsilon_0 \mu_0 \frac{\partial \mathbf{E}}{\partial t} + \mu_0 \mathbf{j}. \quad (311)$$

Here \mathbf{E} is the *electric field* and \mathbf{B} the *magnetic field*. In many cases it is convenient to represent them using potentials

$$\mathbf{E} = -\nabla\varphi - \frac{\partial \mathbf{A}}{\partial t}, \quad (312)$$

$$\mathbf{B} = \nabla \times \mathbf{A}, \quad (313)$$

where φ is the *scalar potential* and \mathbf{A} the *vector potential*.

The Maxwell's equations contain the *charge density* ρ and *electric current density* \mathbf{j} . In studying electromagnetic phenomena in materials, it is useful to separate the charges arising from the polarization and other charges. Correspondingly, the electric current can be divided to the current arising from magnetization and polarization of the medium, and to other currents. Without proper justification we claim that this can be written as the following equations

$$\rho = \rho_f - \nabla \cdot \mathbf{P}, \quad (314)$$

$$\mathbf{j} = \mathbf{j}_f + \nabla \times \mathbf{M} + \frac{\partial \mathbf{P}}{\partial t}. \quad (315)$$

Here \mathbf{P} is the *electric polarization* and \mathbf{M} the *magnetization* of the medium. Other charges and currents are denoted by index f meaning free. We substitute these relations to Maxwell's equations (308)-(311) and define two new fields

$$\mathbf{D} \equiv \epsilon_0 \mathbf{E} + \mathbf{P}, \quad (316)$$

$$\mathbf{H} \equiv \frac{1}{\mu_0} \mathbf{B} - \mathbf{M}. \quad (317)$$

We get *Maxwell's equations in a medium*

$$\begin{aligned} \nabla \cdot \mathbf{D} &= \rho_f, \\ \nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t}, \\ \nabla \cdot \mathbf{B} &= 0, \\ \nabla \times \mathbf{H} &= \frac{\partial \mathbf{D}}{\partial t} + \mathbf{j}_f. \end{aligned} \quad (318)$$