
Theory of Superconductivity

Carsten Timm

Contents

1	Introduction	5
1.1	Scope	5
1.2	Overview	5
1.3	Books	6
2	Basic experiments	7
2.1	Conventional superconductors	7
2.2	Superfluid helium	9
2.3	Unconventional superconductors	10
2.4	Bose-Einstein condensation in dilute gases	12
3	Bose-Einstein condensation	13
4	Normal metals	18
4.1	Electrons in metals	18
4.2	Semiclassical theory of transport	19
5	Electrodynamics of superconductors	23
5.1	London theory	23
5.2	Rigidity of the superfluid state	25
5.3	Flux quantization	26
5.4	Nonlocal response: Pippard theory	27
6	Ginzburg-Landau theory	30
6.1	Landau theory of phase transitions	30
6.2	Ginzburg-Landau theory for neutral superfluids	33
6.3	Ginzburg-Landau theory for superconductors	37
6.4	Type-I superconductors	41
6.5	Type-II superconductors	44
7	Superfluid and superconducting films	52
7.1	Superfluid films	52
7.2	Superconducting films	64
8	Origin of attractive interaction	70
8.1	Reminder on Green functions	70
8.2	Coulomb interaction	73
8.3	Electron-phonon interaction	77
8.4	Effective interaction between electrons	79

9	Cooper instability and BCS ground state	82
9.1	Cooper instability	82
9.2	The BCS ground state	84
10	BCS theory	89
10.1	BCS mean-field theory	89
10.2	Isotope effect	95
10.3	Specific heat	95
10.4	Density of states and single-particle tunneling	97
10.5	Ultrasonic attenuation and nuclear relaxation	102
10.6	Ginzburg-Landau-Gor'kov theory	106
11	Josephson effects	107
11.1	The Josephson effects in Ginzburg-Landau theory	107
11.2	Dynamics of Josephson junctions	109
11.3	Bogoliubov-de Gennes Hamiltonian	113
11.4	Andreev reflection	114
12	Unconventional pairing	120
12.1	The gap equation for unconventional pairing	120
12.2	Cuprates	123
12.3	Pnictides	130
12.4	Triplet superconductors and He-3	132

1

Introduction

1.1 Scope

Superconductivity is characterized by a vanishing static electrical resistivity and an expulsion of the magnetic field from the interior of a sample. We will discuss these basic experiments in the following chapter, but mainly this course is dealing with the *theory* of superconductivity. We want to understand superconductivity using methods of theoretical physics. Experiments will be mentioned if they motivate certain theoretical ideas or support or contradict theoretical predictions, but a systematic discussion of experimental results will not be given.

Superconductivity is somewhat related to the phenomena of *superfluidity* (in He-3 and He-4) and *Bose-Einstein condensation* (in weakly interacting boson systems). The similarities are found to lie more in the effective low-energy description than in the microscopic details. Microscopically, superfluidity in He-3 is most closely related to superconductivity since both phenomena involve the condensation of *fermions*, whereas in He-4 and, of course, Bose-Einstein condensates it is *bosons* that condense. We will discuss these phenomena briefly.

The course assumes knowledge of the standard material from electrodynamics, quantum mechanics I, and thermodynamics and statistics. We will also use the second-quantization formalism (creation and annihilation operators), which are usually introduced in quantum mechanics II. A prior course on introductory solid state physics would be useful but is not required. Formal training in many-particle theory is not required, necessary concepts and methods will be introduced (or recapitulated) as needed.

1.2 Overview

This is a maximal list of topics to be covered, not necessarily in this order:

- basic experiments
- review of Bose-Einstein condensation
- electrodynamics: London and Pippard theories
- Ginzburg-Landau theory, Anderson-Higgs mechanism
- vortices
- origin of electron-electron attraction
- Cooper instability, BCS theory
- consequences: thermodynamics, tunneling, nuclear relaxation
- Josephson effects, Andreev scattering
- Bogoliubov-de Gennes equation
- cuprate superconductors
- pnictide superconductors
- topological superconductors

1.3 Books

There are many textbooks on superconductivity and it is recommended to browse a few of them. None of them covers all the material of this course. M. Tinkham's *Introduction to Superconductivity* (2nd edition) is well written and probably has the largest overlap with this course.

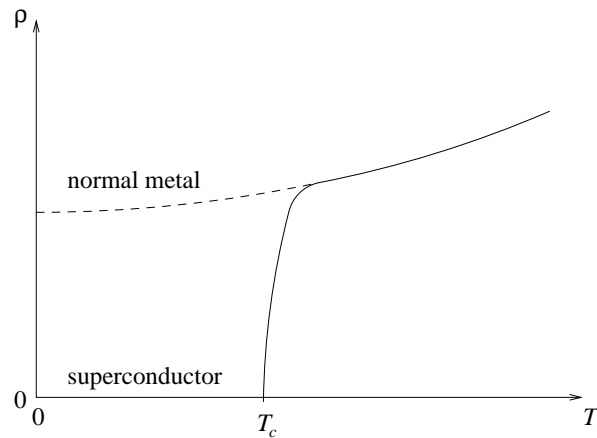
2

Basic experiments

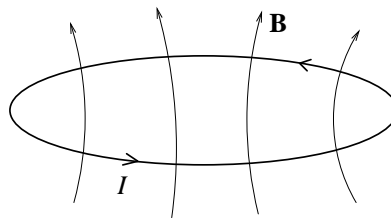
In this chapter we will review the essential experiments that have established the presence of superconductivity, superfluidity, and Bose-Einstein condensation in various materials classes. Experimental observations that have helped to elucidate the detailed properties of the superconductors or superfluids are not covered; some of them are discussed in later chapters.

2.1 Conventional superconductors

After H. Kamerlingh Onnes had managed to liquify Helium, it became for the first time possible to reach temperatures low enough to achieve superconductivity in some chemical elements. In 1911, he found that the static resistivity of mercury abruptly fell to zero at a *critical temperature* T_c of about 4.1 K.

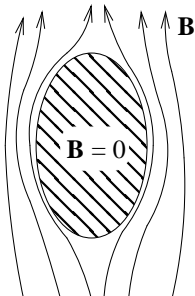


In a normal metal, the resistivity decreases with decreasing temperature but saturates at a finite value for $T \rightarrow 0$. The most stringent bounds on the resistivity can be obtained not from direct measurement but from the decay of persistent currents, or rather from the lack of decay. A current set up (by induction) in a superconducting ring is found to persist without measurable decay after the electromotive force driving the current has been switched off.



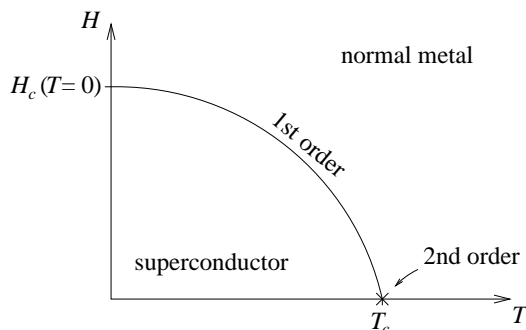
Assuming exponential decay, $I(t) = I(0)e^{-t/\tau}$, a lower bound on the decay time τ is found. From this, an upper bound of $\rho \lesssim 10^{-25} \Omega\text{m}$ has been extracted for the resistivity. For comparison, the resistivity of copper at room temperature is $\rho_{\text{Cu}} \approx 1.7 \times 10^{-8} \Omega\text{m}$.

The *second essential observation* was that superconductors not only prevent a magnetic field from entering but actively *expel* the magnetic field from their interior. This was observed by W. Meißner and R. Ochsenfeld in 1933 and is now called the *Meißner* or *Meißner-Ochsenfeld effect*.



From the materials relation $\mathbf{B} = \mu\mathbf{H}$ with the permeability $\mu = 1 + 4\pi\chi$ and the magnetic susceptibility χ (note that we are using *Gaussian units*) we thus find $\mu = 0$ and $\chi = -1/4\pi$. Superconductors are diamagnetic since $\chi < 0$. What is more, they realize the smallest (most diamagnetic) value of μ consistent with thermodynamic stability. The field is not just diminished but completely expelled. They are thus *perfect diamagnets*.

It costs energy to make the magnetic field nonuniform although the externally applied field is uniform. It is plausible that at some externally applied magnetic field $H_c(T) \equiv B_c(T)$ this cost will be so high that there is no advantage in forming a superconducting state. For typical conventional superconductors, the experimental phase diagram in the temperature-magnetic field plane looks like this:



To specify which superconductors discovered after Hg are *conventional*, we need a definition of what we want to call “conventional superconductors.” There are at least two inequivalent but often coinciding definitions: Conventional superconductors

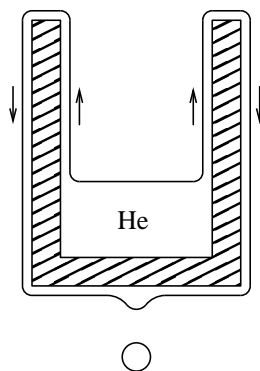
- show a superconducting state of trivial symmetry (we will discuss later what this means),
- result from an attractive interaction between electrons for which phonons play a dominant role.

Conventional superconductivity was observed in quite a lot of elements at low temperatures. The record critical temperature for elements are $T_c = 9.3 \text{ K}$ for Nb under ambient pressure and $T_c = 20 \text{ K}$ for Li under high pressure. Superconductivity is in fact rather common in the periodic table, 53 pure elements show it under some conditions. Many alloys and intermetallic compounds were also found to show conventional superconductivity according to the above criteria. Of these, for a long time Nb_3Ge had the highest known T_c of 23.2 K. But it is now thought that MgB_2 ($T_c = 39 \text{ K}$) and a few related compounds are also conventional superconductors in the above sense. They nevertheless show some interesting properties. The rather high $T_c = 39 \text{ K}$ of MgB_2 is interesting since it is on the order of the maximum $T_c^{\text{max}} \approx 30 \text{ K}$ expected for phonon-driven superconductivity. To increase T_c further, the interaction between electrons and phonons would have to be stronger, which however would make the material unstable towards a charge density wave. MgB_2 would thus be an “optimal” conventional superconductor.

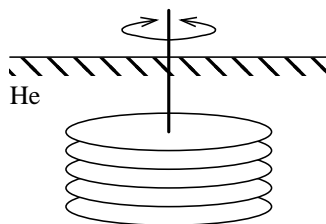
Superconductivity with rather high T_c has also been found in fullerites, i.e., compounds containing fullerene anions. The record T_c in this class is at present $T_c = 38$ K for b.c.c. Cs_3C_{60} under pressure. Superconductivity in fullerites was originally thought to be driven by phonons with strong molecular vibration character but there is recent evidence that it might be unconventional (not phonon-driven).

2.2 Superfluid helium

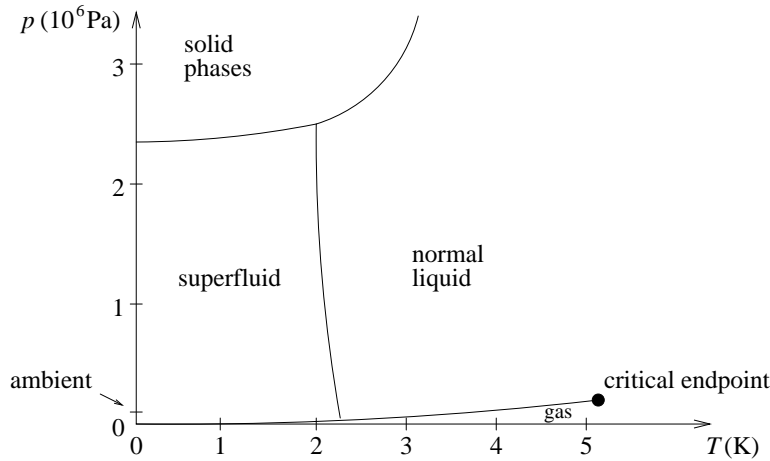
In 1937 P. Kapiza and independently Allen and Misener discovered that helium shows a transition at $T_c = 2.17$ K under ambient pressure, below which it flows through narrow capillaries without resistance. The analogy to superconductivity is obvious but here it was the *viscosity* instead of the resistivity that dropped to zero. The phenomenon was called *superfluidity*. It was also observed that due to the vanishing viscosity an open container of helium would empty itself through a flow in the microscopically thin wetting layer.



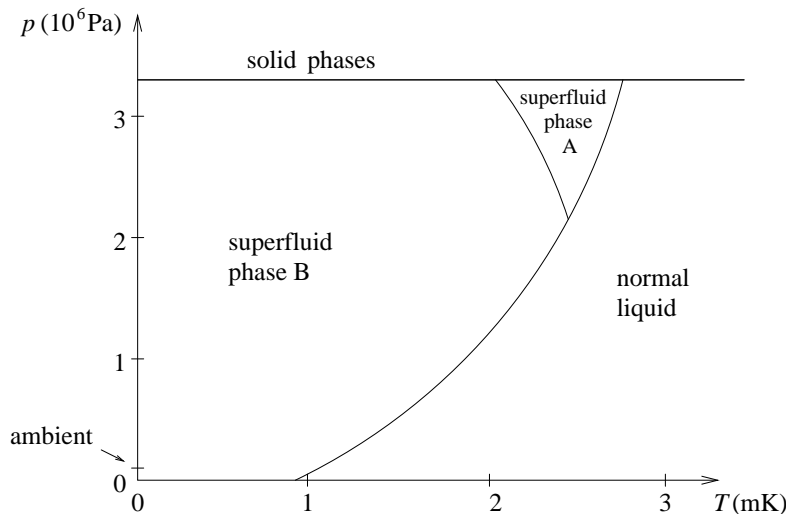
On the other hand, while part of the liquid flows with vanishing viscosity, another part does not. This was shown using torsion pendulums of plates submerged in helium. For $T > 0$ a temperature-dependent *normal* component oscillates with the plates.



Natural atmospheric helium consists of 99.9999% He-4 and only 0.0001% He-3, the only other stable isotope. The observed properties are thus essentially indistinguishable from those of pure He-4. He-4 atoms are bosons since they consist of an even number (six) of fermions. For weakly interacting bosons, A. Einstein predicted in 1925 that a phase transition to a condensed phase should occur (*Bose-Einstein condensation*). The observation of superfluidity in He-4 was thus not a surprise—in contrast to the discovery of superconductivity—but in many details the properties of He-4 were found to be different from the predicted Bose-Einstein condensate. The reason for this is that the interactions between helium atoms are actually quite strong. For completeness, we sketch the temperature-pressure phase diagram of He-4:



The other helium isotope, He-3, consists of *fermionic* atoms so that Bose-Einstein condensation cannot take place. Indeed no superfluid transition was observed in the temperature range of a few Kelvin. It then came as a big surprise when superfluidity was finally observed at much lower temperatures below about 2.6 mK by D. Lee, D. Osheroff, and R. Richardson in 1972. In fact they found *two* new phases at low temperatures. (They originally misinterpreted them as possible magnetic solid phases.) Here is a sketch of the phase diagram, note the temperature scale:

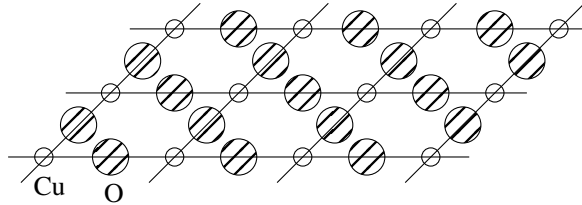


Superfluid He-3 shows the same basic properties as He-4. But unlike in He-4, the superfluid states are sensitive to an applied magnetic field, suggesting that the states have non-trivial magnetic properties.

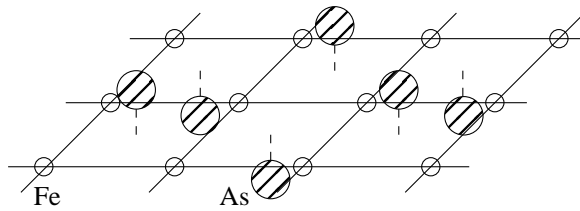
2.3 Unconventional superconductors

By the late 1970's, superconductivity seemed to be a more or less closed subject. It was well understood based on the BCS theory and extensions thereof that dealt with strong interactions. It only occurred at temperatures up to 23.2 K (Nb_3Ge) and thus did not promise widespread technological application. It was restricted to non-magnetic metallic elements and simple compounds. This situation started to change dramatically in 1979. Since then, superconductivity has been observed in various materials classes that are very different from each other and from the typical low- T_c superconductors known previously. In many cases, the superconductivity was unconventional and often T_c was rather high. We now give a brief and incomplete historical overview.

- In 1979, Frank Steglich *et al.* observed superconductivity below $T_c \approx 0.5$ K in CeCu_2Si_2 . This material is not a normal metal in its normal state. Instead it is a *heavy-fermion* metal. The electrons at the Fermi energy have strong Ce f -orbital character. The very strong Coulomb repulsion between electrons in the f -shell leads to a high effective mass $m^* \gg m_e$ at the Fermi energy, hence the name. Since then, superconductivity has been found in various other heavy-fermion compounds. BCS theory cannot explain superconductivity in these highly correlated metals. Nuclear magnetic resonance (discussed below) and other experimental techniques have shown that many of these heavy-fermion superconductors show unconventional symmetry of the superconducting state.
- Also in 1979, D. Jérôme *et al.* (Klaus Bechgaard's group) observed superconductivity in an organic salt called $(\text{TMTSF})_2\text{PF}_6$ with $T_c = 1.1$ K. Superconductivity has since been found in various organic materials with a maximum T_c of about 18 K. The symmetry of the superconducting state is often unconventional. (We do not include fullerenes under organic compounds since they lack hydrogen atoms.)
- While the previously mentioned discoveries showed that superconductivity can occur in unexpected materials classes and probably due to unconventional mechanisms, the T_c values did not surpass the $T_c \approx 23$ K of Nb_3Ge . In 1986, J. G. Bednorz and K. A. Müller observed superconductivity in $\text{La}_{2-x}\text{Ba}_x\text{CuO}_4$ (the layered perovskite cuprate La_2CuO_4 with some Ba substituted for La) with T_c on the order of 35 K. In the following years, many other superconductors based on the same type of nearly flat CuO_2 planes sketched below were discovered. The record transition temperatures for cuprates and for all superconductors are $T_c = 138$ K for $\text{Hg}_{0.8}\text{Tl}_{0.2}\text{Ba}_2\text{Ca}_2\text{Cu}_3\text{O}_{8+\delta}$ at ambient pressure and $T_c = 164$ K for $\text{HgBa}_2\text{Ca}_2\text{Cu}_3\text{O}_{8+\delta}$ under high pressure. The high T_c values as well as many experimental probes show that the cuprates are unconventional superconductors. We will come back to these materials class below.



- In 1991, A. F. Hebard *et al.* found that the fullerite $\text{K}_3\text{C}_{60} = (\text{K}^-)_3\text{C}_{60}^{3-}$ became superconducting below $T_c = 18$ K. T_c in this class has since been pushed to $T_c = 33$ K for $\text{Cs}_2\text{RbC}_{60}$ at ambient pressure and $T_c = 38$ K for (b.c.c., while all the other known superconducting fullerenes are f.c.c.) Cs_3C_{60} under high pressure. The symmetry of the superconducting state appears to be trivial but, as noted above, there is an ongoing debate on whether the pairing is phonon-mediated.
- In 2001, Nagamatsu *et al.* reported superconductivity in MgB_2 with $T_c = 39$ K. The high T_c and the layered crystal structure, reminiscent of cuprates, led to the expectation that superconductivity in MgB_2 is unconventional. However, most experts now think that it is actually conventional, as noted above.
- The most recent series of important discoveries started in 2006, when Kamihara *et al.* (H. Hosono's group) observed superconductivity with $T_c \approx 4$ K in LaFePO , another layered compound. While this result added a new materials class based on Fe^{2+} to the list of superconductors, it did not yet cause much excitement due to the low T_c . However, in 2008, Kamihara *et al.* (the same group) found superconductivity with $T_c \approx 26$ K in $\text{LaFeAsO}_{1-x}\text{F}_x$. Very soon thereafter, the maximum T_c in this iron pnictide class was pushed to 55 K. Superconductivity was also observed in several related materials classes, some of them not containing oxygen (e.g., LiFeAs) and some with the pnictogen (As) replaced by a chalcogen (e.g., FeSe). The common structural element is a flat, square Fe^{2+} layer with a pnictogen or chalcogen sitting alternately above and below the centers of the Fe squares. Superconductivity is thought to be unconventional.



2.4 Bose-Einstein condensation in dilute gases

An important related breakthrough was the realization of a Bose-Einstein condensate (BEC) in a highly diluted and very cold gas of atoms. In 1995, Anderson *et al.* (C. E. Wiman and E. A. Cornell's group) reported condensation in a dilute gas of Rb-87 below $T_c = 170$ nK (!). Only a few months later, Davis *et al.* (W. Ketterle's group) reported a BEC of Na-23 containing many more atoms. About a year later, the same group was able to create two condensates and then merge them. The resulting interference effects showed that the atoms were really in a macroscopic quantum state, i.e., a condensate. All observations are well understood from the picture of a weakly interacting Bose gas. Bose-Einstein condensation will be reviewed in the following chapter.

3

Bose-Einstein condensation

In this short chapter we review the theory of Bose-Einstein condensation. While this is not the correct theory for superconductivity, at least in most superconductors, it is the simplest description of a *macroscopic quantum condensate*. This concept is central also for superconductivity and superfluidity.

We consider an *ideal gas of indistinguishable bosons*. “Ideal” means that we neglect any interaction and also any finite volume of the particles. There are two cases with completely different behavior depending on whether the particle number is conserved or not. Rb-87 atoms are bosons (they consist of 87 nucleons and 37 electrons) with conserved particle number, whereas photons are bosons with non-conserved particle number. Photons can be freely created and destroyed as long as the usual conservation laws (energy, momentum, angular momentum, ...) are satisfied. Bosons without particle-number conservation show a Planck distribution,

$$n_P(E) = \frac{1}{e^{\beta E} - 1} \quad (3.1)$$

with $\beta := 1/k_B T$, for a grand-canonical ensemble in equilibrium. Note the absence of a chemical potential, which is due to the non-conservation of the particle number. This distribution function is an analytical function of temperature and thus does not show any phase transitions.

The situation is different for bosons with conserved particle number. We want to consider the case of a given number N of particles in contact with a heat bath at temperature T . This calls for a *canonical* description (N, T given). However, it is easier to use the *grand-canonical* ensemble with the chemical potential μ given. For large systems, fluctuations of the particle number become small so that the descriptions are equivalent. However, μ must be calculated from the given N .

The grand-canonical partition function is

$$\mathcal{Z} = \prod_i \left(1 + e^{-\beta(\epsilon_i - \mu)} + e^{-2\beta(\epsilon_i - \mu)} + \dots \right) = \prod_i \frac{1}{1 - e^{-\beta(\epsilon_i - \mu)}}, \quad (3.2)$$

where i counts the single-particle states of energy ϵ_i in a volume V . The form of \mathcal{Z} expresses that every state can be occupied not at all, once, twice, etc. For simplicity, we assume the volume to be a cube with periodic boundary conditions. Then the states can be enumerated by wave vectors \mathbf{k} compatible with these boundary conditions. Introducing the *fugacity*

$$y := e^{\beta\mu}, \quad (3.3)$$

we obtain

$$\mathcal{Z} = \prod_{\mathbf{k}} \frac{1}{1 - ye^{-\beta\epsilon_{\mathbf{k}}}} \quad (3.4)$$

$$\Rightarrow \ln \mathcal{Z} = \sum_{\mathbf{k}} \ln \frac{1}{1 - ye^{-\beta\epsilon_{\mathbf{k}}}} = - \sum_{\mathbf{k}} \ln (1 - ye^{-\beta\epsilon_{\mathbf{k}}}). \quad (3.5)$$

The fugacity has to be chosen to give the correct particle number

$$N \stackrel{!}{=} \sum_{\mathbf{k}} \langle n_{\mathbf{k}} \rangle = \sum_{\mathbf{k}} \frac{1}{e^{\beta(\epsilon_{\mathbf{k}} - \mu)} - 1} = \sum_{\mathbf{k}} \frac{1}{y^{-1} e^{\beta \epsilon_{\mathbf{k}}} - 1}. \quad (3.6)$$

Since $\langle n_{\mathbf{k}} \rangle$ must be non-negative, μ must satisfy

$$\mu \leq \epsilon_{\mathbf{k}} \quad \forall \mathbf{k}. \quad (3.7)$$

For a free particle, the lowest possible eigenenergy is $\epsilon_{\mathbf{k}} = 0$ for $\mathbf{k} = 0$ so that we obtain $\mu \leq 0$.

For a large volume V , the allowed vectors \mathbf{k} become dense and we can replace the sums over \mathbf{k} by integrals according to

$$\sum_{\mathbf{k}} \dots \rightarrow V \int \frac{d^3 k}{(2\pi)^3} \dots = \frac{2\pi V}{h^3} (2m)^{3/2} \int_0^{\infty} d\epsilon \sqrt{\epsilon} \dots \quad (3.8)$$

In the last equation we have used the density of states (DOS) of free particles in three dimensions. This replacement contains a fatal mistake, though. The DOS for $\epsilon = 0$ vanishes so that any particles in the state with $\mathbf{k} = 0$, $\epsilon_{\mathbf{k}} = 0$ do not contribute to the results. But that is the ground state! For $T = 0$ *all* bosons should be in this state. We thus expect incorrect results at low temperatures.

Our mistake was that Eq. (3.8) does not hold if the fraction of bosons in the $\mathbf{k} = 0$ ground state is macroscopic, i.e., if $N_0/N := \langle n_0 \rangle / N$ remains finite for large V . To correct this, we treat the $\mathbf{k} = 0$ state explicitly (the same would be necessary for any state with macroscopic occupation). We write

$$\sum_{\mathbf{k}} \dots \rightarrow \frac{2\pi V}{h^3} (2m)^{3/2} \int_0^{\infty} d\epsilon \sqrt{\epsilon} \dots + (\mathbf{k} = 0 \text{ term}). \quad (3.9)$$

Then

$$\ln \mathcal{Z} = -\frac{2\pi V}{h^3} (2m)^{3/2} \int_0^{\infty} d\epsilon \sqrt{\epsilon} \ln(1 - ye^{-\beta\epsilon}) - \ln(1 - y) \stackrel{\text{by parts}}{=} \frac{2\pi V}{h^3} (2m)^{3/2} \frac{2}{3} \beta \int_0^{\infty} d\epsilon \frac{\epsilon^{3/2}}{y^{-1} e^{\beta\epsilon} - 1} - \ln(1 - y) \quad (3.10)$$

with

$$N = \frac{2\pi V}{h^3} (2m)^{3/2} \int_0^{\infty} d\epsilon \sqrt{\epsilon} \frac{1}{y^{-1} e^{-\beta\epsilon} - 1} + \frac{1}{y^{-1} - 1} = \frac{2\pi V}{h^3} (2m)^{3/2} \int_0^{\infty} d\epsilon \frac{\sqrt{\epsilon}}{y^{-1} e^{-\beta\epsilon} - 1} + \frac{y}{1 - y}. \quad (3.11)$$

Defining

$$g_n(y) := \frac{1}{\Gamma(n)} \int_0^{\infty} dx \frac{x^{n-1}}{y^{-1} e^x - 1} \quad (3.12)$$

for $0 \leq y \leq 1$ and $n \in \mathbb{R}$, and the thermal wavelength

$$\lambda := \sqrt{\frac{h^2}{2\pi m k_B T}}, \quad (3.13)$$

we obtain

$$\ln \mathcal{Z} = \frac{V}{\lambda^3} g_{5/2}(y) - \ln(1 - y), \quad (3.14)$$

$$N = \underbrace{\frac{V}{\lambda^3} g_{3/2}(y)}_{=: N_{\epsilon}} + \underbrace{\frac{y}{1 - y}}_{=: N_0}. \quad (3.15)$$

We note the identity

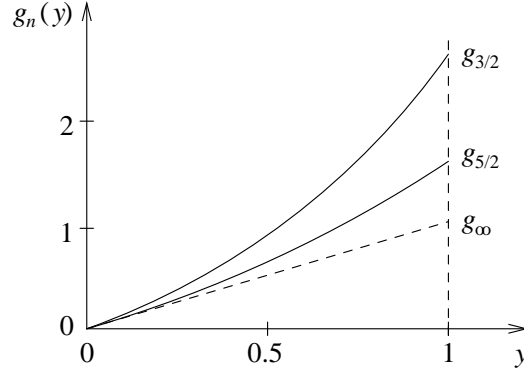
$$g_n(y) = \sum_{k=1}^{\infty} \frac{y^k}{k^n}, \quad (3.16)$$

which implies

$$g_n(0) = 0, \quad (3.17)$$

$$g_n(1) = \sum_{k=1}^{\infty} \frac{1}{k^n} = \zeta(n) \quad \text{for } n > 1 \quad (3.18)$$

with the Riemann zeta function $\zeta(x)$. Furthermore, $g_n(y)$ increases monotonically in y for $y \in [0, 1]$.



We now have to eliminate the fugacity y from Eqs. (3.14) and (3.15) to obtain \mathcal{Z} as a function of the particle number N . In Eq. (3.14), the first term is the number of particles in excited states ($\epsilon_{\mathbf{k}} > 0$), whereas the second term is the number of particles in the ground state. We consider two cases: If y is not very close to unity (specifically, if $1 - y \gg \lambda^3/V$), $N_0 = y/(1 - y)$ is on the order of unity, whereas N_{ϵ} is an *extensive* quantity. Thus N_0 can be neglected and we get

$$N \cong N_{\epsilon} = \frac{V}{\lambda^3} g_{3/2}(y). \quad (3.19)$$

Since $g_{3/2} \leq \zeta(3/2) \approx 2.612$, this equation can only be solved for the fugacity y if the concentration satisfies

$$\frac{N}{V} \leq \frac{\zeta(3/2)}{\lambda^3}. \quad (3.20)$$

To have $1 - y \gg \lambda^3/V$ in the thermodynamic limit we require, more strictly,

$$\frac{N}{V} < \frac{\zeta(3/2)}{\lambda^3}. \quad (3.21)$$

Note that $\lambda^3 \propto T^{-3/2}$ increases with decreasing temperature. Hence, at a critical temperature T_c , the inequality is no longer fulfilled. From

$$\frac{N}{V} \stackrel{!}{=} \frac{\zeta(3/2)}{\left(\frac{h^2}{2\pi m k_B T_c}\right)^{3/2}} \quad (3.22)$$

we obtain

$$k_B T_c = \frac{1}{[\zeta(3/2)]^{2/3}} \frac{h^2}{2\pi m} \left(\frac{N}{V}\right)^{2/3}. \quad (3.23)$$

If, on the other hand, y is very close to unity, N_0 cannot be neglected. Also, in this case we find

$$N_{\epsilon} = \frac{V}{\lambda^3} g_{3/2}(y) = \frac{V}{\lambda^3} g_{3/2}(1 - \mathcal{O}(\lambda^3/V)), \quad (3.24)$$

where $\mathcal{O}(\lambda^3/V)$ is a correction of order $\lambda^3/V \ll 1$. Thus, by Taylor expansion,

$$N_\epsilon = \frac{V}{\lambda^3} g_{3/2}(1) - \mathcal{O}(1) = \frac{V}{\lambda^3} \zeta(3/2) - \mathcal{O}(1). \quad (3.25)$$

The *intensive* term $\mathcal{O}(1)$ can be neglected compared to the extensive one so that

$$N_\epsilon \cong \frac{V}{\lambda^3} \zeta(3/2). \quad (3.26)$$

This is the maximum possible value at temperature T . Furthermore, $N_0 = y/(1-y)$ is solved by

$$y = \frac{N_0}{N_0 - 1} = \frac{1}{1 + 1/N_0}. \quad (3.27)$$

For y to be very close to unity, N_0 must be $N_0 \gg 1$. Since

$$N_0 = N - N_\epsilon \cong N - \frac{V}{\lambda^3} \zeta(3/2) \quad (3.28)$$

must be positive, we require

$$\frac{N}{V} > \frac{\zeta(3/2)}{\lambda^3} \quad (3.29)$$

$$\Rightarrow T < T_c. \quad (3.30)$$

We conclude that the fraction of particles in excited states is

$$\frac{N_\epsilon}{N} \cong \frac{V}{N\lambda^3} \zeta(3/2) = \frac{\lambda^3(T_c)}{\lambda^3(T)} = \left(\frac{T}{T_c}\right)^{3/2}. \quad (3.31)$$

The fraction of particles in the ground state is then

$$\frac{N_0}{N} \cong 1 - \left(\frac{T}{T_c}\right)^{3/2}. \quad (3.32)$$

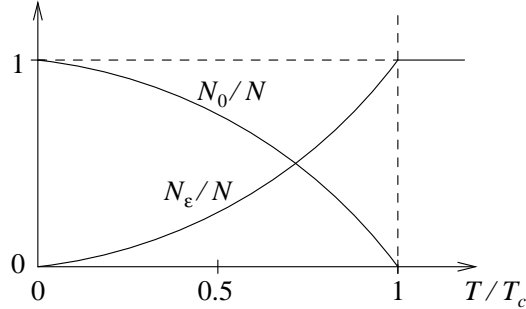
In summery, we find in the thermodynamic limit

(a) for $T > T_c$:

$$\frac{N_\epsilon}{N} \cong 1, \quad \frac{N_0}{N} \ll 1, \quad (3.33)$$

(b) for $T < T_c$:

$$\frac{N_\epsilon}{N} \cong \left(\frac{T}{T_c}\right)^{3/2}, \quad \frac{N_0}{N} \cong 1 - \left(\frac{T}{T_c}\right)^{3/2}. \quad (3.34)$$



We find a phase transition at T_c , below which a *macroscopic fraction* of the particles occupy *the same* single-particle quantum state. This fraction of particles is said to form a *condensate*. While it is remarkable that Bose-Einstein condensation happens in a *non-interacting* gas, the BEC is analogous to the condensate in strongly interacting superfluid He-4 and, with some added twists, in superfluid He-3 and in superconductors.

We can now use the partition function to derive equations of state. As an example, we consider the pressure

$$p = -\frac{\partial\phi}{\partial V} = +\frac{\partial}{\partial V} k_B T \ln \mathcal{Z} = \frac{k_B T}{\lambda^3} g_{5/2}(y) \quad (3.35)$$

(ϕ is the grand-canonical potential). We notice that only the excited states contribute to the pressure. The term $-\ln(1-y)$ from the ground state drops out since it is volume-independent. This is plausible since particles in the condensate have vanishing kinetic energy.

For $T > T_c$, we can find y and thus p numerically. For $T < T_c$ we may set $y = 1$ and obtain

$$p = \frac{k_B T}{\lambda^3} \zeta(5/2) \propto T^{5/2}. \quad (3.36)$$

Remember that for the *classical* ideal gas at constant volume we find

$$p \propto T. \quad (3.37)$$

For the BEC, the pressure drops more rapidly since more and more particles condense and thus no longer contribute to the pressure.

4

Normal metals

To be able to appreciate the remarkable properties of superconductors, it seems useful to review what we know about normal conductors.

4.1 Electrons in metals

Let us ignore electron-electron Coulomb interaction and deviations from a perfectly periodic crystal structure (due to defects or phonons) for now. Then the exact single-particle states are described by *Bloch wavefunctions*

$$\psi_{\alpha\mathbf{k}}(\mathbf{r}) = u_{\alpha\mathbf{k}}(\mathbf{r}) e^{i\mathbf{k}\cdot\mathbf{r}}, \quad (4.1)$$

where $u_{\alpha\mathbf{k}}(\mathbf{r})$ is a lattice-periodic function, α is the band index including the spin, and $\hbar\mathbf{k}$ is the crystal momentum in the first Brillouin zone. Since electrons are fermions, the average occupation number of the state $|\alpha\mathbf{k}\rangle$ with energy $\epsilon_{\alpha\mathbf{k}}$ is given by the Fermi-Dirac distribution function

$$n_F(\epsilon_{\alpha\mathbf{k}}) = \frac{1}{e^{\beta(\epsilon_{\alpha\mathbf{k}} - \mu)} + 1}. \quad (4.2)$$

If the electron number N , and not the chemical potential μ , is given, μ has to be determined from

$$N = \sum_{\alpha\mathbf{k}} \frac{1}{e^{\beta(\epsilon_{\alpha\mathbf{k}} - \mu)} + 1}, \quad (4.3)$$

cf. our discussion for ideal bosons. In the thermodynamic limit we again replace

$$\sum_{\mathbf{k}} \rightarrow V \int \frac{d^3k}{(2\pi)^3}. \quad (4.4)$$

Unlike for bosons, this is harmless for fermions, since any state can at most be occupied once so that macroscopic occupation of the single-particle ground state cannot occur. Thus we find

$$\frac{N}{V} = \sum_{\alpha} \int \frac{dk^3}{(2\pi)^3} \frac{1}{e^{\beta(\epsilon_{\alpha\mathbf{k}} - \mu)} + 1}. \quad (4.5)$$

If we lower the temperature, the Fermi function n_F becomes more and more step-like. For $T \rightarrow 0$, all states with energies $\epsilon_{\alpha\mathbf{k}} \leq E_F := \mu(T \rightarrow 0)$ are occupied (E_F is the Fermi energy), while all states with $\epsilon_{\alpha\mathbf{k}} > E_F$ are empty. This *Fermi sea* becomes fuzzy for energies $\epsilon_{\alpha\mathbf{k}} \approx E_F$ at finite temperatures but remains well defined as long as $k_B T \ll E_F$. This is the case for most materials we will discuss.

The chemical potential, the occupations $n_F(\epsilon_{\alpha\mathbf{k}})$, and thus all thermodynamic variables are *analytic* functions of T and N/V . Thus there is no phase transition, unlike for bosons. Free fermions represent a special case with only a single band with dispersion $\epsilon_{\mathbf{k}} = \hbar^2 k^2 / 2m$. If we replace m by a material-dependent effective mass, this

gives a reasonable approximation for simple metals such as alkali metals. Qualitatively, the conclusions are much more general.

Lattice imperfections and interactions result in the Bloch waves $\psi_{\alpha\mathbf{k}}(\mathbf{r})$ not being exact single-particle eigenstates. (Electron-electron and electron-phonon interactions invalidate the whole idea of single-particle states.) However, if these effects are in some sense small, they can be treated perturbatively in terms of *scattering* of electrons between single-particle states $|\alpha\mathbf{k}\rangle$.

4.2 Semiclassical theory of transport

We now want to derive an expression for the current in the presence of an applied electric field. This is a question about the response of the system to an external perturbation. There are many ways to approach this type of question. If the perturbation is small, the response, in our case the current, is expected to be a linear function of the perturbation. This is the basic assumption of linear-response theory. In the framework of many-particle theory, linear-response theory results in the *Kubo formula* (see lecture notes on many-particle theory). We here take a different route. If the external perturbation changes slowly in time and space on atomic scales, we can use a semiclassical description. Note that the following can be derived cleanly as a limit of many-particle quantum theory.

The idea is to consider the *phase space distribution function* $\rho(\mathbf{r}, \mathbf{k}, t)$. This is a classical concept. From quantum mechanics we know that \mathbf{r} and $\mathbf{p} = \hbar\mathbf{k}$ are subject to the uncertainty principle $\Delta r \Delta p \geq \hbar/2$. Thus distribution functions ρ that are localized in a phase-space volume smaller than on the order of \hbar^3 violate quantum mechanics. On the other hand, if ρ is much broader, quantum effects should be negligible.

The *Liouville theorem* shows that ρ satisfies the continuity equation

$$\frac{\partial \rho}{\partial t} + \dot{\mathbf{r}} \cdot \frac{\partial \rho}{\partial \mathbf{r}} + \dot{\mathbf{k}} \cdot \frac{\partial \rho}{\partial \mathbf{k}} \equiv \frac{d\rho}{dt} = 0 \quad (4.6)$$

(phase-space volume is conserved under the classical time evolution). Assuming for simplicity a free-particle dispersion, we have the canonical (Hamilton) equations

$$\dot{\mathbf{r}} = \frac{\partial H}{\partial \mathbf{p}} = \frac{\mathbf{p}}{m} = \frac{\hbar\mathbf{k}}{m}, \quad (4.7)$$

$$\dot{\mathbf{k}} = \frac{1}{\hbar} \dot{\mathbf{p}} = -\frac{1}{\hbar} \frac{\partial H}{\partial \mathbf{r}} = -\frac{1}{\hbar} \nabla V = \frac{1}{\hbar} \mathbf{F} \quad (4.8)$$

with the Hamiltonian H and the force \mathbf{F} . Thus we can write

$$\left(\frac{\partial}{\partial t} + \frac{\hbar\mathbf{k}}{m} \cdot \frac{\partial}{\partial \mathbf{r}} + \frac{\mathbf{F}}{\hbar} \cdot \frac{\partial}{\partial \mathbf{k}} \right) \rho = 0. \quad (4.9)$$

This equation is appropriate for particles in the absence of any scattering. For electrons in a uniform and time-independent electric field we have

$$\mathbf{F} = -e\mathbf{E}. \quad (4.10)$$

Note that we always use the convention that $e > 0$. It is easy to see that

$$\rho(\mathbf{r}, \mathbf{k}, t) = f\left(\mathbf{k} + \frac{e\mathbf{E}}{\hbar}t\right) \quad (4.11)$$

is a solution of Eq. (4.9) for *any* differentiable function f . This solution is uniform in real space ($\partial\rho/\partial\mathbf{r} \equiv 0$) and shifts to larger and larger momenta $\hbar\mathbf{k}$ for $t \rightarrow \infty$. It thus describes the *free acceleration* of electrons in an electric field. There is no finite conductivity since the current never reaches a stationary value. This is obviously not a correct description of a normal metal.

Scattering will change ρ as a function of time beyond what as already included in Eq. (4.9). We collect all processes not included in Eq. (4.9) into a *scattering term* $\mathcal{S}[\rho]$:

$$\left(\frac{\partial}{\partial t} + \frac{\hbar\mathbf{k}}{m} \cdot \frac{\partial}{\partial \mathbf{r}} + \frac{\mathbf{F}}{\hbar} \cdot \frac{\partial}{\partial \mathbf{k}} \right) \rho = -\mathcal{S}[\rho]. \quad (4.12)$$

This is the famous *Boltzmann equation*. The notation $\mathcal{S}[\rho]$ signifies that the scattering term is a *functional* of ρ . It is generally not simply a function of the local density $\rho(\mathbf{r}, \mathbf{k}, t)$ but depends on ρ everywhere and at all times, to the extent that this is consistent with causality.

While expressions for $\mathcal{S}[\rho]$ can be derived for various cases, for our purposes it is sufficient to employ the simple but common *relaxation-time approximation*. It is based on the observation that $\rho(\mathbf{r}, \mathbf{k}, t)$ should relax to thermal equilibrium if no force is applied. For fermions, the equilibrium is $\rho_0(\mathbf{k}) \propto n_F(\epsilon_{\mathbf{k}})$. This is enforced by the ansatz

$$\mathcal{S}[\rho] = \frac{\rho(\mathbf{r}, \mathbf{k}, t) - \rho_0(\mathbf{k})}{\tau}. \quad (4.13)$$

Here, τ is the relaxation time, which determines how fast ρ relaxes towards ρ_0 .

If there are different scattering mechanisms that act independently, the scattering integral is just a sum of contributions of these mechanisms,

$$\mathcal{S}[\rho] = \mathcal{S}_1[\rho] + \mathcal{S}_2[\rho] + \dots \quad (4.14)$$

Consequently, the relaxation rate $1/\tau$ can be written as (Matthiessen's rule)

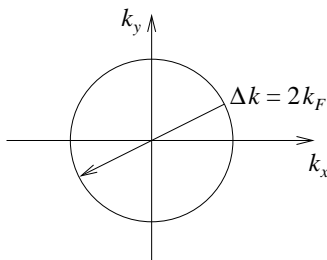
$$\frac{1}{\tau} = \frac{1}{\tau_1} + \frac{1}{\tau_2} + \dots \quad (4.15)$$

There are three main scattering mechanisms:

- Scattering of electrons by disorder: This gives an essentially temperature-independent contribution, which dominates at low temperatures.
- Electron-phonon interaction: This mechanism is strongly temperature-dependent because the available phase space shrinks at low temperatures. One finds a scattering rate

$$\frac{1}{\tau_{\text{e-ph}}} \propto T^3. \quad (4.16)$$

However, this is not the relevant rate for transport calculations. The conductivity is much more strongly affected by scattering that changes the electron momentum $\hbar\mathbf{k}$ by a lot than by processes that change it very little. *Backscattering* across the Fermi sea is most effective.



Since backscattering is additionally suppressed at low T , the relevant *transport scattering rate* scales as

$$\frac{1}{\tau_{\text{e-ph}}^{\text{trans}}} \propto T^5. \quad (4.17)$$

- Electron-electron interaction: For a parabolic free-electron band its contribution to the resistivity is actually zero since Coulomb scattering conserves the total momentum of the two scattering electrons and therefore does not degrade the current. However, in a real metal, *umklapp* scattering can take place that conserves the total momentum only modulo a reciprocal lattice vector. Thus the electron system can transfer momentum to the crystal as a whole and thereby degrade the current. The temperature dependence is typically

$$\frac{1}{\tau_{\text{e-e}}^{\text{umklapp}}} \propto T^2. \quad (4.18)$$

We now consider the force $\mathbf{F} = -e\mathbf{E}$ and calculate the current density

$$\mathbf{j}(\mathbf{r}, t) = -e \int \frac{d^3k}{(2\pi)^3} \frac{\hbar\mathbf{k}}{m} \rho(\mathbf{r}, \mathbf{k}, t). \quad (4.19)$$

To that end, we have to solve the Boltzmann equation

$$\left(\frac{\partial}{\partial t} + \frac{\hbar\mathbf{k}}{m} \cdot \frac{\partial}{\partial \mathbf{r}} - \frac{e\mathbf{E}}{\hbar} \cdot \frac{\partial}{\partial \mathbf{k}} \right) \rho = -\frac{\rho - \rho_0}{\tau}. \quad (4.20)$$

We are interested in the *stationary* solution ($\partial\rho/\partial t = 0$), which, for a uniform field, we assume to be spacially uniform ($\partial\rho/\partial\mathbf{r} = 0$). This gives

$$-\frac{e\mathbf{E}}{\hbar} \cdot \frac{\partial}{\partial \mathbf{k}} \rho(\mathbf{k}) = \frac{\rho_0(\mathbf{k}) - \rho(\mathbf{k})}{\tau} \quad (4.21)$$

$$\Rightarrow \rho(\mathbf{k}) = \rho_0(\mathbf{k}) + \frac{e\mathbf{E}\tau}{\hbar} \cdot \frac{\partial}{\partial \mathbf{k}} \rho(\mathbf{k}). \quad (4.22)$$

We iterate this equation by inserting it again into the final term:

$$\rho(\mathbf{k}) = \rho_0(\mathbf{k}) + \frac{e\mathbf{E}\tau}{\hbar} \cdot \frac{\partial}{\partial \mathbf{k}} \rho_0(\mathbf{k}) + \left(\frac{e\mathbf{E}\tau}{\hbar} \cdot \frac{\partial}{\partial \mathbf{k}} \right) \frac{e\mathbf{E}\tau}{\hbar} \cdot \frac{\partial}{\partial \mathbf{k}} \rho(\mathbf{k}). \quad (4.23)$$

To make progress, we assume that the applied field \mathbf{E} is small so that the response \mathbf{j} is *linear* in \mathbf{E} . Under this assumption we can truncate the iteration after the linear term,

$$\rho(\mathbf{k}) = \rho_0(\mathbf{k}) + \frac{e\mathbf{E}\tau}{\hbar} \cdot \frac{\partial}{\partial \mathbf{k}} \rho_0(\mathbf{k}). \quad (4.24)$$

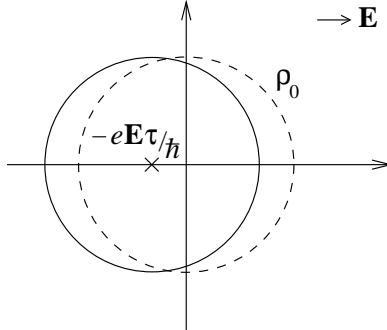
By comparing this to the Taylor expansion

$$\rho_0\left(\mathbf{k} + \frac{e\mathbf{E}\tau}{\hbar}\right) = \rho_0(\mathbf{k}) + \frac{e\mathbf{E}\tau}{\hbar} \cdot \frac{\partial}{\partial \mathbf{k}} \rho_0(\mathbf{k}) + \dots \quad (4.25)$$

we see that the solution is, to linear order in \mathbf{E} ,

$$\rho(\mathbf{k}) = \rho_0\left(\mathbf{k} + \frac{e\mathbf{E}\tau}{\hbar}\right) \propto n_F(\epsilon_{\mathbf{k}+e\mathbf{E}\tau/\hbar}). \quad (4.26)$$

Thus the distribution function is simply shifted in \mathbf{k} -space by $-e\mathbf{E}\tau/\hbar$. Since electrons carry negative charge, the distribution is shifted in the direction *opposite* to the applied electric field.



The current density now reads

$$\mathbf{j} = -e \int \frac{d^3k}{(2\pi)^3} \frac{\hbar\mathbf{k}}{m} \rho_0\left(\mathbf{k} + \frac{e\mathbf{E}\tau}{\hbar}\right) \cong \underbrace{-e \int \frac{d^3k}{(2\pi)^3} \frac{\hbar\mathbf{k}}{m} \rho_0(\mathbf{k})}_{=0} - e \int \frac{d^3k}{(2\pi)^3} \frac{\hbar\mathbf{k}}{m} \frac{e\mathbf{E}\tau}{\hbar} \cdot \frac{\partial \rho_0}{\partial \mathbf{k}}. \quad (4.27)$$

The first term is the current density in equilibrium, which vanishes. In components, we have

$$j_\alpha = -\frac{e^2\tau}{m} \sum_\beta E_\beta \int \frac{d^3k}{(2\pi)^3} k_\alpha \frac{\partial \rho_0}{\partial k_\beta} \stackrel{\text{by parts}}{=} +\frac{e^2\tau}{m} \sum_\beta E_\beta \int \frac{d^3k}{(2\pi)^3} \underbrace{\frac{\partial k_\alpha}{\partial k_\beta}}_{=\delta_{\alpha\beta}} \rho_0 = \frac{e^2\tau}{m} E_\alpha \underbrace{\int \frac{d^3k}{(2\pi)^3} \rho_0}_{=n}. \quad (4.28)$$

Here, the integral is the concentration of electrons in real space, $n := N/V$. We finally obtain

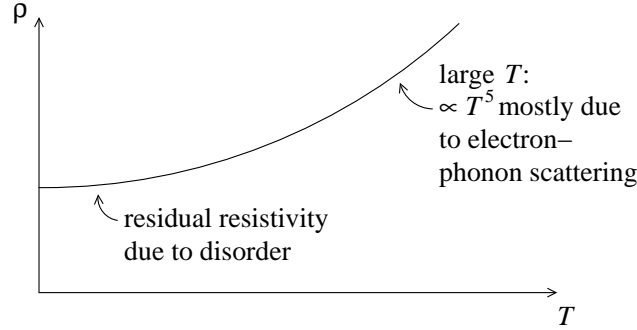
$$\mathbf{j} = \frac{e^2 n \tau}{m} \mathbf{E} \stackrel{!}{=} \sigma \mathbf{E} \quad (4.29)$$

so that the conductivity is

$$\sigma = \frac{e^2 n \tau}{m}. \quad (4.30)$$

This is the famous *Drude formula*. For the resistivity $\rho = 1/\sigma$ we get, based on our discussion of scattering mechanisms,

$$\rho = \frac{m}{e^2 n \tau} = \frac{m}{e^2 n} \left(\frac{1}{\tau_{\text{dis}}} + \frac{1}{\tau_{\text{e-ph}}^{\text{transport}}} + \frac{1}{\tau_{\text{e-e}}^{\text{umklapp}}} \right). \quad (4.31)$$



5

Electrodynamics of superconductors

Superconductors are defined by electrodynamic properties—ideal conduction and magnetic-field expulsion. It is thus appropriate to ask how these materials can be described within the formal framework of electrodynamics.

5.1 London theory

In 1935, F. and H. London proposed a phenomenological theory for the electrodynamic properties of superconductors. It is based on a *two-fluid picture*: For unspecified reasons, the electrons form a normal fluid of concentration n_n and a superfluid of concentration n_s , where $n_n + n_s = n = N/V$. Such a picture seemed quite plausible based on Einstein's theory of Bose-Einstein condensation, although nobody understood how the fermionic electrons could form a superfluid. The normal fluid is postulated to behave normally, i.e., to carry an ohmic current

$$\mathbf{j}_n = \sigma_n \mathbf{E} \quad (5.1)$$

governed by the Drude law

$$\sigma_n = \frac{e^2 n_n \tau}{m}. \quad (5.2)$$

The superfluid is assumed to be insensitive to scattering. As noted in section 4.2, this leads to a free acceleration of the charges. With the supercurrent

$$\mathbf{j}_s = -e n_s \mathbf{v}_s \quad (5.3)$$

and Newton's equation of motion

$$\frac{d}{dt} \mathbf{v}_s = \frac{\mathbf{F}}{m} = -\frac{e\mathbf{E}}{m}, \quad (5.4)$$

we obtain

$$\frac{\partial \mathbf{j}_s}{\partial t} = \frac{e^2 n_s}{m} \mathbf{E}. \quad (5.5)$$

This is the *First London Equation*. We have assumed that n_n and n_s are both *uniform* (constant in space) and *stationary* (constant in time). This is a serious restriction of London theory, which will only be overcome by Ginzburg-Landau theory.

Note that the curl of the First London Equation is

$$\frac{\partial}{\partial t} \nabla \times \mathbf{j}_s = \frac{e^2 n_s}{m} \nabla \times \mathbf{E} = -\frac{e^2 n_s}{mc} \frac{\partial \mathbf{B}}{\partial t}. \quad (5.6)$$

This can be integrated in time to give

$$\nabla \times \mathbf{j}_s = -\frac{e^2 n_s}{mc} \mathbf{B} + \mathbf{C}(\mathbf{r}), \quad (5.7)$$

where the last term represents a constant of integration at each point \mathbf{r} inside the superconductor. $\mathbf{C}(\mathbf{r})$ should be determined from the initial conditions. If we start from a superconducting body in zero applied magnetic

field, we have $\mathbf{j}_s \equiv 0$ and $\mathbf{B} \equiv 0$ initially so that $\mathbf{C}(\mathbf{r}) = 0$. To describe the Meißner-Ochsenfeld effect, we have to consider the case of a body becoming superconducting (by cooling) in a non-zero applied field. However, this case cannot be treated within London theory since we here assume the superfluid density n_s to be constant in time.

To account for the flux expulsion, the Londons *postulated* that $\mathbf{C} \equiv 0$ regardless of the history of the system. This leads to

$$\nabla \times \mathbf{j}_s = -\frac{e^2 n_s}{mc} \mathbf{B}, \quad (5.8)$$

the *Second London Equation*.

Together with Ampère's Law

$$\nabla \times \mathbf{B} = \frac{4\pi}{c} \mathbf{j}_s + \frac{4\pi}{c} \mathbf{j}_n \quad (5.9)$$

(there is no displacement current in the stationary state) we get

$$\nabla \times \nabla \times \mathbf{B} = -\frac{4\pi e^2 n_s}{mc^2} \mathbf{B} + \frac{4\pi}{c} \sigma_n \nabla \times \mathbf{E} = -\frac{4\pi e^2 n_s}{mc^2} \mathbf{B} - \frac{4\pi}{c} \sigma_n \frac{\partial \mathbf{B}}{\partial t}. \quad (5.10)$$

We drop the last term since we are interested in the stationary state and use an identity from vector calculus,

$$-\nabla(\nabla \cdot \mathbf{B}) + \nabla^2 \mathbf{B} = \frac{4\pi e^2 n_s}{mc^2} \mathbf{B}. \quad (5.11)$$

Introducing the *London penetration depth*

$$\lambda_L := \sqrt{\frac{mc^2}{4\pi e^2 n_s}}, \quad (5.12)$$

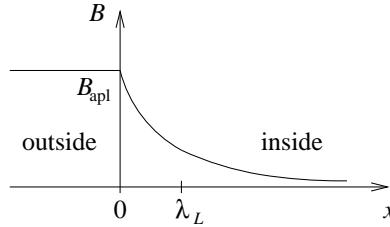
this equation assumes the simple form

$$\nabla^2 \mathbf{B} = \frac{1}{\lambda_L^2} \mathbf{B}. \quad (5.13)$$

Let us consider a semi-infinite superconductor filling the half space $x > 0$. A magnetic field $\mathbf{B}_{\text{apl}} = \mathbf{H}_{\text{apl}} = B_{\text{apl}} \hat{y}$ is applied parallel to the surface. One can immediately see that the equation is solved by

$$\mathbf{B}(x) = B_{\text{apl}} \hat{y} e^{-x/\lambda_L} \quad \text{for } x \geq 0. \quad (5.14)$$

The magnetic field thus decreases exponentially with the distance from the surface of the superconductor. In bulk we indeed find $\mathbf{B} \rightarrow 0$.



The Second London Equation

$$\nabla \times \mathbf{j}_s = -\frac{c}{4\pi\lambda_L^2} \mathbf{B} \quad (5.15)$$

and the continuity equation

$$\nabla \cdot \mathbf{j}_s = 0 \quad (5.16)$$

can now be solved to give

$$\mathbf{j}_s(x) = -\frac{c}{4\pi\lambda_L} B_{\text{apl}} \hat{z} e^{-x/\lambda_L} \quad \text{for } x \geq 0. \quad (5.17)$$

Thus the supercurrent flows in the direction parallel to the surface and perpendicular to \mathbf{B} and decreases into the bulk on the same scale λ_L . \mathbf{j}_s can be understood as the *screening current* required to keep the magnetic field out of the bulk of the superconductor.

The two London equations (5.5) and (5.8) can be summarized using the vector potential:

$$\mathbf{j}_s = -\frac{e^2 n_s}{mc} \mathbf{A}. \quad (5.18)$$

This equation is evidently not gauge-invariant since a change of gauge

$$\mathbf{A} \rightarrow \mathbf{A} + \nabla\chi \quad (5.19)$$

changes the supercurrent. (The whole London theory is gauge-invariant since it is expressed in terms of \mathbf{E} and \mathbf{B} .) Charge conservation requires $\nabla \cdot \mathbf{j}_s = 0$ and thus the vector potential must be transverse,

$$\nabla \cdot \mathbf{A} = 0. \quad (5.20)$$

This is called the *London gauge*. Furthermore, the supercurrent through the surface of the superconducting region is proportional to the normal component A_\perp . For a simply connected region these conditions uniquely determine $\mathbf{A}(\mathbf{r})$. For a multiply connected region this is not the case; we will return to this point below.

5.2 Rigidity of the superfluid state

F. London has given a quantum-mechanical justification of the London equations. If the many-body wavefunction of the electrons forming the superfluid is $\Psi_s(\mathbf{r}_1, \mathbf{r}_2, \dots)$ then the supercurrent in the presence of a vector potential \mathbf{A} is

$$\mathbf{j}_s(\mathbf{r}) = -e \frac{1}{2m} \sum_j \int d^3r_1 d^3r_2 \cdots \delta(\mathbf{r} - \mathbf{r}_j) \left[\Psi_s^* \left(\mathbf{p}_j + \frac{e}{c} \mathbf{A}(\mathbf{r}_j) \right) \Psi_s + \Psi_s \left(\mathbf{p}_j^\dagger + \frac{e}{c} \mathbf{A}(\mathbf{r}_j) \right) \Psi_s^* \right]. \quad (5.21)$$

Here, j sums over all electrons in the superfluid, which have position \mathbf{r}_j and momentum operator

$$\mathbf{p}_j = \frac{\hbar}{i} \frac{\partial}{\partial \mathbf{r}_j}. \quad (5.22)$$

Making \mathbf{p}_j explicit and using the London gauge, we obtain

$$\begin{aligned} \mathbf{j}_s(\mathbf{r}) &= -\frac{e\hbar}{2mi} \sum_j \int d^3r_1 d^3r_2 \cdots \delta(\mathbf{r} - \mathbf{r}_j) \left[\Psi_s^* \frac{\partial}{\partial \mathbf{r}_j} \Psi_s - \Psi_s \frac{\partial}{\partial \mathbf{r}_j} \Psi_s^* \right] \\ &\quad - \frac{e^2}{mc} \mathbf{A}(\mathbf{r}) \sum_j \int d^3r_1 d^3r_2 \cdots \delta(\mathbf{r} - \mathbf{r}_j) \Psi_s^* \Psi_s \\ &= -\frac{e\hbar}{2mi} \sum_j \int d^3r_1 d^3r_2 \cdots \delta(\mathbf{r} - \mathbf{r}_j) \left[\Psi_s^* \frac{\partial}{\partial \mathbf{r}_j} \Psi_s - \Psi_s \frac{\partial}{\partial \mathbf{r}_j} \Psi_s^* \right] - \frac{e^2 n_s}{mc} \mathbf{A}(\mathbf{r}). \end{aligned} \quad (5.23)$$

Now London proposed that the wavefunction Ψ_s is *rigid* under the application of a transverse vector potential. More specifically, he suggested that Ψ_s does not contain a term of first order in \mathbf{A} , provided $\nabla \cdot \mathbf{A} = 0$. Then, to first order in \mathbf{A} , the first term on the right-hand side in Eq. (5.23) contains the unperturbed wavefunction one would obtain for $\mathbf{A} = 0$. The first term is thus the supercurrent for $\mathbf{A} \equiv 0$, which should vanish due to Ampère's law. Consequently, to first order in \mathbf{A} we obtain the London equation

$$\mathbf{j}_s = -\frac{e^2 n_s}{mc} \mathbf{A}. \quad (5.24)$$

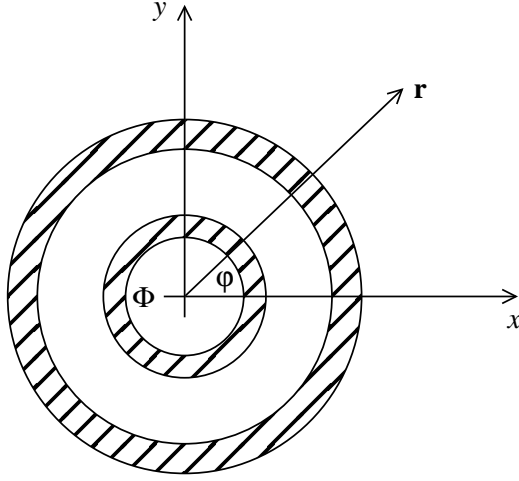
The rigidity of Ψ_s was later understood in the framework of BCS theory as resulting from the existence of a non-zero energy gap for excitations out of the superfluid state.

5.3 Flux quantization

We now consider two concentric superconducting cylinders that are thick compared to the London penetration depth λ_L . A magnetic flux

$$\Phi = \int_{\odot} d^2r B_{\perp} \quad (5.25)$$

penetrates the inner hole and a thin surface layer on the order of λ_L of the inner cylinder. The only purpose of the inner cylinder is to prevent the magnetic field from touching the outer cylinder, which we are really interested in. The outer cylinder is completely field-free. We want to find the possible values of the flux Φ .



Although the region outside of the inner cylinder has $\mathbf{B} = 0$, the vector potential does not vanish. The relation $\nabla \times \mathbf{A} = \mathbf{B}$ implies

$$\oint ds \cdot \mathbf{A} = \iint d^2r \mathbf{B} = \Phi. \quad (5.26)$$

By symmetry, the tangential part of \mathbf{A} is

$$A_{\varphi} = \frac{\Phi}{2\pi r}. \quad (5.27)$$

The London gauge requires this to be the only non-zero component. Thus outside of the inner cylinder we have, in cylindrical coordinates,

$$\mathbf{A} = \frac{\Phi}{2\pi r} \hat{\varphi} = \nabla \frac{\Phi\varphi}{2\pi}. \quad (5.28)$$

Since this is a pure gradient, we can get from $\mathbf{A} = 0$ to $\mathbf{A} = (\Phi/2\pi r) \hat{\varphi}$ by a gauge transformation

$$\mathbf{A} \rightarrow \mathbf{A} + \nabla\chi \quad (5.29)$$

with

$$\chi = \frac{\Phi\varphi}{2\pi}. \quad (5.30)$$

χ is continuous but multivalued outside of the inner cylinder. We recall that a gauge transformation of \mathbf{A} must be accompanied by a transformation of the wavefunction,

$$\Psi_s \rightarrow \exp\left(-\frac{ie}{\hbar c} \sum_j \chi(\mathbf{r}_j)\right) \Psi_s. \quad (5.31)$$

This is most easily seen by noting that this guarantees the current in Sec. 5.2 to remain invariant under gauge transformations. Thus the wavefunction at $\Phi = 0$ ($\mathbf{A} = 0$) and at non-zero flux Φ are related by

$$\Psi_s^{\Phi} = \exp\left(-\frac{ie}{\hbar c} \sum_j \frac{\Phi\varphi_j}{2\pi}\right) \Psi_s^0 = \exp\left(-i\frac{e}{\hbar c} \Phi \sum_j \varphi_j\right) \Psi_s^0, \quad (5.32)$$

where φ_j is the polar angle of electron j . For Ψ_s^Φ as well as Ψ_s^0 to be single-valued and continuous, the exponential factor must not change for $\varphi_j \rightarrow \varphi_j + 2\pi$ for any j . This is the case if

$$\frac{e}{hc} \Phi \in \mathbb{Z} \quad \Leftrightarrow \quad \Phi = n \frac{hc}{e} \quad \text{with } n \in \mathbb{Z}. \quad (5.33)$$

We find that the magnetic flux Φ is quantized in units of hc/e . Note that the inner cylinder can be dispensed with: Assume we are heating it enough to become normal-conducting. Then the flux Φ will fill the whole interior of the *outer* cylinder plus a thin (on the order of λ_L) layer on its inside. But if the outer cylinder is much thicker than λ_L , this should not affect Ψ_s appreciably, away from this thin layer.

The quantum hc/e is actually not correct. Based in the idea that *two* electrons could form a boson that could Bose-Einstein condense, Onsager suggested that the relevant charge is $2e$ instead of e , leading to the superconducting flux quantum

$$\Phi_0 := \frac{hc}{2e}, \quad (5.34)$$

which is indeed found in experiments.

5.4 Nonlocal response: Pippard theory

Experiments often find a magnetic penetration depth λ that is significantly larger than the Londons' prediction λ_L , in particular in dirty samples with large scattering rates $1/\tau$ in the normal state. Pippard explained this on the basis of a nonlocal electromagnetic response of the superconductor. The underlying idea is that the quantum state of the electrons forming the superfluid cannot be arbitrarily localized. The typical energy scale of superconductivity is expected to be $k_B T_c$. Only electrons with energies ϵ within $\sim k_B T_c$ of the Fermi energy can contribute appreciably. This corresponds to a momentum range Δp determined by

$$\frac{k_B T_c}{\Delta p} \stackrel{!}{=} \left. \frac{\partial \epsilon}{\partial p} \right|_{\epsilon=E_F} = \left. \frac{p}{m} \right|_{\epsilon=E_F} = v_F \quad (5.35)$$

$$\Rightarrow \quad \Delta p = \frac{k_B T_c}{v_F} \quad (5.36)$$

with the Fermi velocity v_F . From this, we can estimate that the electrons cannot be localized on scales smaller than

$$\Delta x \approx \frac{\hbar}{\Delta p} = \frac{\hbar v_F}{k_B T_c}. \quad (5.37)$$

Therefore, Pippard introduced the *coherence length*

$$\xi_0 = \alpha \frac{\hbar v_F}{k_B T_c} \quad (5.38)$$

as a measure of the minimum extent of electronic wavepackets. α is a numerical constant of order unity. BCS theory predicts $\alpha \approx 0.180$. Pippard proposed to replace the local equation

$$\mathbf{j}_s = -\frac{e^2 n_s}{mc} \mathbf{A} \quad (5.39)$$

from London theory by the nonlocal

$$\mathbf{j}_s(\mathbf{r}) = -\frac{3}{4\pi\xi_0} \frac{e^2 n_s}{mc} \int d^3 r' \frac{\boldsymbol{\rho}(\boldsymbol{\rho} \cdot \mathbf{A}(\mathbf{r}'))}{\rho^4} e^{-\rho/\xi_0} \quad (5.40)$$

with

$$\boldsymbol{\rho} = \mathbf{r} - \mathbf{r}'. \quad (5.41)$$

The special form of this equation was motivated by an earlier nonlocal generalization of Ohm's law. The main point is that electrons within a distance ξ_0 of the point \mathbf{r}' where the field \mathbf{A} acts have to respond to it because of the stiffness of the wavefunction. If \mathbf{A} does not change appreciably on the scale of ξ_0 , we obtain

$$\mathbf{j}_s(\mathbf{r}) \cong -\frac{3}{4\pi\xi_0} \frac{e^2 n_s}{mc} \int d^3 r' \frac{\boldsymbol{\rho}(\boldsymbol{\rho} \cdot \mathbf{A}(\mathbf{r}))}{\rho^4} e^{-\rho/\xi_0} = -\frac{3}{4\pi\xi_0} \frac{e^2 n_s}{mc} \int d^3 \rho \frac{\boldsymbol{\rho}(\boldsymbol{\rho} \cdot \mathbf{A}(\mathbf{r}))}{\rho^4} e^{-\rho/\xi_0}. \quad (5.42)$$

The result has to be parallel to $\mathbf{A}(\mathbf{r})$ by symmetry (since it is a vector depending on a single vector $\mathbf{A}(\mathbf{r})$). Thus

$$\begin{aligned} \mathbf{j}_s(\mathbf{r}) &= -\frac{3}{4\pi\xi_0} \frac{e^2 n_s}{mc} \mathbf{A}(\mathbf{r}) \int d^3 \rho \frac{(\boldsymbol{\rho} \cdot \hat{\mathbf{A}}(\mathbf{r}))^2}{\rho^4} e^{-\rho/\xi_0} \\ &= -\frac{3}{4\pi\xi_0} \frac{e^2 n_s}{mc} \mathbf{A}(\mathbf{r}) 2\pi \int_{-1}^1 d(\cos \theta) \int_0^\infty d\rho \rho^2 \frac{\rho^2 \cos^2 \theta}{\rho^4} e^{-\rho/\xi_0} \\ &= -\frac{3}{4\pi\xi_0} \frac{e^2 n_s}{mc} \mathbf{A}(\mathbf{r}) 2\pi \frac{2}{3} \xi_0 = -\frac{e^2 n_s}{mc} \mathbf{A}(\mathbf{r}). \end{aligned} \quad (5.43)$$

We recover the local London equation. However, for many conventional superconductors, ξ_0 is much larger than λ . Then $\mathbf{A}(\mathbf{r})$ drops to zero on a length scale of $\lambda \ll \xi_0$. According to Pippard's equation, the electrons respond to the vector potential averaged over regions of size ξ_0 . This averaged field is much smaller than \mathbf{A} at the surface so that the screening current \mathbf{j}_s is strongly reduced and the magnetic field penetrates much deeper than predicted by London theory, i.e., $\lambda \gg \lambda_L$.

The above motivation for the coherence length ξ_0 relied on having a clean system. In the presence of strong scattering the electrons can be localized on the scale of the mean free path $l := v_F \tau$. Pippard phenomenologically generalized the equation for \mathbf{j}_s by introducing a new length ξ where

$$\frac{1}{\xi} = \frac{1}{\xi_0} + \frac{1}{\beta l} \quad (5.44)$$

(β is a numerical constant of order unity) and writing

$$\mathbf{j}_s = -\frac{3}{4\pi\xi_0} \frac{e^2 n_s}{mc} \int d^3 r' \frac{\boldsymbol{\rho}(\boldsymbol{\rho} \cdot \mathbf{A}(\mathbf{r}'))}{\rho^4} e^{-\rho/\xi}. \quad (5.45)$$

Note that ξ_0 appears in the prefactor but ξ in the exponential. This expression is in good agreement with experiments for series of samples with varying disorder. It is essentially the same as the result of BCS theory. Also note that in the *dirty limit* $l \ll \xi_0, \lambda_L$ we again recover the local London result following the same argument as above, but since the integral gives a factor of ξ , which is not canceled by the prefactor $1/\xi_0$, the current is reduced to

$$\mathbf{j}_s = -\frac{e^2 n_s}{mc} \frac{\xi}{\xi_0} \mathbf{A} \cong -\frac{e^2 n_s}{mc} \frac{\beta l}{\xi_0} \mathbf{A}. \quad (5.46)$$

Taking the curl, we obtain

$$\nabla \times \mathbf{j}_s = -\frac{e^2 n_s}{mc} \frac{\beta l}{\xi_0} \mathbf{B} \quad (5.47)$$

$$\Rightarrow \nabla \times \nabla \times \mathbf{B} = -\frac{4\pi e^2 n_s}{mc^2} \frac{\beta l}{\xi_0} \mathbf{B} \quad (5.48)$$

$$\Rightarrow \nabla^2 \mathbf{B} = \frac{4\pi e^2 n_s}{mc^2} \frac{\beta l}{\xi_0} \mathbf{B}, \quad (5.49)$$

in analogy to the derivation in Sec. 5.1. This equation is of the form

$$\nabla^2 \mathbf{B} = \frac{1}{\lambda^2} \mathbf{B} \quad (5.50)$$

with the penetration depth

$$\lambda = \sqrt{\frac{mc^2}{4\pi e^2 n_s}} \sqrt{\frac{\xi_0}{\beta l}} = \lambda_L \sqrt{\frac{\xi_0}{\beta l}}. \quad (5.51)$$

Thus the penetration depth is increased by a factor of order $\sqrt{\xi_0/l}$ in the dirty limit, $l \ll \xi_0$.

6

Ginzburg-Landau theory

Within London and Pippard theory, the superfluid density n_s is treated as given. There is no way to understand the dependence of n_s on, for example, temperature or applied magnetic field within these theories. Moreover, n_s has been assumed to be constant in time and uniform in space—an assumption that is expected to fail close to the surface of a superconductor.

These deficiencies are cured by the Ginzburg-Landau theory put forward in 1950. Like Ginzburg and Landau we ignore complications due to the nonlocal electromagnetic response. Ginzburg-Landau theory is developed as a generalization of London theory, not of Pippard theory. The starting point is the much more general and very powerful Landau theory of phase transitions, which we will review first.

6.1 Landau theory of phase transitions

Landau introduced the concept of the *order parameter* to describe phase transitions. In this context, an order parameter is a thermodynamic variable that is zero on one side of the transition and non-zero on the other. In ferromagnets, the magnetization \mathbf{M} is the order parameter. The theory neglects fluctuations, which means that the order parameter is assumed to be constant in time and space. Landau theory is thus a *mean-field theory*. Now the appropriate thermodynamic potential can be written as a function of the order parameter, which we call Δ , and certain other thermodynamic quantities such as pressure or volume, magnetic field, etc. We will always call the potential the *free energy* F , but whether it really is a free energy, a free enthalpy, or something else depends on which quantities are given (pressure vs. volume etc.). Hence, we write

$$F = F(\Delta, T), \tag{6.1}$$

where T is the temperature, and further variables have been suppressed. The equilibrium state at temperature T is the one that minimizes the free energy. Generally, we do not know $F(\Delta, T)$ explicitly. Landau's idea was to expand F in terms of Δ , including only those terms that are allowed by the symmetry of the system and keeping the minimum number of the simplest terms required to get non-trivial results.

For example, in an isotropic ferromagnet, the order parameter is the three-component vector \mathbf{M} . The free energy must be invariant under rotations of \mathbf{M} because of isotropy. Furthermore, since we want to minimize F as a function of \mathbf{M} , F should be differentiable in \mathbf{M} . Then the leading terms, apart from a trivial constant, are

$$F \cong \alpha \mathbf{M} \cdot \mathbf{M} + \frac{\beta}{2} (\mathbf{M} \cdot \mathbf{M})^2 + \mathcal{O}((\mathbf{M} \cdot \mathbf{M})^3). \tag{6.2}$$

Denoting the coefficients by α and $\beta/2$ is just convention. α and β are functions of temperature (and pressure etc.).

What is the corresponding expansion for a superconductor or superfluid? Lacking a microscopic theory, Ginzburg and Landau assumed based on the analogy with Bose-Einstein condensation that the superfluid part is described by a single one-particle wave function $\Psi_s(\mathbf{r})$. They imposed the plausible normalization

$$\int d^3r |\Psi_s(\mathbf{r})|^2 = N_s = n_s V; \tag{6.3}$$

N_s is the total number of particles in the condensate. They then chose the complex amplitude ψ of $\Psi_s(\mathbf{r})$ as the order parameter, with the normalization

$$|\psi|^2 \propto n_s. \quad (6.4)$$

Thus the order parameter in this case is a complex number. They thereby neglect the spatial variation of $\Psi_s(\mathbf{r})$ on an atomic scale.

The free energy must not depend on the global phase of $\Psi_s(\mathbf{r})$ because the global phase of quantum states is not observable. Thus we obtain the expansion

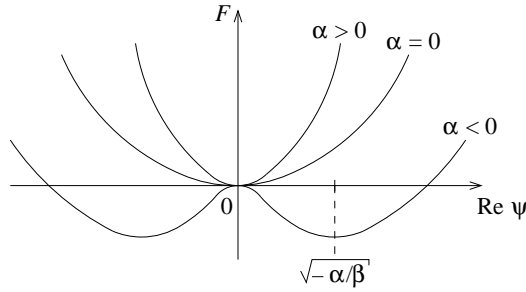
$$F = \alpha |\psi|^2 + \frac{\beta}{2} |\psi|^4 + \mathcal{O}(|\psi|^6). \quad (6.5)$$

Only the absolute value $|\psi|$ appears since F is real. Odd powers are excluded since they are not differentiable at $\psi = 0$. If $\beta > 0$, which is not guaranteed by symmetry but is the case for superconductors and superfluids, we can neglect higher order terms since $\beta > 0$ then makes sure that $F(\psi)$ is bounded from below. Now there are two cases:

- If $\alpha \geq 0$, $F(\psi)$ has a single minimum at $\psi = 0$. Thus the equilibrium state has $n_s = 0$. This is clearly a normal metal ($n_n = n$) or a normal fluid.
- If $\alpha < 0$, $F(\psi)$ has a ring of minima with equal amplitude (modulus) $|\psi|$ but arbitrary phase. We easily see

$$\frac{\partial F}{\partial |\psi|} = 2\alpha |\psi| + 2\beta |\psi|^3 = 0 \quad \Rightarrow \quad |\psi| = 0 \text{ (this is a maximum) or } |\psi| = \sqrt{-\frac{\alpha}{\beta}} \quad (6.6)$$

Note that the radicand is positive.



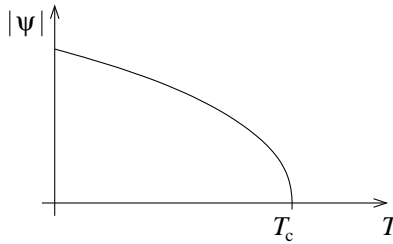
Imagine this figure rotated around the vertical axis to find F as a function of the complex ψ . $F(\psi)$ for $\alpha < 0$ is often called the “Mexican-hat potential.” In Landau theory, the phase transition clearly occurs when $\alpha = 0$. Since then $T = T_c$ by definition, it is useful to expand α and β to leading order in T around T_c . Hence,

$$\alpha \cong \alpha'(T - T_c), \quad \alpha' > 0, \quad (6.7)$$

$$\beta \cong \text{const.} \quad (6.8)$$

Then the order parameter below T_c satisfies

$$|\psi| = \sqrt{-\frac{\alpha'(T - T_c)}{\beta}} = \sqrt{\frac{\alpha'}{\beta}} \sqrt{T - T_c}. \quad (6.9)$$



Note that the expansion of F up to fourth order is only justified as long as ψ is small. The result is thus limited to temperatures not too far below T_c . The scaling $|\psi| \propto (T - T_c)^{1/2}$ is characteristic for mean-field theories. All solutions with this value of $|\psi|$ minimize the free energy. In a given experiment, only one of them is realized. This equilibrium state has an order parameter $\psi = |\psi| e^{i\phi}$ with some fixed phase ϕ . This state is clearly not invariant under rotations of the phase. We say that the *global* U(1) *symmetry* of the system is spontaneously broken since the free energy F has it but the particular equilibrium state does not. It is called U(1) symmetry since the group U(1) of unitary 1×1 matrices just contains phase factors $e^{i\phi}$.

Specific heat

Since we now know the mean-field free energy as a function of temperature, we can calculate further thermodynamic variables. In particular, the free entropy is

$$S = -\frac{\partial F}{\partial T}. \quad (6.10)$$

Since the expression for F used above only includes the contributions of superconductivity or superfluidity, the entropy calculated from it will also only contain these contributions. For $T \geq T_c$, the mean-field free energy is $F(\psi = 0) = 0$ and thus we find $S = 0$. For $T < T_c$ we instead obtain

$$\begin{aligned} S &= -\frac{\partial}{\partial T} F \left(\sqrt{-\frac{\alpha}{\beta}} \right) = -\frac{\partial}{\partial T} \left(-\frac{\alpha^2}{\beta} + \frac{1}{2} \frac{\alpha^2}{\beta} \right) = \frac{\partial}{\partial T} \frac{\alpha^2}{2\beta} \\ &\cong \frac{\partial}{\partial T} \frac{(\alpha')^2}{2\beta} (T - T_c)^2 = \frac{(\alpha')^2}{\beta} (T - T_c) = -\frac{(\alpha')^2}{\beta} (T_c - T) < 0. \end{aligned} \quad (6.11)$$

We find that the entropy is continuous at $T = T_c$. By definition, this means that the phase transition is *continuous*, i.e., not of first order. The heat capacity of the superconductor or superfluid is

$$C = T \frac{\partial S}{\partial T}, \quad (6.12)$$

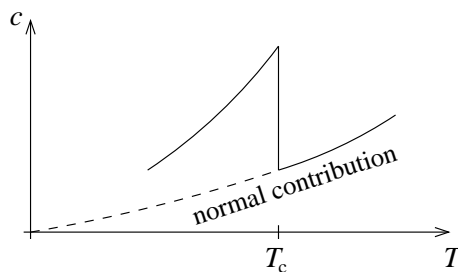
which equals zero for $T \geq T_c$ but is

$$C = \frac{(\alpha')^2}{\beta} T \quad (6.13)$$

for $T < T_c$. Thus the heat capacity has a *jump discontinuity* of

$$\Delta C = -\frac{(\alpha')^2}{\beta} T_c \quad (6.14)$$

at T_c . Adding the other contributions, which are analytic at T_c , the specific heat $c := C/V$ is sketched here:



Recall that Landau theory only works close to T_c .

6.2 Ginzburg-Landau theory for neutral superfluids

To be able to describe also spatially non-uniform situations, Ginzburg and Landau had to go beyond the Landau description for a constant order parameter. To do so, they included *gradients*. We will first discuss the simpler case of a superfluid of electrically neutral particles (think of He-4). We define a macroscopic condensate wave function $\psi(\mathbf{r})$, which is essentially given by $\Psi_s(\mathbf{r})$ averaged over length scales large compared to atomic distances. We expect any spatial changes of $\psi(\mathbf{r})$ to cost energy—this is analogous to the energy of domain walls in magnetic systems. In the spirit of Landau theory, Ginzburg and Landau included the simplest term containing gradients of ψ and allowed by symmetry into the free energy

$$F[\psi] = \int d^3r \left[\alpha |\psi|^2 + \frac{\beta}{2} |\psi|^4 + \gamma (\nabla\psi)^* \cdot \nabla\psi \right], \quad (6.15)$$

where we have changed the definitions of α and β slightly. We require $\gamma > 0$ so that the system does not spontaneously become highly non-uniform. The above expression is a *functional* of ψ , also called the “*Landau functional*.” Calling it a “free energy” is really an abuse of language since the free energy proper is only the value assumed by $F[\psi]$ at its minimum.

If we interpret $\psi(\mathbf{r})$ as the (coarse-grained) condensate wavefunction, it is natural to identify the gradient term as a kinetic energy by writing

$$F[\psi] \cong \int d^3r \left[\alpha |\psi|^2 + \frac{\beta}{2} |\psi|^4 + \frac{1}{2m^*} \left| \frac{\hbar}{i} \nabla\psi \right|^2 \right], \quad (6.16)$$

where m^* is an effective mass of the particles forming the condensate.

From $F[\psi]$ we can derive a differential equation for $\psi(\mathbf{r})$ minimizing F . The derivation is very similar to the derivation of the Lagrange equation (of the second kind) from Hamilton’s principle $\delta S = 0$ known from classical mechanics. Here, we start from the extremum principle $\delta F = 0$. We write

$$\psi(\mathbf{r}) = \psi_0(\mathbf{r}) + \eta(\mathbf{r}), \quad (6.17)$$

where $\psi_0(\mathbf{r})$ is the as yet unknown solution and $\eta(\mathbf{r})$ is a small displacement. Then

$$\begin{aligned} F[\psi_0 + \eta] &= F[\psi_0] \\ &+ \int d^3r \left[\alpha\psi_0^*\eta + \alpha\eta^*\psi_0 + \beta\psi_0^*\psi_0^*\eta + \beta\psi_0^*\eta^*\psi_0\psi_0 + \frac{\hbar^2}{2m^*}(\nabla\psi_0)^* \cdot \nabla\eta + \frac{\hbar^2}{2m^*}(\nabla\eta)^* \cdot \nabla\psi_0 \right] \\ &+ \mathcal{O}(\eta, \eta^*)^2 \\ &\stackrel{\text{by parts}}{=} F[\psi_0] \\ &+ \int d^3r \left[\alpha\psi_0^*\eta + \alpha\eta^*\psi_0 + \beta\psi_0^*\psi_0^*\eta + \beta\psi_0^*\eta^*\psi_0\psi_0 - \frac{\hbar^2}{2m^*}(\nabla^2\psi_0)^*\eta - \frac{\hbar^2}{2m^*}\eta^*\nabla^2\psi_0 \right] \\ &+ \mathcal{O}(\eta, \eta^*)^2. \end{aligned} \quad (6.18)$$

If $\psi_0(\mathbf{r})$ minimizes F , the terms linear in η and η^* must vanish for any $\eta(\mathbf{r})$, $\eta^*(\mathbf{r})$. Noting that η and η^* are linearly independent, this requires the prefactors of η and of η^* to vanish for all \mathbf{r} . Thus we conclude that

$$\alpha\psi_0 + \beta\psi_0^*\psi_0\psi_0 - \frac{\hbar^2}{2m^*}\nabla^2\psi_0 = \alpha\psi_0 + \beta|\psi_0|^2\psi_0 - \frac{\hbar^2}{2m^*}\nabla^2\psi_0 = 0. \quad (6.19)$$

Dropping the subscript and rearranging terms we find

$$-\frac{\hbar^2}{2m^*}\nabla^2\psi + \alpha\psi + \beta|\psi|^2\psi = 0. \quad (6.20)$$

This equation is very similar to the time-independent Schrödinger equation but, interestingly, it contains a non-linear term. We now apply it to find the variation of $\psi(\mathbf{r})$ close to the surface of a superfluid filling the half space $x > 0$. We impose the boundary condition $\psi(x=0) = 0$ and assume that the solution only depends on x . Then

$$-\frac{\hbar^2}{2m^*}\psi''(x) + \alpha\psi(x) + \beta|\psi(x)|^2\psi(x) = 0. \quad (6.21)$$

Since all coefficients are real, the solution can be chosen real. For $x \rightarrow \infty$ we should obtain the uniform solution

$$\lim_{x \rightarrow \infty} \psi(x) = \sqrt{-\frac{\alpha}{\beta}}. \quad (6.22)$$

Writing

$$\psi(x) = \sqrt{-\frac{\alpha}{\beta}} f(x) \quad (6.23)$$

we obtain

$$-\underbrace{\frac{\hbar^2}{2m^*\alpha}}_{>0} f''(x) + f(x) - f(x)^3 = 0. \quad (6.24)$$

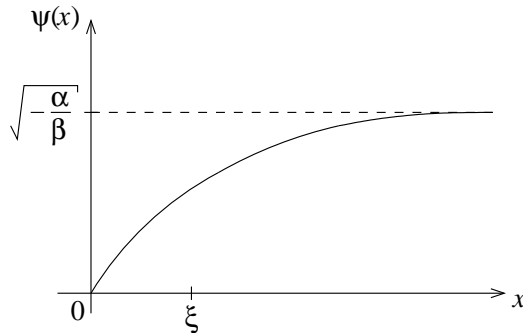
This equation contains a characteristic length ξ with

$$\xi^2 = -\frac{\hbar^2}{2m^*\alpha} \cong \frac{\hbar^2}{2m^*\alpha'(T_c - T)} > 0, \quad (6.25)$$

which is called the *Ginzburg-Landau coherence length*. It is not the same quantity as the the Pippard coherence length ξ_0 or ξ . The Ginzburg-Landau ξ has a strong temperature dependence and actually diverges at $T = T_c$, whereas the Pippard ξ has at most a weak temperature dependence. Microscopic BCS theory reveals how the two quantities are related, though. Equation (6.24) can be solved analytically. It is easy to check that

$$f(x) = \tanh \frac{x}{\sqrt{2}\xi} \quad (6.26)$$

is a solution satisfying the boundary conditions at $x = 0$ and $x \rightarrow \infty$. (In the general, three-dimensional case, the solution can only be given in terms of Jacobian elliptic functions.) The one-dimensional tanh solution is sketched here:



Fluctuations for $T > T_c$

So far, we have only considered the state ψ_0 of the system that minimizes the Landau functional $F[\psi]$. This is the *mean-field state*. At nonzero temperatures, the system will fluctuate about ψ_0 . For a bulk system we have

$$\psi(\mathbf{r}, t) = \psi_0 + \delta\psi(\mathbf{r}, t) \quad (6.27)$$

with uniform ψ_0 . We consider the cases $T > T_c$ and $T < T_c$ separately.

For $T > T_c$, the mean-field solution is just $\psi_0 = 0$. $F[\psi] = F[\delta\psi]$ gives the energy of excitations, we can thus write the partition function \mathcal{Z} as a sum over all possible states of Boltzmann factors containing this energy,

$$\mathcal{Z} = \int D^2\psi e^{-F[\psi]/k_B T}. \quad (6.28)$$

The notation $D^2\psi$ expresses that the integral is over uncountably many complex variables, namely the values of $\psi(\mathbf{r})$ for all \mathbf{r} . This means that \mathcal{Z} is technically a functional integral. The mathematical details go beyond the

scope of this course, though. The integral is difficult to evaluate. A common approximation is to restrict $F[\psi]$ to second-order terms, which is reasonable for $T > T_c$ since the fourth-order term $(\beta/2) |\psi|^4$ is not required to stabilize the theory (i.e., to make $F \rightarrow \infty$ for $|\psi| \rightarrow \infty$). This is called the *Gaussian approximation*. It allows \mathcal{Z} to be evaluated by Fourier transformation: Into

$$F[\psi] \cong \int d^3r \left[\alpha \psi^*(\mathbf{r}) \psi(\mathbf{r}) + \frac{1}{2m^*} \left(\frac{\hbar}{i} \nabla \psi(\mathbf{r}) \right)^* \cdot \frac{\hbar}{i} \nabla \psi(\mathbf{r}) \right] \quad (6.29)$$

we insert

$$\psi(\mathbf{r}) = \frac{1}{\sqrt{V}} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}} \psi_{\mathbf{k}}, \quad (6.30)$$

which gives

$$\begin{aligned} F[\psi] &\cong \frac{1}{V} \sum_{\mathbf{k}\mathbf{k}'} \underbrace{\int d^3r e^{-i\mathbf{k}\cdot\mathbf{r}+i\mathbf{k}'\cdot\mathbf{r}}}_{V\delta_{\mathbf{k}\mathbf{k}'}} \left[\alpha \psi_{\mathbf{k}}^* \psi_{\mathbf{k}'} + \frac{1}{2m^*} (\hbar\mathbf{k}\psi_{\mathbf{k}})^* \cdot \hbar\mathbf{k}'\psi_{\mathbf{k}'} \right] \\ &= \sum_{\mathbf{k}} \left(\alpha \psi_{\mathbf{k}}^* \psi_{\mathbf{k}} + \frac{\hbar^2 k^2}{2m^*} \psi_{\mathbf{k}}^* \psi_{\mathbf{k}} \right) = \sum_{\mathbf{k}} \left(\alpha + \frac{\hbar^2 k^2}{2m^*} \right) \psi_{\mathbf{k}}^* \psi_{\mathbf{k}}. \end{aligned} \quad (6.31)$$

Thus

$$\begin{aligned} \mathcal{Z} &\cong \int \prod_{\mathbf{k}} d^2\psi_{\mathbf{k}} \exp \left(-\frac{1}{k_B T} \sum_{\mathbf{k}} \left(\alpha + \frac{\hbar^2 k^2}{2m^*} \right) \psi_{\mathbf{k}}^* \psi_{\mathbf{k}} \right) \\ &= \prod_{\mathbf{k}} \left\{ \int d^2\psi_{\mathbf{k}} \exp \left(-\frac{1}{k_B T} \left(\alpha + \frac{\hbar^2 k^2}{2m^*} \right) \psi_{\mathbf{k}}^* \psi_{\mathbf{k}} \right) \right\}. \end{aligned} \quad (6.32)$$

The integral is now of Gaussian type (hence ‘‘Gaussian approximation’’) and can be evaluated exactly:

$$\mathcal{Z} \cong \prod_{\mathbf{k}} \frac{\pi k_B T}{\alpha + \frac{\hbar^2 k^2}{2m^*}}. \quad (6.33)$$

From this, we can obtain the thermodynamic variables. For example, the heat capacity is

$$C = T \frac{\partial S}{\partial T} = -T \frac{\partial^2 F}{\partial T^2} = T \frac{\partial^2}{\partial T^2} k_B T \ln \mathcal{Z} \cong k_B T \frac{\partial^2}{\partial T^2} T \sum_{\mathbf{k}} \ln \frac{\pi k_B T}{\alpha + \frac{\hbar^2 k^2}{2m^*}}. \quad (6.34)$$

We only retain the term that is singular at T_c . It stems from the temperature dependence of α , not from the explicit factors of T . This term is

$$\begin{aligned} C_{\text{crit}} &\cong -k_B T^2 \frac{\partial^2}{\partial T^2} \sum_{\mathbf{k}} \ln \left(\alpha + \frac{\hbar^2 k^2}{2m^*} \right) = -k_B T^2 \frac{\partial}{\partial T} \sum_{\mathbf{k}} \frac{\alpha'}{\alpha + \frac{\hbar^2 k^2}{2m^*}} \\ &= k_B T^2 \sum_{\mathbf{k}} \frac{(\alpha')^2}{\left(\alpha + \frac{\hbar^2 k^2}{2m^*} \right)^2} = k_B T^2 \left(\frac{2m^*}{\hbar^2} \right)^2 (\alpha')^2 \sum_{\mathbf{k}} \frac{1}{\left(k^2 + \frac{2m^* \alpha}{\hbar^2} \right)^2}. \end{aligned} \quad (6.35)$$

Going over to an integral over \mathbf{k} , corresponding to the thermodynamic limit $V \rightarrow \infty$, we obtain

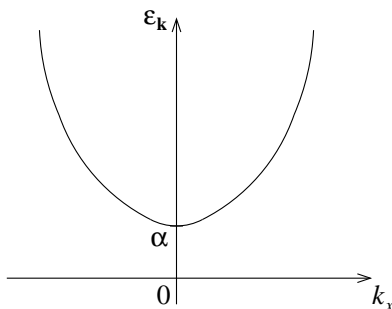
$$\begin{aligned} C_{\text{crit}} &\cong k_B T^2 \left(\frac{2m^*}{\hbar^2} \right)^2 (\alpha')^2 V \int \frac{d^3k}{(2\pi)^3} \frac{1}{\left(k^2 + \frac{2m^* \alpha}{\hbar^2} \right)^2} \\ &= \frac{k_B T^2 (m^*)^2}{2\pi \hbar^4} (\alpha')^2 V \frac{1}{\sqrt{\frac{2m^* \alpha}{\hbar^2}}} = \frac{k_B T^2 (m^*)^{3/2}}{2\sqrt{2}\pi \hbar^3} (\alpha')^2 V \frac{1}{\sqrt{\alpha}} \\ &\sim \frac{1}{\sqrt{T - T_c}}. \end{aligned} \quad (6.36)$$

Recall that at the mean-field level, C just showed a step at T_c . Including fluctuations, we obtain a divergence of the form $1/\sqrt{T - T_c}$ for $T \rightarrow T_c$ from above. This is due to superfluid fluctuations in the normal state. The system “notices” that superfluid states exist at relatively low energies, while the mean-field state is still normal.

We note for later that the derivation has shown that for any \mathbf{k} we have two fluctuation modes with dispersion

$$\epsilon_{\mathbf{k}} = \alpha + \frac{\hbar^2 k^2}{2m^*}. \quad (6.37)$$

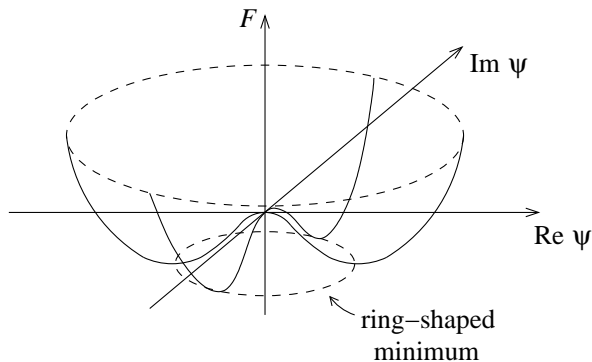
The two modes correspond for example to the real and the imaginary part of $\psi_{\mathbf{k}}$. Since $\alpha = \alpha'(T - T_c) > 0$, the dispersion has an *energy gap* α .



In the language of field theory, one also says that the superfluid above T_c has two degenerate *massive* modes, with the mass proportional to the energy gap.

Fluctuations for $T < T_c$

Below T_c , the situation is a bit more complex due to the Mexican-hat potential: All states with amplitude $|\psi_0| = \sqrt{-\alpha/\beta}$ are equally good mean-field solutions.



It is plausible that fluctuations of the phase have low energies since *global* changes of the phase do not increase F . To see this, we write

$$\psi = (\psi_0 + \delta\psi) e^{i\phi}, \quad (6.38)$$

where ψ_0 and $\delta\psi$ are now real. The Landau functional becomes

$$F[\psi] \cong \int d^3r \left[\alpha(\psi_0 + \delta\psi)^2 + \frac{\beta}{2}(\psi_0 + \delta\psi)^4 + \frac{1}{2m^*} \left(\frac{\hbar}{i}(\nabla\delta\psi)e^{i\phi} + \hbar(\psi_0 + \delta\psi)(\nabla\phi)e^{i\phi} \right)^* \cdot \left(\frac{\hbar}{i}(\nabla\delta\psi)e^{i\phi} + \hbar(\psi_0 + \delta\psi)(\nabla\phi)e^{i\phi} \right) \right]. \quad (6.39)$$

As above, we keep only terms up to second order in the fluctuations (Gaussian approximation), i.e., terms proportional to $\delta\psi^2$, $\delta\psi\phi$, or ϕ^2 . We get

$$F[\delta\psi, \phi] \cong \int d^3r \left[\cancel{2\alpha\psi_0\delta\psi + 2\beta\psi_0^3\delta\psi} + \alpha\delta\psi^2 + 3\beta\psi_0^3\delta\psi^2 + \frac{1}{2m^*} \left(\frac{\hbar}{i} \nabla \delta\psi \right)^* \cdot \frac{\hbar}{i} \nabla \delta\psi \right. \\ \left. + \frac{1}{2m^*} \left(\left(\frac{\hbar}{i} \nabla \delta\psi \right)^* \cdot \hbar\psi_0 \nabla \phi + (\hbar\psi_0 \nabla \phi)^* \cdot \frac{\hbar}{i} \nabla \delta\psi \right) + \frac{1}{2m^*} (\hbar\psi_0 \nabla \phi)^* \cdot \hbar\psi_0 \nabla \phi \right] (+ \text{const}). \quad (6.40)$$

Note that the first-order terms cancel since we are expanding about a minimum. Furthermore, up to second order there are no terms mixing amplitude fluctuations $\delta\psi$ and phase fluctuations ϕ . We can simplify the expression:

$$F[\delta\psi, \phi] \cong \int d^3r \left[-2\alpha\delta\psi^2 + \frac{1}{2m^*} \left(\frac{\hbar}{i} \nabla \delta\psi \right)^* \cdot \frac{\hbar}{i} \nabla \delta\psi - \frac{\hbar^2}{2m^*} \frac{\alpha}{\beta} (\nabla \phi)^* \nabla \phi \right] \\ = \sum_{\mathbf{k}} \left[\underbrace{\left(-2\alpha + \frac{\hbar^2 k^2}{2m^*} \right)}_{>0} \delta\psi_{\mathbf{k}}^* \delta\psi_{\mathbf{k}} - \underbrace{\frac{\alpha}{\beta} \frac{\hbar^2 k^2}{2m^*}}_{>0} \phi_{\mathbf{k}}^* \phi_{\mathbf{k}} \right], \quad (6.41)$$

in analogy to the case $T > T_c$. We see that amplitude fluctuations are gapped (massive) with an energy gap $-2\alpha = \alpha'(T_c - T)$. They are not degenerate. Phase fluctuations on the other hand are *ungapped (massless)* with quadratic dispersion

$$\epsilon_{\mathbf{k}}^{\phi} = -\frac{\alpha}{\beta} \frac{\hbar^2 k^2}{2m^*}. \quad (6.42)$$

The appearance of ungapped so-called *Goldstone modes* is characteristic for systems with spontaneously broken continuous symmetries. We state without derivation that the heat capacity diverges like $C \sim \sqrt{T_c - T}$ for $T \rightarrow T_c^-$, analogous to the case $T > T_c$.

6.3 Ginzburg-Landau theory for superconductors

To describe superconductors, we have to take the charge q of the particles forming the condensate into account. We allow for the possibility that q is not the electron charge $-e$. Then there are two additional terms in the Landau functional:

- The canonical momentum has to be replaced by the kinetic momentum:

$$\frac{\hbar}{i} \nabla \rightarrow \frac{\hbar}{i} \nabla - \frac{q}{c} \mathbf{A}, \quad (6.43)$$

where \mathbf{A} is the vector potential.

- The energy density of the magnetic field, $B^2/8\pi$, has to be included.

Thus we obtain the functional

$$F[\psi, \mathbf{A}] \cong \int d^3r \left[\alpha |\psi|^2 + \frac{\beta}{2} |\psi|^4 + \frac{1}{2m^*} \left| \left(\frac{\hbar}{i} \nabla - \frac{q}{c} \mathbf{A} \right) \psi \right|^2 + \frac{B^2}{8\pi} \right]. \quad (6.44)$$

Minimizing this free-energy functional with respect to ψ , we obtain, in analogy to the previous section,

$$\frac{1}{2m^*} \left(\frac{\hbar}{i} \nabla - \frac{q}{c} \mathbf{A} \right)^2 \psi + \alpha \psi + \beta |\psi|^2 \psi = 0. \quad (6.45)$$

To minimize F with respect to \mathbf{A} , we write $\mathbf{A}(\mathbf{r}) = \mathbf{A}_0(\mathbf{r}) + \mathbf{a}(\mathbf{r})$ and inset this into the terms containing \mathbf{A} :

$$\begin{aligned}
F[\psi, \mathbf{A}_0 + \mathbf{a}] &= F[\psi, \mathbf{A}_0] + \int d^3r \left[\frac{1}{2m^*} \left(\left[\frac{\hbar}{i} \nabla - \frac{q}{c} \mathbf{A}_0 \right] \psi \right)^* \cdot \left(-\frac{q}{c} \right) \mathbf{a} \psi \right. \\
&\quad \left. + \frac{1}{2m^*} \left(-\frac{q}{c} \mathbf{a} \psi \right)^* \cdot \left(\frac{\hbar}{i} \nabla - \frac{q}{c} \mathbf{A}_0 \right) \psi + \frac{1}{4\pi} \underbrace{(\nabla \times \mathbf{A}_0) \cdot (\nabla \times \mathbf{a})}_{=\mathbf{a} \cdot (\nabla \times \nabla \times \mathbf{A}_0) - \nabla \cdot ((\nabla \times \mathbf{A}_0) \times \mathbf{a})} \right] + \mathcal{O}(a^2) \\
&= F[\psi, \mathbf{A}_0] + \int d^3r \left[-\frac{q}{2m^*c} \left(\left[\frac{\hbar}{i} \nabla \psi \right]^* \psi + \psi^* \frac{\hbar}{i} \nabla \psi \right) \cdot \mathbf{a} \right. \\
&\quad \left. + \frac{q^2}{m^*c^2} |\psi|^2 \mathbf{A}_0 \cdot \mathbf{a} + \frac{1}{4\pi} (\nabla \times \mathbf{B}_0) \cdot \mathbf{a} \right] + \mathcal{O}(a^2), \tag{6.46}
\end{aligned}$$

where we have used $\nabla \times \mathbf{A}_0 = \mathbf{B}_0$ and Gauss' theorem. At a minimum, the coefficient of the linear term must vanish. Dropping the subscript we obtain

$$-i \frac{q\hbar}{2m^*c} ([\nabla\psi^*]\psi - \psi^*\nabla\psi) + \frac{q^2}{m^*c^2} |\psi|^2 \mathbf{A} + \frac{1}{4\pi} \nabla \times \mathbf{B} = 0. \tag{6.47}$$

With Ampère's Law we find

$$\mathbf{j} = \frac{c}{4\pi} \nabla \times \mathbf{B} = i \frac{q\hbar}{2m^*} ([\nabla\psi^*]\psi - \psi^*\nabla\psi) - \frac{q^2}{m^*c} |\psi|^2 \mathbf{A}, \tag{6.48}$$

where we have dropped the subscript “ s ” of \mathbf{j} since we assume that any normal current is negligible. Equations (6.45) and (6.48) are called the *Ginzburg-Landau equations*.

In the limit of uniform $\psi(\mathbf{r})$, Eq. (6.48) simplifies to

$$\mathbf{j} = -\frac{q^2 |\psi|^2}{m^*c} \mathbf{A}. \tag{6.49}$$

This should reproduce the London equation

$$\mathbf{j} = -\frac{e^2 n_s}{mc} \mathbf{A}, \tag{6.50}$$

which obviously requires

$$\frac{q^2 |\psi|^2}{m^*} = \frac{e^2 n_s}{m}. \tag{6.51}$$

As noted in Sec. 5.3, based on flux-quantization experiments and on the analogy to Bose-Einstein condensation, it is natural to set $q = -2e$. For m^* one might then insert twice the effective electron (band) mass of the metal. However, it turns out to be difficult to measure m^* independently of the superfluid density n_s and it is therefore common to set $m^* = 2m \equiv 2m_e$. All system-specific properties are thus absorbed into

$$n_s = \frac{m q^2 |\psi|^2}{e^2 m^*} = \frac{m \psi e^2 |\psi|^2}{e^2 2m} = 2 |\psi|^2. \tag{6.52}$$

With this, we can write the penetration depth as

$$\lambda = \sqrt{\frac{mc^2}{4\pi e^2 n_s}} = \sqrt{\frac{mc^2}{8\pi e^2 |\psi|^2}} \cong \sqrt{\frac{mc^2}{8\pi e^2 (-\frac{\alpha}{\beta})}}. \tag{6.53}$$

We have found two characteristic length scales:

- λ : penetration of magnetic field,
- ξ : variation of superconducting coarse-grained wave function (order parameter).

Within the mean-field theories we have so far employed, both quantities scale as

$$\lambda, \xi \propto \frac{1}{\sqrt{T_c - T}} \quad (6.54)$$

close to T_c . Thus their dimensionless ratio

$$\kappa := \frac{\lambda(T)}{\xi(T)} \quad (6.55)$$

is roughly temperature-independent. It is called the *Ginzburg-Landau parameter* and turns out to be very important for the behaviour of superconductors in an applied magnetic field. For elemental superconductors, κ is small compared to one.

Fluctuations and the Anderson-Higgs mechanism

When discussing fluctuations about the mean-field state of a superconductor, we have to take the coupling to the electromagnetic field into account. We proceed analogously to the case of a neutral superfluid and employ the Gaussian approximation throughout. For $T > T_c$, we note that the kinetic term

$$\frac{1}{2m^*} \left| \left(\frac{\hbar}{i} \nabla - \frac{q}{c} \mathbf{A} \right) \psi \right|^2 \quad (6.56)$$

is explicitly of second order in the fluctuations $\delta\psi = \psi$ so that electromagnetic field fluctuations appear only in higher orders. Thus to second order, the order-parameter fluctuations decouple from the electromagnetic fluctuations,

$$F[\psi, \mathbf{B}] \cong \int d^3r \left[\alpha \psi^*(\mathbf{r}) \psi(\mathbf{r}) + \frac{1}{2m^*} \left| \frac{\hbar}{i} \nabla \psi \right|^2 + \frac{B^2}{8\pi} \right] = \sum_{\mathbf{k}} \left(\alpha + \frac{\hbar^2 k^2}{2m^*} \right) \psi_{\mathbf{k}}^* \psi_{\mathbf{k}} + \sum_{\mathbf{k}} \frac{\mathbf{B}_{\mathbf{k}}^* \cdot \mathbf{B}_{\mathbf{k}}}{8\pi}. \quad (6.57)$$

The superconducting fluctuations consist of two degenerate massive modes with dispersion $\epsilon_{\mathbf{k}} = \alpha + \hbar^2 k^2 / 2m^*$, as for the neutral superfluid.

The case $T < T_c$ is more interesting. Writing, as above, $\psi = (\psi_0 + \delta\psi) e^{i\phi}$, we obtain to second order

$$\begin{aligned} F[\delta\psi, \phi, \mathbf{A}] &= \int d^3r \left[\cancel{2\alpha\psi_0\delta\psi + 2\beta\psi_0^3\delta\psi} + \alpha\delta\psi^2 + 3\beta\psi_0^2\delta\psi^2 + \frac{1}{2m^*} \left(\frac{\hbar}{i} \nabla \delta\psi \right)^* \cdot \frac{\hbar}{i} \nabla \delta\psi \right. \\ &\quad + \frac{1}{2m^*} \left(\left(\frac{\hbar}{i} \nabla \delta\psi \right)^* \cdot \hbar\psi_0 \nabla \phi + \cancel{(\hbar\psi_0 \nabla \phi)^* \cdot \frac{\hbar}{i} \nabla \delta\psi} \right) + \frac{1}{2m^*} (\hbar\psi_0 \nabla \phi)^* \cdot \hbar\psi_0 \nabla \phi \\ &\quad + \frac{1}{2m^*} \left(\left(\frac{\hbar}{i} \nabla \delta\psi \right)^* \cdot \left(-\frac{q}{c} \right) \mathbf{A} \psi_0 + \cancel{\left(-\frac{q}{c} \mathbf{A} \psi_0 \right)^* \cdot \frac{\hbar}{i} \nabla \delta\psi} \right) \\ &\quad + \frac{1}{2m^*} \left((\hbar\psi_0 \nabla \phi)^* \cdot \left(-\frac{q}{c} \right) \mathbf{A} \psi_0 + \left(-\frac{q}{c} \mathbf{A} \psi_0 \right)^* \cdot \hbar\psi_0 \nabla \phi \right) \\ &\quad \left. + \frac{1}{2m^*} \left(-\frac{q}{c} \mathbf{A} \psi_0 \right)^* \cdot \left(-\frac{q}{c} \right) \mathbf{A} \psi_0 + \frac{1}{8\pi} (\nabla \times \mathbf{A})^* \cdot (\nabla \times \mathbf{A}) \right] (+ \text{const}) \\ &= \int d^3r \left[-2\alpha\delta\psi^2 + \frac{1}{2m^*} \left(\frac{\hbar}{i} \nabla \delta\psi \right)^* \cdot \frac{\hbar}{i} \nabla \delta\psi \right. \\ &\quad \left. - \frac{\hbar^2}{2m^*} \frac{\alpha}{\beta} \left(\nabla \phi - \frac{q}{\hbar c} \mathbf{A} \right)^* \cdot \left(\nabla \phi - \frac{q}{\hbar c} \mathbf{A} \right) + \frac{1}{8\pi} (\nabla \times \mathbf{A})^* \cdot (\nabla \times \mathbf{A}) \right]. \quad (6.58) \end{aligned}$$

Note that the phase ϕ of the macroscopic wave function and the vector potential appear in the combination

$\nabla\phi - (q/\hbar c)\mathbf{A}$. Physics is invariant under the gauge transformation

$$\mathbf{A} \rightarrow \mathbf{A} + \nabla\chi, \quad (6.59)$$

$$\Phi \rightarrow \Phi - \frac{1}{c}\dot{\chi}, \quad (6.60)$$

$$\psi \rightarrow e^{iq\chi/\hbar c}\psi, \quad (6.61)$$

where Φ is the scalar electric potential and $\chi(\mathbf{r}, t)$ is an arbitrary scalar field. We make use of this gauge invariance by choosing

$$\chi = -\frac{\hbar c}{q}\phi. \quad (6.62)$$

Under this transformation, we get

$$\mathbf{A} \rightarrow \mathbf{A} - \frac{\hbar c}{q}\nabla\phi =: \mathbf{A}', \quad (6.63)$$

$$\psi = (\psi_0 + \delta\psi) e^{i\phi} \rightarrow (\psi_0 + \delta\psi) e^{i(-\phi+\phi)} = \psi_0 + \delta\psi. \quad (6.64)$$

The macroscopic wave function becomes purely real (and positive). The Landau functional thus transforms into

$$\begin{aligned} F[\delta\psi, \mathbf{A}'] \rightarrow \int d^3r \left[-2\alpha\delta\psi^2 + \frac{1}{2m^*} \left(\frac{\hbar}{i}\nabla\delta\psi \right)^* \cdot \frac{\hbar}{i}\nabla\delta\psi \right. \\ \left. - \frac{\alpha}{\beta} \frac{q^2}{2m^*c^2} (\mathbf{A}')^* \cdot \mathbf{A}' + \frac{1}{8\pi} (\nabla \times \mathbf{A}')^* \cdot (\nabla \times \mathbf{A}') \right] \end{aligned} \quad (6.65)$$

(note that $\nabla \times \mathbf{A}' = \nabla \times \mathbf{A}$). Thus the phase no longer appears in F ; it has been *absorbed into the vector potential*. Furthermore, dropping the prime,

$$\begin{aligned} F[\delta\psi, \mathbf{A}] &\cong \sum_{\mathbf{k}} \left[\left(-2\alpha + \frac{\hbar^2 k^2}{2m^*} \right) \delta\psi_{\mathbf{k}}^* \delta\psi_{\mathbf{k}} - \frac{\alpha}{\beta} \frac{q^2}{2m^*c^2} \mathbf{A}_{\mathbf{k}}^* \cdot \mathbf{A}_{\mathbf{k}} + \frac{1}{8\pi} (\mathbf{k} \times \mathbf{A}_{\mathbf{k}})^* \cdot (\mathbf{k} \times \mathbf{A}_{\mathbf{k}}) \right] \\ &= \sum_{\mathbf{k}} \left[\left(-2\alpha + \frac{\hbar^2 k^2}{2m^*} \right) \delta\psi_{\mathbf{k}}^* \delta\psi_{\mathbf{k}} - \frac{\alpha}{\beta} \frac{q^2}{2m^*c^2} \mathbf{A}_{\mathbf{k}}^* \cdot \mathbf{A}_{\mathbf{k}} + \frac{k^2}{8\pi} \mathbf{A}_{\mathbf{k}}^* \cdot \mathbf{A}_{\mathbf{k}} - \frac{1}{8\pi} (\mathbf{k} \cdot \mathbf{A}_{\mathbf{k}}^*) (\mathbf{k} \cdot \mathbf{A}_{\mathbf{k}}) \right]. \end{aligned} \quad (6.66)$$

Obviously, amplitude fluctuations decouple from electromagnetic fluctuations and behave like for a neutral superfluid. We discuss the electromagnetic fluctuations further. The term proportional to α is due to superconductivity. Without it, we would have the free-field functional

$$F_{\text{free}}[\mathbf{A}] = \sum_{\mathbf{k}} \frac{1}{8\pi} [k^2 \mathbf{A}_{\mathbf{k}}^* \cdot \mathbf{A}_{\mathbf{k}} - (\mathbf{k} \cdot \mathbf{A}_{\mathbf{k}}^*) (\mathbf{k} \cdot \mathbf{A}_{\mathbf{k}})]. \quad (6.67)$$

Decomposing \mathbf{A} into longitudinal and transverse components,

$$\mathbf{A}_{\mathbf{k}} = \underbrace{\hat{k}(\hat{k} \cdot \mathbf{A}_{\mathbf{k}})}_{=: \mathbf{A}_{\mathbf{k}}^{\parallel}} + \underbrace{\mathbf{A}_{\mathbf{k}} - \hat{k}(\hat{k} \cdot \mathbf{A}_{\mathbf{k}})}_{=: \mathbf{A}_{\mathbf{k}}^{\perp}} \quad (6.68)$$

with $\hat{k} := \mathbf{k}/k$, we obtain

$$\begin{aligned} F_{\text{free}}[\mathbf{A}] &= \sum_{\mathbf{k}} \frac{1}{8\pi} \left[k^2 \mathbf{A}_{\mathbf{k}}^{\parallel*} \cdot \mathbf{A}_{\mathbf{k}}^{\parallel} + k^2 \mathbf{A}_{\mathbf{k}}^{\perp*} \cdot \mathbf{A}_{\mathbf{k}}^{\perp} - (\mathbf{k} \cdot \mathbf{A}_{\mathbf{k}}^{\parallel*}) (\mathbf{k} \cdot \mathbf{A}_{\mathbf{k}}^{\parallel}) \right] \\ &= \sum_{\mathbf{k}} \frac{1}{8\pi} \left[\cancel{k^2 \mathbf{A}_{\mathbf{k}}^{\parallel*} \cdot \mathbf{A}_{\mathbf{k}}^{\parallel}} + k^2 \mathbf{A}_{\mathbf{k}}^{\perp*} \cdot \mathbf{A}_{\mathbf{k}}^{\perp} - \cancel{k^2 \mathbf{A}_{\mathbf{k}}^{\parallel*} \cdot \mathbf{A}_{\mathbf{k}}^{\parallel}} \right] = \sum_{\mathbf{k}} \frac{1}{8\pi} k^2 \mathbf{A}_{\mathbf{k}}^{\perp*} \cdot \mathbf{A}_{\mathbf{k}}^{\perp}. \end{aligned} \quad (6.69)$$

Thus only the *transverse* components appear—only they are degrees of freedom of the free electromagnetic field. They do not have an energy gap.

Including the superconducting contribution, we get

$$F[\mathbf{A}] = \sum_{\mathbf{k}} \left[-\frac{\alpha}{\beta} \frac{q^2}{2m^*c^2} \mathbf{A}_{\mathbf{k}}^{\parallel*} \cdot \mathbf{A}_{\mathbf{k}}^{\parallel} - \frac{\alpha}{\beta} \frac{q^2}{2m^*c^2} \mathbf{A}_{\mathbf{k}}^{\perp*} \cdot \mathbf{A}_{\mathbf{k}}^{\perp} + \frac{1}{8\pi} k^2 \mathbf{A}_{\mathbf{k}}^{\perp*} \cdot \mathbf{A}_{\mathbf{k}}^{\perp} \right]. \quad (6.70)$$

All *three* components of \mathbf{A} appear now, the longitudinal one has been introduced by absorbing the phase $\phi(\mathbf{r})$. Even more importantly, all components obtain a term with a constant coefficient, i.e., a *mass term*. Thus the electromagnetic field inside a superconductor becomes massive. This is the famous *Anderson-Higgs mechanism*. The same general idea is also thought to explain the masses of elementary particles, although in a more complicated way. The ‘‘Higgs bosons’’ in our case are the amplitude-fluctuation modes described by $\delta\psi$. Contrary to what is said in popular discussions, they are not *responsible* for giving mass to the field \mathbf{A} . Rather, they are left over when the phase fluctuations are eaten by the field \mathbf{A} .

The mass term in the superconducting case can be thought of as leading to the Meißner effect (finite penetration depth λ). Indeed, we can write

$$F[\delta\psi, \mathbf{A}] = \sum_{\mathbf{k}} \left(-2\alpha + \frac{\hbar^2 k^2}{2m^*} \right) \delta\psi_{\mathbf{k}}^* \delta\psi_{\mathbf{k}} + \sum_{\mathbf{k}} \frac{1}{8\pi} \left[\frac{1}{\lambda^2} \mathbf{A}_{\mathbf{k}}^* \cdot \mathbf{A}_{\mathbf{k}} + k^2 \mathbf{A}_{\mathbf{k}}^* \cdot \mathbf{A}_{\mathbf{k}} - (\mathbf{k} \cdot \mathbf{A}_{\mathbf{k}}^*)(\mathbf{k} \cdot \mathbf{A}_{\mathbf{k}}) \right]. \quad (6.71)$$

The photon mass is proportional to $1/\lambda$. (To see that the dispersion relation is $\hbar^2\omega^2 = m^2c^4 + p^2c^2$, we would have to consider the full action including a temporal integral over F and terms containing time-derivatives.)

Elitzur’s theorem

One sometimes reads that in the superconducting state gauge symmetry is broken. This is not correct. Gauge symmetry is the invariance under local gauge transformations. S. Elitzur showed in 1975 that a local gauge symmetry cannot be spontaneously broken. Rather, superconductors and superfluids spontaneously break a *global* $U(1)$ symmetry in that the ordered state has a preferred *macroscopic phase*, as noted above.

6.4 Type-I superconductors

Superconductors with small Ginzburg-Landau parameter κ are said to be of *type I*. The exact condition is

$$\kappa < \frac{1}{\sqrt{2}}. \quad (6.72)$$

It turns out that these superconductors have a uniform state in an applied magnetic field, at least for simple geometries such as a thin cylinder parallel to the applied field.

The appropriate thermodynamic potential for describing a superconductor in an applied magnetic field is the Gibbs free energy G (natural variable \mathbf{H}) and not the Helmholtz free energy F (natural variable \mathbf{B}), which we have used so far. They are related by the Legendre transformation

$$G = F - \int d^3r \frac{\mathbf{H} \cdot \mathbf{B}}{4\pi}. \quad (6.73)$$

For a type-I superconductor, the equilibrium state in the bulk is uniform (we will see later why this is not a trivial statement). The order parameter is then $|\psi| = \sqrt{-\alpha/\beta}$ and the magnetic induction \mathbf{B} as well as the vector potential (in the London gauge) vanish. Thus the Gibbs free-energy density is

$$g_s = f_s = \alpha |\psi|^2 + \frac{\beta}{2} |\psi|^4 = -\frac{\alpha^2}{\beta} + \frac{\beta}{2} \frac{\alpha^2}{\beta} = -\frac{\alpha^2}{2\beta}. \quad (6.74)$$

On the other hand, in the normal state ψ vanishes, but the field penetrates the system and $\mathbf{B} \equiv \mathbf{H}$ so that

$$g_n = f_n - \frac{HB}{4\pi} = \frac{H^2}{8\pi} - \frac{H^2}{4\pi} = -\frac{H^2}{8\pi}. \quad (6.75)$$

The system will only become superconducting if this reduces the Gibbs free energy, i.e., if

$$g_s - g_n = -\frac{\alpha^2}{2\beta} + \frac{H^2}{8\pi} < 0. \quad (6.76)$$

Thus in an applied magnetic field of $H \geq H_c$, with the *critical field*

$$H_c(T) := \sqrt{4\pi \frac{\alpha^2}{\beta}}, \quad (6.77)$$

superconductivity does not occur.

We can use the relation (6.77) and

$$\lambda^2(T) = \frac{mc^2}{8\pi e^2 |\psi|^2} = -\frac{mc^2}{8\pi e^2} \frac{\beta}{\alpha} \quad (6.78)$$

to express the phenomenological parameters α and β in terms of the measurable quantities λ and H_c :

$$\alpha = -\frac{2e^2}{mc^2} \lambda^2 H_c^2, \quad (6.79)$$

$$\beta = 4\pi \left(\frac{2e^2}{mc^2} \right)^2 \lambda^4 H_c^2. \quad (6.80)$$

Domain-wall energy and intermediate states

We can now calculate the energy per area of a domain wall between superconducting and normal regions. We assume that $\psi(\mathbf{r})$ is only a function of x and impose the boundary conditions

$$\psi(x) \rightarrow \begin{cases} \psi_\infty := \sqrt{-\alpha/\beta} & \text{for } x \rightarrow \infty, \\ 0 & \text{for } x \rightarrow -\infty. \end{cases} \quad (6.81)$$

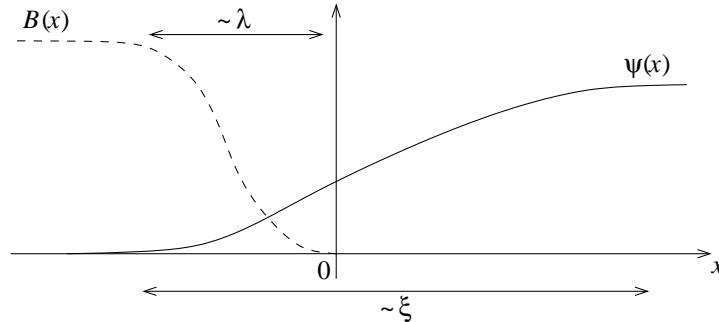
What are reasonable boundary conditions for the local field $B(x)$? In the superconductor we have

$$B(x) \rightarrow 0 \quad \text{for } x \rightarrow \infty. \quad (6.82)$$

At the other end, if we have $\lim_{x \rightarrow \infty} B(x) > H_c$, the bulk superconductor has a positive Gibbs free-energy density, as seen above. Thus superconductivity is not stable. For $\lim_{x \rightarrow \infty} B(x) < H_c$, the Gibbs free-energy density of the superconductor is negative and it eats up the normal phase. Thus a domain wall can only be stable for

$$B(x) \rightarrow H_c = \sqrt{4\pi \frac{\alpha^2}{\beta}} \quad \text{for } x \rightarrow -\infty. \quad (6.83)$$

Technically, we have to solve the Ginzburg-Landau equations under these boundary conditions, which can only be done numerically. The qualitative form of the solution is clear, though:



The Gibbs free energy of the domain wall, per unit area, will be denoted by γ . To derive it, first note that for the given boundary conditions, $g_s(x \rightarrow \infty) = g_n(x \rightarrow -\infty)$, i.e., the superconducting free-energy density deep inside the superconductor equals the normal free-energy density deep inside the normal conductor. This bulk Gibbs free-energy density is, as derived above,

$$g_s(x \rightarrow \infty) = g_n(x \rightarrow -\infty) = f_n(x \rightarrow -\infty) - \frac{H_c^2}{8\pi} = -\frac{H_c^2}{8\pi}. \quad (6.84)$$

The additional free-energy density due to the domain wall is

$$g_s(x) - \left(-\frac{H_c^2}{8\pi}\right) = g_s(x) + \frac{H_c^2}{8\pi}. \quad (6.85)$$

The corresponding free energy per area is

$$\begin{aligned} \gamma &= \int_{-\infty}^{\infty} dx \left[g_s(x) + \frac{H_c^2}{8\pi} \right] = \int_{-\infty}^{\infty} dx \left[f_s(x) - \frac{B(x)H_c}{4\pi} + \frac{H_c^2}{8\pi} \right] \\ &= \int_{-\infty}^{\infty} dx \left[\alpha |\psi|^2 + \frac{\beta}{2} |\psi|^4 + \frac{1}{2m^*} \left| \left(\frac{\hbar}{i} \frac{\partial}{\partial x} - \frac{q}{c} A(x) \right) \psi \right|^2 + \underbrace{\frac{B^2}{8\pi} - \frac{BH_c}{4\pi} + \frac{H_c^2}{8\pi}}_{=(B-H_c)^2/8\pi} \right]. \end{aligned} \quad (6.86)$$

We can simplify this by multiplying the first Ginzburg-Landau equation (6.45) by ψ^* and integrating over x :

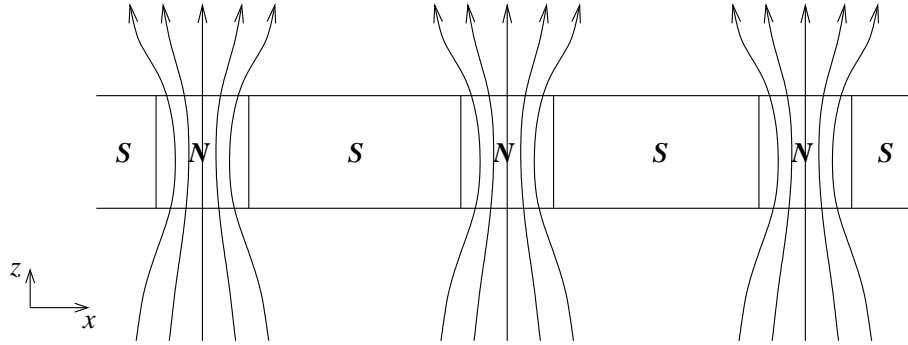
$$\begin{aligned} &\int_{-\infty}^{\infty} dx \left[\alpha |\psi|^2 + \beta |\psi|^4 + \frac{1}{2m^*} \psi^* \left(\frac{\hbar}{i} \frac{\partial}{\partial x} - \frac{q}{c} A(x) \right)^2 \psi \right] \\ &\stackrel{\text{by parts}}{=} \int_{-\infty}^{\infty} dx \left[\alpha |\psi|^2 + \beta |\psi|^4 + \frac{1}{2m^*} \left(-\frac{\hbar}{i} \frac{\partial}{\partial x} - \frac{q}{c} A(x) \right) \psi^*(x) \left(\frac{\hbar}{i} \frac{\partial}{\partial x} - \frac{q}{c} A(x) \right) \psi \right] \\ &= \int_{-\infty}^{\infty} dx \left[\alpha |\psi|^2 + \beta |\psi|^4 + \frac{1}{2m^*} \left| \left(\frac{\hbar}{i} \frac{\partial}{\partial x} - \frac{q}{c} A(x) \right) \psi \right|^2 \right] = 0. \end{aligned} \quad (6.87)$$

Thus

$$\begin{aligned} \gamma &= \int_{-\infty}^{\infty} dx \left[-\frac{\beta}{2} |\psi|^4 + \frac{(B-H_c)^2}{8\pi} \right] = \frac{H_c^2}{8\pi} \int_{-\infty}^{\infty} dx \left[-\beta \frac{\beta}{\alpha^2} |\psi|^4 + \left(1 - \frac{B}{H_c} \right)^2 \right] \\ &= \frac{H_c^2}{8\pi} \int_{-\infty}^{\infty} dx \left[\left(1 - \frac{B}{H_c} \right)^2 - \frac{|\psi|^4}{\psi_\infty^4} \right], \end{aligned} \quad (6.88)$$

where we have drawn out the characteristic energy density $H_c^2/8\pi$. The domain wall energy is given by the difference of the energy *cost* of expelling the magnetic field and the energy *gain* due to superconducting condensation. For strong type-I superconductors, $\xi \gg \lambda$; there is a region of thickness $\xi - \lambda > 0$ in which the first term is already large, while the second only slowly approaches its bulk value (see sketch above). Thus $\gamma > 0$ for type-I superconductors. They therefore tend to *minimize* the total area of domain walls.

Note that even in samples of relatively simple shape, magnetic flux *will* penetrate for non-zero applied field. It will do so in a way that minimizes the total Gibbs free energy, only one contribution to which is the domain-wall energy. For example in a very large slab perpendicular to the applied field, the flux has to penetrate in some manner, since going around the large sample would cost much energy. A careful analysis shows that this will usually happen in the form of normal stripes separated by superconducting stripes, see Tinkham's book. Such a state is called an *intermediate state*. It should not be confused with the vortex state to be discussed later.

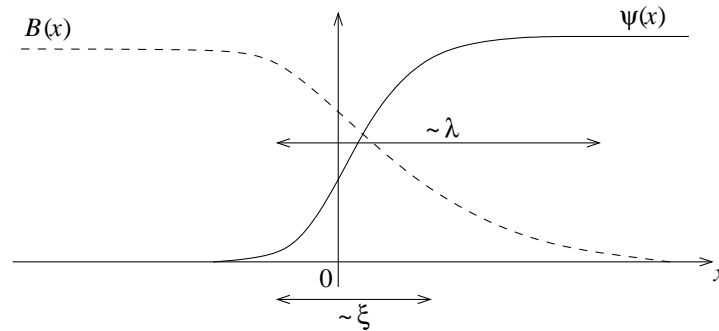


6.5 Type-II superconductors

Type-II superconductors are defined by a large Ginzburg-Landau parameter

$$\kappa > \frac{1}{\sqrt{2}}. \quad (6.89)$$

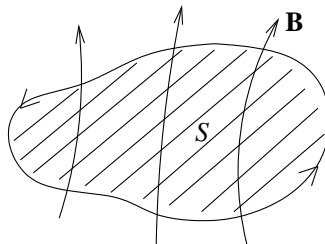
The analysis in the previous section goes through. But now domain walls have a region where the condensate is nearly fully developed but the flux is not completely expelled.



Therefore, the domain-wall energy is *negative*, $\gamma < 0$. Hence, the system tends to *maximize* the total area of domain walls. This tendency should be counterbalanced by some other effect, otherwise the system would become inhomogeneous on microscopic (atomic) length scales and Ginzburg-Landau theory would break down. The effect in question is flux quantization—the penetrating magnetic flux cannot be split into portions smaller than the flux quantum $\Phi_0 = hc/2e$, see Sec. 5.3.

Fluxoid quantization

We revisit the quantization of magnetic flux. Consider an arbitrary closed path ∂S forming the edge of a surface S , where at least the edge ∂S must lie inside a superconductor,



The magnetic flux Φ through this loop is

$$\Phi = \int_S d\mathbf{a} \cdot \mathbf{B} = \int_S d\mathbf{a} \cdot (\nabla \times \mathbf{A}) \stackrel{\text{Stokes}}{=} \oint_{\partial S} d\mathbf{s} \cdot \mathbf{A}. \quad (6.90)$$

With the second Ginzburg-Landau equation

$$\mathbf{j} = i \frac{q\hbar}{2m^*} ([\nabla\psi^*]\psi - \psi^*\nabla\psi) - \frac{q^2}{m^*c} |\psi|^2 \mathbf{A} \quad (6.91)$$

we obtain

$$\begin{aligned} \Phi &= \oint_{\partial S} d\mathbf{s} \cdot \left\{ -\frac{m^*c\mathbf{j}}{q^2|\psi|^2} + i\frac{\hbar c}{2q|\psi|^2} ([\nabla\psi^*]\psi - \psi^*\nabla\psi) \right\} = \oint_{\partial S} d\mathbf{s} \cdot \left\{ -\frac{mc}{e^2n_s}\mathbf{j} + i\frac{\hbar c}{2q}(-2i)\nabla\phi \right\} \\ &= \frac{mc}{e} \oint_{\partial S} d\mathbf{s} \cdot \mathbf{v}_s - \frac{\hbar c}{2e} \oint_{\partial S} d\mathbf{s} \cdot \nabla\phi, \end{aligned} \quad (6.92)$$

where $\mathbf{v}_s := -\mathbf{j}/en_s$ is the superfluid velocity. The last term contains the charge of the phase of the macroscopic wave function along the path, which must be an integer multiple of 2π for $\psi(\mathbf{r})$ to be continuous:

$$\oint_{\partial S} d\mathbf{s} \cdot \nabla\phi = -2\pi n, \quad n \in \mathbb{Z}. \quad (6.93)$$

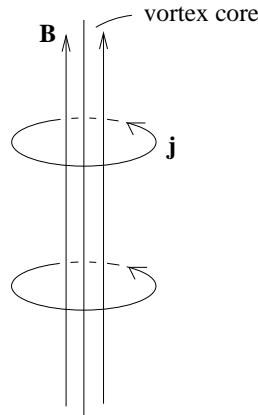
(The minus sign is conventional.) Thus we find for the so-called *fluxoid*

$$\Phi' := \Phi - \frac{mc}{e} \oint_{\partial S} d\mathbf{s} \cdot \mathbf{v}_s = \frac{\hbar c}{2e} 2\pi n = \frac{\hbar c}{2e} n = n\Phi_0, \quad n \in \mathbb{Z}. \quad (6.94)$$

We see that it is not the flux but the fluxoid Φ' that is quantized. However, deep inside a superconducting region the current vanishes and Φ' equals Φ .

Vortices

The smallest amount of fluxoid that can penetrate a superconductor is one flux quantum Φ_0 . It does so as a *vortex* (or *vortex line*).



Following the above arguments, the phase ϕ of ψ changes by -2π as one circles a vortex in the positive direction, where, by convention, the direction of a vortex is the direction of the magnetic field \mathbf{B} . Thus in the center of the vortex line (the *vortex core*), the phase is *undefined*. This is only consistent with a continuous ψ if $\psi = 0$ in the

vortex core. For a vortex along the z -axis we choose cylindrical coordinates ϱ, φ, z . We then have to solve the Ginzburg-Landau equations with the boundary conditions

$$\psi(\varrho = 0) = 0, \quad (6.95)$$

$$|\psi(\varrho \rightarrow \infty)| = \psi_0, \quad (6.96)$$

$$\mathbf{B}(\varrho \rightarrow \infty) = 0. \quad (6.97)$$

From symmetry, we have

$$\psi(\mathbf{r}) = |\psi|(\varrho) e^{i\phi} = |\psi|(\varrho) e^{-i\varphi}, \quad (6.98)$$

$$\mathbf{v}_s(\mathbf{r}) = \hat{\varphi} v_s(\varrho), \quad (6.99)$$

$$\mathbf{B}(\mathbf{r}) = \hat{z} B(\varrho) \quad (6.100)$$

and we can choose

$$\mathbf{A}(\mathbf{r}) = \hat{\varphi} A(\varrho) \quad (6.101)$$

where

$$B(\varrho) = \frac{1}{\varrho} \frac{\partial}{\partial \varrho} \varrho A(\varrho). \quad (6.102)$$

Choosing a circular integration path of radius r centered at the vortex core, the enclosed fluxoid is

$$\Phi'(\mathbf{r}) = \Phi(r) - \frac{mc}{e} \oint_{\partial S} d\mathbf{s} \cdot \mathbf{v}_s = 2\pi r A(r) - \frac{mc}{e} 2\pi r v_s(r) \stackrel{!}{=} \Phi_0$$

$$\Rightarrow A(r) - \frac{mc}{e} v_s(r) = \frac{\Phi_0}{2\pi r}. \quad (6.103)$$

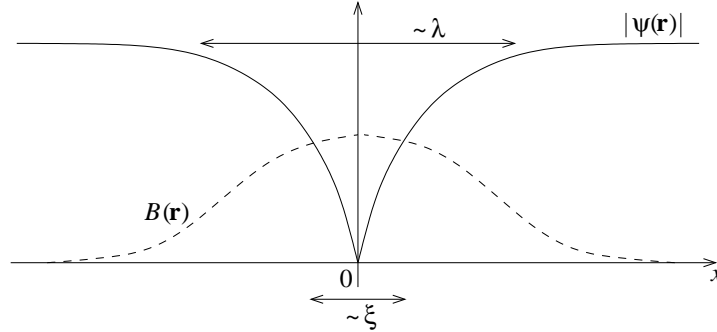
This relation follows from fluxoid quantization and thus ultimately from the second Ginzburg-Landau equation. To obtain $\mathbf{j}(\mathbf{r})$, $\psi(\mathbf{r})$, and $\mathbf{A}(\mathbf{r})$, one also has to solve the first Ginzburg-Landau equation

$$\frac{1}{4m} \left(\frac{\hbar}{i} \nabla + \frac{2e}{c} \mathbf{A} \right)^2 \psi + \alpha \psi + \beta |\psi|^2 \psi = 0 \quad (6.104)$$

and Ampère's Law

$$\mathbf{j} = \frac{c}{4\pi} \nabla \times \mathbf{B}. \quad (6.105)$$

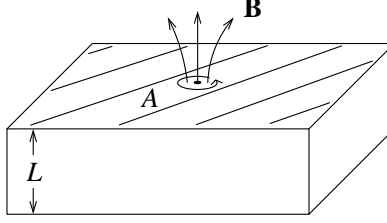
This cannot be done analytically because of the non-linear term $\beta |\psi|^2 \psi$. For small distances ϱ from the core, one can drop this term, thereby linearizing the Ginzburg-Landau equation. The solution is still complicated, the result is that $|\psi|$ increases *linearly* in ϱ . Numerical integration of the full equations gives the results sketched here for a cut through the vortex core:



Another useful quantity is the free energy per unit length of a vortex line (its *line tension*). An analytical expression can be obtained in the strong type-II case of $\kappa \gg 1$. We only give the result here: The vortex line tension is

$$\begin{aligned}\epsilon_v &\approx \left(\frac{\Phi_0}{4\pi\lambda} \right)^2 \ln \kappa \\ &= \frac{H_c^2}{8\pi} 4\pi\xi^2 \ln \kappa.\end{aligned}\tag{6.106}$$

We can now calculate the field for which the first vortex enters the superconductor, the so-called *lower critical field* H_{c1} . This happens when the Gibbs free energy for a superconductor without vortices (Meißner phase) equals the Gibbs free energy in the presence of a single vortex. We assume the sample to have cross section A and thickness L , parallel to the applied field.



Then we have the condition

$$\begin{aligned}0 &= G_{\text{one vortex}} - G_{\text{no vortex}} \\ &= \cancel{\mathcal{F}_s} + L\epsilon_v - \frac{1}{4\pi} \int d^3r \mathbf{H} \cdot \mathbf{B} - \cancel{\mathcal{F}_s} \\ &= L\epsilon_v - \frac{H_{c1}}{4\pi} \int d^3r B(\mathbf{r}) \\ &= L\epsilon_v - \frac{H_{c1}L}{4\pi} \Phi_0.\end{aligned}\tag{6.107}$$

Thus

$$H_{c1} = \frac{4\pi\epsilon_v}{\Phi_0}.\tag{6.108}$$

The line tension in Eq. (6.106) can also be written as

$$\epsilon_v \approx \frac{\Phi_0}{4\pi\lambda} \frac{H_c \xi}{\sqrt{2}} \ln \kappa = \frac{\Phi_0}{4\pi} \frac{H_c}{\sqrt{2}} \frac{\ln \kappa}{\kappa}\tag{6.109}$$

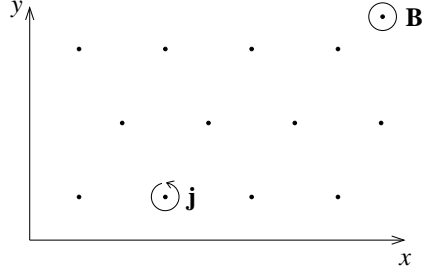
so that

$$H_{c1} = \frac{H_c}{\sqrt{2}} \frac{\ln \kappa}{\kappa}.\tag{6.110}$$

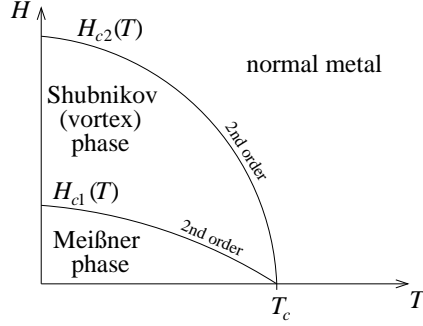
Recall that this expression only holds for $\kappa \gg 1$. H_c is the *thermodynamic critical field* derived above. In a type-II superconductor, nothing interesting happens at $H = H_c$.

The Abrikosov vortex lattice

We have considered the structure of an isolated vortex line. How does a finite magnetic flux penetrate a type-II superconductor? Based on Ginzburg-Landau theory, A. A. Abrikosov proposed in 1957 that flux should enter as a periodic lattice of parallel vortex lines carrying a single flux quantum each. He proposed a square lattice, which was due to a small mistake. The lowest-free-energy state is actually a *triangular lattice*.



As noted above, the magnetic flux starts to penetrate the superconductor at the *lower critical field* H_{c1} . Furthermore, since flux expulsion in type-II superconductors need not be perfect, they can withstand stronger magnetic fields than type-I superconductors, up to an *upper critical field* $H_{c2} > H_c, H_{c1}$.



We will now review the basic ideas of Abrikosov's approach. Abrikosov's results are quantitatively valid only close to H_{c2} since he assumed the magnetic flux density \mathbf{B} to be uniform, which is valid for

$$\lambda \gg l, \quad (6.111)$$

where

$$l = \sqrt{\frac{\Phi_0}{B}} \quad (6.112)$$

is the typical distance between vortices (B/Φ_0 is the two-dimensional concentration of vortex lines). For $\mathbf{B} = B\hat{z} = \text{const} = H\hat{z}$ we can choose the gauge $\mathbf{A} = \hat{y}Hx$. Then the first Ginzburg-Landau equation becomes (note $m^* = 2m$)

$$\frac{1}{2m^*} \left(\frac{\hbar}{i} \nabla + \frac{2eH}{c} \hat{y}x \right)^2 \psi + \alpha \psi + \beta |\psi|^2 \psi = 0. \quad (6.113)$$

Slightly below H_{c2} , $|\psi|$ is expected to be small (this should be checked!) so that we can neglect the non-linear term. Introducing the cyclotron frequency of a superconductor,

$$\omega_c := \frac{2eH}{m^*c}, \quad (6.114)$$

we obtain

$$\left(-\frac{\hbar^2}{2m^*} \nabla^2 - i\hbar\omega_c x \frac{\partial}{\partial y} + \frac{1}{2}m^*\omega_c^2 x^2 \right) \psi(\mathbf{r}) = \underbrace{-\alpha}_{>0} \psi(\mathbf{r}). \quad (6.115)$$

This looks very much like a Schrödinger equation and from the derivation it has to be the Schrödinger equation for a particle of mass m^* and charge $q = -2e$ in a uniform magnetic field H . This well-known problem is solved by the ansatz

$$\psi(x, y) = e^{ik_y y} f(x). \quad (6.116)$$

We obtain

$$\begin{aligned}
& e^{-ik_y y} \left(-\frac{\hbar^2}{2m^*} \nabla^2 - i\hbar\omega_c x \frac{\partial}{\partial y} + \frac{1}{2} m^* \omega_c^2 x^2 \right) e^{ik_y y} f(x) \\
&= -\frac{\hbar^2}{2m^*} \frac{d^2 f}{dx^2} + \frac{\hbar^2 k_y^2}{2m^*} f - i\hbar\omega_c x i k_y f + \frac{1}{2} m^* \omega_c^2 x^2 f \\
&= -\frac{\hbar^2}{2m^*} \frac{d^2 f}{dx^2} + \frac{1}{2} m^* \omega_c^2 \left(x^2 + 2 \frac{\hbar k_y}{m^* \omega_c} + \frac{\hbar^2 k_y^2}{(m^*)^2 \omega_c^2} \right) f = -\alpha f.
\end{aligned} \tag{6.117}$$

Defining $x_0 := -\hbar k_y / m^* \omega_c$, we get

$$-\frac{\hbar^2}{2m^*} \frac{d^2 f}{dx^2} + \frac{1}{2} m^* \omega_c^2 (x - x_0)^2 f = -\alpha f. \tag{6.118}$$

This is the Schrödinger equation for a one-dimensional harmonic oscillator with shifted minimum. Thus we obtain solutions $f_n(x)$ as shifted harmonic-oscillator eigenfunctions for

$$-\alpha = \hbar\omega_c \left(n + \frac{1}{2} \right), \quad n = 0, 1, 2, \dots \tag{6.119}$$

Solution can only exist if

$$-\alpha = -\alpha'(T - T_c) = \alpha'(T_c - T) \geq \frac{\hbar\omega_c}{2} = \frac{\hbar e H}{m^* c}. \tag{6.120}$$

Keeping $T \leq T_c$ fixed, a solution can thus only exist for $H \leq H_{c2}$ with the upper critical field

$$H_{c2} = -\frac{m^* c}{\hbar e} \alpha. \tag{6.121}$$

Using $\xi^2 = -\hbar^2 / 2m^* \alpha$, we obtain

$$H_{c2}(T) = \frac{\hbar c}{2e \xi^2(T)} = \frac{\Phi_0}{2\pi \xi^2(T)}. \tag{6.122}$$

Note that since $\xi \sim 1/\sqrt{T_c - T}$ close to T_c , $H_{c2}(T)$ sets in linearly, as shown in the above sketch. H_{c2} should be compared to the thermodynamic critical field H_c ,

$$\begin{aligned}
H_c^2 &= -\frac{mc^2}{2e^2} \frac{1}{\lambda^2} \alpha = \frac{\hbar^2 c^2}{8e^2} \frac{1}{\lambda^2 \xi^2} = \frac{1}{8\pi^2} \left(\frac{\hbar c}{2e} \right)^2 \frac{1}{\lambda^2 \xi^2} \\
\Rightarrow H_c &= \frac{\Phi_0}{2\pi \sqrt{2} \lambda \xi} = \frac{H_{c2} \xi}{\sqrt{2} \lambda} = \frac{H_{c2}}{\sqrt{2} \kappa} \\
\Rightarrow H_{c2} &= \sqrt{2} \kappa H_c.
\end{aligned} \tag{6.123}$$

For $\kappa = 1/\sqrt{2}$, H_{c2} equals H_c . This is the transition between a type-II and a type-I superconductor.

So far, our considerations have not told us what the state for $H \lesssim H_{c2}$ actually looks like. To find out, one in principle has to solve the full, not the linearized, Ginzburg-Landau equation. (We have seen that the linearized equation is equivalent to the Schrödinger equation for a particle in a uniform magnetic field. The solutions are known: The eigenfunctions have uniform amplitude and the eigenenergies form discrete Landau levels, very different from what is observed in the Shubnikov phase.) Abrikosov did this approximately using a variational approach. He used linear combinations of solutions of the linearized equation as the variational ansatz and *assumed* that the solution is periodic in two dimensions, up to a plane wave.

We call the periods in the x - and y -directions a_x and a_y , respectively. The function

$$\psi_n(x, y) = e^{ik_y y} f_n(x) \tag{6.124}$$

has the period a_y in y if

$$k_y = \frac{2\pi}{a_y} q, \quad q \in \mathbb{Z}. \tag{6.125}$$

Then the harmonic oscillator is centered at

$$x_0 = -\frac{\hbar}{m^*\omega_c} \frac{2\pi}{a_y} q = -\frac{\Phi_0}{Ha_y} q. \quad (6.126)$$

Since the lowest-energy solution is obtained for $n = 0$ (the ground state of the harmonic oscillator), Abrikosov only considered the $n = 0$ solutions

$$\psi_0(x, y) = \exp\left(iq \frac{2\pi y}{a_y}\right) f_0(x) = \underbrace{C}_{\text{normalization}} \exp\left(iq \frac{2\pi y}{a_y}\right) \exp\left(-\frac{1}{2} \frac{m^*\omega_c}{\hbar} \left[x + \underbrace{\frac{\Phi_0}{Ha_y} q}_{-x_0}\right]^2\right). \quad (6.127)$$

In the Gauss function we find the quantity

$$\frac{\hbar}{m^*\omega_c} = \frac{\hbar c}{2eH} \approx \frac{\hbar c}{2eH_{c2}} = -\frac{\hbar^2}{2m^*\alpha} = \xi^2(T), \quad (6.128)$$

as long as $H \approx H_{c2}$. Thus

$$\psi_0(x, y) = C \exp\left(iq \frac{2\pi y}{a_y}\right) \exp\left(-\frac{1}{2\xi^2} \left[x + \frac{\Phi_0}{Ha_y} q\right]^2\right). \quad (6.129)$$

This is a set of functions enumerated by $q \in \mathbb{Z}$. Abrikosov considered linear combinations

$$\psi(x, y) = \sum_{q=-\infty}^{\infty} C_q \exp\left(iq \frac{2\pi y}{a_y}\right) \exp\left(-\frac{1}{2\xi^2} \left[x + \frac{\Phi_0}{Ha_y} q\right]^2\right). \quad (6.130)$$

To be periodic in x with period a_x , up to a plane wave (the corresponding discussion in Tinkham's book is not fully correct), this ansatz has to satisfy

$$\begin{aligned} \psi(x + a_x, y) &= \sum_{q=-\infty}^{\infty} C_q \exp\left(iq \frac{2\pi y}{a_y}\right) \exp\left(-\frac{1}{2\xi^2} \left[x + a_x + \frac{\Phi_0}{Ha_y} q\right]^2\right) \\ &= \sum_{q=-\infty}^{\infty} C_q \exp\left(iq \frac{2\pi y}{a_y}\right) \exp\left(-\frac{1}{2\xi^2} \left[x + \frac{\Phi_0}{Ha_y} \left(q + \frac{Ha_y}{\Phi_0} a_x\right)\right]^2\right) \\ &\stackrel{!}{\propto} \psi(x, y) \quad \forall x, y. \end{aligned} \quad (6.131)$$

This requires

$$\frac{Ha_x a_y}{\Phi_0} =: \nu \in \mathbb{N}. \quad (6.132)$$

(Note that this quantity is positive.) Then

$$\begin{aligned} \psi(x + a_x, y) &= \sum_{q=-\infty}^{\infty} C_q \exp\left(iq \frac{2\pi y}{a_y}\right) \exp\left(-\frac{1}{2\xi^2} \left[x + \frac{\Phi_0}{Ha_y} (q + \nu)\right]^2\right) \\ &= \exp\left(i\nu \frac{2\pi y}{a_y}\right) \sum_{q=-\infty}^{\infty} C_{q-\nu} \exp\left(iq \frac{2\pi y}{a_y}\right) \exp\left(-\frac{1}{2\xi^2} \left[x + \frac{\Phi_0}{Ha_y} q\right]^2\right). \end{aligned} \quad (6.133)$$

This equals $\psi(x, y)$ up to a plane-wave factor if

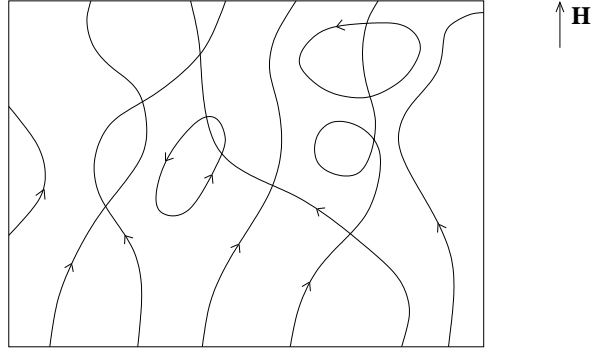
$$C_{q-\nu} = C_q \quad \forall q, \quad (6.134)$$

i.e., if C_q is periodic. Abrikosov considered the case $\nu = 1$, which leads to a square lattice. The lowest-free-energy solution is obtained for $\nu = 2$ and $C_1 = iC_0$, which gives a triangular lattice, sketched above. Note that

$$\nu = \frac{Ha_x a_y}{\Phi_0} \quad (6.135)$$

has a simple interpretation: It is the number of flux quanta passing through a rectangular unit cell of the vortex lattice.

The vortex lattice is a rather complex system: It is a lattice of interacting lines with a line tension ϵ_v . At non-zero temperatures, the vortices fluctuate, which can lead to the *melting* of the vortex lattice. The resulting *vortex liquid* can be pictured as a pot of boiling spaghetti, with the constraint that the vortex lines must either terminate at the surface or form closed loops (!). Moving vortices lead to ohmic resistance, even though most of the sample is still superconducting. To complicate matters, interaction of vortices with defects (“pinning”) plays an important role. There is an extensive research literature on vortex matter, which we cannot review here.



7

Superfluid and superconducting films

Generally, fluctuations are stronger in systems of lower dimension. Indeed, they change the properties of two-dimensional superfluid and superconducting films qualitatively compared to three-dimensional samples. We will consider such films within the Ginzburg-Landau theory.

7.1 Superfluid films

We start from the two-dimensional Landau functional

$$F[\psi] = \int d^2r \left[\alpha |\psi|^2 + \frac{\beta}{2} |\psi|^4 + \gamma (\nabla \psi)^* \cdot \nabla \psi \right]. \quad (7.1)$$

We consider temperatures $T < T_c$. As shown in Sec. 6.2, fluctuations of the amplitude $|\psi|$ then have an energy gap, whereas fluctuations of the phase ϕ are ungapped. Not too close to T_c , phase fluctuations will thus dominate and we neglect amplitude fluctuations, writing

$$\psi(\mathbf{r}) = \psi_0 e^{i\phi(\mathbf{r})} \quad \text{with} \quad \psi_0 = \sqrt{-\alpha/\beta}. \quad (7.2)$$

Thus, up to an irrelevant constant,

$$F[\phi] = \int d^2r \underbrace{\left(-\gamma \frac{\alpha}{\beta} \right)}_{>0} (\nabla \phi)^* \cdot \nabla \phi = \sum_{\mathbf{k}} \left(-\gamma \frac{\alpha}{\beta} \right) k^2 \phi_{\mathbf{k}}^* \phi_{\mathbf{k}}. \quad (7.3)$$

We can now calculate the correlation function

$$\langle \psi(\mathbf{r})^* \psi(0) \rangle = \psi_0^2 \langle e^{-i\phi(\mathbf{r})} e^{i\phi(0)} \rangle = \psi_0^2 \frac{1}{Z} \int D\phi e^{-i\phi(\mathbf{r}) + i\phi(0)} e^{-F[\phi]/k_B T}. \quad (7.4)$$

This is a Gaussian average since $F[\phi]$ is bilinear in $\phi_{\mathbf{k}}$. Thus we can write

$$\begin{aligned}
\langle \psi(\mathbf{r})^* \psi(0) \rangle &= \psi_0^2 \sum_{n=0}^{\infty} \frac{1}{n!} (-i)^n \langle (\phi(\mathbf{r}) - \phi(0))^n \rangle \\
&= \psi_0^2 \sum_{\substack{n=0 \\ n \text{ even}}}^{\infty} \frac{1}{n!} (-i)^n 1 \times 3 \times 5 \times \cdots \times (n-1) \langle (\phi(\mathbf{r}) - \phi(0))^2 \rangle \\
&\equiv \psi_0^2 \sum_{m=0}^{\infty} (-1)^m \frac{1 \times 3 \times 5 \times \cdots \times (2m-1)}{1 \times 2 \times 3 \times \cdots \times (2m)} \langle (\phi(\mathbf{r}) - \phi(0))^2 \rangle \\
&= \psi_0^2 \sum_{m=0}^{\infty} (-1)^m \frac{1}{\underbrace{2 \times 4 \times \cdots \times (2m)}_{=2^m m!}} \langle (\phi(\mathbf{r}) - \phi(0))^2 \rangle \\
&= \psi_0^2 \sum_{m=0}^{\infty} \frac{1}{m!} \left(-\frac{1}{2} \right)^m \langle (\phi(\mathbf{r}) - \phi(0))^2 \rangle = \psi_0^2 \exp \left(-\frac{1}{2} \langle (\phi(\mathbf{r}) - \phi(0))^2 \rangle \right). \tag{7.5}
\end{aligned}$$

Herein, we have

$$\begin{aligned}
\langle (\phi(\mathbf{r}) - \phi(0))^2 \rangle &= \frac{1}{V} \sum_{\mathbf{k}\mathbf{k}'} (e^{-i\mathbf{k}\cdot\mathbf{r}} - 1)(e^{i\mathbf{k}'\cdot\mathbf{r}} - 1) \langle \phi_{\mathbf{k}}^* \phi_{\mathbf{k}'} \rangle \\
&= \frac{1}{V} \frac{1}{2} \sum_{\mathbf{k}\mathbf{k}'} (e^{-i\mathbf{k}\cdot\mathbf{r}} - 1)(e^{i\mathbf{k}'\cdot\mathbf{r}} - 1) \delta_{\mathbf{k}\mathbf{k}'} \frac{k_B T \beta}{-\gamma \alpha} \frac{1}{k^2} \\
&= \frac{1}{V} \frac{1}{2} \sum_{\mathbf{k}} (2 - 2 \cos \mathbf{k} \cdot \mathbf{r}) \frac{k_B T \beta}{-\gamma \alpha} \frac{1}{k^2}. \tag{7.6}
\end{aligned}$$

Going over to the continuum, we obtain

$$\begin{aligned}
\langle (\phi(\mathbf{r}) - \phi(0))^2 \rangle &= \frac{k_B T \beta}{-\gamma \alpha} \int \frac{d^2 k}{(2\pi)^2} \frac{1}{k^2} (1 - \cos \mathbf{k} \cdot \mathbf{r}) \\
&= \frac{k_B T \beta}{-\gamma \alpha} \frac{1}{(2\pi)^2} \int_0^{2\pi} d\varphi \int_0^{\Lambda} dk \frac{1}{k} (1 - \cos(kr \cos \varphi)) \\
&= \frac{k_B T \beta}{-\gamma \alpha} \frac{1}{2\pi} \int_0^{\Lambda} dk \frac{1}{k} (1 - \underbrace{J_0(kr)}_{\text{Bessel function}}). \tag{7.7}
\end{aligned}$$

The k -integral is cut off at Λ , which is the inverse of some microscopic length scale, $\lambda = 1/r_0$. The idea is that physics at shorter length scales $r < r_0$ have been “integrated out” to obtain the Landau functional. For large distances, $r \gg r_0$, we approximate, rather brutally,

$$J_0(kr) \approx \begin{cases} 1 & \text{for } kr < 1 \\ 0 & \text{for } kr \geq 1 \end{cases} \tag{7.8}$$

and obtain

$$\langle (\phi(\mathbf{r}) - \phi(0))^2 \rangle \cong \frac{k_B T \beta}{-\gamma \alpha} \frac{1}{2\pi} \int_{1/r}^{1/r_0} \frac{dk}{k} = \frac{k_B T \beta}{-\gamma \alpha} \frac{1}{2\pi} \ln \frac{r}{r_0} \tag{7.9}$$

and thus

$$\langle \psi(\mathbf{r})^* \psi(0) \rangle \cong \psi_0^2 \exp \left(\frac{1}{2} \frac{k_B T \beta}{\gamma \alpha} \frac{1}{2\pi} \ln \frac{r}{r_0} \right) = \psi_0^2 \left(\frac{r}{r_0} \right)^{-\eta} \tag{7.10}$$

with

$$\eta = -\frac{1}{4\pi} \frac{k_B T \beta}{\gamma \alpha} > 0. \quad (7.11)$$

Thus the correlation function of the order parameters decays like a *power law* of distance in two dimensions. We do not find long-range order, which would imply $\lim_{r \rightarrow \infty} \langle \psi(\mathbf{r})^* \psi(0) \rangle = \text{const}$. This agrees with the Mermin-Wagner theorem, which forbids long-range order for the two-dimensional superfluid. The power-law decay characterizes so-called *quasi-long-range order* (short range order would have an even faster, e.g., exponential, decay).

Isolated vortices

We have argued that fluctuations in the amplitude are less important because they have an energy gap proportional to $-2\alpha > 0$. This is indeed true for *small* amplitude fluctuations. However, there exist variations of the amplitude that are, while energetically costly, very stable once they have been created. These are vortices. In two dimensions, a vortex is a zero-dimensional object; the order parameter goes to zero at a single point at its center. The simplest form of a vortex at the origin can be represented by

$$\psi(\mathbf{r}) = |\psi(\mathbf{r})| e^{i\phi(\mathbf{r})} = |\psi(r)| e^{i(\varphi - \varphi_0)}, \quad (7.12)$$

where r and φ are (planar) polar coordinates of \mathbf{r} . An *antivortex* would be described by

$$\psi(\mathbf{r}) = |\psi(r)| e^{-i(\varphi - \varphi_0)}. \quad (7.13)$$

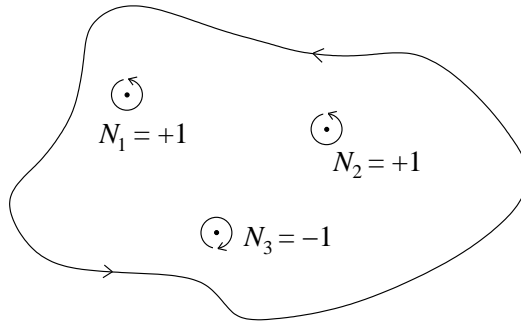
In both cases, $\lim_{r \rightarrow 0} |\psi(r)| = 0$. Note that we have changed the convention for the sign in the exponent compared to superconducting vortices.

In the presence of vortices, the phase $\phi(\mathbf{r})$ of the order parameter is multivalued and, of course, undefined at the vortex centers. On the other hand, the phase gradient $\mathbf{v} = \nabla \phi$ is single-valued (but still undefined at the vortex centers). For any closed loop \mathcal{C} not touching any vortex cores, we have

$$\oint_{\mathcal{C}} ds \cdot \mathbf{v} = \text{total change in phase along } \mathcal{C} = 2\pi N_{\mathcal{C}}, \quad (7.14)$$

where $N_{\mathcal{C}} \in \mathbb{Z}$ is the enclosed *winding number* or *vorticity*. The vorticity can be written as the sum of the vorticities $N_i = \pm 1$ of all vortices and antivortices inside the loop,

$$N_{\mathcal{C}} = \sum_i N_i. \quad (7.15)$$



Defining the vortex concentration by

$$n_v(\mathbf{r}) := \sum_i N_i \delta(\mathbf{r} - \mathbf{R}_i), \quad (7.16)$$

where \mathbf{R}_i is the location of the center of vortex i , we obtain

$$\nabla \times \mathbf{v} \equiv \frac{\partial}{\partial x} v_y - \frac{\partial}{\partial y} v_x = 2\pi n_v(\mathbf{r}). \quad (7.17)$$

Note that in two dimensions the curl is a scalar.

Now it is always possible to decompose a vector field into a rotation-free and a divergence-free component,

$$\mathbf{v} = \mathbf{v}_{\text{ph}} + \mathbf{v}_v \quad (7.18)$$

with

$$\nabla \times \mathbf{v}_{\text{ph}} = 0, \quad (7.19)$$

$$\nabla \cdot \mathbf{v}_v = 0. \quad (7.20)$$

Since the vortex concentration associated with \mathbf{v}_{ph} vanishes, the component \mathbf{v}_{ph} does not contain any vortices. Alternatively, note that the first equation implies that there exists a *single-valued* scalar field $\Omega(\mathbf{r})$ so that

$$\mathbf{v}_{\text{ph}} = \nabla \Omega. \quad (7.21)$$

Ω is a single-valued component of the phase, which cannot be due to vortices. This is the contribution from small phase fluctuations, which we have already discussed. Conversely, \mathbf{v}_v is the vortex part, for which

$$\nabla \times \mathbf{v}_v = 2\pi n_v, \quad (7.22)$$

$$\nabla \cdot \mathbf{v}_v = 0. \quad (7.23)$$

This already suggests an electro-dynamical analogy, but to formulate this analogy it is advantageous to rescale and rotate the field \mathbf{v}_v :

$$\mathbf{E}(\mathbf{r}) := - \underbrace{\sqrt{-2\gamma \frac{\alpha}{\beta}}}_{>0} \hat{\mathbf{z}} \times \mathbf{v}_v(\mathbf{r}). \quad (7.24)$$

Then the energy density far from vortex cores, where $|\psi| \cong \sqrt{-\alpha/\beta}$, is

$$w = -\gamma \frac{\alpha}{\beta} (\nabla \phi_v) \cdot \nabla \phi_v = -\gamma \frac{\alpha}{\beta} \mathbf{v}_v \cdot \mathbf{v}_v = -\gamma \frac{\alpha}{\beta} \left(-2\gamma \frac{\alpha}{\beta} \right)^{-1} \mathbf{E} \cdot \mathbf{E} = \frac{1}{2} \mathbf{E} \cdot \mathbf{E}. \quad (7.25)$$

Also, we find

$$\nabla \cdot \mathbf{E} = -\sqrt{-2\gamma \frac{\alpha}{\beta}} (-\nabla \times \mathbf{v}_v) = 2\pi \sqrt{-2\gamma \frac{\alpha}{\beta}} n_v \quad (7.26)$$

and

$$\nabla \times \mathbf{E} = 0. \quad (7.27)$$

These equations reproduce the fundamental equations of electrostatics if we identify the ‘‘charge density’’ with

$$\rho_v = \sqrt{-2\gamma \frac{\alpha}{\beta}} n_v. \quad (7.28)$$

(The factor in Gauss’ law is 2π instead of 4π since we are considering a two-dimensional system.) We can now derive the pseudo-electric field $\mathbf{E}(\mathbf{r})$ for a single vortex,

$$\oint d\mathbf{a} \cdot \mathbf{E} = 2\pi r E = 2\pi \sqrt{-2\gamma \frac{\alpha}{\beta}} \quad (7.29)$$

$$\Rightarrow E(\mathbf{r}) = \frac{\sqrt{-2\gamma \frac{\alpha}{\beta}}}{r} \quad (7.30)$$

and thus

$$\mathbf{E}(\mathbf{r}) = \frac{\sqrt{-2\gamma \frac{\alpha}{\beta}}}{r} \hat{\mathbf{r}}. \quad (7.31)$$

From this, we obtain the energy of a single vortex,

$$E_1 = E_{\text{core}} + \int d^2r \frac{1}{2} \mathbf{E} \cdot \mathbf{E} = E_{\text{core}} - \frac{1}{2} 2\gamma \frac{\alpha}{\beta} \int d^2r \frac{1}{r^2} = E_{\text{core}} - 2\pi\gamma \frac{\alpha}{\beta} \int dr \frac{1}{r}. \quad (7.32)$$

Now we note that the derivation does not hold for small distances from the vortex center since there $|\psi|^2$ is not close to $-\alpha/\beta$. Thus we cut off the radial integral at the lower end at some ‘‘vortex core radius’’ r_0 and put all energy contributions from the core into E_{core} . r_0 is on the order of the coherence length ξ . But the integral still diverges; if our system has a characteristic size of L , we obtain

$$E_1 = E_{\text{core}} - 2\pi\gamma \frac{\alpha}{\beta} \int_{r_0}^L \frac{dr}{r} = E_{\text{core}} - \underbrace{2\pi\gamma \frac{\alpha}{\beta} \ln \frac{L}{r_0}}_{>0}. \quad (7.33)$$

Thus the energy of a single, isolated vortex diverges logarithmically with the system size. This suggests that isolated vortices will never be present as thermal fluctuations as long as $\alpha < 0$. This is not true, though. The probability of such vortices should be

$$\begin{aligned} p_1 &\propto \frac{1}{r_0^2} e^{-E_1/k_B T} = \frac{1}{r_0^2} e^{-E_{\text{core}}/k_B T} \exp\left(\frac{2\pi}{k_B T} \gamma \frac{\alpha}{\beta} \ln \frac{L}{r_0}\right) \\ &= \frac{1}{r_0^2} e^{-E_{\text{core}}/k_B T} \left(\frac{L}{r_0}\right)^{\frac{2\pi}{k_B T} \gamma \frac{\alpha}{\beta}} = \frac{1}{r_0^2} e^{-E_{\text{core}}/k_B T} \left(\frac{L}{r_0}\right)^{-\frac{1}{2\eta}}. \end{aligned} \quad (7.34)$$

The typical area per vortex is $1/p_1$ and the total number of vortices will be, on average,

$$N_v = \frac{L^2}{1/p_1} = L^2 p_1 = e^{-E_{\text{core}}/k_B T} \left(\frac{L}{r_0}\right)^{2-\frac{1}{2\eta}}. \quad (7.35)$$

For $\eta > 1/4$, N_v diverges in the thermodynamic limit so that infinitely many vortices are present. For $\eta < 1/4$, $N_v \rightarrow 0$ for $L \rightarrow \infty$ and according to our argument, which is essentially due to Kosterlitz and Thouless, there are no vortices. It is plausible and indeed true that free vortices destroy quasi-long-range order and in this sense also superfluidity. Note that

$$\eta = -\frac{1}{4\pi} \frac{k_B T \beta}{\gamma \alpha(T)} \stackrel{!}{=} \frac{1}{4} \quad (7.36)$$

is an equation for a critical temperature for the appearance of free vortices. We thus find that the critical temperature in superfluid films should be reduced from the point where $\alpha = 0$ ($\eta = \infty$) to the one where $\eta = 1/4$ due to vortices appearing as fluctuations of the order parameter. While qualitatively true, our description is still incomplete, though, since we have so far neglected interactions between vortices.

Vortex interaction

The energy of two vortices with vorticities ± 1 can easily be obtained from the electrostatic analogy. We assume that core regions do not overlap, i.e., the separation is $R \geq 2r_0$. The pseudo-electric field of the two vortices, assumed to be located at $\pm \mathbf{R}/2 = \pm R \hat{\mathbf{x}}/2$, is

$$\begin{aligned} \mathbf{E}(\mathbf{r}) &= \sqrt{-2\gamma \frac{\alpha}{\beta}} \frac{\mathbf{r} - \mathbf{R}/2}{|\mathbf{r} - \mathbf{R}/2|^2} - \sqrt{-2\gamma \frac{\alpha}{\beta}} \frac{\mathbf{r} + \mathbf{R}/2}{|\mathbf{r} + \mathbf{R}/2|^2} \\ &= \sqrt{-2\gamma \frac{\alpha}{\beta}} \frac{|\mathbf{r} + \mathbf{R}/2|^2 (\mathbf{r} - \mathbf{R}/2) - |\mathbf{r} - \mathbf{R}/2|^2 (\mathbf{r} + \mathbf{R}/2)}{|\mathbf{r} - \mathbf{R}/2|^2 |\mathbf{r} + \mathbf{R}/2|^2}. \end{aligned} \quad (7.37)$$

Thus the energy is

$$\begin{aligned}
E_2 &= 2E_{\text{core}} + \int d^2r \frac{1}{2} \mathbf{E} \cdot \mathbf{E} \\
&= 2E_{\text{core}} - \frac{1}{2} 2\gamma \frac{\alpha}{\beta} \int d^2r \left(\frac{|\mathbf{r} + \mathbf{R}/2|^2 (\mathbf{r} - \mathbf{R}/2) - |\mathbf{r} - \mathbf{R}/2|^2 (\mathbf{r} + \mathbf{R}/2)}{|\mathbf{r} - \mathbf{R}/2|^2 |\mathbf{r} + \mathbf{R}/2|^2} \right)^2 \\
&= 2E_{\text{core}} - \gamma \frac{\alpha}{\beta} 2 \int_{-\infty}^{\infty} dx \int_0^{\infty} dy \\
&\quad \times \frac{|\mathbf{r} + \mathbf{R}/2|^4 |\mathbf{r} - \mathbf{R}/2|^2 + |\mathbf{r} - \mathbf{R}/2|^4 |\mathbf{r} + \mathbf{R}/2|^2 - 2 |\mathbf{r} + \mathbf{R}/2|^2 |\mathbf{r} - \mathbf{R}/2|^2 (r^2 - R^2/4)}{|\mathbf{r} - \mathbf{R}/2|^4 |\mathbf{r} + \mathbf{R}/2|^4}. \tag{7.38}
\end{aligned}$$

We introduce elliptic coordinates σ, τ according to

$$x = \frac{R}{2} \sigma \tau, \tag{7.39}$$

$$y = \frac{R}{2} \sqrt{\sigma^2 - 1} \sqrt{1 - \tau^2}, \tag{7.40}$$

where $\sigma \in [1, \infty[, \tau \in [-1, 1]$. Then

$$\begin{aligned}
E_2 &= 2E_{\text{core}} - 2\gamma \frac{\alpha}{\beta} \left(\frac{R}{2} \right)^2 \int d\sigma d\tau \frac{\sigma^2 - \tau^2}{\sqrt{\sigma^2 - 1} \sqrt{1 - \tau^2}} \\
&\quad \times \frac{\left(\frac{R}{2} \right)^4 (\sigma + \tau)^4 \left(\frac{R}{2} \right)^2 (\sigma - \tau)^2 + \left(\frac{R}{2} \right)^4 (\sigma - \tau)^4 \left(\frac{R}{2} \right)^2 (\sigma + \tau)^2}{\left(\frac{R}{2} \right)^4 (\sigma - \tau)^4 \left(\frac{R}{2} \right)^4 (\sigma + \tau)^4} \\
&\quad - \frac{2 \left(\frac{R}{2} \right)^2 (\sigma + \tau)^2 \left(\frac{R}{2} \right)^2 (\sigma - \tau)^2 \left(\frac{R}{2} \right)^2 (\sigma^2 \tau^2 + (\sigma^2 - 1)(1 - \tau^2) - 1)}{\left(\frac{R}{2} \right)^4 (\sigma - \tau)^4 \left(\frac{R}{2} \right)^4 (\sigma + \tau)^4} \\
&= 2E_{\text{core}} - 2\gamma \frac{\alpha}{\beta} \int d\sigma d\tau \frac{1}{\sqrt{\sigma^2 - 1} \sqrt{1 - \tau^2}} \frac{(\sigma + \tau)^2 + (\sigma - \tau)^2 - 2(\sigma^2 + \tau^2 - 2)}{\sigma^2 - \tau^2} \\
&= 2E_{\text{core}} - 2\gamma \frac{\alpha}{\beta} \int d\sigma d\tau \frac{1}{\sqrt{\sigma^2 - 1} \sqrt{1 - \tau^2}} \frac{4}{\sigma^2 - \tau^2} \\
&= 2E_{\text{core}} - 8\pi\gamma \frac{\alpha}{\beta} \int d\sigma \frac{1}{\sigma(\sigma^2 - 1)}. \tag{7.41}
\end{aligned}$$

We have to keep in mind that the integrals in real space have a lower cutoff r_0 . The minimum separation from the vortex at $R\hat{\mathbf{x}}/2$ is $(\sigma - 1)R/2$. For this separation to equal r_0 , the lower cutoff for σ must be

$$\sigma_0 = 1 + \frac{2r_0}{R}. \tag{7.42}$$

With this cutoff, we obtain

$$E_2 = 2E_{\text{core}} - 8\pi\gamma \frac{\alpha}{\beta} \frac{1}{2} \ln \frac{(R + 2r_0)^2}{4r_0(R + r_0)}, \tag{7.43}$$

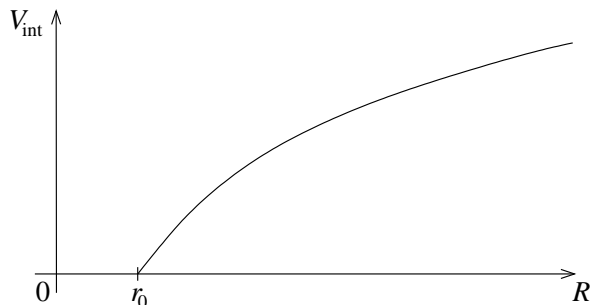
which for $R \gg r_0$ becomes

$$E_2 = 2E_{\text{core}} - 4\pi\gamma \frac{\alpha}{\beta} \ln \frac{R}{4r_0}. \tag{7.44}$$

We absorb an R -independent constant into E_{core} and finally obtain

$$E_2 \equiv 2E_{\text{core}} + V_{\text{int}}(R) = 2E_{\text{core}} - \underbrace{4\pi\gamma \frac{\alpha}{\beta} \ln \frac{R}{r_0}}_{>0}. \tag{7.45}$$

Note that the energy of a vortex-antivortex pair remains finite for large system sizes, $L \rightarrow \infty$. Also, the vortex-antivortex interaction $V_{\text{int}}(R)$ increases with R , i.e., vortex and antivortex *attract* each other. Conversely, one can show that vortices with the same vorticity *repel* each other.



One can also show that for arbitrary vorticities $N_1, N_2 \in \mathbb{Z}$, the interaction reads

$$V_{\text{int}}(R) = \underbrace{4\pi\gamma\frac{\alpha}{\beta}}_{< 0} N_1 N_2 \ln \frac{R}{r_0}. \quad (7.46)$$

Since the equations of (pseudo-) electrostatics are linear, the superposition principle applies and we do not have additional 3-, 4-, etc. body interactions. The energy of a system of vortices is thus

$$E = \sum_i E_{\text{core}} + \frac{1}{2} \sum_{ij, i \neq j} N_i N_j v_{\text{int}}(|\mathbf{r}_i - \mathbf{r}_j|) \quad (7.47)$$

with

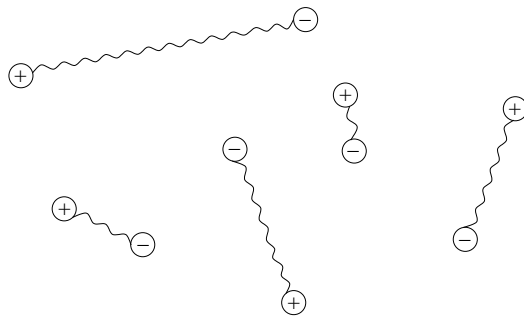
$$v_{\text{int}}(r) := 4\pi\gamma\frac{\alpha}{\beta} \ln \frac{r}{r_0}, \quad (7.48)$$

provided that $\sum_i N_i = 0$. If the total vorticity $\sum_i N_i$ does not vanish, the energy diverges logarithmically with the system size, as we have seen. As long as the total vorticity is zero, the energy per vortex is finite and thus we expect a non-zero concentration of vortices for all temperatures $T > 0$. We now want to understand the consequences of their presence.

Berezinskii-Kosterlitz-Thouless theory

The vortices in a two-dimensional superfluid behave like a *Coulomb gas*—a gas of charged particles with Coulomb interaction, which is logarithmic in 2D. We have seen that the total vorticity (charge) has to vanish. It is therefore possible to group the vortices into *vortex-antivortex pairs*. We do this using the following simple algorithm:

1. find the vortex and the antivortex with the smallest separation r ,
2. mark this vortex and this antivortex as a pair,
3. repeat these steps for all remaining vortices until none are left.



The energy of an isolated vortex-antivortex pair (we will use the term “*vortex pair*” from now on) of size r is

$$E_2(r) = 2E_{\text{core}} - 4\pi\gamma\frac{\alpha}{\beta} \ln \frac{r}{r_0}. \quad (7.49)$$

Thus the probability density of such pairs is

$$p_2(r) = \frac{1}{r_0^4} e^{-E_2(r)/k_B T} = \frac{1}{r_0^4} y_0^2 \exp\left(-2\pi K_0 \ln \frac{r}{r_0}\right) = \frac{1}{r_0^4} y_0^2 \left(\frac{r}{r_0}\right)^{2\pi K_0}, \quad (7.50)$$

where

$$y_0 := N_0 e^{-E_{\text{core}}/k_B T} \quad (7.51)$$

is a vortex *fugacity* (N_0 is a constant of order unity) or, more precisely, y_0^2 is a vortex-pair fugacity, and

$$K_0 := -\frac{1}{k_B T} 2\gamma\frac{\alpha}{\beta} > 0 \quad (7.52)$$

is a dimensionless measure for the interaction strength in units of the thermal energy. K_0 is called the *stiffness*. The crucial idea is now that smaller pairs are *polarized* in the pseudo-electric field of the vortex and antivortex forming a given pair. This leads to the *screening* of the vortex-antivortex interaction and thus reduces the energy of large pairs. Formally, this can be described by renormalization-group (RG) theory (Kosterlitz 1974), which we will summarize in the following. The grand-canonical partition function of the vortex-antivortex system is

$$\mathcal{Z}_0 = \sum_N \frac{1}{(N!)^2} y_0^{2N} \int_{D_1} \frac{d^2 r_1}{r_0^2} \cdots \int_{D_{2N}} \frac{d^2 r_{2N}}{r_0^2} \exp\left(\frac{1}{2} \sum_{i \neq j} 2\pi N_i N_j K_0 \ln \frac{|\mathbf{r}_i - \mathbf{r}_j|}{r_0}\right), \quad (7.53)$$

where we have already implemented the constraint that the number of vortices, N , equals the number of antivortices. The ranges of integration, D_i , comprise the two-dimensional space \mathbb{R}^2 excluding disks of radius r_0 centered at all vortices (and antivortices) with numbers $j < i$. This means that the minimum separation is r_0 . The idea of RG theory is to perform the integrals for the smallest pairs of sizes between r_0 and $r_0 + dr$ and rewrite the result (approximately) in a form identical to \mathcal{Z}_0 but with changed (“renormalized”) parameters. Physically, we thereby omit the smallest pairs and take their effect into account by renormalizing the parameters. In this way, the partition function, the fugacity, and the stiffness become functions of the smallest length scale, which we now denote by r . These “running” quantities are written as \mathcal{Z} , y , and K , respectively. Thus

$$\mathcal{Z} = \sum_N \frac{1}{(N!)^2} y^{2N} \int_{D_1} \frac{d^2 r_1}{r^2} \cdots \int_{D_{2N}} \frac{d^2 r_{2N}}{r^2} \exp\left(\frac{1}{2} \sum_{i \neq j} 2\pi N_i N_j K \ln \frac{|\mathbf{r}_i - \mathbf{r}_j|}{r}\right). \quad (7.54)$$

We now perform the integrals over the smallest separations between r and $r + dr$. A crucial assumption is that this only involves vortex-antivortex pairs, not pairs of equal vorticity. This is plausible since a vortex and an antivortex attract each other. The integration starts by splitting the integrals,

$$\begin{aligned} \int_{D_1} d^2 r_1 \cdots \int_{D_{2N}} d^2 r_{2N} &\cong \int_{D'_1} d^2 r_1 \cdots \int_{D'_{2N}} d^2 r_{2N} \\ &+ \frac{1}{2} \sum_{i \neq j} \delta_{N_i, -N_j} \underbrace{\int_{D'_1} d^2 r_1 \cdots \int_{D'_{2N}} d^2 r_{2N}}_{\text{excluding } i, j} \int_{D_{i,j}} d^2 r_j \int_{r \leq |\mathbf{r}_i - \mathbf{r}_j| < r + dr} d^2 r_i. \end{aligned} \quad (7.55)$$

The D'_i correspond to the D_i but with the minimum separation increased to $r + dr$. $\bar{D}_{i,j}$ is the full \mathbb{R}^2 excluding disks of radius r around all vortices $n \neq i, j$. Inserting this into the expression for \mathcal{Z} , we obtain the integral, from

the second term,

$$\begin{aligned}
& \int_{\bar{D}_{i,j}} d^2 r_j \int_{r \leq |\mathbf{r}_i - \mathbf{r}_j| < r + dr} d^2 r_i \exp \left(2\pi \sum_{n \neq i,j} N_n N_j K \ln \frac{|\mathbf{r}_n - \mathbf{r}_j|}{r} \right. \\
& \quad \left. - 2\pi \sum_{n \neq i,j} N_n N_j K \ln \frac{|\mathbf{r}_n - \mathbf{r}_i|}{r} - 2\pi K \ln \frac{|\mathbf{r}_j - \mathbf{r}_i|}{r} \right) \\
& \cong \int_{\bar{D}_{i,j}} d^2 r_j 2\pi r dr \left(1 + \pi^2 K^2 \sum_{m,n \neq i,j} N_m N_n \frac{(\mathbf{r}_m - \mathbf{r}_j) \cdot (\mathbf{r}_n - \mathbf{r}_j) r^2}{|\mathbf{r}_m - \mathbf{r}_j|^2 |\mathbf{r}_n - \mathbf{r}_j|^2} \right) \\
& = 2\pi r dr \left(v + \pi^2 K^2 \sum_{m,n \neq i,j} N_m N_n r^2 \int_{\bar{D}_{i,j}} d^2 r_j \frac{(\mathbf{r}_m - \mathbf{r}_j) \cdot (\mathbf{r}_n - \mathbf{r}_j)}{|\mathbf{r}_m - \mathbf{r}_j|^2 |\mathbf{r}_n - \mathbf{r}_j|^2} \right). \tag{7.56}
\end{aligned}$$

Herein, we have

$$\int_{\bar{D}_{i,j}} d^2 r_j \frac{(\mathbf{r}_m - \mathbf{r}_j) \cdot (\mathbf{r}_n - \mathbf{r}_j)}{|\mathbf{r}_m - \mathbf{r}_j|^2 |\mathbf{r}_n - \mathbf{r}_j|^2} = 2\pi \ln \frac{L}{r} - 2\pi (1 - \delta_{mn}) \ln \frac{|\mathbf{r}_m - \mathbf{r}_n|}{r}. \tag{7.57}$$

The first term diverges for $L \rightarrow \infty$ but drops out when the sum over m, n is performed, due to overall vanishing vorticity. The result is

$$\cdots = 2\pi r dr \left(v - 2\pi^3 K^2 r^2 \sum_{\substack{m,n \neq i,j \\ m \neq n}} N_m N_n \ln \frac{|\mathbf{r}_m - \mathbf{r}_n|}{r} \right) \tag{7.58}$$

and with the sum over i, j :

$$\frac{1}{2} \sum_{i \neq j} \cdots = 2\pi r dr \left(N^2 V - 2\pi^3 K^2 r^2 \sum_{m \neq n} (N-1)^2 N_m N_n \ln \frac{|\mathbf{r}_m - \mathbf{r}_n|}{r} \right), \tag{7.59}$$

neglecting some terms of order N^0 . Inserting everything into \mathcal{Z} , we obtain two terms: The first corresponds to \mathcal{Z} with D_i replaced by D'_i and the second reads

$$\begin{aligned}
& \sum_N \frac{1}{(N!)^2} \left(\frac{y}{r^2} \right)^{2N} \underbrace{\int_{D'_1} d^2 r_1 \cdots \int_{D'_{2N}} d^2 r_{2N} 2\pi r dr}_{2N-2 \text{ integrals}} \\
& \times \left(N^2 V - 2\pi^3 K^2 r^2 \sum_{m \neq n} (N-1)^2 N_m N_n \ln \frac{|\mathbf{r}_m - \mathbf{r}_n|}{r} \right) \exp \left(\underbrace{\frac{1}{2} 2\pi K \sum_{i \neq j} N_i N_j \ln \frac{|\mathbf{r}_m - \mathbf{r}_n|}{r}}_{(2N-2)(2N-3) \text{ terms}} \right). \tag{7.60}
\end{aligned}$$

We rename the summation index N as $N+1$ in this second term. Then both terms contain $2N$ integrals under the sum over N . We also put all terms containing dr into the exponent using $1 + a dr = e^{a dr}$. We obtain

$$\begin{aligned}
\mathcal{Z}' & = \exp \left(2\pi \left(\frac{y}{r^2} \right)^2 r dr V \right) \sum_N \frac{1}{(N!)^2} \left(\frac{y}{r^2} \right)^{2N} \int_{D'_1} d^2 r_1 \cdots \int_{D'_{2N}} d^2 r_{2N} \\
& \times \exp \left[-\frac{1}{2} \sum_{i \neq j} \left(-2\pi K + 8\pi^4 y^2 K^2 \frac{dr}{r} \right) N_i N_j \ln \frac{|\mathbf{r}_i - \mathbf{r}_j|}{r} \right]. \tag{7.61}
\end{aligned}$$

Next, we have to express \mathcal{Z}' in terms of the new length scale $r' := r + dr$. This is only relevant in expressions not already linear in dr . This applies to, on the one hand,

$$\frac{1}{r^2} = \frac{1}{(r' - dr)^2} = \frac{1 + 2 \frac{dr}{r'}}{(r')^2}, \quad (7.62)$$

and, on the other,

$$\begin{aligned} \exp \left(-\frac{1}{2} \sum_{i \neq j} 2\pi K N_i N_j \ln r \right) &= \exp \left(-\frac{1}{2} \sum_{i \neq j} 2\pi K N_i N_j \ln(r' - dr) \right) \\ &= \exp \left(-\frac{1}{2} \sum_{i \neq j} 2\pi K N_i N_j \ln r' \right) \exp \left(\frac{1}{2} \sum_{i \neq j} 2\pi K N_i N_j \frac{dr}{r'} \right) \\ &= \exp \left(-\frac{1}{2} \sum_{i \neq j} 2\pi K N_i N_j \ln r' \right) \left(1 - \pi K \frac{dr}{r'} \right)^{2N}, \end{aligned} \quad (7.63)$$

neglecting terms of order N^0 compared to N . The renormalized partition function is finally

$$\begin{aligned} \mathcal{Z}' &= \underbrace{\exp \left[2\pi \left(\frac{y}{(r')^2} \right)^2 r' dr V \right]}_{=: \mathcal{Z}_{\text{pair}}} \sum_N \frac{1}{(N!)^2} \left(\frac{y}{(r')^2} \right)^{2N} \left[1 + (2 - \pi K) \frac{dr}{r'} \right]^{2N} \\ &\quad \times \int_{D'_1} d^2 r_1 \cdots \int_{D'_{2N}} d^2 r_{2N} \exp \left[-\frac{1}{2} \sum_{i \neq j} \left(-2\pi K + 8\pi^4 y^2 K^2 \frac{dr}{r'} \right) N_i N_j \ln \frac{|\mathbf{r}_i - \mathbf{r}_j|}{r'} \right]. \end{aligned} \quad (7.64)$$

Here, $\mathcal{Z}_{\text{pair}}$ is the partition function of the small pair we have integrated out. It is irrelevant for the renormalization of y and K . Apart from this factor, \mathcal{Z}' is identical to \mathcal{Z} if we set

$$y(r') = \left[1 + (2 - \pi K(r)) \frac{dr}{r'} \right] y(r) \quad (7.65)$$

$$K(r') = K(r) - 4\pi^3 y^2 K^2(r) \frac{dr}{r'}. \quad (7.66)$$

Introducing the logarithmic length scale

$$l := \ln \frac{r}{r_0} \quad \Rightarrow \quad dl = \frac{dr}{r}, \quad (7.67)$$

we obtain the Kosterlitz RG *flow equations*

$$\frac{dy}{dl} = (2 - \pi K) y, \quad (7.68)$$

$$\frac{dK}{dl} = -4\pi^3 y^2 K^2. \quad (7.69)$$

The initial conditions are

$$y(l=0) = y_0 = e^{-E_{\text{core}}/k_B T}, \quad (7.70)$$

$$K(l=0) = K_0 = -\frac{1}{k_B T} 2\gamma \frac{\alpha}{\beta}, \quad (7.71)$$

i.e., the parameters assume their “bare” values at $r = r_0$ ($l = 0$).

We will now discuss the physics encoded by the RG flow equations. First, note that the quantity

$$C := 2\pi^2 y^2 - \frac{2}{\pi K} - \ln K \quad (7.72)$$

is invariant under the RG flow:

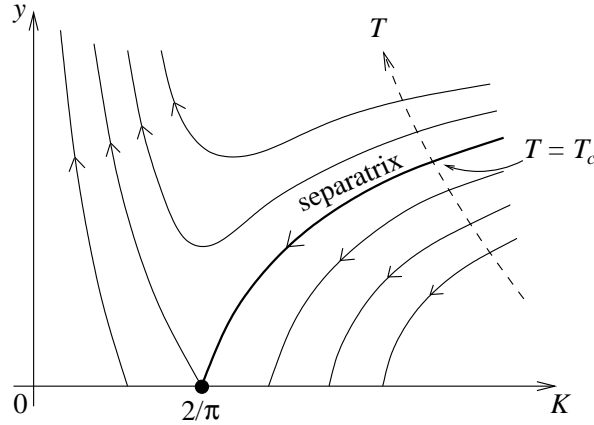
$$\frac{dC}{dl} = 4\pi^2 y \frac{dy}{dl} - \frac{2}{\pi K^2} \frac{dK}{dl} - \frac{1}{K} \frac{dK}{dl} = 4\pi^2 y^2 (2 - \pi K) - 8\pi^2 y^2 + 4\pi^3 y^2 K = 0. \quad (7.73)$$

Thus C is a first integral of the flow equations. We can calculate C from the initial values y_0, K_0 and obtain

$$2\pi^2 y^2 = 2\pi^2 y_0^2 + \frac{2}{\pi} \left(\frac{1}{K} - \frac{1}{K_0} \right) + \ln \frac{K}{K_0} \quad (7.74)$$

$$\Rightarrow y = \sqrt{y_0^2 + \frac{1}{\pi^3} \left(\frac{1}{K} - \frac{1}{K_0} \right) + \frac{1}{2\pi^2} \ln \frac{K}{K_0}}. \quad (7.75)$$

The RG flow is along curves described by this expression, where the curves are specified by y_0, K_0 . These parameters change with temperature as given in Eqs. (7.70) and (7.71). These initial conditions are sketched as a dashed line in the figure. We see that there are two distinct cases:



- For $T < T_c$, K flows to some finite value $K(l \rightarrow \infty) > 2/\pi$. This means that even infinitely large pairs feel a logarithmic attraction, i.e., are bound. Moreover, the fugacity y flows to zero, $y(l \rightarrow \infty) = 0$. Thus large pairs are very rare, which is consistent with their (logarithmically) diverging energy.
- For $T > T_c$, K flows to $K(l \rightarrow \infty) = 0$. Thus the interaction between a vortex and an antivortex that are far apart is completely screened. Large pairs become unbound. Also, y diverges on large length scales, which means that these unbound vortices proliferate. This divergence is an artifact of keeping only the leading order in y in the derivation. It is cut off at finite y if we count vortex-antivortex pairs consistently. But the limit $K \rightarrow 0$ remains valid.

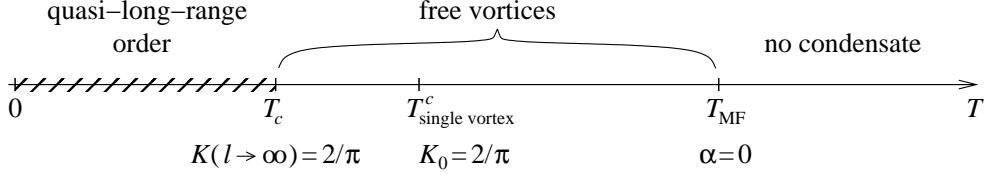
At $T = T_c$ we thus find a phase transition at which vortex-antivortex pairs unbind, forming free vortices. It is called the *Berezinskii-Kosterlitz-Thouless (BKT) transition*. In two-dimensional films, vortex interactions thus suppress the temperature where free vortices appear and quasi-long-range order is lost from the point $T = T_{\text{single vortex}}^c$ where

$$\eta = -\frac{1}{4\pi} \frac{k_B T \beta}{\gamma \alpha} = \frac{1}{2\pi} \frac{1}{K_0} \stackrel{!}{=} \frac{1}{4} \Rightarrow K_0 = \frac{2}{\pi} \quad (7.76)$$

to the one where $K(l \rightarrow \infty) = 2/\pi$ and y_0, K_0 lie on the “separatrix” between the two phases,

$$C = 2\pi^2 y_0^2 - \frac{2}{\pi K_0} - \ln K_0 \stackrel{!}{=} 0 - 1 - \ln \frac{2}{\pi} \Rightarrow \frac{2}{\pi K_0} + \ln K_0 = 1 + \ln \frac{2}{\pi} + 2\pi^2 y_0^2. \quad (7.77)$$

Clearly the two criteria agree if $y_0 = 0$. This makes sense since for $y_0 = 0$ there are no vortex pairs to screen the interaction. In addition, a third temperature scale is given by the mean-field transition temperature T_{MF} , where $\alpha \stackrel{!}{=} 0$.



In the low-temperature phase, the largest pairs determine the decay of the correlation function $\langle \psi^*(\mathbf{r})\psi(0) \rangle$ for large r . Thus we find

$$\langle \psi^*(\mathbf{r})\psi(0) \rangle \cong \psi_0^2 \left(\frac{r}{r_0} \right)^{-\eta} \quad (7.78)$$

with

$$\eta = \frac{1}{2\pi} \frac{1}{K(l \rightarrow \infty)}. \quad (7.79)$$

We note that the exponent η changes with temperature (one could say that the whole low-temperature phase is critical) but assumes a universal value at T_c : There,

$$\lim_{l \rightarrow \infty} K = \frac{2}{\pi} \quad \Rightarrow \quad \eta(T_c) = \frac{1}{4}. \quad (7.80)$$

Due to the screening of the vortex interaction, the condensate is less stiff than it would be in the absence of vortices. This is described by the renormalization of K from K_0 to $K(l \rightarrow \infty)$. It is customary but somewhat misleading to express this as a renormalization of the superfluid density n_s , which is after all proportional to K_0 . Assuming that the temperature dependence of α and thus of the bare superfluid density

$$n_s^0 \propto \psi_0^2 = -\frac{\alpha}{\beta} \propto K_0 \quad (7.81)$$

is negligible close to T_c , the renormalized superfluid density

$$n_s(T) := \frac{K}{K_0} n_s^0 \quad (7.82)$$

obtains its temperature dependence exclusively from K/K_0 . (The argument $l \rightarrow \infty$ is implied here and in the following.) For $T < T_c$ but close to the BKT transition we have

$$K = \frac{2}{\pi} + \Delta K. \quad (7.83)$$

The invariant $C = 2\pi^2 y_0^2 - \frac{2}{\pi K_0} - \ln K_0$ is an analytic function of temperature and its value at T_c is

$$C_c = \left[2\pi^2 y^2 - \frac{2}{\pi K} - \ln K \right]_{T=T_c} = -1 - \ln \frac{2}{\pi}. \quad (7.84)$$

Thus we can write, close to T_c ,

$$C \cong -1 - \ln \frac{2}{\pi} + b(T - T_c) \quad (7.85)$$

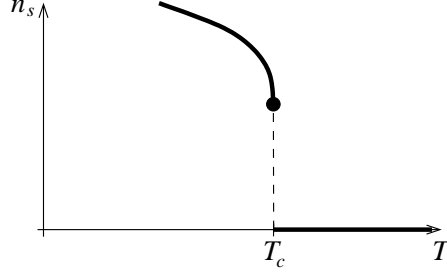
with some constant $b := dC/dT|_{T=T_c}$. On the other hand,

$$\begin{aligned} C &= \left[2\pi^2 y^2 - \frac{2}{\pi K} - \ln K \right]_T \\ &= -\frac{2}{\pi \left(\frac{2}{\pi} + \Delta K \right)} - \ln \left(\frac{2}{\pi} + \Delta K \right) \\ &= -\frac{1}{1 + \frac{\pi}{2} \Delta K} - \ln \frac{2}{\pi} - \ln \left(1 + \frac{\pi}{2} \Delta K \right) \\ &\cong -1 + \frac{\pi}{2} \cancel{\Delta K} - \frac{\pi^2}{4} \Delta K^2 - \ln \frac{2}{\pi} - \frac{\pi}{2} \cancel{\Delta K} + \frac{1}{2} \frac{\pi^2}{4} \Delta K^2 \\ &= -1 - \ln \frac{2}{\pi} - \frac{\pi^2}{8} \Delta K^2. \end{aligned} \quad (7.86)$$

Thus

$$-\frac{\pi^2}{8} \Delta K^2 \cong b(T - T_c) \quad \Rightarrow \quad \Delta K \cong \underbrace{\frac{2}{\pi} \sqrt{2b}}_{\text{const}} \sqrt{T_c - T}. \quad (7.87)$$

Consequently, $n_s = (K/K_0) n_s^0$ jumps to a finite value at T_c and then increases like a square root.



This behavior was indeed measured in torsion-pendulum experiments on He-4 films (Bishop and Reppy, 1980).

7.2 Superconducting films

In this section we concentrate on what is different for charged superconductors compared to neutral superfluids. Recall that for a superfluid the gradient term in the Landau functional reads

$$\int d^2 r \gamma (\nabla \psi)^* \cdot \nabla \psi. \quad (7.88)$$

If we move around a vortex once, the phase has to change by 2π , regardless of the distance from the vortex core. This phase change leads to an unavoidable contribution to the gradient term of

$$\begin{aligned} \int d^2 r \gamma (\nabla \psi)^* \cdot \nabla \psi &\cong -\gamma \frac{\alpha}{\beta} \int dr d\varphi r \left(\frac{1}{r} \frac{\partial}{\partial \varphi} e^{i\phi} \right)^* \frac{1}{r} \frac{\partial}{\partial \varphi} e^{i\phi} \\ &= -\gamma \frac{\alpha}{\beta} \int dr d\varphi \frac{1}{r} \left(\frac{\partial \phi}{\partial \varphi} \right)^2 = -\gamma \frac{\alpha}{\beta} \int dr d\varphi \frac{1}{r} = -2\pi \gamma \frac{\alpha}{\beta} \int \frac{dr}{r}, \end{aligned} \quad (7.89)$$

which diverges logarithmically with system size. On the other hand, for a superconducting film the gradient term reads

$$\int d^2 r \frac{1}{2m^*} \left| \left(\frac{\hbar}{i} \nabla - \frac{q}{c} \mathbf{A} \right) \psi \right|^2. \quad (7.90)$$

Again, the phase of ψ winds by 2π around a vortex, but the associated gradient can, in principle, be compensated by the vector potential. Pearl (1964) showed that this indeed leads to a *finite* energy of a single vortex in a superconducting film. In addition, the free energy contains the magnetic-field term

$$\int d^3 r \frac{B^2(\mathbf{r})}{8\pi}. \quad (7.91)$$

Note that the integral is *three-dimensional*—the field is present in all space, also outside of the film.

We first note that the film has a new effective length scale: In the London gauge, Ampère's law reads

$$\nabla \times \nabla \times \mathbf{A} = \nabla(\underbrace{\nabla \cdot \mathbf{A}}_{=0}) - \nabla^2 \mathbf{A} = \frac{4\pi}{c} \mathbf{j} \quad \Rightarrow \quad \nabla^2 \mathbf{A} = -\frac{4\pi}{c} \mathbf{j}. \quad (7.92)$$

On the other hand, from the London equation we have

$$\mathbf{j} = -\frac{c}{4\pi\lambda^2} \mathbf{A}, \quad (7.93)$$

which is valid within the film and away from vortex cores so that $|\psi|^2 \cong -\alpha/\beta$. However, the current is confined to the thin film. We can write

$$\mathbf{j}(\mathbf{r}) = \mathbf{K}(x, y) \delta(z) \quad (7.94)$$

with a surface current density \mathbf{K} . Thus

$$\mathbf{K}(x, y) = \int_{-\infty}^{\infty} dz \mathbf{j}(\mathbf{r}) = \int_{\text{film}} dz \left(-\frac{c}{4\pi\lambda^2} \right) \mathbf{A}(\mathbf{r}). \quad (7.95)$$

If the thickness is d and \mathbf{A} is approximately constant across the thickness, we have

$$\mathbf{K}(x, y) = -\frac{c}{4\pi} \frac{d}{\lambda^2} \mathbf{A}(x, y, 0) \quad (7.96)$$

$$\Rightarrow \mathbf{j}(\mathbf{r}) = -\frac{c}{4\pi} \frac{d}{\lambda^2} \mathbf{A}(\mathbf{r}) \delta(z) \quad (7.97)$$

and finally

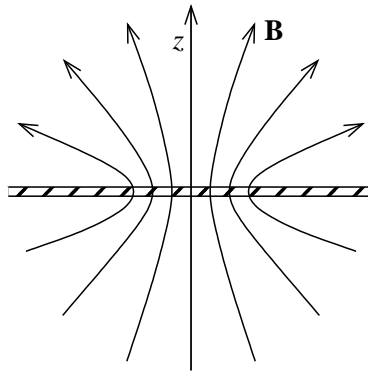
$$\nabla^2 \mathbf{A} = \frac{d}{\lambda^2} \mathbf{A} \delta(z). \quad (7.98)$$

This result exhibits the new length scale that controls the spatial variation of \mathbf{A} and thus of the current \mathbf{j} ,

$$\lambda_{\perp} := \frac{\lambda^2}{d}, \quad (7.99)$$

which is large, $\lambda_{\perp} \gg \lambda$, for a thin film. We see that in thin films, λ_{\perp} assumes the role of the penetration depth λ . Since λ_{\perp} is large for thin films, one could say that thin films are always effectively of type II.

We here do not discuss the full derivation of Pearl but only consider the “far field” for large $r \gg \lambda_{\perp}$ and show that it does not lead to a diverging free energy for a single vortex. We assume that most of the magnetic flux of Φ_0 penetrates the film for $\varrho = \sqrt{x^2 + y^2} \lesssim \lambda_{\perp}$. Then the magnetic field above (and below) the film looks like a monopole field for $r \gg \lambda_{\perp}$.



This field is easy to obtain from symmetry:

$$\int_{\substack{\text{half space} \\ \text{for } z > 0}} d\mathbf{a} \cdot \mathbf{B} = \Phi_0 \quad \Rightarrow \quad 2\pi r^2 B = \Phi_0 \quad (7.100)$$

$$\Rightarrow B(r) = \frac{\Phi_0}{2\pi r^2} \quad \text{and} \quad \mathbf{B}(\mathbf{r}) = \frac{\Phi_0}{2\pi r^2} \hat{\mathbf{r}} \quad \text{for } z > 0. \quad (7.101)$$

By symmetry,

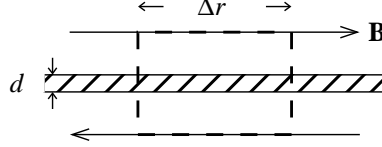
$$\mathbf{B}(\mathbf{r}) = \text{sgn } z \frac{\Phi_0}{2\pi r^2} \hat{\mathbf{r}}. \quad (7.102)$$

This gives a contribution to the field energy of

$$\int d^3r \frac{B^2}{8\pi} = \frac{\Phi_0^2}{32\pi^3} \int d^3r \frac{1}{r^4} = \frac{\Phi_0^2}{32\pi^3} 4\pi \int_{\lambda_\perp}^{\infty} \frac{dr}{r^2} = \frac{\Phi_0^2}{8\pi^2} \frac{1}{\lambda_\perp}, \quad (7.103)$$

which is finite for an infinity film. The value of the lower cutoff does not matter for this. The cutoff has to be present, since the monopole field is not a valid approximation for small r .

We next obtain the current from Ampère's law in integral form:



$$\oint d\mathbf{r} \cdot \mathbf{B} = \frac{4\pi}{c} \int d\mathbf{a} \cdot \mathbf{j} \quad (7.104)$$

$$\Rightarrow 2 \int_r^{r+\Delta r} dr' \frac{\Phi_0}{2\pi(r')^2} = \frac{4\pi}{c} \Delta r dj(r) \quad (7.105)$$

$$\Rightarrow \frac{4\pi}{c} \Delta r dj(r) = \frac{\Phi_0}{\pi} \left(\frac{1}{r} - \frac{1}{r+\Delta r} \right) \cong \frac{\Phi_0}{\pi} \frac{\Delta r}{r^2} \quad (7.106)$$

$$\Rightarrow j(r) = \frac{\Phi_0}{4\pi^2} \frac{1}{r^2 d} \quad (7.107)$$

and with the vector character restored

$$\mathbf{j}(\mathbf{r}) = \frac{\Phi_0}{4\pi^2} \frac{\hat{\varphi}}{r^2 d}. \quad (7.108)$$

Note that the sheet current is thus

$$\mathbf{K}(\mathbf{r}) = \frac{\Phi_0}{4\pi^2} \frac{\hat{\varphi}}{r^2}. \quad (7.109)$$

For large r we have $|\psi| = \psi_0 = \sqrt{-\alpha/\beta}$ (note that typically $\lambda_\perp \gg \xi$). Then the second Ginzburg-Landau equation gives

$$\mathbf{j} = i \frac{q\hbar}{2m^*} (-i\nabla\phi - i\nabla\phi)\psi_0^2 - \frac{q^2}{m^*c} \psi_0^2 \mathbf{A} = \psi_0^2 \frac{q}{m^*} \left(\hbar\nabla\phi - \frac{q}{c} \mathbf{A} \right). \quad (7.110)$$

Thus we can rewrite the gradient term in the free energy as

$$d \int d^2r \frac{1}{2m^*} \psi_0^2 \left(\hbar\nabla\phi - \frac{q}{c} \mathbf{A} \right)^2 = d \int d^2r \frac{\psi_0^2 (m^*)^2}{2m^* \psi_0^4 q^2} \mathbf{j} \cdot \mathbf{j} = d \int d^2r \frac{m^*}{2q^2 \psi_0^2} \mathbf{j} \cdot \mathbf{j}. \quad (7.111)$$

The contribution from large r is

$$\frac{dm^*}{2q^2 \psi_0^2} \int d^2r \frac{\Phi_0^2}{16\pi^4} \frac{1}{r^4 d^2} = \frac{m^* \Phi_0^2}{32\pi^4 q^2 \psi_0^2 d} 2\pi \int_{\lambda_\perp}^{\infty} \frac{dr}{r^3} = \frac{m^* \Phi_0^2}{8\pi^3 q^2 \psi_0^2 d} \frac{1}{\lambda_\perp^2}, \quad (7.112)$$

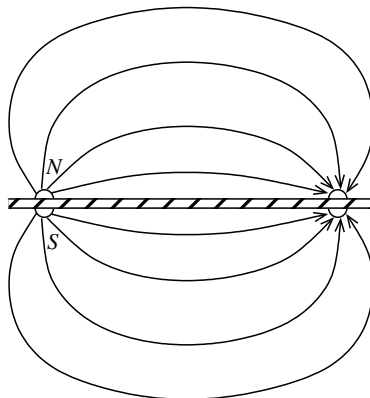
which is also finite for an infinite film.

We conclude that the free energy of an isolated vortex is finite in a superconducting film. It is then plausible and indeed true that the interaction energy of a vortex-antivortex pair does not diverge for large separations r but saturates for $r \gg \lambda_\perp$. Since for the far field of a single vortex the magnetic-field energy, Eq. (7.103), dominates over the energy due to the gradient term, we expect the large- r interaction to be dominated by the Coulomb-type

attraction of the magnetic monopoles in the upper and lower half spaces. This supposition is borne out by a proper analysis. Consequently, for large r the interaction behaves like

$$V_{\text{int}} = \text{const} - \left(\frac{\Phi_0}{2\pi}\right)^2 \frac{1}{r} \quad (7.113)$$

($\Phi_0/2\pi$ is the monopole strength, according to Eq. (7.102)).



All this shows that, strictly speaking, there will be a non-zero concentration of free vortices at any temperature $T > 0$. Thus there is no quasi-long-range order. However, the relevant length scale is $\lambda_{\perp} = \lambda^2/d$, which can be very large for thin films, even compared to the lateral size L of the sample. In this case the large- r limit is experimentally irrelevant. But for vortex separations $r \ll \lambda_{\perp}$, the magnetic-field expulsion on the scale r is very weak since λ_{\perp} is the effective penetration depth. Then the fact that the condensate is charged is irrelevant and we obtain the same logarithmic interaction as for a neutral superfluid.

Thus for thin films of typical size we can use the previously discussed BKT theory. For superconducting films we even have the advantage of an additional observable, namely the *voltage* for given current. We give a hand-waving derivation of $V(I)$. The idea is that a current exerts a *Magnus force* on a vortex, in the direction perpendicular to the current. The force is opposite for vortices and antivortices and is thus able to break vortex-antivortex pairs. As noted above, free vortices lead to dissipation. A vortex moving through the sample in the orthogonal direction between source and drain contacts leads to a change of the phase difference $\Delta\phi$ by $\pm 2\pi$. We will see in the chapter on Josephson effects why this corresponds to a non-zero voltage. Since free vortices act independently, it is plausible to assume that the resistance is

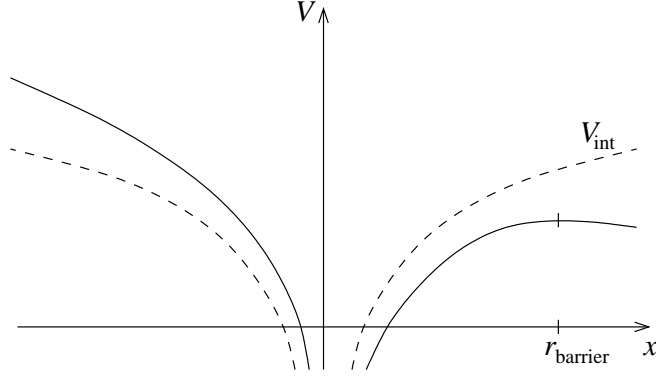
$$R \propto n_v, \quad (7.114)$$

where n_v now denotes the concentration of free vortices. To find it, note that the total potential energy due to vortex-antivortex attraction and Magnus force can be written as

$$V = V_{\text{int}} - 2F_{\text{Magnus}} r \quad (7.115)$$

with

$$V_{\text{int}} = 2\pi k_B T K \ln \frac{r}{r_0}. \quad (7.116)$$



There is a finite barrier for vortex-antivortex unbinding at a separation r_{barrier} determined from $\partial V/\partial r = 0$. This gives

$$r_{\text{barrier}} = \frac{2\pi k_B T K_{\text{barrier}}}{2F_{\text{Magnus}}}, \quad (7.117)$$

where $K_{\text{barrier}} := K(r_{\text{barrier}})$. The barrier height is

$$\begin{aligned} \Delta E &:= V(r_{\text{barrier}}) - V(r_0) \cong V(r_{\text{barrier}}) \\ &= 2\pi k_B T K_{\text{barrier}} \ln \frac{r_{\text{barrier}}}{r_0} - 2F_{\text{Magnus}} r_{\text{barrier}} \\ &= 2\pi k_B T K_{\text{barrier}} \left(\ln \frac{r_{\text{barrier}}}{r_0} - 1 \right). \end{aligned} \quad (7.118)$$

For small currents we have $r_{\text{barrier}} \gg r_0$ and thus

$$\Delta E \cong 2\pi k_B T \underbrace{K(l \rightarrow \infty)}_{\equiv K} \ln \frac{r_{\text{barrier}}}{r_0}. \quad (7.119)$$

The rate at which free vortices are generated is

$$R_{\text{gen}} \propto e^{-\beta \Delta E} \cong \left(\frac{r_{\text{barrier}}}{r_0} \right)^{-2\pi K}. \quad (7.120)$$

The recombination rate of two vortices to form a pair is

$$R_{\text{rec}} \propto n_v^2, \quad (7.121)$$

since two vortices must meet. In the stationary state we have

$$R_{\text{gen}} = R_{\text{rec}} \quad \Rightarrow \quad n_v \propto \sqrt{R_{\text{gen}}} \propto \left(\frac{r_{\text{barrier}}}{r_0} \right)^{-\pi K} \quad (7.122)$$

and thus a resistance of

$$R \propto n_v \propto \left(\frac{r_{\text{barrier}}}{r_0} \right)^{-\pi K}. \quad (7.123)$$

Since the Magnus force is $F_{\text{Magnus}} \propto I$ we have

$$r_{\text{barrier}} \propto \frac{1}{F_{\text{Magnus}}} \propto \frac{1}{I} \quad (7.124)$$

so that

$$R \propto \left(\frac{1}{I} \right)^{-\pi K} = I^{\pi K}. \quad (7.125)$$

Finally, the voltage measured for a current I is

$$V = RI \propto I^{\pi K} I = I^{1+\pi K}, \quad (7.126)$$

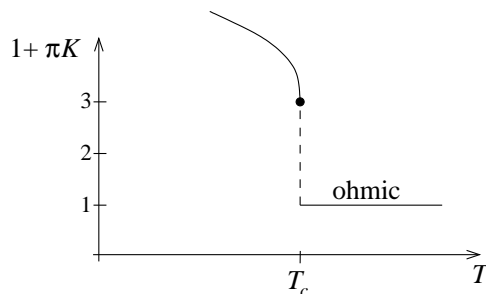
where $K \equiv K(l \rightarrow \infty)$ is the renormalized stiffness. Since

$$K = \frac{2}{\pi} + \Delta K \cong \frac{2}{\pi} \left(1 + \sqrt{2b} \sqrt{T_c - T}\right) \quad (7.127)$$

we find for the exponent

$$1 + \pi K \cong 3 + \underbrace{2\sqrt{2b}}_{=\text{const}} \sqrt{T_c - T} \quad (7.128)$$

for $T \lesssim T_c$.



Above T_c we have $K = 0$ and thus ohmic resistance, $V \propto I$, as expected. Below T_c , the voltage is *sub-ohmic*, i.e., the voltage is finite for finite current but rises more slowly than linearly for small currents. This behavior has been observed for thin superconducting films.

8

Origin of attractive interaction

In the following chapters we turn to the microscopic theory of superconductivity. While the BCS theory is reasonably easy to understand if one assumes an attractive interaction between electrons in a superconductor, it is far from obvious where such an interaction should come from. The *only* fundamental interaction that is relevant for (non-radioactive) solids is the electromagnetic one, which naively gives a *repulsive* interaction of a typical strength of several eV between nearby electrons. How can this lead to an attraction at the low energy scale $k_B T_c \sim 1$ meV? We will see that the lattice of ion cores (nuclei with tightly bound inner electrons) plays an important role. We will use Feynman diagrams for Green functions to describe the physics. Unfortunately, we do not have the time to introduce these concepts rigorously; this is done in many good textbooks on many-particle physics as well as in the lecture notes on Vielteilchentheorie (in German), which are available online. For those familiar with Feynman diagrams, they rigorously represent mathematical expressions, for the others they should at least be useful as cartoons of the relevant processes.

8.1 Reminder on Green functions

Nevertheless, we start by briefly summarizing some properties of Green functions. In many-particle physics, we usefully express the Hamiltonian in terms of an electronic *field operator* $\Psi_\sigma(\mathbf{r}, t)$, where $\sigma = \uparrow, \downarrow$ is the electron spin. The field operator can be expanded in any convenient basis of single-particle states characterized by wavefunctions $\varphi_{\nu\sigma}(\mathbf{r})$, where ν represents all relevant quantum numbers,

$$\Psi_\sigma(\mathbf{r}, t) = \sum_{\nu} \varphi_{\nu\sigma}(\mathbf{r}) c_{\nu\sigma}(t), \quad (8.1)$$

$$\Psi_\sigma^\dagger(\mathbf{r}, t) = \sum_{\nu} \varphi_{\nu\sigma}^*(\mathbf{r}) c_{\nu\sigma}^\dagger(t). \quad (8.2)$$

$c_{\nu\sigma}$ and $c_{\nu\sigma}^\dagger$ are annihilation and creation operators of electrons in the single-particle state, respectively. Note that we are working in the Heisenberg picture, i.e., the wavefunctions $\varphi_{\nu\sigma}(\mathbf{r})$ are independent of time, whereas the time dependence is carried by the operators.

For example, the Hamiltonian for free electrons reads

$$H = \int d^3r \sum_{\sigma} \Psi_\sigma^\dagger(\mathbf{r}) \frac{-\hbar^2 \nabla^2}{2m} \Psi_\sigma(\mathbf{r}). \quad (8.3)$$

In this case it is useful to expand into plane waves,

$$\Psi_\sigma(\mathbf{r}) = \frac{1}{\sqrt{V}} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}} c_{\mathbf{k}\sigma} \quad (8.4)$$

$$\begin{aligned} \Rightarrow H &= \frac{1}{V} \sum_{\mathbf{k}\mathbf{k}'} \int d^3r \sum_{\sigma} e^{-i\mathbf{k}\cdot\mathbf{r}+i\mathbf{k}'\cdot\mathbf{r}} c_{\mathbf{k}\sigma}^\dagger \frac{\hbar^2(k')^2}{2m} c_{\mathbf{k}'\sigma} \\ &= \sum_{\mathbf{k}\sigma} \frac{\hbar^2 k^2}{2m} c_{\mathbf{k}\sigma}^\dagger c_{\mathbf{k}\sigma}. \end{aligned} \quad (8.5)$$

Two types of Green functions are defined as follows:

- *greater Green function*:

$$G^>(\mathbf{r}\sigma t, \mathbf{r}'\sigma' t') := -i \left\langle \Psi_\sigma(\mathbf{r}, t) \Psi_{\sigma'}^\dagger(\mathbf{r}', t') \right\rangle, \quad (8.6)$$

where $\langle \dots \rangle$ is the equilibrium average

$$\langle A \rangle = \text{Tr} A \rho_{\text{eq}} = \frac{\text{Tr} A e^{-\beta H}}{\text{Tr} e^{-\beta H}} \equiv \frac{1}{\mathcal{Z}} \text{Tr} A e^{-\beta H}, \quad (8.7)$$

where $\beta = 1/k_B T$ is the inverse temperature and H is the many-particle Hamiltonian. For non-interacting electrons at temperature $T \ll E_F/k_B$, ρ_{eq} describes the Fermi sea. $G^>$ describes the *conditional probability amplitude* for an *electron* created at point \mathbf{r}' with spin σ' at time t' to be found at point \mathbf{r} with spin σ at time t .

- *lesser Green function*:

$$G^<(\mathbf{r}\sigma t, \mathbf{r}'\sigma' t') := +i \left\langle \Psi_{\sigma'}^\dagger(\mathbf{r}', t') \Psi_\sigma(\mathbf{r}, t) \right\rangle; \quad (8.8)$$

it describes the propagation of a *hole* from time t to time t' .

These are not the most useful definitions since in quantum theory propagation of an electron forward in time cannot be separated from propagation of a hole backward in time. It is also useful to distinguish between the cases $t > t'$ and $t < t'$. This is accomplished by these definitions:

- *retarded Green function*:

$$G^R(\mathbf{r}\sigma t, \mathbf{r}'\sigma' t') := -i \Theta(t - t') \left\langle \left\{ \Psi_\sigma(\mathbf{r}, t), \Psi_{\sigma'}^\dagger(\mathbf{r}', t') \right\} \right\rangle, \quad (8.9)$$

where $\{A, B\} := AB + BA$ is the anti-commutator (appropriate for fermions) and the step function

$$\Theta(t - t') = \begin{cases} 1 & \text{for } t > t' \\ 0 & \text{for } t < t' \end{cases} \quad (8.10)$$

selects only contributions with $t > t'$ (“forward in time”).

- *advanced Green function*:

$$G^A(\mathbf{r}\sigma t, \mathbf{r}'\sigma' t') := +i \Theta(t' - t) \left\langle \left\{ \Psi_\sigma(\mathbf{r}, t), \Psi_{\sigma'}^\dagger(\mathbf{r}', t') \right\} \right\rangle, \quad (8.11)$$

this Green function analogously contains only contributions “backward in time.”

The thermal averages in the Green functions introduce operators $e^{-\beta H}$, whereas the time evolution of operators introduces time-evolution operators according to

$$A(t) = e^{iHt/\hbar} A e^{-iHt/\hbar}. \quad (8.12)$$

Specifically, we have

$$G^R(\mathbf{r}\sigma t, \mathbf{r}'\sigma' t') = -i\Theta(t-t') \frac{1}{Z} \text{Tr} e^{-\beta H} \left(e^{iHt/\hbar} \Psi_\sigma(\mathbf{r}) e^{-iH(t-t')/\hbar} \Psi_{\sigma'}^\dagger(\mathbf{r}') e^{-iHt'/\hbar} + e^{iHt'/\hbar} \Psi_{\sigma'}^\dagger(\mathbf{r}') e^{-iH(t'-t)/\hbar} \Psi_\sigma(\mathbf{r}) e^{-iHt/\hbar} \right). \quad (8.13)$$

In practice, we will want to write the Hamiltonian in the form $H = H_0 + V$ and treat V in perturbation theory. It is not surprising that this is complicated due to the presence of H in several exponential factors. However, these factors are of a similar form, only in some the prefactor of H is imaginary and in one it is the real inverse temperature. Can one simplify calculations by making all prefactors real? This is indeed possible by formally replacing $t \rightarrow -i\tau$, $t' \rightarrow -i\tau'$, which is the main idea behind the imaginary-time formalism. We cannot discuss it here but only state a few relevant results.

It turns out to be useful to consider the *Matsubara* (or thermal) *Green function*

$$\mathcal{G}(\mathbf{r}\sigma t, \mathbf{r}'\sigma' t') := -\left\langle T_\tau \Psi_\sigma(\mathbf{r}, t) \Psi_{\sigma'}^\dagger(\mathbf{r}', t') \right\rangle, \quad (8.14)$$

where for any operator

$$A(\tau) := e^{H\tau/\hbar} A e^{-H\tau/\hbar}. \quad (8.15)$$

and T_τ is a *time-ordering directive*:

$$T_\tau A(\tau) B(\tau') = \begin{cases} A(\tau) B(\tau') & \text{for } \tau > \tau' \\ \pm B(\tau') A(\tau) & \text{for } \tau < \tau' \end{cases} \quad (8.16)$$

(upper/lower sign for bosonic/fermionic operators). For time-independent Hamiltonians, the Green function only depends on the difference $\tau - \tau'$. One can then show that the resulting Green function $\mathcal{G}(\mathbf{r}\sigma, \mathbf{r}'\sigma', \tau)$ is defined only for $\tau \in [-\hbar\beta, \hbar\beta]$ and satisfies

$$\mathcal{G}(\mathbf{r}\sigma, \mathbf{r}'\sigma', \tau + \beta) = -\mathcal{G}(\mathbf{r}\sigma, \mathbf{r}'\sigma', \tau) \quad (8.17)$$

for fermions. This implies that the Fourier transform is a discrete sum over the fermionic *Matsubara frequencies*

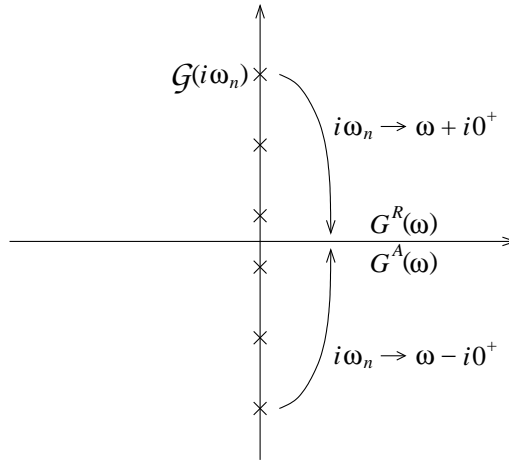
$$\omega_n := \frac{(2n+1)\pi}{\hbar\beta}, \quad n \in \mathbb{Z}. \quad (8.18)$$

The imaginary-time formalism is useful mainly because \mathcal{G} is easier to obtain or approximate than the other Green functions and these can be calculated from \mathcal{G} based on the following theorem: The retarded Green function $G^R(\omega)$ in Fourier space is obtained from $\mathcal{G}(i\omega_n)$ by means of replacing $i\omega_n$ by $\omega + i0^+$, where $i0^+$ is an infinitesimal positive imaginary part,

$$G^R(\omega) = \mathcal{G}(i\omega_n \rightarrow \omega + i0^+). \quad (8.19)$$

This is called “analytic continuation.” Analogously,

$$G^A(\omega) = \mathcal{G}(i\omega_n \rightarrow \omega - i0^+). \quad (8.20)$$



It will be useful to write the electronic Green function in \mathbf{k} space. In the cases we are interested in, momentum and also spin are conserved so that the Green function can be written as

$$\mathcal{G}_{\mathbf{k}\sigma}(\tau) = -\left\langle T_\tau c_{\mathbf{k}\sigma}(\tau) c_{\mathbf{k}\sigma}^\dagger(0) \right\rangle. \quad (8.21)$$

Analogously, the bosonic Matsubara Green function of phonons can be written as

$$\mathcal{D}_{\mathbf{q}\lambda}(\tau) = -\left\langle T_\tau b_{\mathbf{q}\lambda}(\tau) b_{\mathbf{q}\lambda}^\dagger(0) \right\rangle, \quad (8.22)$$

where λ enumerates the three polarizations of acoustic phonons. We also note that for bosons we have

$$\mathcal{D}(\tau + \beta) = +\mathcal{D}(\tau), \quad (8.23)$$

i.e., the opposite sign compared to fermions. Therefore, in the Fourier transform only the bosonic Matsubara frequencies

$$\nu_n := \frac{2\pi n}{\hbar\beta}, \quad n \in \mathbb{Z} \quad (8.24)$$

occur.

8.2 Coulomb interaction

We now discuss the effect of the electron-electron Coulomb interaction. We will mainly do that at the level of Feynman diagrams but it should be kept in mind that these represent mathematical expressions that can be evaluated if needed. The Coulomb interaction can be written in second-quantized form as

$$V_{\text{int}} = \frac{1}{2} \sum_{\sigma_1 \sigma_2} \int d^3 r_1 d^3 r_2 \Psi_{\sigma_1}^\dagger(\mathbf{r}_1) \Psi_{\sigma_2}^\dagger(\mathbf{r}_2) V_C(|\mathbf{r}_2 - \mathbf{r}_1|) \Psi_{\sigma_2}(\mathbf{r}_2) \Psi_{\sigma_1}(\mathbf{r}_1) \quad (8.25)$$

with

$$V_C(r) = \frac{e^2}{r}. \quad (8.26)$$

In momentum space we have

$$\Psi_\sigma(\mathbf{r}) = \frac{1}{\sqrt{V}} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}} c_{\mathbf{k}\sigma} \quad (8.27)$$

so that

$$\begin{aligned} V_{\text{int}} &= \frac{1}{2V^2} \sum_{\mathbf{k}\mathbf{k}'\mathbf{k}''\mathbf{k}'''} \sum_{\sigma_1 \sigma_2} \int d^3 r_1 d^3 r_2 e^{-i\mathbf{k}\cdot\mathbf{r}_1 - i\mathbf{k}'\cdot\mathbf{r}_2 + i\mathbf{k}''\cdot\mathbf{r}_2 + i\mathbf{k}'''\cdot\mathbf{r}_1} c_{\mathbf{k}\sigma_1}^\dagger c_{\mathbf{k}'\sigma_2}^\dagger V_C(|\mathbf{r}_2 - \mathbf{r}_1|) c_{\mathbf{k}''\sigma_2} c_{\mathbf{k}'''\sigma_1} \\ &= \frac{1}{2V^2} \sum_{\mathbf{k}\mathbf{k}'\mathbf{k}''\mathbf{k}'''} \sum_{\sigma_1 \sigma_2} \int d^3 R d^3 \rho e^{i(-\mathbf{k}-\mathbf{k}'+\mathbf{k}''+\mathbf{k}''')\cdot\mathbf{R}} e^{i(\mathbf{k}-\mathbf{k}'+\mathbf{k}''-\mathbf{k}''')\cdot\rho/2} c_{\mathbf{k}\sigma_1}^\dagger c_{\mathbf{k}'\sigma_2}^\dagger V_C(\rho) c_{\mathbf{k}''\sigma_2} c_{\mathbf{k}'''\sigma_1}, \end{aligned} \quad (8.28)$$

where

$$\mathbf{R} = \frac{\mathbf{r}_1 + \mathbf{r}_2}{2}, \quad (8.29)$$

$$\boldsymbol{\rho} = \mathbf{r}_2 - \mathbf{r}_1. \quad (8.30)$$

We can now perform the integral over \mathbf{R} :

$$\begin{aligned} V_{\text{int}} &= \frac{1}{2V} \sum_{\mathbf{k}\mathbf{k}'\mathbf{k}''\mathbf{k}'''} \sum_{\sigma_1 \sigma_2} \int d^3 \rho \delta_{\mathbf{k}+\mathbf{k}'-\mathbf{k}''-\mathbf{k}''',0} e^{i(\mathbf{k}-\mathbf{k}'+\mathbf{k}''-\mathbf{k}''')\cdot\rho/2} c_{\mathbf{k}\sigma_1}^\dagger c_{\mathbf{k}'\sigma_2}^\dagger V_C(\rho) c_{\mathbf{k}''\sigma_2} c_{\mathbf{k}'''\sigma_1} \\ &= \frac{1}{2V} \sum_{\mathbf{k}\mathbf{k}'\mathbf{k}''} \sum_{\sigma_1 \sigma_2} \int d^3 \rho e^{-i(\mathbf{k}'-\mathbf{k}'')\cdot\rho} V_C(\rho) c_{\mathbf{k}\sigma_1}^\dagger c_{\mathbf{k}'\sigma_2}^\dagger c_{\mathbf{k}''\sigma_2} c_{\mathbf{k}+\mathbf{k}'-\mathbf{k}'',\sigma_1}. \end{aligned} \quad (8.31)$$

Substituting new momentum variables

$$\mathbf{k}_1 = \mathbf{k} + \mathbf{k}' - \mathbf{k}'', \quad (8.32)$$

$$\mathbf{k}_2 = \mathbf{k}'', \quad (8.33)$$

$$\mathbf{q} = \mathbf{k}' - \mathbf{k}'', \quad (8.34)$$

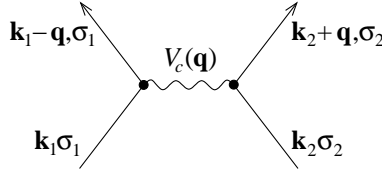
we obtain

$$V_{\text{int}} = \frac{1}{2V} \sum_{\mathbf{k}_1 \mathbf{k}_2 \mathbf{q}} \sum_{\sigma_1 \sigma_2} V_C(\mathbf{q}) c_{\mathbf{k}_1 - \mathbf{q}, \sigma_1}^\dagger c_{\mathbf{k}_2 + \mathbf{q}, \sigma_2}^\dagger c_{\mathbf{k}_2 \sigma_2} c_{\mathbf{k}_1 \sigma_1} \quad (8.35)$$

with

$$V_C(\mathbf{q}) = \int d^3 \rho e^{-i\mathbf{q} \cdot \rho} V_C(\rho). \quad (8.36)$$

The interaction V_{int} is a sum over all processes in which two electrons come in with momenta \mathbf{k}_1 and \mathbf{k}_2 , a momentum of \mathbf{q} is transferred from one to the other through the Coulomb interaction, and the electrons fly out with momenta $\mathbf{k}_1 - \mathbf{q}$ and $\mathbf{k}_2 + \mathbf{q}$.



$V_C(\mathbf{q})$ is obviously the Fourier transform of the Coulomb interaction. It is most easily obtained by Fourier transforming the Poisson equation for a point charge,

$$\nabla^2 \phi(\mathbf{r}) = -4\pi \rho(\mathbf{r}) = -4\pi Q \delta(\mathbf{r}) \quad (8.37)$$

$$\Rightarrow \int d^3 r e^{-i\mathbf{q} \cdot \mathbf{r}} \nabla^2 \phi(\mathbf{r}) = -4\pi Q \quad (8.38)$$

$$\text{by parts} \Rightarrow \int d^3 r (-i\mathbf{q})^2 e^{-i\mathbf{q} \cdot \mathbf{r}} \phi(\mathbf{r}) = -4\pi Q \quad (8.39)$$

$$\Rightarrow q^2 \phi(\mathbf{q}) = 4\pi Q \quad (8.40)$$

$$\Rightarrow \phi(\mathbf{q}) = 4\pi \frac{Q}{q^2} \quad (8.41)$$

so that

$$V_C(\mathbf{q}) = 4\pi \frac{e^2}{q^2}. \quad (8.42)$$

Screening and RPA

The bare Coulomb interaction V_C is strongly repulsive, as noted above. But if one (indirectly) measures the interaction between charges in a metal, one does not find V_C but a reduced interaction. First of all, there will be a dielectric function from the polarizability of the ion cores,

$$V_C(\mathbf{q}) \rightarrow \frac{4\pi}{\epsilon} \frac{e^2}{q^2}, \quad (8.43)$$

but this only leads to a quantitative change, not a qualitative one. We absorb the factor $1/\epsilon$ into e^2 from now on. More importantly, a test charge in a metal is *screened* by a cloud of opposite charge so that from far away the effective charge is strongly reduced. Diagrammatically, the effective Coulomb interaction $V_C^{\text{full}}(\mathbf{q}, i\nu_n)$ is given by the sum of all connected diagrams with two external legs that represent the Coulomb interaction. With the representations

$$-V_C \equiv \text{~~~~~} \quad (8.44)$$

(the minus sign is conventional) and

$$-V_C^{\text{full}} \equiv \text{~~~~~} \quad (8.45)$$

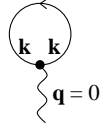
as well as

$$\mathcal{G}^0 \equiv \text{---} \leftarrow \text{---} \quad (8.46)$$

for the bare electronic Green function one would find for the non-interacting Hamiltonian H_0 , we obtain

$$\begin{aligned} \text{~~~~~} &= \text{~~~~~} \\ &+ \text{~~~~~} + \text{~~~~~} + \text{~~~~~} \\ &+ \text{~~~~~} + \text{~~~~~} + \text{~~~~~} + \text{~~~~~} \\ &+ \dots \end{aligned} \quad (8.47)$$

This is an expansion in powers of e^2 since V_C contributes a factor of e^2 . We have exhibited all diagrams up to order e^6 . If we try to evaluate this sum term by term, we encounter a problem: At each vertex \bullet , momentum and energy (frequency) must be conserved. Thus in partial diagrams of the form



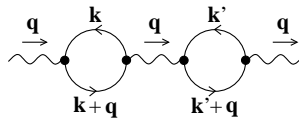
the Coulomb interaction carries momentum $\mathbf{q} = 0$. But

$$V_C(0) = 4\pi \frac{e^2}{0^2} \quad (8.48)$$

is infinite. However, one can show that the closed \mathcal{G} loop corresponds to the average electron density so that the diagram signifies the Coulomb interaction with the average electronic charge density (i.e., the *Hartree energy*). But this is compensated by the average charge density of the nuclei. Thus we can omit all diagrams containing the “tadpole” diagram shown above.

Since we still cannot evaluate the sum in closed form we need an approximation. We first consider the limiting cases of small and large \mathbf{q} in $V_C^{\text{full}}(\mathbf{q}, i\nu_n)$.

- For large \mathbf{q} , corresponding to small distances, the first diagram is proportional to $1/q^2$, whereas all the others are at least of order $1/q^4$ and are thus suppressed. We should recover the bare Coulomb interaction for large \mathbf{q} or small distances, which is plausible since the polarization of the electron gas cannot efficiently screen the interaction between two test charges that are close together.
- For small \mathbf{q} we find that higher-order terms contain higher and higher powers of $1/q^2$ and thus become very large. This is alarming. The central idea of our approximation is to keep only the dominant term (diagram) at each order in e^2 . The dominant term is the one with the highest power in $1/q^2$. Only the V_C lines forming the backbone of the diagrams (drawn horizontally) carry the external momentum \mathbf{q} due to momentum conservation at the vertices. Thus the dominant terms are the ones with *all* V_C lines in the backbone:



Summing up these dominant terms we obtain the approximate effective interaction

$$\begin{aligned}
-V_C^{\text{RPA}} := & \text{~~~~~} + \text{~~~~~} \\
& + \text{~~~~~} + \text{~~~~~} \\
& + \dots
\end{aligned} \tag{8.49}$$

This is called the *random phase approximation* (RPA) for historical reasons that do not concern us here, or the Thomas-Fermi approximation. The most important part of RPA diagrams is clearly the bubble diagram

$$\Pi_0 \equiv \text{~~~~~}, \tag{8.50}$$

which stands for

$$\begin{aligned}
\Pi_0(\mathbf{q}, i\nu_n) &= -\frac{1}{\beta} \sum_{i\omega_n} \frac{1}{V} \sum_{\mathbf{k}\sigma} \mathcal{G}_{\mathbf{k}+\mathbf{q},\sigma}^0(i\omega_n + i\nu_n) \mathcal{G}_{\mathbf{k}\sigma}^0(i\omega_n) \\
&= -\frac{2}{\beta} \sum_{i\omega_n} \sum_{\mathbf{k}} \frac{1}{i\omega_n + i\nu_n - \xi_{\mathbf{k}+\mathbf{q}}} \frac{1}{i\omega_n - \xi_{\mathbf{k}}}
\end{aligned} \tag{8.51}$$

(the factor of 2 is due to the spin). We can write the diagrammatic series also as

$$\begin{aligned}
-V_C^{\text{RPA}}(\mathbf{q}, i\nu_n) &= -V_C(\mathbf{q}) + V_C(\mathbf{q}) \Pi_0(\mathbf{q}, i\nu_n) V_C(\mathbf{q}) - V_C(\mathbf{q}) \Pi_0(\mathbf{q}, i\nu_n) V_C(\mathbf{q}) \Pi_0(\mathbf{q}, i\nu_n) V_C(\mathbf{q}) + \dots \\
&= -V_C(\mathbf{q}) [1 - \Pi_0(\mathbf{q}, i\nu_n) V_C(\mathbf{q}) + \Pi_0(\mathbf{q}, i\nu_n) V_C(\mathbf{q}) \Pi_0(\mathbf{q}, i\nu_n) V_C(\mathbf{q}) - \dots].
\end{aligned} \tag{8.52}$$

This is a geometric series, which we can sum up,

$$V_C^{\text{RPA}}(\mathbf{q}, i\nu_n) = \frac{V_C(\mathbf{q})}{1 + V_C(\mathbf{q}) \Pi_0(\mathbf{q}, i\nu_n)}. \tag{8.53}$$

Π_0 can be evaluated, it is essentially -1 times the susceptibility of the free electron gas. We cannot explain this here, but it is plausible that the susceptibility, which controls the electric polarization of the electron gas, should enter into a calculation of the screened Coulomb interaction. Note that V_C^{RPA} has a frequency dependence since Π_0 (or the susceptibility) has one.

In the static limit $i\nu \rightarrow 0 + i0^+$ and at low temperatures $T \ll E_F/k_B$, one can show that

$$\Pi_0(\mathbf{q}, 0) \cong \text{const} = N(E_F), \tag{8.54}$$

where $N(E_F)$ is the electronic density of states at the Fermi energy, including a factor of two for the spin. Thus we obtain

$$\begin{aligned}
V_C^{\text{RPA}}(\mathbf{q}) &= \frac{4\pi \frac{e^2}{q^2}}{1 + 4\pi \frac{e^2}{q^2} \Pi_0} = 4\pi \frac{e^2}{q^2 + 4\pi e^2 \Pi_0} \\
&\cong 4\pi \frac{e^2}{q^2 + 4\pi e^2 N(E_F)} \\
&= 4\pi \frac{e^2}{q^2 + \kappa_s^2}
\end{aligned} \tag{8.55}$$

with

$$\kappa_s := \sqrt{4\pi e^2 N(E_F)}. \tag{8.56}$$

Note that summing up the more and more strongly diverging terms has led to a regular result in the limit $\mathbf{q} \rightarrow 0$. The result can be Fourier-transformed to give

$$V_C^{\text{RPA}}(\mathbf{r}) = e^2 \frac{e^{-\kappa_s r}}{r} \quad (8.57)$$

(not too much confusion should result from using the same symbol “ e ” for the elementary charge and the base of the exponential function). This is the *Yukawa potential*, which is exponentially suppressed beyond the screening length $1/\kappa_s$. This length is on the order of $10^{-8} \text{ m} = 10 \text{ nm}$ in typical metals. While we thus find a strong suppression of the repulsive interaction at large distances, there is no sign of it becoming attractive.

On the other hand, for very high frequencies $\nu \gtrsim E_F/\hbar$ the electron gas cannot follow the perturbation, the susceptibility and Π_0 go to zero, and we obtain the bare interaction,

$$V_C^{\text{RPA}}(\mathbf{q}, \nu \rightarrow \infty) \cong V_C(\mathbf{q}) = 4\pi \frac{e^2}{q^2}. \quad (8.58)$$

8.3 Electron-phonon interaction

The nuclei (or ion cores) in a crystal oscillate about their equilibrium positions. The quanta of these lattice vibrations are the phonons. The many-particle Hamiltonian including the phonons has the form

$$H = H_{\text{el}} + H_{\text{ph}} + H_{\text{el-ph}}, \quad (8.59)$$

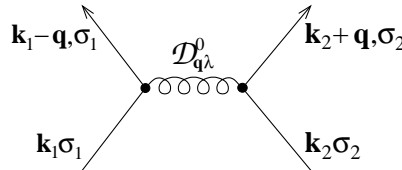
where we have discussed the electronic part H_{el} before,

$$H_{\text{ph}} = \sum_{\mathbf{q}\lambda} \Omega_{\mathbf{q}\lambda} \left(b_{\mathbf{q}\lambda}^\dagger b_{\mathbf{q}\lambda} + \frac{1}{2} \right) \quad (8.60)$$

is the bare Hamiltonian of phonons with dispersion $\Omega_{\mathbf{q}\lambda}$, and

$$H_{\text{el-ph}} = \frac{1}{V} \sum_{\mathbf{k}\sigma} \sum_{\mathbf{q}\lambda} g_{\mathbf{q}\lambda} c_{\mathbf{k}+\mathbf{q},\sigma}^\dagger c_{\mathbf{k}\sigma} \left(b_{\mathbf{q}\lambda} + b_{-\mathbf{q},\lambda}^\dagger \right) \quad (8.61)$$

describes the electron-phonon coupling. $g_{\mathbf{q}\lambda}$ is the coupling strength. Physically, an electron can absorb ($b_{\mathbf{q}\lambda}$) or emit ($b_{-\mathbf{q},\lambda}^\dagger$) a phonon under conservation of momentum. Hence, electrons can interact with one another by exchanging phonons. Diagrammatically, we draw the simplest possible process as



In detail, we define, in analogy to the Coulomb interaction

$$-V_C(\mathbf{q}) \equiv \text{~~~~~}, \quad (8.62)$$

the interaction due to phonon exchange,

$$-V_{\text{ph}}(\mathbf{q}, \lambda, i\nu_n) \equiv -\frac{1}{V} |g_{\mathbf{q}\lambda}|^2 \mathcal{D}_{\mathbf{q}\lambda}^0(i\nu_n) \equiv \text{~~~~~}. \quad (8.63)$$

We quote the expression for the bare phonon Green function, i.e., the one obtained from H_{ph} alone:

$$\mathcal{D}_{\mathbf{q}\lambda}^0(i\nu_n) = \frac{1}{i\nu_n - \Omega_{\mathbf{q}\lambda}} + \frac{1}{-i\nu_n - \Omega_{\mathbf{q}\lambda}} = \frac{2\Omega_{\mathbf{q}\lambda}}{(i\nu_n)^2 - \Omega_{\mathbf{q}\lambda}^2}. \quad (8.64)$$

The phonon-mediated interaction is thus frequency-dependent, whereas the Coulomb interaction is static. However, we could write the Coulomb interaction in a very similar form as the exchange of *photons*. Since the speed of light is so much larger than the speed of sound, we can neglect the dynamics for photon exchange but not for phonon exchange.

Jellium phonons

For our discussions we need a specific model for phonons. We use the simplest one, based on the *jellium approximation* for the nuclei (or ion cores). In this approximation we describe the nuclei by a smooth positive charge density $\rho_+(\mathbf{r}, t)$. In equilibrium, this charge density is uniform, $\rho_+(\mathbf{r}, t) = \rho_+^0$. We consider small deviations

$$\rho_+(\mathbf{r}, t) = \rho_+^0 + \delta\rho_+(\mathbf{r}, t). \quad (8.65)$$

Gauss' law reads

$$\nabla \cdot \mathbf{E} = 4\pi \delta\rho_+(\mathbf{r}, t), \quad (8.66)$$

since ρ_+^0 is compensated by the average electronic charge density. The density of force acting on ρ_+ is $\mathbf{f} = \rho_+ \mathbf{E} \cong \rho_+^0 \mathbf{E}$ to leading order in $\delta\rho_+$. Thus

$$\nabla \cdot \mathbf{f} \cong 4\pi \rho_+^0 \delta\rho_+. \quad (8.67)$$

The conservation of charge is expressed by the continuity equation

$$\frac{\partial}{\partial t} \rho_+ + \nabla \cdot \underbrace{\rho_+ \mathbf{v}}_{=\mathbf{j}_+} = 0. \quad (8.68)$$

To leading order this reads

$$\frac{\partial}{\partial t} \delta\rho_+ + \rho_+^0 \nabla \cdot \mathbf{v} \cong 0 \quad (8.69)$$

$$\Rightarrow \frac{\partial^2}{\partial t^2} \delta\rho_+ \cong -\rho_+^0 \nabla \cdot \frac{\partial}{\partial t} \mathbf{v} \stackrel{\text{Newton}}{=} -\rho_+^0 \nabla \cdot \frac{\mathbf{f}}{\rho_m}, \quad (8.70)$$

where ρ_m is the *mass* density of the nuclei. With the nuclear charge Ze and mass M we obtain

$$\frac{\partial^2}{\partial t^2} \delta\rho_+ \cong -\frac{Ze}{M} \nabla \cdot \mathbf{f} = -\frac{Ze}{M} 4\pi \rho_+^0 \delta\rho_+, \quad (8.71)$$

which is solved by

$$\delta\rho_+(\mathbf{r}, t) = \delta\rho_+(\mathbf{r}) e^{-i\Omega t} \quad (8.72)$$

with

$$\Omega = \sqrt{4\pi \frac{Ze}{M} \rho_+^0} = \sqrt{4\pi \frac{Z^2 e^2}{M} n_+^0}, \quad (8.73)$$

where n_+^0 is the concentration of nuclei (or ions). We thus obtain *optical* phonons with completely flat dispersion, i.e., we find the same frequency Ω for all vibrations.

One can also obtain the coupling strength $g_{\mathbf{q}}$. It is clear that it will be controlled by the Coulomb interaction between electrons and fluctuations $\delta\rho_+$ in the jellium charge density. We refer to the lecture notes on many-particle theory and only give the result:

$$\frac{1}{V} |g_{\mathbf{q}}|^2 = \frac{\Omega}{2} V_C(\mathbf{q}). \quad (8.74)$$

Consequently, the electron-electron interaction due to phonon exchange becomes

$$V_{\text{ph}}(\mathbf{q}, i\nu_n) = \frac{1}{V} |g_{\mathbf{q}}|^2 \mathcal{D}_{\mathbf{q}}^0(i\nu_n) = \frac{\Omega}{2} V_C(\mathbf{q}) \frac{2\Omega}{(i\nu_n)^2 - \Omega^2} = V_C(\mathbf{q}) \frac{\Omega^2}{(i\nu_n)^2 - \Omega^2}. \quad (8.75)$$

It is thus proportional to the bare Coulomb interaction, with an additional frequency-dependent factor. The retarded form reads

$$V_{\text{ph}}^R(\mathbf{q}, \nu) = V_{\text{ph}}(\mathbf{q}, i\nu_n \rightarrow \nu + i0^+) = V_C(\mathbf{q}) \frac{\Omega^2}{(\nu + i0^+)^2 - \Omega^2} = V_C(\mathbf{q}) \frac{\Omega^2}{\nu^2 - \Omega^2 + i0^+ \text{sgn } \nu}, \quad (8.76)$$

where we have used $2\nu i0^+ = i0^+ \text{sgn } \nu$ and have neglected the square of infinitesimal quantities.

8.4 Effective interaction between electrons

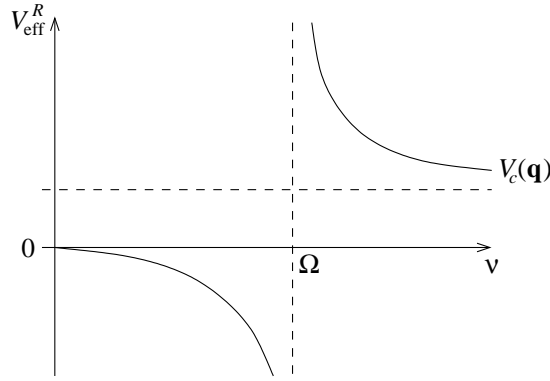
Combining the bare Coulomb interaction and the bare interaction due to phonon exchange, calculated for the jellium model, we obtain the bare effective interaction between electrons,

$$\begin{aligned}
 V_{\text{eff}}(\mathbf{q}, i\nu_n) &:= V_C(\mathbf{q}) + V_{\text{ph}}(\mathbf{q}, i\nu_n) \\
 &= V_C(\mathbf{q}) + V_C(\mathbf{q}) \frac{\Omega^2}{(i\nu_n)^2 - \Omega^2} \\
 &= V_C(\mathbf{q}) \frac{(i\nu_n)^2}{(i\nu_n)^2 - \Omega^2}.
 \end{aligned} \tag{8.77}$$

The retarded form is

$$\begin{aligned}
 V_{\text{eff}}^R(\mathbf{q}, \nu) &= V_{\text{eff}}(\mathbf{q}, i\nu_n \rightarrow \nu + i0^+) \\
 &= V_C(\mathbf{q}) \frac{\nu^2}{\nu^2 - \Omega^2 + i0^+ \text{sgn } \nu}.
 \end{aligned} \tag{8.78}$$

This expression is real except at $\nu = \Omega$ and has a pole there. Moreover, V_{eff}^R is proportional to V_C with a *negative* prefactor as long as $\nu < \Omega$.



The effective interaction is thus *attractive* for $\nu < \Omega$. The exchange of phonons overcompensates the repulsive Coulomb interaction. On the other hand, for $\nu \rightarrow 0$, the effective interaction vanishes. This means that in a quasi-static situation the electrons do not see each other at all.

What happens physically is that the electrons polarize the (jellium) charge density of the nuclei. The nuclei have a high inertial mass, their reaction to a perturbation has a typical time scale of $1/\Omega$ or a frequency scale of Ω . For processes slow compared to Ω , the nuclei can completely screen the electron charge, forming a *polaron*, which is charge-neutral. For frequencies $\nu > 0$, we have to think in terms of the response of the system to a test electron oscillating with frequency ν . The jellium acts as an oscillator with eigenfrequency Ω . At the present level of approximation it is an undamped oscillator. The jellium oscillator is excited at the frequency ν . For $0 < \nu < \Omega$, it is driven below its eigenfrequency and thus oscillates in phase with the test electron. The amplitude, i.e., the jellium polarization, is enhanced compared to the $\nu = 0$ limit simply because the system is closer to the resonance at Ω . Therefore, the oscillating electron charge is *overscreened*. On the other hand, for $\nu > \Omega$ the jellium oscillator is driven above its eigenfrequency and thus follows the test electron with a phase difference of π . Thus the electron charge is not screened at all but rather enhanced and the interaction is more strongly repulsive than the pure Coulomb interaction.

Screening of the effective interaction

From our discussion of the Coulomb interaction we know that the real interaction between two electrons in a metal is strongly screened at all except very short distances. This screening is well described within the RPA. We now apply the RPA to the effective interaction derived above. We define

$$-V_{\text{eff}} \equiv \text{~~~~~} = \text{~~~~~} + \text{~~~~~} \tag{8.79}$$

and

$$-V_{\text{eff}}^{\text{RPA}} \equiv \text{~~~~~} := \text{~~~~~} + \text{~~~~~} + \text{~~~~~} + \dots \quad (8.80)$$

or

$$\begin{aligned} -V_{\text{eff}}^{\text{RPA}}(\mathbf{q}, i\nu_n) &= -V_{\text{eff}}(\mathbf{q}, i\nu_n) + V_{\text{eff}}(\mathbf{q}, i\nu_n) \Pi_0(\mathbf{q}, i\nu_n) V_{\text{eff}}(\mathbf{q}, i\nu_n) \\ &\quad - V_{\text{eff}}(\mathbf{q}, i\nu_n) \Pi_0(\mathbf{q}, i\nu_n) V_{\text{eff}}(\mathbf{q}, i\nu_n) \Pi_0(\mathbf{q}, i\nu_n) V_{\text{eff}}(\mathbf{q}, i\nu_n) + \dots \end{aligned} \quad (8.81)$$

As above, we can sum this up,

$$\begin{aligned} V_{\text{eff}}^{\text{RPA}}(\mathbf{q}, i\nu_n) &= \frac{V_{\text{eff}}(\mathbf{q}, i\nu_n)}{1 + V_{\text{eff}}(\mathbf{q}, i\nu_n) \Pi_0(\mathbf{q}, i\nu_n)} = V_C(\mathbf{q}) \frac{\frac{(i\nu_n)^2}{(i\nu_n)^2 - \Omega^2}}{1 + V_C(\mathbf{q}) \frac{(i\nu_n)^2}{(i\nu_n)^2 - \Omega^2} \Pi_0(\mathbf{q}, i\nu_n)} \\ &= V_C(\mathbf{q}) \frac{(i\nu_n)^2}{(i\nu_n)^2 - \Omega^2 + (i\nu_n)^2 V_C(\mathbf{q}) \Pi_0(\mathbf{q}, i\nu_n)} \\ &= \frac{V_C(\mathbf{q})}{\underbrace{1 + V_C(\mathbf{q}) \Pi_0(\mathbf{q}, i\nu_n)}_{= V_C^{\text{RPA}}(\mathbf{q}, i\nu_n)}} \frac{(i\nu_n)^2 + (i\nu_n)^2 V_C(\mathbf{q}) \Pi_0(\mathbf{q}, i\nu_n)}{(i\nu_n)^2 - \Omega^2 + (i\nu_n)^2 V_C(\mathbf{q}) \Pi_0(\mathbf{q}, i\nu_n)} \\ &= V_C^{\text{RPA}}(\mathbf{q}, i\nu_n) \frac{(i\nu_n)^2}{(i\nu_n)^2 - \frac{\Omega^2}{1 + V_C(\mathbf{q}) \Pi_0(\mathbf{q}, i\nu_n)}} \\ &= V_C^{\text{RPA}}(\mathbf{q}, i\nu_n) \frac{(i\nu_n)^2}{(i\nu_n)^2 - \omega_{\mathbf{q}}^2(i\nu_n)} \end{aligned} \quad (8.82)$$

with the *renormalized phonon frequency*

$$\omega_{\mathbf{q}}(i\nu_n) := \frac{\Omega}{\sqrt{1 + V_C(\mathbf{q}) \Pi_0(\mathbf{q}, i\nu_n)}}. \quad (8.83)$$

To see that this is a reasonable terminology, compare $V_{\text{eff}}^{\text{RPA}}$ to the bare effective interaction

$$V_{\text{eff}}(\mathbf{q}, i\nu_n) = V_C(\mathbf{q}) \frac{(i\nu_n)^2}{(i\nu_n)^2 - \Omega^2}. \quad (8.84)$$

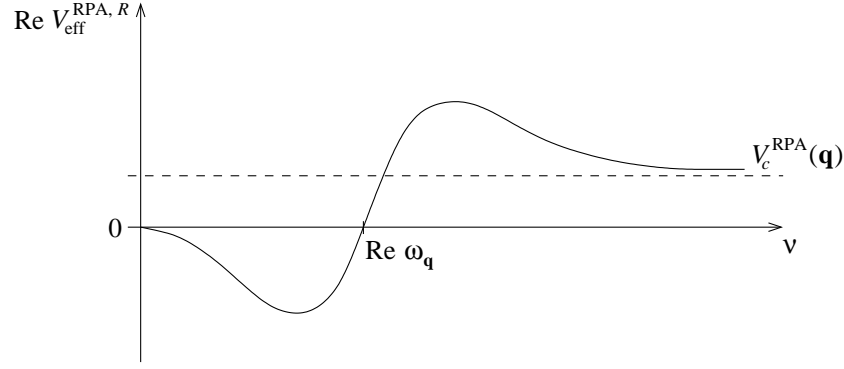
Evidently, screening leads to the replacements $V_C \rightarrow V_C^{\text{RPA}}$ and $\Omega \rightarrow \omega_{\mathbf{q}}$.

For small momenta and frequencies, we have $\Pi_0 \rightarrow N(E_F)$, the density of states at E_F . In this limit we thus obtain

$$\omega_{\mathbf{q}} \cong \frac{\Omega}{\sqrt{1 + 4\pi \frac{e^2}{q^2} N(E_F)}} = \frac{\Omega}{\sqrt{1 + \frac{\kappa_s^2}{q^2}}} \cong \frac{\Omega}{\sqrt{\frac{\kappa_s^2}{q^2}}} = \frac{\Omega}{\kappa_s} q. \quad (8.85)$$

Due to screening we thus find an *acoustic* dispersion of jellium phonons. This is of course much more realistic than an optical Einstein mode.

Beyond the low-frequency limit it is important that Π_0 and thus $\omega_{\mathbf{q}}$ obtains a sizable imaginary part. It smears out the pole in the retarded interaction $V_{\text{eff}}^{\text{RPA}}(\mathbf{q}, \nu)$ or rather moves it away from the real-frequency axis—the lattice vibrations are now damped. The real part of the retarded interaction is sketched here for fixed \mathbf{q} :



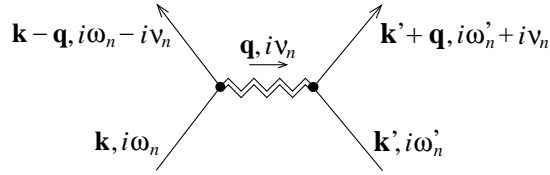
Note that

- the interaction still vanishes in the static limit $\nu \rightarrow 0$,
- the interaction is attractive for $0 < \nu < \text{Re } \omega_{\mathbf{q}}$, where $\text{Re } \omega_{\mathbf{q}} \sim q$.

It is important that the static interaction is not attractive but zero. Hence, we do not expect *static* bound states of two electrons.

To obtain analytical results, it is necessary to simplify the interaction. The main property required for superconductivity is that the interaction is attractive for frequencies below some typical phonon frequency. The typical phonon frequency is the material specific *Debye frequency* ω_D . We write the effective RPA interaction in terms of the incoming and transferred momenta and frequencies,

$$V_{\text{eff}}^{\text{RPA}} = V_{\text{eff}}^{\text{RPA}}(\mathbf{k}, i\omega_n; \mathbf{k}', i\omega'_n; \mathbf{q}, i\nu_n). \quad (8.86)$$



We then approximate the interaction (very crudely) by a constant $-V_0 < 0$ if both incoming frequencies are smaller than ω_D and by zero otherwise,

$$V_{\text{eff}}^{\text{RPA}} \approx \begin{cases} -V_0 & \text{for } |i\omega_n|, |i\omega'_n| < \omega_D, \\ 0 & \text{otherwise.} \end{cases} \quad (8.87)$$

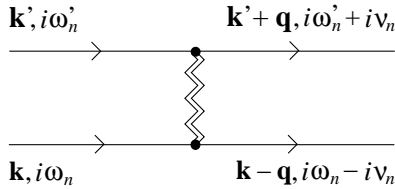
9

Cooper instability and BCS ground state

In this chapter we will first show that the attractive effective interaction leads to an instability of the normal state, i.e., of the Fermi sea. Then we will discuss the new state that takes its place.

9.1 Cooper instability

Let us consider the scattering of two electrons due to the effective interaction. A single scattering event is represented by the diagram



Electrons can also scatter multiple times:

An instability occurs if this series diverges since then the scattering becomes infinitely strong. In this case the perturbative expansion in the interaction strength V_0 represented by the diagrams breaks down. This means that the true equilibrium state cannot be obtained from the equilibrium state for $V_0 = 0$, namely the noninteracting Fermi gas, by perturbation theory. A state that is perturbatively connected to the free Fermi gas is called a Landau Fermi liquid. It is an appropriate description for normal metals. Conversely, a scattering instability signals that the equilibrium state is no longer a Fermi liquid.

Like in the RPA, it turns out to be sufficient to consider the dominant diagrams at each order. These are the *ladder diagrams*, which do not contain crossing interaction lines. Moreover, the instability occurs first for the scattering of two electrons with opposite momentum, frequency, and spin. We thus restrict ourselves to the

diagrams describing this situation. We define the *scattering vertex* Λ by

$$\begin{aligned}
-\Lambda &\equiv \text{Diagram with a shaded box } \Lambda \text{ and external lines } (-\mathbf{k} \downarrow -i\omega_n, -\mathbf{p} \downarrow -i\Omega_n, \mathbf{k} \uparrow i\omega_n, \mathbf{p} \uparrow i\Omega_n) \\
&:= \text{Diagram with a wavy line and external lines } (-\mathbf{k} \downarrow -i\omega_n, -\mathbf{p} \downarrow -i\Omega_n, \mathbf{k} \uparrow i\omega_n, \mathbf{p} \uparrow i\Omega_n) \\
&\quad + \text{Diagram with two wavy lines and external lines } (-\mathbf{k} \downarrow -i\omega_n, -\mathbf{k}_1 \downarrow -i\omega_n^1, -\mathbf{p} \downarrow -i\Omega_n, \mathbf{k} \uparrow i\omega_n, \mathbf{k}_1 \uparrow i\omega_n^1, \mathbf{p} \uparrow i\Omega_n) \\
&\quad + \dots
\end{aligned} \tag{9.2}$$

This is a geometric series, which we can sum up,

$$\text{Diagram with shaded box } \Lambda = \text{Diagram with wavy line} \left[1 + \text{Diagram with two wavy lines} + \text{Diagram with three wavy lines} + \dots \right] = \frac{\text{Diagram with wavy line}}{1 - \text{Diagram with two wavy lines}} \tag{9.3}$$

With our approximation

$$V_{\text{eff}}^{\text{RPA}} \approx \begin{cases} -V_0 & \text{for } |\omega_n| < \omega_D, \\ 0 & \text{otherwise,} \end{cases} \tag{9.4}$$

we obtain

$$-\Lambda(i\omega_n) \approx \frac{+V_0}{1 - V_0 \frac{1}{\beta} \sum_{i\omega_n^1, |\omega_n^1| < \omega_D} \frac{1}{V} \sum_{\mathbf{k}_1} \mathcal{G}_{\mathbf{k}_1 \uparrow}^0(i\omega_n^1) \mathcal{G}_{-\mathbf{k}_1 \downarrow}^0(-i\omega_n^1)} \tag{9.5}$$

for $|\omega_n| < \omega_D$ and zero otherwise. Thus

$$\Lambda(i\omega_n) \approx \begin{cases} \frac{-V_0}{1 - V_0 \frac{1}{\beta} \sum_{i\omega_n^1, |\omega_n^1| < \omega_D} \frac{1}{V} \sum_{\mathbf{k}_1} \mathcal{G}_{\mathbf{k}_1 \uparrow}^0(i\omega_n^1) \mathcal{G}_{-\mathbf{k}_1 \downarrow}^0(-i\omega_n^1)} & \text{for } |\omega_n| < \omega_D, \\ 0 & \text{otherwise.} \end{cases} \tag{9.6}$$

We see that the scattering vertex Λ diverges if

$$V_0 \frac{1}{\beta} \sum_{\substack{i\omega_n^1 \\ |\omega_n^1| < \omega_D}} \frac{1}{V} \sum_{\mathbf{k}_1} \mathcal{G}_{\mathbf{k}_1 \uparrow}^0(i\omega_n^1) \mathcal{G}_{-\mathbf{k}_1 \downarrow}^0(-i\omega_n^1) = 1. \tag{9.7}$$

This expression depends on temperature. We now evaluate it explicitly:

$$V_0 k_B T \sum_{\substack{i\omega_n^1 \\ |\omega_n^1| < \omega_D}} \frac{1}{V} \sum_{\mathbf{k}_1} \frac{1}{i\omega_n^1 - \xi_{\mathbf{k}_1}} \frac{1}{-i\omega_n^1 - \xi_{\mathbf{k}_1}} = V_0 k_B T \sum_{\substack{i\omega_n^1 \\ |\omega_n^1| < \omega_D}} \int_{-\infty}^{\infty} d\xi D(\mu + \xi) \frac{1}{(\omega_n^1)^2 + \xi^2} \tag{9.8}$$

with the density of states per spin direction and per unit cell, $D(\epsilon)$. Assuming the density of states to be approximately constant close to the Fermi energy we get (with $\hbar = 1$)

$$\begin{aligned}
\cdots &\approx V_0 k_B T \sum_{\substack{i\omega_n^1 \\ |i\omega_n^1| < \omega_D}} D(E_F) \underbrace{\int_{-\infty}^{\infty} \frac{d\xi}{(\omega_n^1)^2 + \xi^2}}_{= \pi/|\omega_n^1|} \\
&= V_0 k_B T D(E_F) \pi 2 \sum_{n=0}^{\beta\omega_D/2\pi} \frac{1}{\frac{(2n+1)\pi}{\beta}} = V_0 k_B T D(E_F) \beta \sum_{n=0}^{\beta\omega_D/2\pi} \frac{1}{n + \frac{1}{2}} \\
&\cong V_0 D(E_F) \left[\gamma + \ln \left(4 \frac{\beta\omega_D}{2\pi} \right) \right]. \tag{9.9}
\end{aligned}$$

In the last step we have used an approximation for the sum over n that is valid for $\beta\omega_D \gg 1$, i.e., if the sum has many terms. Since T_c for superconductors is typically small compared to the Debye temperature ω_D/k_B (a few hundred Kelvin), this is justified. $\gamma \approx 0.577216$ is the Euler constant. Altogether, we find

$$\Lambda \approx \frac{-V_0}{1 - V_0 D(E_F) \left(\gamma + \ln \frac{2\beta\omega_D}{\pi} \right)} \tag{9.10}$$

for $|i\omega_n| < \omega_D$. Coming from high temperatures, but still satisfying $k_B T \ll \omega_D$, multiple scattering enhances Λ . Λ diverges at $T = T_c$, where

$$V_0 D(E_F) \left(\gamma + \ln \frac{2\omega_D}{\pi k_B T_c} \right) = 1 \tag{9.11}$$

$$\Rightarrow \ln \frac{2e^\gamma \omega_D}{\pi k_B T_c} = \frac{1}{V_0 D(E_F)} \tag{9.12}$$

$$\Rightarrow k_B T_c = \underbrace{\frac{2e^\gamma}{\pi}}_{\approx 1.13387 \sim 1} \omega_D \exp \left(-\frac{1}{V_0 D(E_F)} \right). \tag{9.13}$$

This is the *Cooper instability*. Its characteristic temperature scale appears to be the Debye temperature of a few hundred Kelvin. This is disturbing since we do not observe an instability at such high temperatures. However, the exponential factor tends to be on the order of 1/100 so that we obtain T_c of a few Kelvin. It is important that $k_B T_c$ is not analytic in V_0 at $V_0 = 0$ (the function has an essential singularity there). Thus $k_B T_c$ cannot be expanded into a Taylor series around the non-interacting limit. This means that we cannot obtain $k_B T_c$ in perturbation theory in V_0 to any finite order. BCS theory is indeed non-perturbative.

9.2 The BCS ground state

We have seen that the Fermi sea becomes unstable due to the scattering of electrons in states $|\mathbf{k}, \uparrow\rangle$ and $|\mathbf{-k}, \downarrow\rangle$. Bardeen, Cooper, and Schrieffer (BCS) have proposed an ansatz for the new ground state. It is based on the idea that electrons from the states $|\mathbf{k}, \uparrow\rangle$ and $|\mathbf{-k}, \downarrow\rangle$ form (so-called Cooper) pairs and that the ground state is a superposition of states built up of such pairs. The ansatz reads

$$|\psi_{\text{BCS}}\rangle = \prod_{\mathbf{k}} \left(u_{\mathbf{k}} + v_{\mathbf{k}} c_{\mathbf{k}\uparrow}^\dagger c_{\mathbf{-k},\downarrow}^\dagger \right) |0\rangle, \tag{9.14}$$

where $|0\rangle$ is the vacuum state without any electrons and $u_{\mathbf{k}}, v_{\mathbf{k}}$ are as yet unknown complex coefficients. Normalization requires

$$\begin{aligned}
1 &= \langle \psi_{\text{BCS}} | \psi_{\text{BCS}} \rangle \\
&= \langle 0 | \prod_{\mathbf{k}} (u_{\mathbf{k}}^* + v_{\mathbf{k}}^* c_{-\mathbf{k},\downarrow} c_{\mathbf{k}\uparrow}) \prod_{\mathbf{k}'} (u_{\mathbf{k}'} + v_{\mathbf{k}'} c_{\mathbf{k}'\uparrow}^\dagger c_{-\mathbf{k}',\downarrow}^\dagger) | 0 \rangle \\
&= \langle 0 | \prod_{\mathbf{k}} (|u_{\mathbf{k}}|^2 + u_{\mathbf{k}}^* v_{\mathbf{k}} c_{\mathbf{k}\uparrow}^\dagger c_{-\mathbf{k},\downarrow}^\dagger + u_{\mathbf{k}} v_{\mathbf{k}}^* c_{-\mathbf{k},\downarrow} c_{\mathbf{k}\uparrow} + |v_{\mathbf{k}}|^2) | 0 \rangle \\
&= \prod_{\mathbf{k}} (|u_{\mathbf{k}}|^2 + |v_{\mathbf{k}}|^2). \tag{9.15}
\end{aligned}$$

This is certainly satisfied if we demand $|u_{\mathbf{k}}|^2 + |v_{\mathbf{k}}|^2 = 1$ for all \mathbf{k} , which we will do from now on. Note that the occupations of $|\mathbf{k}, \uparrow\rangle$ and $|\mathbf{k}, \downarrow\rangle$ are maximally correlated; either both are occupied or both are empty. Also, $|\psi_{\text{BCS}}\rangle$ is peculiar in that it is a superposition of states with different total electron numbers. (We could imagine the superconductor to be entangled with a much larger electron reservoir so that the total electron number in superconductor and reservoir is fixed.) As a consequence, expressions containing unequal numbers of electronic creation and annihilation operators can have non-vanishing expectation values. For example,

$$\begin{aligned}
\langle c_{\mathbf{k}\uparrow}^\dagger c_{-\mathbf{k},\downarrow}^\dagger \rangle_{\text{BCS}} &\equiv \langle \psi_{\text{BCS}} | c_{\mathbf{k}\uparrow}^\dagger c_{-\mathbf{k},\downarrow}^\dagger | \psi_{\text{BCS}} \rangle \\
&= \langle 0 | \prod_{\mathbf{k}'} (u_{\mathbf{k}'}^* + v_{\mathbf{k}'}^* c_{-\mathbf{k}',\downarrow} c_{\mathbf{k}'\uparrow}) c_{\mathbf{k}\uparrow}^\dagger c_{-\mathbf{k},\downarrow}^\dagger \prod_{\mathbf{k}''} (u_{\mathbf{k}''} + v_{\mathbf{k}''} c_{\mathbf{k}''\uparrow}^\dagger c_{-\mathbf{k}'',\downarrow}^\dagger) | 0 \rangle \\
&= \langle 0 | v_{\mathbf{k}}^* u_{\mathbf{k}} \underbrace{\prod_{\mathbf{k}' \neq \mathbf{k}} (|u_{\mathbf{k}'}|^2 + |v_{\mathbf{k}'}|^2)}_{=1} | 0 \rangle = v_{\mathbf{k}}^* u_{\mathbf{k}}. \tag{9.16}
\end{aligned}$$

The coefficients $u_{\mathbf{k}}, v_{\mathbf{k}}$ are chosen so as to minimize the expectation value of the energy, $\langle \psi_{\text{BCS}} | H | \psi_{\text{BCS}} \rangle$, under the constraint $|u_{\mathbf{k}}|^2 + |v_{\mathbf{k}}|^2 = 1$ for all \mathbf{k} . $|\psi_{\text{BCS}}\rangle$ is thus a *variational ansatz*.

We write the Hamiltonian as

$$H = \sum_{\mathbf{k}\sigma} \xi_{\mathbf{k}} c_{\mathbf{k}\sigma}^\dagger c_{\mathbf{k}\sigma} + V_{\text{int}} \tag{9.17}$$

and, in the spirit of the previous section, choose the simplest non-trivial approximation for V_{int} that takes into account that

1. only electrons with energies $|\xi_{\mathbf{k}}| \lesssim \omega_D$ relative to the Fermi energy are important and
2. the instability is due to the scattering between electrons in single-particle states $|\mathbf{k}, \uparrow\rangle$ and $|\mathbf{k}, \downarrow\rangle$.

This leads to

$$V_{\text{int}} = \frac{1}{N} \sum_{\mathbf{k}\mathbf{k}'} V_{\mathbf{k}\mathbf{k}'} c_{\mathbf{k}\uparrow}^\dagger c_{-\mathbf{k},\downarrow}^\dagger c_{-\mathbf{k}',\downarrow} c_{\mathbf{k}'\uparrow} \tag{9.18}$$

with

$$V_{\mathbf{k}\mathbf{k}'} = \begin{cases} -V_0 & \text{for } |\xi_{\mathbf{k}}| < \omega_D \text{ and } |\xi_{\mathbf{k}'}| < \omega_D, \\ 0 & \text{otherwise.} \end{cases} \tag{9.19}$$

We assume that scattering without momentum transfer, $\mathbf{k}' = \mathbf{k}$, contributes negligibly compared to $\mathbf{k}' \neq \mathbf{k}$ since there are many more scattering channels for $\mathbf{k}' \neq \mathbf{k}$. We also assume that $u_{-\mathbf{k}} = u_{\mathbf{k}}, v_{-\mathbf{k}} = v_{\mathbf{k}}$, which is, at worst,

a restriction of our variational ansatz. Then we obtain

$$\begin{aligned}
\langle \psi_{\text{BCS}} | H | \psi_{\text{BCS}} \rangle &= \sum_{\mathbf{k}\sigma} \xi_{\mathbf{k}} \langle 0 | \prod_{\mathbf{q}} (u_{\mathbf{q}}^* + v_{\mathbf{q}}^* c_{-\mathbf{q},\downarrow} c_{\mathbf{q}\uparrow}) c_{\mathbf{k}\sigma}^\dagger c_{\mathbf{k}\sigma} \prod_{\mathbf{q}'} (u_{\mathbf{q}'} + v_{\mathbf{q}'} c_{\mathbf{q}'\uparrow}^\dagger c_{-\mathbf{q}',\downarrow}^\dagger) | 0 \rangle \\
&\quad + \frac{1}{N} \sum_{\mathbf{k}\mathbf{k}'} V_{\mathbf{k}\mathbf{k}'} \langle 0 | \prod_{\mathbf{q}} (u_{\mathbf{q}}^* + v_{\mathbf{q}}^* c_{-\mathbf{q},\downarrow} c_{\mathbf{q}\uparrow}) c_{\mathbf{k}\uparrow}^\dagger c_{-\mathbf{k},\downarrow}^\dagger c_{-\mathbf{k}'\downarrow} c_{\mathbf{k}'\uparrow} \prod_{\mathbf{q}'} (u_{\mathbf{q}'} + v_{\mathbf{q}'} c_{\mathbf{q}'\uparrow}^\dagger c_{-\mathbf{q}',\downarrow}^\dagger) | 0 \rangle \\
&= \sum_{\mathbf{k}} \xi_{\mathbf{k}} \langle 0 | |v_{\mathbf{k}}|^2 c_{-\mathbf{k},\downarrow} c_{\mathbf{k}\uparrow} c_{\mathbf{k}\uparrow}^\dagger c_{-\mathbf{k},\downarrow}^\dagger | 0 \rangle + \sum_{\mathbf{k}} \xi_{\mathbf{k}} \langle 0 | |v_{-\mathbf{k}}|^2 c_{\mathbf{k}\downarrow} c_{-\mathbf{k},\uparrow} c_{\mathbf{k}\downarrow}^\dagger c_{-\mathbf{k},\uparrow}^\dagger | 0 \rangle \\
&\quad + \frac{1}{N} \sum_{\mathbf{k}\mathbf{k}'} V_{\mathbf{k}\mathbf{k}'} \langle 0 | v_{\mathbf{k}}^* u_{\mathbf{k}'}^* u_{\mathbf{k}} v_{\mathbf{k}'} c_{-\mathbf{k},\downarrow} c_{\mathbf{k}\uparrow}^\dagger c_{-\mathbf{k},\downarrow}^\dagger c_{-\mathbf{k}',\downarrow} c_{\mathbf{k}'\uparrow}^\dagger c_{-\mathbf{k}',\downarrow}^\dagger | 0 \rangle \\
&= \sum_{\mathbf{k}} 2\xi_{\mathbf{k}} |v_{\mathbf{k}}|^2 + \frac{1}{N} \sum_{\mathbf{k}\mathbf{k}'} V_{\mathbf{k}\mathbf{k}'} v_{\mathbf{k}}^* u_{\mathbf{k}} u_{\mathbf{k}'}^* v_{\mathbf{k}'} =: E_{\text{BCS}}. \tag{9.20}
\end{aligned}$$

This energy should be minimized with respect to the $u_{\mathbf{k}}, v_{\mathbf{k}}$. For E_{BCS} to be real, the phases of $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ must be the same. But since E_{BCS} is invariant under

$$u_{\mathbf{k}} \rightarrow u_{\mathbf{k}} e^{i\varphi_{\mathbf{k}}}, \quad v_{\mathbf{k}} \rightarrow v_{\mathbf{k}} e^{i\varphi_{\mathbf{k}}}, \tag{9.21}$$

we can choose all $u_{\mathbf{k}}, v_{\mathbf{k}}$ real. The constraint from normalization then reads $u_{\mathbf{k}}^2 + v_{\mathbf{k}}^2 = 1$ and we can parametrize the coefficients by

$$u_{\mathbf{k}} = \cos \theta_{\mathbf{k}}, \quad v_{\mathbf{k}} = \sin \theta_{\mathbf{k}}. \tag{9.22}$$

Then

$$\begin{aligned}
E_{\text{BCS}} &= \sum_{\mathbf{k}} 2\xi_{\mathbf{k}} \sin^2 \theta_{\mathbf{k}} + \frac{1}{N} \sum_{\mathbf{k}\mathbf{k}'} V_{\mathbf{k}\mathbf{k}'} \sin \theta_{\mathbf{k}} \cos \theta_{\mathbf{k}} \sin \theta_{\mathbf{k}'} \cos \theta_{\mathbf{k}'} \\
&= \sum_{\mathbf{k}} \xi_{\mathbf{k}} (1 - \cos 2\theta_{\mathbf{k}}) + \frac{1}{N} \sum_{\mathbf{k}\mathbf{k}'} \frac{V_{\mathbf{k}\mathbf{k}'}}{4} \sin 2\theta_{\mathbf{k}} \sin 2\theta_{\mathbf{k}'}. \tag{9.23}
\end{aligned}$$

We obtain the minimum from

$$\begin{aligned}
\frac{\partial E_{\text{BCS}}}{\partial \theta_{\mathbf{q}}} &= 2\xi_{\mathbf{q}} \sin 2\theta_{\mathbf{q}} + \frac{1}{N} \sum_{\mathbf{k}'} \frac{V_{\mathbf{q}\mathbf{k}'}}{2} \cos 2\theta_{\mathbf{q}} \sin 2\theta_{\mathbf{k}'} + \frac{1}{N} \sum_{\mathbf{k}} \frac{V_{\mathbf{k}\mathbf{q}}}{2} \sin 2\theta_{\mathbf{k}} \cos 2\theta_{\mathbf{q}} \\
&= 2\xi_{\mathbf{q}} \sin 2\theta_{\mathbf{q}} + \frac{1}{N} \sum_{\mathbf{k}'} V_{\mathbf{q}\mathbf{k}'} \cos 2\theta_{\mathbf{q}} \sin 2\theta_{\mathbf{k}'} \stackrel{!}{=} 0. \tag{9.24}
\end{aligned}$$

We replace \mathbf{q} by \mathbf{k} and parametrize $\theta_{\mathbf{k}}$ by

$$\sin 2\theta_{\mathbf{k}} =: \frac{\Delta_{\mathbf{k}}}{\sqrt{\xi_{\mathbf{k}}^2 + \Delta_{\mathbf{k}}^2}} \tag{9.25}$$

and write

$$\cos 2\theta_{\mathbf{k}} = \frac{\xi_{\mathbf{k}}}{\sqrt{\xi_{\mathbf{k}}^2 + \Delta_{\mathbf{k}}^2}}. \tag{9.26}$$

The last equality is only determined by the previous one up to the sign. We could convince ourselves that the other possible choice does not lead to a lower E_{BCS} . Equation (9.24) now becomes

$$2 \frac{\xi_{\mathbf{k}} \Delta_{\mathbf{k}}}{\sqrt{\xi_{\mathbf{k}}^2 + \Delta_{\mathbf{k}}^2}} + \frac{1}{N} \sum_{\mathbf{k}'} V_{\mathbf{k}\mathbf{k}'} \frac{\xi_{\mathbf{k}} \Delta_{\mathbf{k}'}}{\sqrt{\xi_{\mathbf{k}}^2 + \Delta_{\mathbf{k}}^2} \sqrt{\xi_{\mathbf{k}'}^2 + \Delta_{\mathbf{k}'}^2}} = 0 \tag{9.27}$$

$$\Rightarrow \Delta_{\mathbf{k}} = -\frac{1}{N} \sum_{\mathbf{k}'} V_{\mathbf{k}\mathbf{k}'} \frac{\Delta_{\mathbf{k}'}}{2\sqrt{\xi_{\mathbf{k}'}^2 + \Delta_{\mathbf{k}'}^2}}. \tag{9.28}$$

This is called the *BCS gap equation* for reasons to be discussed below. When we have solved it, it is easy to obtain the original variational parameters in terms of $\Delta_{\mathbf{k}}$,

$$u_{\mathbf{k}}^2 = \frac{1}{2} \left(1 + \frac{\xi_{\mathbf{k}}}{\sqrt{\xi_{\mathbf{k}}^2 + \Delta_{\mathbf{k}}^2}} \right), \quad (9.29)$$

$$v_{\mathbf{k}}^2 = \frac{1}{2} \left(1 - \frac{\xi_{\mathbf{k}}}{\sqrt{\xi_{\mathbf{k}}^2 + \Delta_{\mathbf{k}}^2}} \right), \quad (9.30)$$

$$u_{\mathbf{k}} v_{\mathbf{k}} = \frac{\Delta_{\mathbf{k}}}{2\sqrt{\xi_{\mathbf{k}}^2 + \Delta_{\mathbf{k}}^2}}. \quad (9.31)$$

The *relative* sign of $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ is thus the sign of $\Delta_{\mathbf{k}}$. The *absolute* sign of, say, $u_{\mathbf{k}}$ is irrelevant because of the invariance of E_{BCS} under simultaneous phase rotations of $u_{\mathbf{k}}, v_{\mathbf{k}}$ (consider a phase factor of $e^{i\pi} = -1$).

For our special interaction

$$V_{\mathbf{k}\mathbf{k}'} = \begin{cases} -V_0 & \text{for } |\xi_{\mathbf{k}}|, |\xi_{\mathbf{k}'}| < \omega_D, \\ 0 & \text{otherwise,} \end{cases} \quad (9.32)$$

the BCS gap equation becomes

$$\Delta_{\mathbf{k}} = \begin{cases} -\frac{1}{N} \sum_{\mathbf{k}', |\xi_{\mathbf{k}'}| < \omega_D} (-V_0) \frac{\Delta_{\mathbf{k}'}}{2\sqrt{\xi_{\mathbf{k}'}^2 + \Delta_{\mathbf{k}'}^2}} & \text{for } |\xi_{\mathbf{k}}| < \omega_D, \\ 0 & \text{otherwise,} \end{cases} \quad (9.33)$$

which can be solved with the ansatz

$$\Delta_{\mathbf{k}} = \begin{cases} \Delta_0 > 0 & \text{for } |\xi_{\mathbf{k}}| < \omega_D, \\ 0 & \text{otherwise.} \end{cases} \quad (9.34)$$

We obtain

$$\Delta_0 = \frac{V_0}{N} \sum_{\mathbf{k}', |\xi_{\mathbf{k}'}| < \omega_D} \frac{\Delta_0}{2\sqrt{\xi_{\mathbf{k}'}^2 + \Delta_0^2}} \quad (9.35)$$

$$\Rightarrow 1 = \frac{V_0}{N} \sum_{\mathbf{k}', |\xi_{\mathbf{k}'}| < \omega_D} \frac{1}{2\sqrt{\xi_{\mathbf{k}'}^2 + \Delta_0^2}} = V_0 \int_{-\omega_D}^{\omega_D} d\xi D(\mu + \xi) \frac{1}{2\sqrt{\xi^2 + \Delta_0^2}}, \quad (9.36)$$

where $D(\epsilon)$ is the density of states per spin direction and per unit cell. If the density of states is approximately constant within $\pm\omega_D$ of the Fermi energy, we obtain

$$1 = V_0 D(E_F) \frac{1}{2} \int_{-\omega_D}^{\omega_D} \frac{d\xi}{\sqrt{\xi^2 + \Delta_0^2}} = V_0 D(E_F) \text{Arsinh} \frac{\omega_D}{\Delta_0} \quad (9.37)$$

$$\Rightarrow \sinh \frac{1}{V_0 D(E_F)} = \frac{\omega_D}{\Delta_0} \quad (9.38)$$

$$\Rightarrow \Delta_0 = \omega_D \frac{1}{\sinh \frac{1}{V_0 D(E_F)}}. \quad (9.39)$$

In the so-called *weak-coupling limit* of small $V_0 D(E_F)$, this result simplifies to

$$\Delta_0 \cong 2\omega_D \exp\left(-\frac{1}{V_0 D(E_F)}\right). \quad (9.40)$$

Interestingly, apart from a numerical factor, the value of Δ_0 agrees with $k_B T_c$ for the Cooper instability. We will return to this observation below. The non-analyticity of the function $\Delta_0(V_0)$ means that we cannot obtain Δ_0 and thus $|\psi_{\text{BCS}}\rangle$ within perturbation theory in V_0 .

We can now find the energy gain due to the superconducting state, i.e., the condensation energy. For this, we insert $u_{\mathbf{k}}, v_{\mathbf{k}}$ into E_{BCS} ,

$$E_{\text{BCS}} = \sum_{\mathbf{k}} \cancel{2} \xi_{\mathbf{k}} \frac{1}{\cancel{2}} \left(1 - \frac{\xi_{\mathbf{k}}}{\sqrt{\xi_{\mathbf{k}}^2 + \Delta_{\mathbf{k}}^2}} \right) + \frac{1}{N} \sum_{\mathbf{k}\mathbf{k}'} V_{\mathbf{k}\mathbf{k}'} \frac{\Delta_{\mathbf{k}} \Delta_{\mathbf{k}'}}{4\sqrt{\xi_{\mathbf{k}}^2 + \Delta_{\mathbf{k}}^2} \sqrt{\xi_{\mathbf{k}'}^2 + \Delta_{\mathbf{k}'}^2}}. \quad (9.41)$$

We use the simple form of $V_{\mathbf{k}\mathbf{k}'}$ and assume that $D(\epsilon)$ is constant within $\pm\omega_D$ of the Fermi energy but not outside of this interval. This gives

$$\begin{aligned} E_{\text{BCS}} &= N \int_{-\infty}^{-\omega_D} d\xi D(\mu + \xi) 2\xi + N \int_{-\omega_D}^{\omega_D} d\xi D(\mu + \xi) \xi \left(1 - \frac{\xi}{\sqrt{\xi^2 + \Delta_0^2}} \right) \\ &\quad + N \int_{\omega_D}^{\infty} d\xi D(\mu + \xi) 0 + N \int_{-\omega_D}^{\omega_D} d\xi \int_{-\omega_D}^{\omega_D} d\xi' D(\mu + \xi) D(\mu + \xi') (-V_0) \frac{\Delta_0^2}{4\sqrt{\xi^2 + \Delta_0^2} \sqrt{(\xi')^2 + \Delta_0^2}} \\ &\cong 2N \int_{-\infty}^{-\omega_D} d\xi D(\mu + \xi) \xi + N D(E_F) \left(-\omega_D \sqrt{\omega_D^2 + \Delta_0^2} + \Delta_0^2 \text{Arsinh} \frac{\omega_D}{\Delta_0} \right) \\ &\quad - N V_0 D^2(E_F) \Delta_0^2 \text{Arsinh}^2 \frac{\omega_D}{\Delta_0}. \end{aligned} \quad (9.42)$$

With the gap equation

$$1 = V_0 D(E_F) \text{Arsinh} \frac{\omega_D}{\Delta_0} \quad (9.43)$$

this simplifies to

$$E_{\text{BCS}} \cong 2N \int_{-\infty}^{-\omega_D} d\xi D(\mu + \xi) \xi - N D(E_F) \omega_D \sqrt{\omega_D^2 + \Delta_0^2}. \quad (9.44)$$

The normal-state energy should be recovered by taking $\Delta_0 \rightarrow 0$. The energy difference is

$$\Delta E_{\text{BCS}} := E_{\text{BCS}} - E_{\text{BCS}}|_{\Delta_0 \rightarrow 0} = -N D(E_F) \omega_D \sqrt{\omega_D^2 + \Delta_0^2} + N D(E_F) \omega_D^2. \quad (9.45)$$

Since for weak coupling we have $\Delta_0 \ll \omega_D$, we can expand this in Δ_0/ω_D ,

$$\Delta E_{\text{BCS}} = -N D(E_F) \omega_D^2 \sqrt{1 + \left(\frac{\Delta_0}{\omega_D} \right)^2} + N D(E_F) \omega_D^2 \cong -\frac{1}{2} N D(E_F) \Delta_0^2. \quad (9.46)$$

The condensation-energy density is thus (counted positively)

$$e_{\text{BCS}} \cong \frac{1}{2} \frac{N}{V} D(E_F) \Delta_0^2. \quad (9.47)$$

Type-I superconductivity is destroyed if e_{BCS} equals the energy density required for magnetic-field expulsion. At $H = H_c$, this energy is $H_c^2/8\pi$, as we have seen above. We thus conclude that

$$H_c \cong \sqrt{4\pi \frac{N}{V} D(E_F) \Delta_0}. \quad (9.48)$$

This prediction of BCS theory is in reasonably good agreement with experiments for simple superconductors.

10

BCS theory

The variational ansatz of Sec. 9.2 has given us an approximation for the many-particle ground state $|\psi_{\text{BCS}}\rangle$. While this is interesting, it does not yet allow predictions of thermodynamic properties, such as the critical temperature. We will now consider superconductors at non-zero temperatures within mean-field theory, which will also provide a new perspective on the BCS gap equation and on the meaning of $\Delta_{\mathbf{k}}$.

10.1 BCS mean-field theory

We start again from the Hamiltonian

$$H = \sum_{\mathbf{k}\sigma} \xi_{\mathbf{k}} c_{\mathbf{k}\sigma}^\dagger c_{\mathbf{k}\sigma} + \frac{1}{N} \sum_{\mathbf{k}\mathbf{k}'} V_{\mathbf{k}\mathbf{k}'} c_{\mathbf{k}\uparrow}^\dagger c_{-\mathbf{k},\downarrow}^\dagger c_{-\mathbf{k}',\downarrow} c_{\mathbf{k}'\uparrow}. \quad (10.1)$$

A *mean-field approximation* consists of replacing products of operators A, B according to

$$AB \cong \langle A \rangle B + A \langle B \rangle - \langle A \rangle \langle B \rangle. \quad (10.2)$$

Note that the error introduced by this replacement is

$$AB - \langle A \rangle B - A \langle B \rangle + \langle A \rangle \langle B \rangle = (A - \langle A \rangle)(B - \langle B \rangle), \quad (10.3)$$

i.e., it is of second order in the deviations of A and B from their averages. A well-known mean-field approximation is the Hartree or Stoner approximation, which for our Hamiltonian amounts to the choice $A = c_{\mathbf{k}\uparrow}^\dagger c_{\mathbf{k}'\uparrow}$, $B = c_{-\mathbf{k},\downarrow}^\dagger c_{-\mathbf{k}',\downarrow}$. However, Bardeen, Cooper, and Schrieffer realized that superconductivity can be understood with the help of a different choice, namely $A = c_{\mathbf{k}\uparrow}^\dagger c_{-\mathbf{k},\downarrow}^\dagger$, $B = c_{-\mathbf{k}',\downarrow} c_{\mathbf{k}'\uparrow}$. This leads to the mean-field BCS Hamiltonian

$$H_{\text{BCS}} = \sum_{\mathbf{k}\sigma} \xi_{\mathbf{k}} c_{\mathbf{k}\sigma}^\dagger c_{\mathbf{k}\sigma} + \frac{1}{N} \sum_{\mathbf{k}\mathbf{k}'} V_{\mathbf{k}\mathbf{k}'} \left(\langle c_{\mathbf{k}\uparrow}^\dagger c_{-\mathbf{k},\downarrow}^\dagger \rangle c_{-\mathbf{k}',\downarrow} c_{\mathbf{k}'\uparrow} + c_{\mathbf{k}\uparrow}^\dagger c_{-\mathbf{k},\downarrow}^\dagger \langle c_{-\mathbf{k}',\downarrow} c_{\mathbf{k}'\uparrow} \rangle - \langle c_{\mathbf{k}\uparrow}^\dagger c_{-\mathbf{k},\downarrow}^\dagger \rangle \langle c_{-\mathbf{k}',\downarrow} c_{\mathbf{k}'\uparrow} \rangle \right). \quad (10.4)$$

We define

$$\Delta_{\mathbf{k}} := -\frac{1}{N} \sum_{\mathbf{k}'} V_{\mathbf{k}\mathbf{k}'} \langle c_{-\mathbf{k}',\downarrow} c_{\mathbf{k}'\uparrow} \rangle \quad (10.5)$$

so that

$$\Delta_{\mathbf{k}}^* = -\frac{1}{N} \sum_{\mathbf{k}'} V_{\mathbf{k}\mathbf{k}'} \langle c_{\mathbf{k}'\uparrow}^\dagger c_{-\mathbf{k}',\downarrow}^\dagger \rangle. \quad (10.6)$$

At this point it is not obvious that the quantity $\Delta_{\mathbf{k}}$ is the same as the one introduced in Sec. 9.2 for the special case of the ground state. Since this will turn out to be the case, we nevertheless use the same symbol from the start. We can now write

$$H_{\text{BCS}} = \sum_{\mathbf{k}\sigma} \xi_{\mathbf{k}} c_{\mathbf{k}\sigma}^\dagger c_{\mathbf{k}\sigma} - \sum_{\mathbf{k}} \Delta_{\mathbf{k}}^* c_{-\mathbf{k},\downarrow} c_{\mathbf{k}\uparrow} - \sum_{\mathbf{k}} \Delta_{\mathbf{k}} c_{\mathbf{k}\uparrow}^\dagger c_{-\mathbf{k},\downarrow}^\dagger + \text{const.} \quad (10.7)$$

The constant is irrelevant for the following derivation and is omitted from now on. Since H_{BCS} is bilinear in c, c^\dagger it describes a non-interacting effective system. But what is unusual is that H_{BCS} contains terms of the form cc and $c^\dagger c^\dagger$, which do not conserve the electron number. We thus expect that the eigenstates of H_{BCS} do not have a sharp electron number. We had already seen that the BCS ground state has this property. This is a bit strange since superpositions of states with different electron numbers are never observed. One can formulate the theory of superconductivity in terms of states with fixed electron number, but this formulation is cumbersome and we will not pursue it here.

To diagonalize H_{BCS} , we introduce new fermionic operators, which are linear combinations of electron creation and annihilation operators,

$$\begin{pmatrix} \gamma_{\mathbf{k}\uparrow} \\ \gamma_{-\mathbf{k},\downarrow}^\dagger \end{pmatrix} = \begin{pmatrix} u_{\mathbf{k}}^* & -v_{\mathbf{k}} \\ v_{\mathbf{k}}^* & u_{\mathbf{k}} \end{pmatrix} \begin{pmatrix} c_{\mathbf{k}\uparrow} \\ c_{-\mathbf{k},\downarrow}^\dagger \end{pmatrix}. \quad (10.8)$$

This mapping is called *Bogoliubov* (or *Bogoliubov-Valatin*) *transformation*. Again, it is not clear yet that $u_{\mathbf{k}}, v_{\mathbf{k}}$ are related to the previously introduced quantities denoted by the same symbols. For the γ to satisfy fermionic anticommutation relations, we require

$$\begin{aligned} \{\gamma_{\mathbf{k}\uparrow}, \gamma_{\mathbf{k}\uparrow}^\dagger\} &= \gamma_{\mathbf{k}\uparrow} \gamma_{\mathbf{k}\uparrow}^\dagger + \gamma_{\mathbf{k}\uparrow}^\dagger \gamma_{\mathbf{k}\uparrow} \\ &= u_{\mathbf{k}}^* u_{\mathbf{k}} c_{\mathbf{k}\uparrow} c_{\mathbf{k}\uparrow}^\dagger - u_{\mathbf{k}}^* v_{\mathbf{k}}^* c_{\mathbf{k}\uparrow} c_{-\mathbf{k},\downarrow} - v_{\mathbf{k}} u_{\mathbf{k}} c_{-\mathbf{k},\downarrow}^\dagger c_{\mathbf{k}\uparrow}^\dagger + v_{\mathbf{k}} v_{\mathbf{k}}^* c_{-\mathbf{k},\downarrow}^\dagger c_{-\mathbf{k},\downarrow} \\ &\quad + u_{\mathbf{k}} u_{\mathbf{k}}^* c_{\mathbf{k}\uparrow}^\dagger c_{\mathbf{k}\uparrow} - u_{\mathbf{k}} v_{\mathbf{k}} c_{\mathbf{k}\uparrow}^\dagger c_{-\mathbf{k},\downarrow}^\dagger - v_{\mathbf{k}}^* u_{\mathbf{k}}^* c_{-\mathbf{k},\downarrow} c_{\mathbf{k}\uparrow} + v_{\mathbf{k}}^* v_{\mathbf{k}} c_{-\mathbf{k},\downarrow} c_{-\mathbf{k},\downarrow}^\dagger \\ &= |u_{\mathbf{k}}|^2 \underbrace{\{c_{\mathbf{k}\uparrow}, c_{\mathbf{k}\uparrow}^\dagger\}}_{=1} - u_{\mathbf{k}}^* v_{\mathbf{k}}^* \underbrace{\{c_{\mathbf{k}\uparrow}, c_{-\mathbf{k},\downarrow}\}}_{=0} - v_{\mathbf{k}} u_{\mathbf{k}} \underbrace{\{c_{-\mathbf{k},\downarrow}^\dagger, c_{\mathbf{k}\uparrow}^\dagger\}}_{=0} + |v_{\mathbf{k}}|^2 \underbrace{\{c_{-\mathbf{k},\downarrow}^\dagger, c_{-\mathbf{k},\downarrow}\}}_{=1} \\ &= |u_{\mathbf{k}}|^2 + |v_{\mathbf{k}}|^2 \stackrel{!}{=} 1. \end{aligned} \quad (10.9)$$

Using this constraint, we find the inverse transformation,

$$\begin{pmatrix} c_{\mathbf{k}\uparrow} \\ c_{-\mathbf{k},\downarrow}^\dagger \end{pmatrix} = \begin{pmatrix} u_{\mathbf{k}} & v_{\mathbf{k}} \\ -v_{\mathbf{k}}^* & u_{\mathbf{k}}^* \end{pmatrix} \begin{pmatrix} \gamma_{\mathbf{k}\uparrow} \\ \gamma_{-\mathbf{k},\downarrow}^\dagger \end{pmatrix}. \quad (10.10)$$

Insertion into H_{BCS} yields

$$\begin{aligned} H_{\text{BCS}} &= \sum_{\mathbf{k}} \left\{ \xi_{\mathbf{k}} \left(u_{\mathbf{k}}^* \gamma_{\mathbf{k}\uparrow}^\dagger + v_{\mathbf{k}}^* \gamma_{-\mathbf{k},\downarrow} \right) \left(u_{\mathbf{k}} \gamma_{\mathbf{k}\uparrow} + v_{\mathbf{k}} \gamma_{-\mathbf{k},\downarrow}^\dagger \right) + \xi_{\mathbf{k}} \left(-v_{\mathbf{k}}^* \gamma_{\mathbf{k}\uparrow} + u_{\mathbf{k}}^* \gamma_{-\mathbf{k},\downarrow}^\dagger \right) \left(-v_{\mathbf{k}} \gamma_{\mathbf{k}\uparrow}^\dagger + u_{\mathbf{k}} \gamma_{-\mathbf{k},\downarrow} \right) \right. \\ &\quad \left. - \Delta_{\mathbf{k}}^* \left(-v_{\mathbf{k}} \gamma_{\mathbf{k}\uparrow}^\dagger + u_{\mathbf{k}} \gamma_{-\mathbf{k},\downarrow} \right) \left(u_{\mathbf{k}} \gamma_{\mathbf{k}\uparrow} + v_{\mathbf{k}} \gamma_{-\mathbf{k},\downarrow}^\dagger \right) - \Delta_{\mathbf{k}} \left(u_{\mathbf{k}}^* \gamma_{\mathbf{k}\uparrow}^\dagger + v_{\mathbf{k}}^* \gamma_{-\mathbf{k},\downarrow} \right) \left(-v_{\mathbf{k}}^* \gamma_{\mathbf{k}\uparrow} + u_{\mathbf{k}}^* \gamma_{-\mathbf{k},\downarrow}^\dagger \right) \right\} \\ &= \sum_{\mathbf{k}} \left\{ \left(\xi_{\mathbf{k}} |u_{\mathbf{k}}|^2 - \xi_{\mathbf{k}} |v_{\mathbf{k}}|^2 + \Delta_{\mathbf{k}}^* v_{\mathbf{k}} u_{\mathbf{k}} + \Delta_{\mathbf{k}} u_{\mathbf{k}}^* v_{\mathbf{k}} \right) \gamma_{\mathbf{k}\uparrow}^\dagger \gamma_{\mathbf{k}\uparrow} \right. \\ &\quad + \left(-\xi_{\mathbf{k}} |v_{\mathbf{k}}|^2 + \xi_{\mathbf{k}} |u_{\mathbf{k}}|^2 + \Delta_{\mathbf{k}}^* u_{\mathbf{k}} v_{\mathbf{k}} + \Delta_{\mathbf{k}} v_{\mathbf{k}}^* u_{\mathbf{k}}^* \right) \gamma_{-\mathbf{k},\downarrow}^\dagger \gamma_{-\mathbf{k},\downarrow} \\ &\quad + \left(\xi_{\mathbf{k}} u_{\mathbf{k}}^* v_{\mathbf{k}} + \xi_{\mathbf{k}} u_{\mathbf{k}}^* v_{\mathbf{k}} + \Delta_{\mathbf{k}}^* v_{\mathbf{k}}^2 - \Delta_{\mathbf{k}} (u_{\mathbf{k}}^*)^2 \right) \gamma_{\mathbf{k}\uparrow}^\dagger \gamma_{-\mathbf{k},\downarrow}^\dagger \\ &\quad \left. + \left(\xi_{\mathbf{k}} v_{\mathbf{k}}^* u_{\mathbf{k}} + \xi_{\mathbf{k}} v_{\mathbf{k}}^* u_{\mathbf{k}} - \Delta_{\mathbf{k}}^* u_{\mathbf{k}}^2 + \Delta_{\mathbf{k}} (v_{\mathbf{k}})^2 \right) \gamma_{-\mathbf{k},\downarrow} \gamma_{\mathbf{k}\uparrow} \right\} + \text{const.} \end{aligned} \quad (10.11)$$

The coefficients $u_{\mathbf{k}}, v_{\mathbf{k}}$ should now be chosen such that the $\gamma\gamma$ and $\gamma^\dagger\gamma^\dagger$ terms vanish. This requires

$$2\xi_{\mathbf{k}} u_{\mathbf{k}}^* v_{\mathbf{k}} + \Delta_{\mathbf{k}}^* v_{\mathbf{k}}^2 - \Delta_{\mathbf{k}} (u_{\mathbf{k}}^*)^2 = 0. \quad (10.12)$$

Writing

$$\Delta_{\mathbf{k}} = |\Delta_{\mathbf{k}}| e^{i\phi_{\mathbf{k}}}, \quad (10.13)$$

$$u_{\mathbf{k}} = |u_{\mathbf{k}}| e^{i\alpha_{\mathbf{k}}}, \quad (10.14)$$

$$v_{\mathbf{k}} = |v_{\mathbf{k}}| e^{i\beta_{\mathbf{k}}}, \quad (10.15)$$

we obtain

$$2\xi_{\mathbf{k}} |u_{\mathbf{k}}| |v_{\mathbf{k}}| e^{i(\beta_{\mathbf{k}} - \alpha_{\mathbf{k}})} + |\Delta_{\mathbf{k}}| \left(|v_{\mathbf{k}}|^2 e^{i(2\beta_{\mathbf{k}} - \phi_{\mathbf{k}})} - |u_{\mathbf{k}}|^2 e^{i(\phi_{\mathbf{k}} - 2\alpha_{\mathbf{k}})} \right) = 0. \quad (10.16)$$

A special solution of this equation (we do not require the general solution) is given by

$$\alpha_{\mathbf{k}} = 0, \quad (10.17)$$

$$\beta_{\mathbf{k}} = \phi_{\mathbf{k}}, \quad (10.18)$$

$$2\xi_{\mathbf{k}} |u_{\mathbf{k}}| |v_{\mathbf{k}}| + |\Delta_{\mathbf{k}}| \left(|v_{\mathbf{k}}|^2 - |u_{\mathbf{k}}|^2 \right) = 0. \quad (10.19)$$

From the last equation we obtain

$$4\xi_{\mathbf{k}}^2 |u_{\mathbf{k}}|^2 |v_{\mathbf{k}}|^2 = |\Delta_{\mathbf{k}}|^2 \left(|v_{\mathbf{k}}|^4 - 2|v_{\mathbf{k}}|^2 |u_{\mathbf{k}}|^2 + |u_{\mathbf{k}}|^4 \right) \quad (10.20)$$

$$\Rightarrow 4 \left(\xi_{\mathbf{k}}^2 + |\Delta_{\mathbf{k}}|^2 \right) |u_{\mathbf{k}}|^2 |v_{\mathbf{k}}|^2 = |\Delta_{\mathbf{k}}|^2 \left(|v_{\mathbf{k}}|^4 + 2|v_{\mathbf{k}}|^2 |u_{\mathbf{k}}|^2 + |u_{\mathbf{k}}|^4 \right) = |\Delta_{\mathbf{k}}|^2 \left(|v_{\mathbf{k}}|^2 + |u_{\mathbf{k}}|^2 \right)^2 = |\Delta_{\mathbf{k}}|^2 \quad (10.21)$$

$$\Rightarrow |u_{\mathbf{k}}| |v_{\mathbf{k}}| = \frac{|\Delta_{\mathbf{k}}|}{2\sqrt{\xi_{\mathbf{k}}^2 + |\Delta_{\mathbf{k}}|^2}} \quad (10.22)$$

so that

$$|u_{\mathbf{k}}|^2 - |v_{\mathbf{k}}|^2 = \frac{2\xi_{\mathbf{k}} |u_{\mathbf{k}}| |v_{\mathbf{k}}|}{|\Delta_{\mathbf{k}}|} = \frac{\xi_{\mathbf{k}}}{\sqrt{\xi_{\mathbf{k}}^2 + |\Delta_{\mathbf{k}}|^2}}. \quad (10.23)$$

Together with $|u_{\mathbf{k}}|^2 + |v_{\mathbf{k}}|^2 = 1$ we thus find

$$|u_{\mathbf{k}}|^2 = \frac{1}{2} \left(1 + \frac{\xi_{\mathbf{k}}}{\sqrt{\xi_{\mathbf{k}}^2 + |\Delta_{\mathbf{k}}|^2}} \right), \quad (10.24)$$

$$|v_{\mathbf{k}}|^2 = \frac{1}{2} \left(1 - \frac{\xi_{\mathbf{k}}}{\sqrt{\xi_{\mathbf{k}}^2 + |\Delta_{\mathbf{k}}|^2}} \right). \quad (10.25)$$

Restoring the phases in Eq. (10.22), we also conclude that

$$u_{\mathbf{k}} v_{\mathbf{k}} = \frac{\Delta_{\mathbf{k}}}{2\sqrt{\xi_{\mathbf{k}}^2 + |\Delta_{\mathbf{k}}|^2}}. \quad (10.26)$$

The BCS Hamiltonian now reads, ignoring a constant,

$$\begin{aligned} H_{\text{BCS}} &= \sum_{\mathbf{k}} \left(\frac{\xi_{\mathbf{k}}^2}{\sqrt{\xi_{\mathbf{k}}^2 + |\Delta_{\mathbf{k}}|^2}} + \frac{|\Delta_{\mathbf{k}}|^2}{\sqrt{\xi_{\mathbf{k}}^2 + |\Delta_{\mathbf{k}}|^2}} \right) \left(\gamma_{\mathbf{k}\uparrow}^\dagger \gamma_{\mathbf{k}\uparrow} + \gamma_{-\mathbf{k},\downarrow}^\dagger \gamma_{-\mathbf{k},\downarrow} \right) \\ &= \sum_{\mathbf{k}} \sqrt{\xi_{\mathbf{k}}^2 + |\Delta_{\mathbf{k}}|^2} \left(\gamma_{\mathbf{k}\uparrow}^\dagger \gamma_{\mathbf{k}\uparrow} + \gamma_{-\mathbf{k},\downarrow}^\dagger \gamma_{-\mathbf{k},\downarrow} \right). \end{aligned} \quad (10.27)$$

Using $\xi_{-\mathbf{k}} = \xi_{\mathbf{k}}$ and the plausible assumption $|\Delta_{-\mathbf{k}}| = |\Delta_{\mathbf{k}}|$, we obtain the simple form

$$H_{\text{BCS}} = \sum_{\mathbf{k}\sigma} E_{\mathbf{k}} \gamma_{\mathbf{k}\sigma}^\dagger \gamma_{\mathbf{k}\sigma} \quad (10.28)$$

with the dispersion

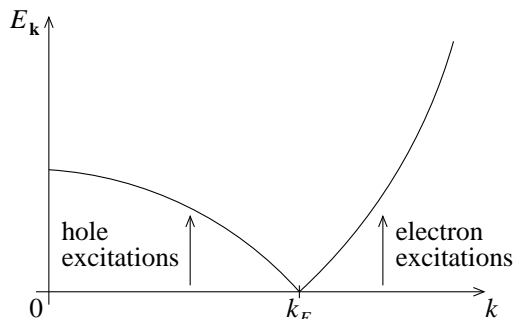
$$E_{\mathbf{k}} := \sqrt{\xi_{\mathbf{k}}^2 + |\Delta_{\mathbf{k}}|^2}. \quad (10.29)$$

It is instructive to first consider the normal state, for which $\Delta_{\mathbf{k}} \rightarrow 0$. Then

$$|u_{\mathbf{k}}|^2 = \frac{1}{2} \left(1 + \frac{\xi_{\mathbf{k}}}{|\xi_{\mathbf{k}}|} \right) = \begin{cases} 0 & \text{for } \xi_{\mathbf{k}} < 0, \\ 1 & \text{for } \xi_{\mathbf{k}} > 0, \end{cases} \quad (10.30)$$

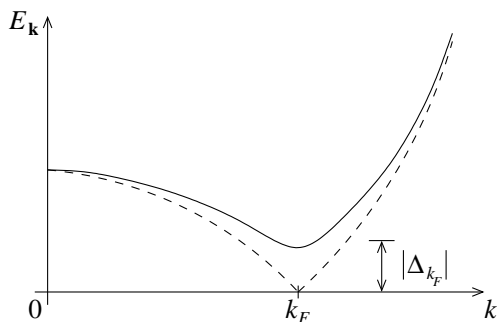
$$|v_{\mathbf{k}}|^2 = \frac{1}{2} \left(1 - \frac{\xi_{\mathbf{k}}}{|\xi_{\mathbf{k}}|} \right) = \begin{cases} 1 & \text{for } \xi_{\mathbf{k}} < 0, \\ 0 & \text{for } \xi_{\mathbf{k}} > 0. \end{cases} \quad (10.31)$$

We see that the Bogoliubov quasiparticles described by γ, γ^\dagger , are *holes* for energies below the Fermi energy ($\xi_{\mathbf{k}} < 0$) and *electrons* for energies above ($\xi_{\mathbf{k}} > 0$). Their dispersion is $E_{\mathbf{k}} = |\xi_{\mathbf{k}}|$. For a parabolic normal dispersion $\xi_{\mathbf{k}}$:



The excitation energies $E_{\mathbf{k}}$ are always *positive* except at the Fermi surface—it costs energy to create a hole in the Fermi sea and also to insert an electron into an empty state outside of the Fermi sea.

Superconductivity changes the dispersion to $E_{\mathbf{k}} = \sqrt{\xi_{\mathbf{k}}^2 + |\Delta_{\mathbf{k}}|^2}$:

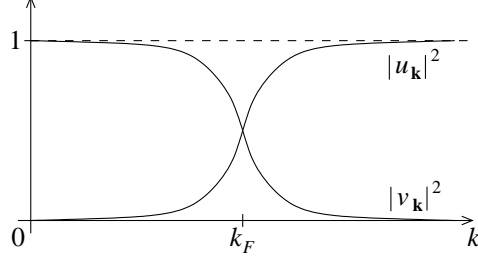


Superconductivity evidently opens an *energy gap* of magnitude $|\Delta_{k_F}|$ in the excitation spectrum.

We should recall that in deriving H_{BCS} we have ignored a constant, which we now reinsert,

$$H_{\text{BCS}} = E_{\text{BCS}} + \sum_{\mathbf{k}\sigma} E_{\mathbf{k}} \gamma_{\mathbf{k}\sigma}^\dagger \gamma_{\mathbf{k}\sigma}. \quad (10.32)$$

The energy of the system is E_{BCS} if no quasiparticles are present and is increased (by at least $|\Delta_{k_F}|$) if quasiparticles are excited. The state without any quasiparticles is the pure condensate. Since E_{BCS} depends on temperature through $\langle c_{-\mathbf{k},\downarrow} c_{\mathbf{k}\uparrow} \rangle$, the condensate is not generally the BCS ground state discussed previously. However, one can show that it agrees with the ground state in the limit $T \rightarrow 0$.



We also find $0 < |u_{\mathbf{k}}|^2 < 1$ and $0 < |v_{\mathbf{k}}|^2 < 1$, i.e., the Bogoliubov quasiparticles are superpositions of particles and holes. Deep inside the Fermi sea, the quasiparticles are mostly hole-like, while far above E_F they are mostly electron-like. But right at the Fermi surface we find, for example for spin $\sigma = \uparrow$,

$$\gamma_{\mathbf{k}\uparrow} = \frac{1}{\sqrt{2}} c_{\mathbf{k}\uparrow} - \frac{1}{\sqrt{2}} e^{i\phi_{\mathbf{k}}} c_{-\mathbf{k},\downarrow}^\dagger. \quad (10.33)$$

The quasiparticles here consist of electrons and holes with the same amplitude. This means that they are electrically neutral on average.

So far, we have not determined the gap function $\Delta_{\mathbf{k}}$. This can be done by inserting the Bogoliubov transformation into the definition

$$\Delta_{\mathbf{k}} = -\frac{1}{N} \sum_{\mathbf{k}'} V_{\mathbf{k}\mathbf{k}'} \langle c_{-\mathbf{k}',\downarrow} c_{\mathbf{k}'\uparrow} \rangle, \quad (10.34)$$

which yields

$$\begin{aligned} \Delta_{\mathbf{k}} &= -\frac{1}{N} \sum_{\mathbf{k}'} V_{\mathbf{k}\mathbf{k}'} \left\langle \left(-v_{\mathbf{k}'} \gamma_{\mathbf{k}'\uparrow}^\dagger + u_{\mathbf{k}'} \gamma_{-\mathbf{k}',\downarrow} \right) \left(u_{\mathbf{k}'} \gamma_{\mathbf{k}'\uparrow} + v_{\mathbf{k}'} \gamma_{-\mathbf{k}',\downarrow}^\dagger \right) \right\rangle \\ &= -\frac{1}{N} \sum_{\mathbf{k}'} V_{\mathbf{k}\mathbf{k}'} \left\{ -v_{\mathbf{k}'} u_{\mathbf{k}'} \langle \gamma_{\mathbf{k}'\uparrow}^\dagger \gamma_{\mathbf{k}'\uparrow} \rangle - v_{\mathbf{k}'}^2 \langle \gamma_{\mathbf{k}'\uparrow}^\dagger \gamma_{-\mathbf{k}',\downarrow}^\dagger \rangle + u_{\mathbf{k}'}^2 \langle \gamma_{-\mathbf{k}',\downarrow} \gamma_{\mathbf{k}'\uparrow} \rangle + u_{\mathbf{k}'} v_{\mathbf{k}'} \langle \gamma_{-\mathbf{k}',\downarrow} \gamma_{-\mathbf{k}',\downarrow}^\dagger \rangle \right\}. \end{aligned} \quad (10.35)$$

For selfconsistency, the averages have to be evaluated with the BCS Hamiltonian H_{BCS} . This gives

$$\langle \gamma_{\mathbf{k}'\uparrow}^\dagger \gamma_{\mathbf{k}'\uparrow} \rangle = n_F(E_{\mathbf{k}'}), \quad (10.36)$$

$$\langle \gamma_{\mathbf{k}'\uparrow}^\dagger \gamma_{-\mathbf{k}',\downarrow}^\dagger \rangle = 0, \quad (10.37)$$

$$\langle \gamma_{-\mathbf{k}',\downarrow} \gamma_{\mathbf{k}'\uparrow} \rangle = 0, \quad (10.38)$$

$$\langle \gamma_{-\mathbf{k}',\downarrow} \gamma_{-\mathbf{k}',\downarrow}^\dagger \rangle = 1 - n_F(E_{\mathbf{k}'}), \quad (10.39)$$

and we obtain the *BCS gap equation*, now at arbitrary temperature,

$$\begin{aligned} \Delta_{\mathbf{k}} &= -\frac{1}{N} \sum_{\mathbf{k}'} V_{\mathbf{k}\mathbf{k}'} u_{\mathbf{k}'} v_{\mathbf{k}'} [1 - 2n_F(E_{\mathbf{k}'})] = -\frac{1}{N} \sum_{\mathbf{k}'} V_{\mathbf{k}\mathbf{k}'} \frac{\Delta_{\mathbf{k}'}}{2\sqrt{\xi_{\mathbf{k}'}^2 + |\Delta_{\mathbf{k}'}|^2}} [1 - 2n_F(E_{\mathbf{k}'})] \\ &\equiv -\frac{1}{N} \sum_{\mathbf{k}'} V_{\mathbf{k}\mathbf{k}'} \frac{\Delta_{\mathbf{k}'}}{2E_{\mathbf{k}'}} [1 - 2n_F(E_{\mathbf{k}'})]. \end{aligned} \quad (10.40)$$

We see that $n_F(E_{\mathbf{k}'}) \rightarrow 0$ for $E_{\mathbf{k}'} > 0$ and $T \rightarrow 0$ so that the zero-temperature BCS gap equation (9.28) is recovered as a limiting case.

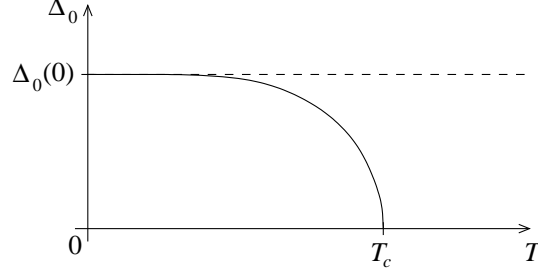
For our model interaction and assuming a \mathbf{k} -independent real gap, we obtain, in analogy to the ground-state derivation,

$$1 = V_0 \int_{-\omega_D}^{\omega_D} d\xi D(\mu + \xi) \frac{1 - 2n_F(\sqrt{\xi^2 + \Delta_0^2})}{2\sqrt{\xi^2 + \Delta_0^2}} = V_0 \int_{-\omega_D}^{\omega_D} d\xi D(\mu + \xi) \frac{\tanh \frac{\beta}{2} \sqrt{\xi^2 + \Delta_0^2}}{2\sqrt{\xi^2 + \Delta_0^2}}. \quad (10.41)$$

Note that this only works for $\Delta_0 \neq 0$ since we have divided by Δ_0 . If the density of states is approximately constant close to E_F , the equation simplifies to

$$1 \cong V_0 D(E_F) \int_{-\omega_D}^{\omega_D} d\xi \frac{\tanh \frac{\beta}{2} \sqrt{\xi^2 + \Delta_0^2}}{2\sqrt{\xi^2 + \Delta_0^2}}. \quad (10.42)$$

The integral is easily evaluated numerically, leading to the temperature dependence of Δ_0 :



For weak coupling we have already seen that

$$\Delta_0(0) \cong 2\omega_D \exp\left(-\frac{1}{V_0 D(E_F)}\right). \quad (10.43)$$

We can also obtain an analytical expression for T_c : If T approaches T_c from below, we can take the limit $\Delta_0 \rightarrow 0$ in the gap equation,

$$\begin{aligned} 1 &\cong V_0 D(E_F) \int_{-\omega_D}^{\omega_D} d\xi \frac{\tanh \frac{\beta}{2} |\xi|}{2|\xi|} = V_0 D(E_F) \int_0^{\omega_D} d\xi \frac{\tanh \frac{\beta}{2} \xi}{\xi} = V_0 D(E_F) \int_0^{\beta\omega_D/2} dx \frac{\tanh x}{x} \\ &\stackrel{\text{by parts}}{\cong} V_0 D(E_F) \left\{ \ln \frac{\beta\omega_D}{2} \tanh \frac{\beta\omega_D}{2} - \int_0^{\beta\omega_D/2} dx \frac{\ln x}{\cosh^2 x} \right\}. \end{aligned} \quad (10.44)$$

In the weak-coupling limit we have $\beta\omega_D = \omega_D/k_B T \gg 1$ (this assertion should be checked *a-posteriori*). Since the integrand of the last integral decays exponentially for large x , we can send the upper limit to infinity,

$$1 \cong V_0 D(E_F) \left\{ \ln \frac{\beta\omega_D}{2} \underbrace{\tanh \frac{\beta\omega_D}{2}}_{\cong 1} - \int_0^{\infty} dx \frac{\ln x}{\cosh^2 x} \right\} \cong V_0 D(E_F) \left(\ln \frac{\beta\omega_D}{2} + \gamma - \ln \frac{\pi}{4} \right), \quad (10.45)$$

where γ is again the Euler gamma constant. This implies

$$\exp\left(\frac{1}{V_0 D(E_F)}\right) \cong \frac{2\beta\omega_D}{\pi} e^\gamma \quad (10.46)$$

$$\Rightarrow k_B T_c \cong \frac{2e^\gamma}{\pi} \omega_D \exp\left(-\frac{1}{V_0 D(E_F)}\right). \quad (10.47)$$

This is exactly the same expression we have found above for the critical temperature of the Cooper instability. Since the approximations used are quite different, this agreement is not trivial. The gap at zero temperature and the critical temperature thus have a *universal* ratio in BCS theory,

$$\frac{2\Delta_0(0)}{k_B T_c} \cong \frac{2\pi}{e^\gamma} \approx 3.528. \quad (10.48)$$

This ratio is close to the result measured for simple elementary superconductors. For example, for tin one finds $2\Delta_0(0)/k_B T_c \approx 3.46$. For superconductors with stonger coupling, such as mercury, and for unconventional superconductors the agreement is not good, though.

10.2 Isotope effect

How can one check that superconductivity is indeed governed by a phonon-mediated interaction? BCS theory predicts

$$k_B T_c, \Delta_0 \propto \omega_D \exp\left(-\frac{1}{V_0 D(E_F)}\right). \quad (10.49)$$

It would be ideal to compare $k_B T_c$ or Δ_0 for superconductors that only differ in the Debye frequency ω_D , not in V_0 of $D(E_F)$. This is at least approximately possible by using samples containing different isotopes (or different fractions of isotopes) of the same elements.

The eigenfrequency of a harmonic oscillator scales with the mass like

$$\omega \sim \frac{1}{\sqrt{m}}. \quad (10.50)$$

The entire phonon dispersion, and thus in particular the Debye frequency, also scales like

$$\omega_{\mathbf{q}\lambda} \sim \frac{1}{\sqrt{M}}, \quad \omega_D \sim \frac{1}{\sqrt{M}} \quad (10.51)$$

with the atomic mass M for an elementary superconductor. The same scaling has been found above for the jellium model, see Eq. (8.73). Consequently, for elementary BCS weak-coupling superconductors,

$$k_B T_c, \Delta_0 \sim M^{-\alpha} \quad \text{with} \quad \alpha = \frac{1}{2}. \quad (10.52)$$

This is indeed found for simple superconductors. The exponent is found to be smaller or even negative for materials that are not in the weak-coupling regime $V D(E_F) \ll 1$ or that are not phonon-mediated superconductors. In particular, if the relevant interaction has nothing to do with phonons, we expect $\alpha = 0$. This is observed for optimally doped (highest T_c) cuprate high-temperature superconductors.

10.3 Specific heat

We now discuss further predictions following from BCS theory. We start by revisiting the heat capacity or specific heat. The BCS Hamiltonian

$$H_{\text{BCS}} = E_{\text{BCS}} + \sum_{\mathbf{k}\sigma} E_{\mathbf{k}} \gamma_{\mathbf{k}\sigma}^\dagger \gamma_{\mathbf{k}\sigma} \quad (10.53)$$

with

$$E_{\mathbf{k}} = \sqrt{\xi_{\mathbf{k}}^2 + |\Delta_{\mathbf{k}}|^2} \quad (10.54)$$

leads to the internal energy

$$U = \langle H_{\text{BCS}} \rangle = E_{\text{BCS}} + 2 \sum_{\mathbf{k}} E_{\mathbf{k}} n_F(E_{\mathbf{k}}). \quad (10.55)$$

However, this is inconvenient for the calculation of the heat capacity $C = dU/dT$ since the condensate energy E_{BCS} depends on temperature through $\langle c_{-\mathbf{k},\downarrow} c_{\mathbf{k}\uparrow} \rangle$. We better consider the entropy, which has no contribution from the condensate. It reads

$$S = -k_B \sum_{\mathbf{k}\sigma} [(1 - n_F) \ln(1 - n_F) + n_F \ln n_F], \quad (10.56)$$

where $n_F \equiv n_F(E_{\mathbf{k}})$. From the entropy, we obtain the heat capacity

$$\begin{aligned} C &= T \frac{dS}{dT} = -\beta \frac{dS}{d\beta} \\ &= 2k_B \beta \sum_{\mathbf{k}} \frac{d}{d\beta} [(1 - n_F) \ln(1 - n_F) + n_F \ln n_F] = 2k_B \beta \sum_{\mathbf{k}} \underbrace{[-\ln(1 - n_F) - 1 + \ln n_F + 1]}_{= \ln \frac{n_F}{1 - n_F} = \ln e^{-\beta E_{\mathbf{k}}} = -\beta E_{\mathbf{k}}} \frac{dn_F}{d\beta} \\ &= -2k_B \beta^2 \sum_{\mathbf{k}} E_{\mathbf{k}} \frac{dn_F(E_{\mathbf{k}})}{d\beta}. \end{aligned} \quad (10.57)$$

Note that $n_F(E_{\mathbf{k}})$ depends on β or temperature both explicitly and through the temperature dependence of $E_{\mathbf{k}} = \sqrt{\xi_{\mathbf{k}}^2 + |\Delta_{\mathbf{k}}(T)|^2}$:

$$C = -2k_B\beta^2 \sum_{\mathbf{k}} E_{\mathbf{k}} \left(\underbrace{\frac{\partial n_F}{\partial \beta}}_{= \frac{E_{\mathbf{k}}}{\beta} \frac{\partial n_F}{\partial E_{\mathbf{k}}}} + \frac{\partial n_F}{\partial E_{\mathbf{k}}} \frac{1}{2E_{\mathbf{k}}} \frac{d|\Delta_{\mathbf{k}}|^2}{d\beta} \right) = -2k_B\beta \sum_{\mathbf{k}} \frac{\partial n_F}{\partial E_{\mathbf{k}}} \left(E_{\mathbf{k}}^2 + \frac{1}{2} \beta \frac{d|\Delta_{\mathbf{k}}|^2}{d\beta} \right). \quad (10.58)$$

The first term is due to the explicit β dependence, i.e., to the change of occupation of quasiparticle states with temperature. The second term results from the temperature dependence of the quasiparticle spectrum and is absent for $T \geq T_c$, where $|\Delta_{\mathbf{k}}|^2 = 0 = \text{const.}$ The sum over \mathbf{k} contains the factor

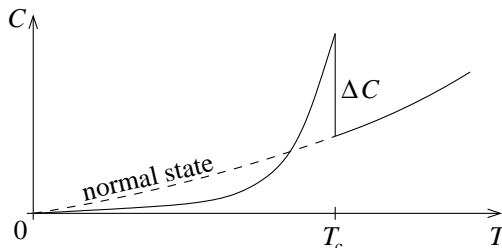
$$\frac{\partial n_F}{\partial E_{\mathbf{k}}} = \beta \frac{\partial}{\partial \beta E_{\mathbf{k}}} \frac{1}{e^{\beta E_{\mathbf{k}}} + 1} = -\beta \frac{e^{\beta E_{\mathbf{k}}}}{(e^{\beta E_{\mathbf{k}}} + 1)^2} = -\beta n_F(E_{\mathbf{k}})[1 - n_F(E_{\mathbf{k}})] \quad (10.59)$$

so that

$$C = 2k_B\beta^2 \sum_{\mathbf{k}} n_F(1 - n_F) \left(E_{\mathbf{k}}^2 + \frac{1}{2} \beta \frac{d|\Delta_{\mathbf{k}}|^2}{d\beta} \right). \quad (10.60)$$

Here, $n_F(1 - n_F)$ is exponentially small for $E_{\mathbf{k}} \gg k_B T$. This means that for $k_B T \ll \Delta_{\min}$, where Δ_{\min} is the minimum superconducting gap, *all* terms in the sum are exponentially suppressed since $k_B T \ll \Delta_{\min} \leq E_{\mathbf{k}}$. Thus the heat capacity is exponentially small at low temperatures. This result is not specific to superconductors—all systems with an energy gap for excitations show this behavior.

For the simple interaction used above, the heat capacity can be obtained in terms of an integral over energy. The numerical evaluation gives the following result:



We find a downward jump at T_c , reproducing a result obtained from Landau theory in section 6.1. The jump occurs for any mean-field theory describing a second-order phase transition for a complex order parameter. Since BCS theory is such a theory, it recovers the result.

The height of the jump can be found as follows: The $E_{\mathbf{k}}^2$ term in Eq. (10.60) is continuous for $T \rightarrow T_c^-$ ($\Delta_0 \rightarrow 0$). Thus the jump is given by

$$\Delta C = \frac{1}{k_B^2 T_c^3} \sum_{\mathbf{k}} n_F(\xi_{\mathbf{k}}) [1 - n_F(\xi_{\mathbf{k}})] \left. \frac{d\Delta_0^2}{d\beta} \right|_{T \rightarrow T_c^-}. \quad (10.61)$$

To obtain Δ_0 close to T_c , we have to solve the gap equation

$$1 = V_0 D(E_F) \int_{-\omega_D}^{\omega_D} d\xi \frac{\tanh \frac{\beta}{2} \sqrt{\xi^2 + \Delta_0^2}}{2\sqrt{\xi^2 + \Delta_0^2}} = V_0 D(E_F) \int_0^{\omega_D} d\xi \frac{\tanh \frac{\beta}{2} \sqrt{\xi^2 + \Delta_0^2}}{\sqrt{\xi^2 + \Delta_0^2}} \quad (10.62)$$

for small Δ_0 . Writing

$$\beta = \frac{1}{k_B T} = \frac{1}{k_B(T_c - \Delta T)} \quad (10.63)$$

and expanding for small ΔT and small Δ_0 , we obtain

$$\begin{aligned}
1 &\cong \underbrace{V_0 D(E_F) \int_0^{\omega_D} d\xi \frac{\tanh \frac{\xi}{2k_B T_c}}{\xi}}_{=1 \text{ (gap equation)}} + \frac{V_0 D(E_F)}{2k_B T_c^2} \Delta T \int_0^{\omega_D} \frac{d\xi}{\cosh^2 \frac{\xi}{2k_B T_c}} \\
&+ \frac{V_0 D(E_F)}{4k_B T_c} \Delta_0^2 \int_0^{\omega_D} \frac{d\xi}{\xi^2} \left(\frac{1}{\cosh^2 \frac{\xi}{2k_B T_c}} - \frac{2k_B T_c \tanh \frac{\xi}{2k_B T_c}}{\xi} \right). \tag{10.64}
\end{aligned}$$

In the weak-coupling limit, $\omega_D \gg k_B T_c$, we can extend the integrals to infinity, which yields

$$0 \cong \frac{V_0 D(E_F)}{2k_B T_c^2} \Delta T 2k_B T_c - \frac{V_0 D(E_F)}{4k_B T_c} \Delta_0^2 \frac{7\zeta(3)}{2k_B T_c} = V_0 D(E_F) \frac{\Delta T}{T_c} - V_0 D(E_F) \frac{7\zeta(3)}{8\pi^2} \frac{\Delta_0^2}{(k_B T_c)^2} \tag{10.65}$$

$$\Rightarrow \Delta_0^2 \cong \frac{8\pi^2}{7\zeta(3)} k_B^2 T_c \Delta T. \tag{10.66}$$

Thus

$$\frac{d\Delta_0^2}{d\beta} \Big|_{T \rightarrow T_c^-} = \frac{d\Delta T}{d\beta} \frac{d\Delta_0^2}{d\Delta T} \Big|_{\Delta T \rightarrow 0} = k_B T_c^2 \frac{d\Delta_0^2}{d\Delta T} \Big|_{\Delta T \rightarrow 0} \cong \frac{8\pi^2}{7\zeta(3)} k_B^3 T_c^3 \tag{10.67}$$

$$\begin{aligned}
\Rightarrow \Delta C &= \frac{1}{k_B^2 T_c^3} \sum_{\mathbf{k}} n_F(\xi_{\mathbf{k}}) [1 - n_F(\xi_{\mathbf{k}})] \frac{8\pi^2}{7\zeta(3)} k_B^3 T_c^3 = \frac{8\pi^2}{7\zeta(3)} k_B N D(E_F) \underbrace{\int_{-\infty}^{\infty} d\xi n_F(\xi) [1 - n_F(\xi)]}_{=k_B T_c} \\
&= \frac{8\pi^2}{7\zeta(3)} N D(E_F) k_B^2 T_c, \tag{10.68}
\end{aligned}$$

where N is the number of unit cells. The specific-heat jump is

$$\Delta c = \frac{\Delta C}{V} = \frac{8\pi^2}{7\zeta(3)} d(E_F) k_B^2 T_c, \tag{10.69}$$

where $d(E_F) = D(E_F) N/V$ is the density of states per volume.

10.4 Density of states and single-particle tunneling

Detailed experimental information on the excitation spectrum in a superconductor can be obtained from single-particle tunneling between a superconductor and either a normal metal or another superconductor. We discuss this in the following assuming, for simplicity, that the density of states $D(E)$ in the normal state is approximately constant close to the Fermi energy. We restrict ourselves to single-electron tunneling; pair tunneling, which leads to the Josephson effects, will be discussed later.

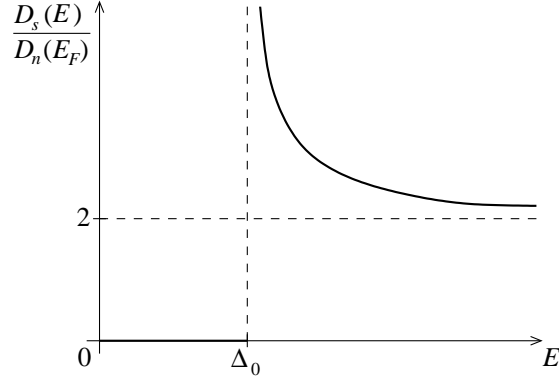
Quasiparticle density of states

The density of states per spin of the Bogoliubov quasiparticles created by γ^\dagger is easily obtained from their dispersion:

$$D_s(E) = \frac{1}{N} \sum_{\mathbf{k}} \delta(E - E_{\mathbf{k}}) = \frac{1}{N} \sum_{\mathbf{k}} \delta \left(E - \sqrt{\xi_{\mathbf{k}}^2 + |\Delta_{\mathbf{k}}|^2} \right). \tag{10.70}$$

The subscript “s” of $D_s(E)$ stands for “superconducting.” For the case of approximately constant normal-state density of states D_n and constant gap Δ_0 , we find

$$\begin{aligned}
D_s(E) &= \int_{-\infty}^{\infty} d\xi D_n(\mu - \xi) \delta\left(E - \sqrt{\xi^2 + \Delta_0^2}\right) \cong D_n(E_F) \int_{-\infty}^{\infty} d\xi \delta\left(E - \sqrt{\xi^2 + \Delta_0^2}\right) \\
&= D_n(E_F) \int_{-\infty}^{\infty} d\xi \left(\frac{\delta\left(\xi + \sqrt{E^2 - \Delta_0^2}\right)}{\frac{|\xi|}{\sqrt{\xi^2 + \Delta_0^2}}} + \frac{\delta\left(\xi - \sqrt{E^2 - \Delta_0^2}\right)}{\frac{\xi}{\sqrt{\xi^2 + \Delta_0^2}}} \right) \\
&= \begin{cases} D_n(E_F) \frac{2E}{\sqrt{E^2 - \Delta_0^2}} & \text{for } E > \Delta_0, \\ 0 & \text{for } E < \Delta_0. \end{cases} \tag{10.71}
\end{aligned}$$



There are of course no states in the energy gap. Importantly, we find a *divergence* at the gap edge at $E = \Delta_0$. In the normal-state limit, $\Delta_0 \rightarrow 0$, we obtain $D_s(E)/D_n(E_F) \rightarrow 2$, deviating from the result of unity given in Tinkham’s book. The origin is that E is an excitation energy relative to the Fermi sea, i.e., both electron and hole excitations contribute to the density of states at positive E .

Single-particle tunneling

It is plausible that the density of states $D_s(E)$ can be mapped out by tunneling experiments, for example using a normal-metal/insulator/superconductor structure. However, we have to keep in mind that the particles tunneling out of or into a normal metal are real electrons, whereas the quasiparticles in a superconductor are superpositions of electrons and holes. To study tunneling effects theoretically, we employ a *tunneling Hamiltonian* of the form

$$H = H_L + H_R + H_T, \tag{10.72}$$

where H_L and H_R describe the materials to the “left” and “right” of the tunneling region. Either can be a normal metal or a superconductor and is assumed to be unaffected by the presence of the tunneling region. For example, while translational invariance is necessarily broken in a tunneling device, we nevertheless write $H_{L,R}$ in terms of lattice-momentum states. H_T describes the tunneling between the two materials,

$$H_T = \sum_{\mathbf{k}\mathbf{q}\sigma} t_{\mathbf{k}\mathbf{q}} c_{\mathbf{k}\sigma}^\dagger d_{\mathbf{q}\sigma} + \text{h.c.} \tag{10.73}$$

Here, c and d are electronic operators referring to the two sides of the tunneling barrier and $t_{\mathbf{k}\mathbf{q}}$ is a tunneling matrix element, which might depend on the momenta of the incoming and outgoing electron. We assume a non-magnetic tunneling barrier so that the electron spin is conserved and $t_{\mathbf{k}\mathbf{q}}$ does not depend on it. Note that in the presence of interactions H_T is an approximation valid for weak tunneling. We treat the two bulk materials in the mean-field approximation so that $H_{L,R}$ are effectively non-interacting mean-field Hamiltonians. We further assume constant normal-state densities of states and momentum-independent tunneling matrix elements $t = t_{\mathbf{k}\mathbf{q}}$.

For tunneling between *two normal metals*, we can calculate the current for an applied voltage V using the *Landauer formula*:

$$I_{nn} = \underbrace{I_{L \rightarrow R}}_{\text{left to right}} - \underbrace{I_{R \rightarrow L}}_{\text{right to left}} \quad (10.74)$$

with

$$I_{L \rightarrow R} \propto \frac{2e}{h} \int_{-\infty}^{\infty} d\xi D_L^n(\mu + \xi) D_R^n(\mu + \xi + eV) |t|^2 n_F(\xi) [1 - n_F(\xi + eV)], \quad (10.75)$$

$$I_{R \rightarrow L} \propto \frac{2e}{h} \int_{-\infty}^{\infty} d\xi D_L^n(\mu + \xi) D_R^n(\mu + \xi + eV) |t|^2 n_F(\xi + eV) [1 - n_F(\xi)], \quad (10.76)$$

where $D_{L,R}^n$ is the density of states per spin direction in the left/right normal metal. The factors of $D_L^n(\mu + \xi) n_F(\xi)$ etc. can be understood as the probabilities that the relevant initial states exist and are occupied and the final states exist and are empty. We get

$$\begin{aligned} I_{nn} &\propto \frac{2e}{h} \int_{-\infty}^{\infty} d\xi D_L^n(\mu + \xi) D_R^n(\mu + \xi + eV) |t|^2 [n_F(\xi) - n_F(\xi + eV)] \\ &\cong \frac{2e}{h} D_L^n(E_F) D_R^n(E_F) |t|^2 \int_{-\infty}^{\infty} d\xi [n_F(\xi) - n_F(\xi + eV)] \\ &= \frac{2e^2}{h} D_L^n(E_F) D_R^n(E_F) |t|^2 V. \end{aligned} \quad (10.77)$$

We thus find ohmic behaviour,

$$I_{nn} = G_{nn} V. \quad (10.78)$$

We next consider the case that one material is *normal* while the other (without loss of generality the left) is *superconducting*. Now there are additional factors because the quasiparticles in the superconductor are not pure electrons or holes. Let us say an electron is tunneling out of the superconductor with energy $\omega > 0$. First of all, this is only possible if $\omega \geq \Delta_0$ because of the energy gap. Now the electron can either come from an electron-like quasiparticle ($k > k_F$), which contains an electron portion of

$$|u_{\mathbf{k}}|^2 = \frac{1}{2} \left(1 + \frac{\xi_{\mathbf{k}}}{\sqrt{\xi_{\mathbf{k}}^2 + \Delta_0^2}} \right) \Big|_{\xi_{\mathbf{k}} = \sqrt{\omega^2 + \Delta_0^2}} = \frac{1}{2} \left(1 + \frac{\sqrt{\omega^2 - \Delta_0^2}}{\omega} \right), \quad (10.79)$$

or from a hole-like quasiparticle ($k' < k_F$) with an electron portion of

$$|u_{\mathbf{k}'}|^2 = \frac{1}{2} \left(1 + \frac{\xi_{\mathbf{k}'}}{\sqrt{\xi_{\mathbf{k}'}^2 + \Delta_0^2}} \right) \Big|_{\xi_{\mathbf{k}'} = -\sqrt{\omega^2 + \Delta_0^2}} = \frac{1}{2} \left(1 - \frac{\sqrt{\omega^2 - \Delta_0^2}}{\omega} \right). \quad (10.80)$$

On the other hand, an electron tunneling out with energy $\omega < 0$ is best described as a hole tunneling in with energy $-\omega > 0$. The relevant factors are

$$|v_{\mathbf{k}}|^2 = \frac{1}{2} \left(1 - \frac{\sqrt{\omega^2 - \Delta_0^2}}{|\omega|} \right) \quad (10.81)$$

and

$$|v_{\mathbf{k}'}|^2 = \frac{1}{2} \left(1 + \frac{\sqrt{\omega^2 - \Delta_0^2}}{|\omega|} \right). \quad (10.82)$$

Thus the current flowing from left to right is

$$\begin{aligned}
I_{L \rightarrow R} &\propto \frac{2e}{h} \int_0^{\infty} d\omega D_L^s(\omega) D_R^n(\mu + \omega + eV) |t|^2 n_F(\omega) [1 - n_F(\omega + eV)] \\
&\quad \times \underbrace{\left\{ \frac{1}{2} \left(1 + \frac{\sqrt{\omega^2 - \Delta_0^2}}{\omega} \right) + \frac{1}{2} \left(1 - \frac{\sqrt{\omega^2 - \Delta_0^2}}{\omega} \right) \right\}}_{=1} \\
&\quad + \frac{2e}{h} \int_{-\infty}^0 d\omega D_L^s(|\omega|) D_R^n(\mu + \omega + eV) |t|^2 n_F(\omega) [1 - n_F(\omega + eV)] \\
&\quad \times \underbrace{\left\{ \frac{1}{2} \left(1 - \frac{\sqrt{\omega^2 - \Delta_0^2}}{|\omega|} \right) + \frac{1}{2} \left(1 + \frac{\sqrt{\omega^2 - \Delta_0^2}}{|\omega|} \right) \right\}}_{=1} \\
&= \frac{2e}{h} \int_{-\infty}^{\infty} d\omega D_L^s(|\omega|) D_R^n(\mu + \omega + eV) |t|^2 n_F(\omega) [1 - n_F(\omega + eV)], \tag{10.83}
\end{aligned}$$

where $D_L^s(\omega)$ now is the superconducting density of states neither containing a spin factor of 2 nor the factor of 2 due describing both electrons and holes as excitations with positive energy—positive and negative energies are here treated explicitly and separately. We see that the electron-hole mixing does not lead to any additional factors beyond the changed density of states. With the analogous expression

$$I_{R \rightarrow L} \propto \frac{2e}{h} \int_{-\infty}^{\infty} d\omega D_L^s(|\omega|) D_R^n(\mu + \omega + eV) |t|^2 n_F(\omega + eV) [1 - n_F(\omega)] \tag{10.84}$$

we obtain

$$\begin{aligned}
I_{sn} &\propto \frac{2e}{h} \int_{-\infty}^{\infty} d\omega D_L^s(|\omega|) D_R^n(\mu + \omega + eV) |t|^2 [n_F(\omega) - n_F(\omega + eV)] \\
&\cong \frac{2e}{h} D_L^n(E_F) D_R^n(E_F) |t|^2 \int_{-\infty}^{\infty} d\omega \frac{D_L^s(|\omega|)}{D_L^n(E_F)} [n_F(\omega) - n_F(\omega + eV)], \tag{10.85}
\end{aligned}$$

where

$$\frac{D_L^s(|\omega|)}{D_L^n(E_F)} = \begin{cases} \frac{|\omega|}{\sqrt{\omega^2 - \Delta_0^2}} & \text{for } |\omega| > \Delta_0, \\ 0 & \text{for } |\omega| < \Delta_0. \end{cases} \tag{10.86}$$

Thus

$$I_{sn} = \frac{G_{nn}}{e} \int_{-\infty}^{\infty} d\omega \frac{D_L^s(|\omega|)}{D_L^n(E_F)} [n_F(\omega) - n_F(\omega + eV)]. \tag{10.87}$$

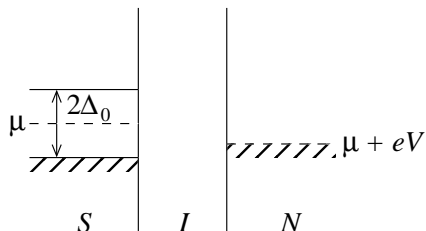
It is useful to consider the *differential conductance*

$$\begin{aligned}
G_{sn} &:= \frac{dI_{sn}}{dV} = G_{nn} \int_{-\infty}^{\infty} d\omega \frac{D_L^s(|\omega|)}{D_L^n(E_F)} \left(-\frac{\partial n_F(\omega + eV)}{\partial \omega} \right) \\
&= G_{nn} \int_{-\infty}^{\infty} d\omega \frac{D_L^s(|\omega|)}{D_L^n(E_F)} \beta n_F(\omega + eV) [1 - n_F(\omega + eV)]. \tag{10.88}
\end{aligned}$$

In the limit $k_B T \rightarrow 0$ this becomes

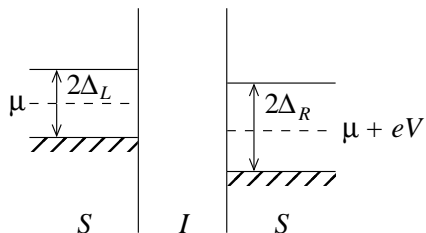
$$G_{sn} = G_{nn} \int_{-\infty}^{\infty} d\omega \frac{D_L^s(|\omega|)}{D_L^n(E_F)} \delta(\omega + eV) = G_{nn} \frac{D_L^s(|eV|)}{D_L^n(E_F)}. \quad (10.89)$$

Thus low-temperature tunneling directly measures the superconducting density of states. At non-zero temperatures, the features are smeared out over an energy scale of $k_B T$. In a band picture we can see that the Fermi energy in the normal material is used to scan the density of states in the superconductor.

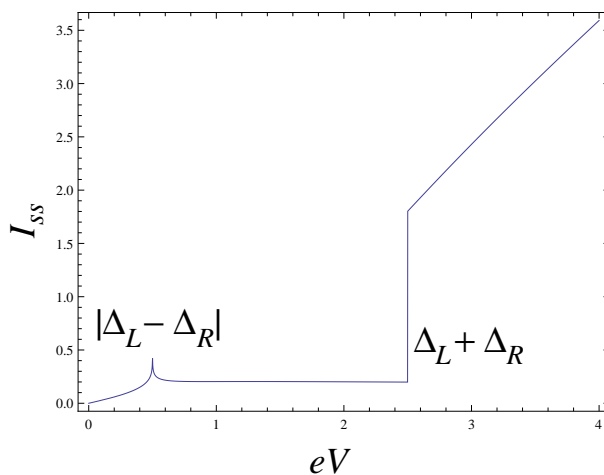


For a *superconductor-insulator-superconductor* contact, we only present the result for the current without derivation. The result is very plausible in view of the previous cases:

$$I_{ss} = \frac{G_{nn}}{e} \int_{-\infty}^{\infty} d\omega \frac{D_L^s(|\omega|)}{D_L^n(E_F)} \frac{D_R^s(|\omega + eV|)}{D_R^n(E_F)} [n_F(\omega) - n_F(\omega + eV)]. \quad (10.90)$$



Now we find a large change in the current when the voltage V is chosen such that the lower gap edge at one side is aligned with the upper gap edge at the other side. This is the case for $|eV| = \Delta_L + \Delta_R$. This feature will remain sharp at non-zero temperatures since the densities of states retain their divergences at the gap edges as long as superconductivity is not destroyed. Effectively, we are using the density-of-states singularity of one superconductor to scan the density of states of the other. A numerical evaluation of I_{ss} gives the following typical behavior (here for $\Delta_L = 2k_B T$, $\Delta_R = 3k_B T$):



10.5 Ultrasonic attenuation and nuclear relaxation

To conclude the brief survey of experimental consequences of BCS theory, we discuss the effects of time-dependent perturbations. They will be exemplified by ultrasonic attenuation and nuclear relaxation, which represent two distinct ways in which quasiparticle interference comes into play. Quite generally, we write the perturbation part of the Hamiltonian as

$$H_1 = \sum_{\mathbf{k}\mathbf{k}'} \sum_{\sigma\sigma'} B_{\mathbf{k}'\sigma'\mathbf{k}\sigma} c_{\mathbf{k}'\sigma'}^\dagger c_{\mathbf{k}\sigma}, \quad (10.91)$$

where $B_{\mathbf{k}'\sigma'\mathbf{k}\sigma}$ are matrix elements of the perturbation between single-electron states of the non-interacting system. In the superconducting state, we have to express c, c^\dagger in terms of γ, γ^\dagger ,

$$c_{\mathbf{k}\uparrow} = u_{\mathbf{k}}\gamma_{\mathbf{k}\uparrow} + v_{\mathbf{k}}\gamma_{-\mathbf{k},\downarrow}^\dagger, \quad (10.92)$$

$$c_{\mathbf{k}\downarrow} = u_{\mathbf{k}}\gamma_{\mathbf{k}\downarrow} - v_{\mathbf{k}}\gamma_{-\mathbf{k},\uparrow}^\dagger, \quad (10.93)$$

where we have assumed $u_{-\mathbf{k}} = u_{\mathbf{k}}, v_{-\mathbf{k}} = v_{\mathbf{k}}$. Thus, with $\sigma, \sigma' = \pm 1$,

$$\begin{aligned} H_1 &= \sum_{\mathbf{k}\mathbf{k}'} \sum_{\sigma\sigma'} B_{\mathbf{k}'\sigma'\mathbf{k}\sigma} \left(u_{\mathbf{k}'}^* \gamma_{\mathbf{k}'\sigma'}^\dagger + \sigma' v_{\mathbf{k}'}^* \gamma_{-\mathbf{k}',-\sigma'} \right) \left(u_{\mathbf{k}} \gamma_{\mathbf{k}\sigma} + \sigma v_{\mathbf{k}} \gamma_{-\mathbf{k},-\sigma}^\dagger \right) \\ &= \sum_{\mathbf{k}\mathbf{k}'} \sum_{\sigma\sigma'} B_{\mathbf{k}'\sigma'\mathbf{k}\sigma} \left(u_{\mathbf{k}'}^* u_{\mathbf{k}} \gamma_{\mathbf{k}'\sigma'}^\dagger \gamma_{\mathbf{k}\sigma} + \sigma u_{\mathbf{k}'}^* v_{\mathbf{k}} \gamma_{\mathbf{k}'\sigma'}^\dagger \gamma_{-\mathbf{k},-\sigma}^\dagger \right. \\ &\quad \left. + \sigma' v_{\mathbf{k}'}^* u_{\mathbf{k}} \gamma_{-\mathbf{k}',-\sigma'} \gamma_{\mathbf{k}\sigma} + \sigma \sigma' v_{\mathbf{k}'}^* v_{\mathbf{k}} \gamma_{-\mathbf{k}',-\sigma'} \gamma_{-\mathbf{k},-\sigma}^\dagger \right). \end{aligned} \quad (10.94)$$

It is useful to combine the terms containing $B_{\mathbf{k}'\sigma'\mathbf{k}\sigma}$ and $B_{-\mathbf{k},-\sigma,-\mathbf{k}',-\sigma'}$ since both refer to processes that change momentum by $\mathbf{k}' - \mathbf{k}$ and spin by $\sigma' - \sigma$. If the perturbation couples to the electron concentration, which is the case for ultrasound, one finds simply

$$B_{-\mathbf{k},-\sigma,-\mathbf{k}',-\sigma'} = B_{\mathbf{k}'\sigma'\mathbf{k}\sigma}. \quad (10.95)$$

This is often called *case I*. Furthermore, spin is conserved by the coupling to ultrasound, thus

$$B_{\mathbf{k}'\sigma'\mathbf{k}\sigma} = \delta_{\sigma\sigma'} B_{\mathbf{k}'\sigma\mathbf{k}\sigma}. \quad (10.96)$$

Adding the two terms, we obtain

$$\begin{aligned} H_{\text{ultra}} &= \frac{1}{2} \sum_{\mathbf{k}\mathbf{k}'} \sum_{\sigma} B_{\mathbf{k}'\sigma\mathbf{k}\sigma} \left(u_{\mathbf{k}'}^* u_{\mathbf{k}} \gamma_{\mathbf{k}'\sigma}^\dagger \gamma_{\mathbf{k}\sigma} + \sigma u_{\mathbf{k}'}^* v_{\mathbf{k}} \gamma_{\mathbf{k}'\sigma}^\dagger \gamma_{-\mathbf{k},-\sigma}^\dagger \right. \\ &\quad \left. + \sigma v_{\mathbf{k}'}^* u_{\mathbf{k}} \gamma_{-\mathbf{k}',-\sigma} \gamma_{\mathbf{k}\sigma} + v_{\mathbf{k}'}^* v_{\mathbf{k}} \gamma_{-\mathbf{k}',-\sigma} \gamma_{-\mathbf{k},-\sigma}^\dagger + u_{\mathbf{k}}^* u_{\mathbf{k}'} \gamma_{-\mathbf{k},-\sigma}^\dagger \gamma_{-\mathbf{k}',-\sigma} \right. \\ &\quad \left. - \sigma u_{\mathbf{k}}^* v_{\mathbf{k}'} \gamma_{-\mathbf{k},-\sigma}^\dagger \gamma_{\mathbf{k}'\sigma}^\dagger - \sigma v_{\mathbf{k}}^* u_{\mathbf{k}'} \gamma_{\mathbf{k}\sigma} \gamma_{-\mathbf{k}',-\sigma} + v_{\mathbf{k}}^* v_{\mathbf{k}'} \gamma_{\mathbf{k}\sigma} \gamma_{\mathbf{k}'\sigma}^\dagger \right). \end{aligned} \quad (10.97)$$

Assuming $u_{\mathbf{k}}, v_{\mathbf{k}}, \Delta_{\mathbf{k}} \in \mathbb{R}$ for simplicity, we get, up to a constant,

$$\begin{aligned} H_{\text{ultra}} &= \frac{1}{2} \sum_{\mathbf{k}\mathbf{k}'} \sum_{\sigma} B_{\mathbf{k}'\sigma\mathbf{k}\sigma} \left[(u_{\mathbf{k}'} u_{\mathbf{k}} - v_{\mathbf{k}'} v_{\mathbf{k}}) \left(\gamma_{\mathbf{k}'\sigma}^\dagger \gamma_{\mathbf{k}\sigma} + \gamma_{-\mathbf{k},-\sigma}^\dagger \gamma_{-\mathbf{k}',-\sigma} \right) \right. \\ &\quad \left. + \sigma (u_{\mathbf{k}'} v_{\mathbf{k}} + v_{\mathbf{k}'} u_{\mathbf{k}}) \left(\gamma_{\mathbf{k}'\sigma}^\dagger \gamma_{-\mathbf{k},-\sigma}^\dagger + \gamma_{-\mathbf{k}',-\sigma} \gamma_{\mathbf{k}\sigma} \right) \right]. \end{aligned} \quad (10.98)$$

We thus find effective matrix elements $B_{\mathbf{k}'\sigma\mathbf{k}\sigma} (u_{\mathbf{k}'} u_{\mathbf{k}} - v_{\mathbf{k}'} v_{\mathbf{k}})$ for quasiparticle scattering and $B_{\mathbf{k}'\sigma\mathbf{k}\sigma} \sigma (u_{\mathbf{k}'} v_{\mathbf{k}} + v_{\mathbf{k}'} u_{\mathbf{k}})$ for creation and annihilation of two quasiparticles. Transition rates calculated from Fermi's golden rule contain the absolute values squared of matrix elements. Thus the following two *coherence factors* will be impor-

tant. The first one is

$$\begin{aligned}
(u_{\mathbf{k}'}u_{\mathbf{k}} - v_{\mathbf{k}'}v_{\mathbf{k}})^2 &= \frac{1}{4} \left\{ \left(1 + \frac{\xi_{\mathbf{k}'}}{\sqrt{\xi_{\mathbf{k}'}}^2 + \Delta_{\mathbf{k}'}}^2} \right) \left(1 + \frac{\xi_{\mathbf{k}}}{\sqrt{\xi_{\mathbf{k}}}}^2 + \Delta_{\mathbf{k}}}}^2} \right) \right. \\
&\quad \left. - 2 \frac{\xi_{\mathbf{k}'}}{\sqrt{\xi_{\mathbf{k}'}}^2 + \Delta_{\mathbf{k}'}}^2} \frac{\xi_{\mathbf{k}}}{\sqrt{\xi_{\mathbf{k}}}}^2 + \Delta_{\mathbf{k}}}}^2} + \left(1 - \frac{\xi_{\mathbf{k}'}}{\sqrt{\xi_{\mathbf{k}'}}^2 + \Delta_{\mathbf{k}'}}^2} \right) \left(1 - \frac{\xi_{\mathbf{k}}}{\sqrt{\xi_{\mathbf{k}}}}^2 + \Delta_{\mathbf{k}}}}^2} \right) \right\} \\
&= \frac{1}{2} \left(1 + \frac{\xi_{\mathbf{k}}\xi_{\mathbf{k}'}}{E_{\mathbf{k}}E_{\mathbf{k}'}} - \frac{\Delta_{\mathbf{k}}\Delta_{\mathbf{k}'}}{E_{\mathbf{k}}E_{\mathbf{k}'}} \right)
\end{aligned} \tag{10.99}$$

with $E_{\mathbf{k}} = \sqrt{\xi_{\mathbf{k}}^2 + \Delta_{\mathbf{k}}^2}$. If the matrix elements B and the normal-state density of states are even functions of the normal-state energy relative to the Fermi energy, $\xi_{\mathbf{k}}$, the second term, which is odd in $\xi_{\mathbf{k}}$ and $\xi_{\mathbf{k}'}$, will drop out under the sum $\sum_{\mathbf{k}\mathbf{k}'}$. Hence, this term is usually omitted, giving the coherence factor

$$F_-(\mathbf{k}, \mathbf{k}') := \frac{1}{2} \left(1 - \frac{\Delta_{\mathbf{k}}\Delta_{\mathbf{k}'}}{E_{\mathbf{k}}E_{\mathbf{k}'}} \right) \tag{10.100}$$

relevant for quasiparticle scattering in ultrasound experiments. Analogously, we obtain the second coherence factor

$$\underbrace{\sigma^2}_{=1} (u_{\mathbf{k}'}v_{\mathbf{k}} + v_{\mathbf{k}'}u_{\mathbf{k}})^2 = \left(1 + \frac{\Delta_{\mathbf{k}}\Delta_{\mathbf{k}'}}{E_{\mathbf{k}}E_{\mathbf{k}'}} \right) =: F_+(\mathbf{k}, \mathbf{k}') \tag{10.101}$$

for quasiparticle creation and annihilation.

We can gain insight into the temperature dependence of ultrasound attenuation by making the rather crude approximation that the matrix element B is independent of \mathbf{k}, \mathbf{k}' and thus of energy. Furthermore, typical ultrasound frequencies Ω satisfy $\hbar\Omega \ll \Delta_0$ and $\hbar\Omega \ll k_B T$. Then only scattering of quasiparticles by phonons but not their creation is important since the phonon energy is not sufficient for quasiparticle creation.

The rate of ultrasound absorption (attenuation) can be written in a plausible form analogous to the current in the previous section:

$$\alpha_s \propto \int_{-\infty}^{\infty} d\omega D_s(|\omega|) D_s(|\omega + \Omega|) |B|^2 F_-(\omega, \omega + \Omega) [n_F(\omega) - n_F(\omega + \Omega)], \tag{10.102}$$

where now

$$F_-(\omega, \omega') = \frac{1}{2} \left(1 - \frac{\Delta_0^2}{|\omega||\omega'|} \right). \tag{10.103}$$

With the approximations introduced above we get

$$\alpha_s \propto D_n^2(E_F) |B|^2 \int_{-\infty}^{\infty} d\omega \frac{D_s(|\omega|)}{D_n(E_F)} \frac{D_s(|\omega + \Omega|)}{D_n(E_F)} \frac{1}{2} \left(1 - \frac{\Delta_0^2}{|\omega||\omega + \Omega|} \right) [n_F(\omega) - n_F(\omega + \Omega)]. \tag{10.104}$$

The normal-state attenuation rate is found by letting $\Delta_0 \rightarrow 0$:

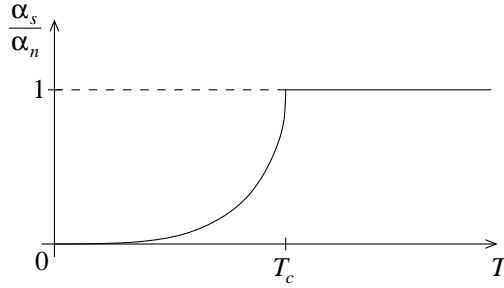
$$\alpha_n \propto D_n^2(E_F) |B|^2 \int_{-\infty}^{\infty} d\omega \frac{n_F(\omega) - n_F(\omega + \Omega)}{2} = \frac{1}{2} D_n^2(E_F) |B|^2 \Omega \tag{10.105}$$

$$\Rightarrow \frac{\alpha_s}{\alpha_n} = \frac{1}{\Omega} \int_{-\infty}^{\infty} d\omega \frac{D_s(|\omega|)}{D_n(E_F)} \frac{D_s(|\omega + \Omega|)}{D_n(E_F)} \frac{|\omega||\omega + \Omega| - \Delta_0^2}{|\omega||\omega + \Omega|} [n_F(\omega) - n_F(\omega + \Omega)]. \tag{10.106}$$

Since Ω is small, we can expand the integral up to linear order in Ω ,

$$\begin{aligned}
\frac{\alpha_s}{\alpha_n} &\cong \frac{1}{\Omega} \int_{-\infty}^{\infty} d\omega \left[\frac{D_s(|\omega|)}{D_n(E_F)} \right]^2 \frac{\omega^2 - \Delta_0^2}{\omega^2} (-\Omega) \frac{\partial n_F}{\partial \omega} \\
&= - \left(\int_{-\infty}^{-\Delta_0} + \int_{\Delta_0}^{\infty} \right) d\omega \left[\frac{|\omega|}{\sqrt{\omega^2 - \Delta_0^2}} \right]^2 \frac{\omega^2 - \Delta_0^2}{\omega^2} \frac{\partial n_F}{\partial \omega} = - \left(\int_{-\infty}^{-\Delta_0} + \int_{\Delta_0}^{\infty} \right) d\omega \frac{\partial n_F}{\partial \omega} \\
&= \underbrace{-n_F(\infty)}_{=0} + n_F(\Delta_0) - n_F(-\Delta_0) + \underbrace{n_F(-\infty)}_{=1} = n_F(\Delta_0) + 1 - n_F(-\Delta_0) \\
&= 2 n_F(\Delta_0) = \frac{2}{e^{\beta \Delta_0} + 1}.
\end{aligned} \tag{10.107}$$

Inserting the BCS prediction for $\Delta_0(T)$, we can plot α_s/α_n vs. temperature:



We now turn to the relaxation of nuclear spins due to their coupling to the electrons. We note without derivation that the hyperfine interaction relevant for this process is odd in momentum if the electron spin is not changed but is even if the electron spin is flipped, i.e.,

$$B_{-\mathbf{k}, -\sigma, -\mathbf{k}', -\sigma'} = -\sigma\sigma' B_{\mathbf{k}'\sigma'\mathbf{k}\sigma} \tag{10.108}$$

(recall $\sigma\sigma' = \pm 1$). This is called *case II*. The perturbation Hamiltonian now reads

$$\begin{aligned}
H_{\text{NMR}} &= \frac{1}{2} \sum_{\mathbf{k}\mathbf{k}'} \sum_{\sigma\sigma'} B_{\mathbf{k}'\sigma'\mathbf{k}\sigma} \left(u_{\mathbf{k}'}^* u_{\mathbf{k}} \gamma_{\mathbf{k}'\sigma'}^\dagger \gamma_{\mathbf{k}\sigma} + \sigma u_{\mathbf{k}'}^* v_{\mathbf{k}} \gamma_{\mathbf{k}'\sigma'}^\dagger \gamma_{-\mathbf{k}, -\sigma}^\dagger \right. \\
&\quad + \sigma' v_{\mathbf{k}'}^* u_{\mathbf{k}} \gamma_{-\mathbf{k}', -\sigma'} \gamma_{\mathbf{k}\sigma} + \sigma\sigma' v_{\mathbf{k}'}^* v_{\mathbf{k}} \gamma_{-\mathbf{k}', -\sigma'} \gamma_{-\mathbf{k}, -\sigma}^\dagger - \sigma\sigma' u_{\mathbf{k}}^* u_{\mathbf{k}'} \gamma_{-\mathbf{k}, -\sigma}^\dagger \gamma_{-\mathbf{k}', -\sigma'} \\
&\quad \left. + \sigma u_{\mathbf{k}}^* v_{\mathbf{k}'} \gamma_{-\mathbf{k}, -\sigma}^\dagger \gamma_{\mathbf{k}'\sigma'}^\dagger + \sigma' v_{\mathbf{k}}^* u_{\mathbf{k}'} \gamma_{\mathbf{k}\sigma} \gamma_{-\mathbf{k}', -\sigma'} - v_{\mathbf{k}}^* v_{\mathbf{k}'} \gamma_{\mathbf{k}\sigma} \gamma_{\mathbf{k}'\sigma'}^\dagger \right).
\end{aligned} \tag{10.109}$$

Assuming $u_{\mathbf{k}}, v_{\mathbf{k}}, \Delta_{\mathbf{k}} \in \mathbb{R}$, we get, up to a constant,

$$\begin{aligned}
H_{\text{NMR}} &= \frac{1}{2} \sum_{\mathbf{k}\mathbf{k}'} \sum_{\sigma\sigma'} B_{\mathbf{k}'\sigma'\mathbf{k}\sigma} \left[(u_{\mathbf{k}'} u_{\mathbf{k}} + v_{\mathbf{k}'} v_{\mathbf{k}}) \left(\gamma_{\mathbf{k}'\sigma'}^\dagger \gamma_{\mathbf{k}\sigma} - \sigma\sigma' \gamma_{-\mathbf{k}, -\sigma}^\dagger \gamma_{-\mathbf{k}', -\sigma'} \right) \right. \\
&\quad \left. + \sigma (u_{\mathbf{k}'} v_{\mathbf{k}} - v_{\mathbf{k}'} u_{\mathbf{k}}) \left(\gamma_{\mathbf{k}'\sigma'}^\dagger \gamma_{-\mathbf{k}, -\sigma}^\dagger - \sigma\sigma' \gamma_{-\mathbf{k}', -\sigma'} \gamma_{\mathbf{k}\sigma} \right) \right].
\end{aligned} \tag{10.110}$$

Compared to ultrasonic attenuation (case I) there is thus a change of sign in both coherence factors. An analogous derivation now gives the interchanged coherence factors

$$F_+(\mathbf{k}, \mathbf{k}') = \frac{1}{2} \left(1 + \frac{\Delta_{\mathbf{k}} \Delta_{\mathbf{k}'}}{E_{\mathbf{k}} E_{\mathbf{k}'}} \right) \tag{10.111}$$

for quasiparticle scattering and

$$F_-(\mathbf{k}, \mathbf{k}') = \frac{1}{2} \left(1 - \frac{\Delta_{\mathbf{k}} \Delta_{\mathbf{k}'}}{E_{\mathbf{k}} E_{\mathbf{k}'}} \right) \tag{10.112}$$

for quasiparticle creation and annihilation. The relevant energy $\hbar\Omega$ is the Zeeman energy of a nuclear spin in the applied uniform magnetic field and is small compared to the gap Δ_0 . Thus we can again restrict ourselves to the small- Ω limit. The derivation is initially analogous to case I, but with F_- replaced by F_+ . The nuclear-spin relaxation rate is

$$\alpha_s \propto \int_{-\infty}^{\infty} d\omega D_s(|\omega|) D_s(|\omega + \Omega|) |B|^2 F_+(\omega, \omega + \Omega) [n_F(\omega) - n_F(\omega + \Omega)] \quad (10.113)$$

with

$$F_+(\omega, \omega') = \frac{1}{2} \left(1 + \frac{\Delta_0^2}{|\omega||\omega'|} \right). \quad (10.114)$$

Thus

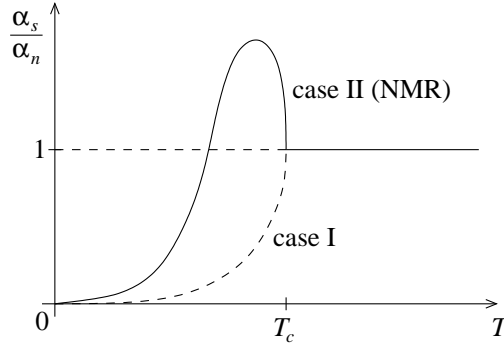
$$\alpha_s \propto D_n^2(E_F) |B|^2 \int_{-\infty}^{\infty} d\omega \frac{D_s(|\omega|)}{D_n(E_F)} \frac{D_s(|\omega + \Omega|)}{D_n(E_F)} \frac{1}{2} \left(1 + \frac{\Delta_0^2}{|\omega||\omega + \Omega|} \right) [n_F(\omega) - n_F(\omega + \Omega)] \quad (10.115)$$

$$\Rightarrow \frac{\alpha_s}{\alpha_n} = \frac{1}{\Omega} \int_{-\infty}^{\infty} d\omega \frac{D_s(|\omega|)}{D_n(E_F)} \frac{D_s(|\omega + \Omega|)}{D_n(E_F)} \frac{|\omega||\omega + \Omega| + \Delta_0^2}{|\omega||\omega + \Omega|} [n_F(\omega) - n_F(\omega + \Omega)]. \quad (10.116)$$

If we now expand the integral for small Ω as above, we encounter a problem:

$$\frac{\alpha_s}{\alpha_n} \cong \frac{1}{\Omega} 2 \int_{\Delta_0}^{\infty} d\omega \left[\frac{|\omega|}{\sqrt{\omega^2 - \Delta_0^2}} \right]^2 \frac{\omega^2 + \Delta_0^2}{\omega^2} (-\Omega) \frac{\partial n_F}{\partial \omega} = -2 \int_{\Delta_0}^{\infty} d\omega \frac{\omega^2 + \Delta_0^2}{\omega^2 - \Delta_0^2} \frac{\partial n_F}{\partial \omega}. \quad (10.117)$$

This integral diverges logarithmically at the lower limit. Keeping a non-zero but realistically small Ω removes the divergence. However, the calculated α_s/α_n is still too large compared to experiments. The origin of this problem is the strong singularity in the superconducting density of states. A \mathbf{k} -dependent gap $\Delta_{\mathbf{k}}$ removes the problem; in a realistic theory $\Delta_{\mathbf{k}}$ always has a \mathbf{k} dependence since it cannot have higher symmetry than the underlying normal dispersion $\xi_{\mathbf{k}}$. Introducing some broadening of the density of states by hand, we numerically find the following temperature dependence:



There is a large maximum below the transition temperature, called the *Hebel-Slichter peak*. It results from the factor $D_s^2(|\omega|)$ (for $\Omega \rightarrow 0$) in the integrand,

$$D_s^2(|\omega|) \cong D_n^2(E_F) \frac{\omega^2}{\omega^2 - \Delta_0^2}, \quad (10.118)$$

which for nuclear relaxation is not canceled by the coherence factor F_+ , whereas for ultrasonic attenuation it is canceled by F_- . Physically, the strong enhancement below T_c of the density of states of both initial and final states at $\omega \gtrsim \Delta_0$ leads to increased nuclear relaxation.

10.6 Ginzburg-Landau-Gor'kov theory

We conclude this chapter by remarking that Lev Gor'kov managed, two years after the publication of BCS theory, to derive Ginzburg-Landau theory from BCS theory. The correspondence is perfect if the gap is sufficiently small, i.e., T is close to T_c , and the electromagnetic field varies slowly on the scale of the Pippard coherence length ξ_0 (see Sec. 5.4). These are indeed the conditions under which Ginzburg and Landau expected their theory to be valid.

Gor'kov used equations of motion for electronic Green functions, which he decoupled with a mean-field-like approximation, which allowed for spatial variations of the decoupling term $\Delta(\mathbf{r})$. The derivation is given in Schrieffer's book and we omit it here. Gor'kov found that in order to obtain the Ginzburg-Landau equations, he had to take

$$q = -2e, \tag{10.119}$$

$$m^* = 2m, \tag{10.120}$$

as anticipated, and (using our conventions)

$$\psi(\mathbf{r}) = \frac{\sqrt{7\zeta(3)}}{4\pi} \sqrt{n_s^0} \frac{\Delta(\mathbf{r})}{k_B T_c}, \tag{10.121}$$

where

$$n_s^0 := \left. \frac{n_s}{1 - \frac{T}{T_c}} \right|_{T \rightarrow T_c^-}. \tag{10.122}$$

Recall that $n_s \sim 1 - T/T_c$ close to T_c in Ginzburg-Landau theory. The spatially dependent gap is thus locally proportional to the Ginzburg-Landau “condensate wavefunction” or order parameter $\psi(\mathbf{r})$.

Since we have already found that London theory is a limiting case of Ginzburg-Landau theory, it is also a limiting case of BCS theory. But London theory predicts the two central properties of superconductors: Ideal conduction and flux expulsion. Thus Gor'kov's derivation also shows that BCS theory indeed describes a superconducting state. (Historically, this has been shown by BCS before Gor'kov established the formal relationship between the various theories.)

11

Josephson effects

Brian Josephson made two important predictions for the current flowing through a tunneling barrier between two superconductors. The results have later been extended to various other systems involving two superconducting electrodes, such as superconductor/normal-metal/superconductor heterostructures and superconducting weak links. Rather generally, for *vanishing* applied voltage a supercurrent I_s is flowing which is related to the phase difference $\Delta\phi$ of the two condensates by

$$I_s = I_c \sin \left(\Delta\phi - \frac{2\pi}{\Phi_0} \int ds \cdot \mathbf{A} \right). \quad (11.1)$$

We will discuss the *critical current* I_c presently. We consider the case without magnetic field so that we can choose the gauge $\mathbf{A} \equiv 0$. Then the Josephson relation simplifies to

$$I_s = I_c \sin \Delta\phi. \quad (11.2)$$

It should be noted that this *DC Josephson effect* is an *equilibrium* phenomenon since no bias voltage is applied. The current thus continues to flow as long as the phase difference $\Delta\phi$ is maintained.

Secondly, Josephson predicted that in the presence of a constant bias voltage V , the phase difference would evolve according to

$$\frac{d}{dt} \Delta\phi = -\frac{2e}{\hbar} V \quad (11.3)$$

(recall that we use the convention $e > 0$) so that an alternating current would flow,

$$I_s(t) = I_c \sin \left(\Delta\phi_0 - \frac{2e}{\hbar} Vt \right). \quad (11.4)$$

This is called the *AC Josephson effect*. The frequency

$$\omega_J := \frac{2eV}{\hbar} \quad (11.5)$$

of the current is called the *Josephson frequency*. The AC Josephson effect relates frequencies (or times) to voltages, which makes it important for metrology.

11.1 The Josephson effects in Ginzburg-Landau theory

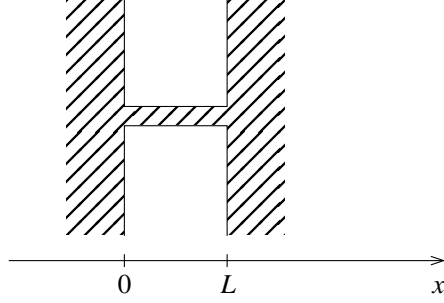
We consider a weak link between two identical bulk superconductors. The weak link is realized by a short wire of length $L \ll \xi$ and cross section A made from the same material as the bulk superconductors. We choose this setup since it is the easiest to treat in Ginzburg-Landau theory since the parameters α and β are uniform, but the only property that really matters is that the phase ϕ of the order parameter $\psi(\mathbf{r})$ only changes within the weak

link. We employ the first Ginzburg-Landau equation for $\mathbf{A} \equiv 0$ assuming $\psi(\mathbf{r})$ to depend only on the coordinate x along the wire,

$$\xi^2 f''(x) + f(x) - f^3(x) = 0 \quad (11.6)$$

with

$$f(x) = \frac{\psi(x)}{|\psi(\pm\infty)|} = \sqrt{-\frac{\alpha}{\beta}} \psi(x). \quad (11.7)$$



We assume the two bulk superconductors to be uniform and to have a relative phase of $\Delta\phi$. This allows us to write

$$f(x) = \begin{cases} 1 & \text{for } x \leq 0, \\ e^{i\Delta\phi} & \text{for } x \geq L. \end{cases} \quad (11.8)$$

For the wire we have to solve Eq. (11.6) with the boundary conditions

$$f(0) = 1, \quad f(L) = e^{i\Delta\phi}. \quad (11.9)$$

Since $L \ll \xi$, the first term in Eq. (11.6) is larger than the other two by a factor of order ξ^2/L^2 , unless $\Delta\phi = 0$, in which case the solution is trivially $f \equiv 1$. It is thus sufficient to solve $f''(x) = 0$, which has the solution

$$f(x) = \frac{L-x}{L} + \frac{x}{L} e^{i\Delta\phi}. \quad (11.10)$$

Inserting $f(x)$ into the second Ginzburg-Landau equation (with $\mathbf{A} \equiv 0$), we obtain

$$\begin{aligned} j_s &= i \frac{q\hbar}{2m^*} [(\psi')^* \psi - \psi^* \psi'] = -i \frac{e\hbar}{2m} \left(-\frac{\alpha}{\beta} \right) [(f')^* f - f^* f'] \\ &= -i \frac{e\hbar}{2m} 2n_s \left[\left(-\frac{1}{L} + \frac{1}{L} e^{-i\Delta\phi} \right) \left(\frac{L-x}{L} + \frac{x}{L} e^{i\Delta\phi} \right) - \left(\frac{L-x}{L} + \frac{x}{L} e^{-i\Delta\phi} \right) \left(-\frac{1}{L} + \frac{1}{L} e^{i\Delta\phi} \right) \right] \\ &= -i \frac{e\hbar n_s}{m} \left[-\frac{x}{L^2} (e^{i\Delta\phi} - e^{-i\Delta\phi}) - \frac{L-x}{L^2} (e^{i\Delta\phi} - e^{-i\Delta\phi}) \right] = 2 \frac{e\hbar n_s}{m} \left[-\frac{x}{L^2} \sin \Delta\phi - \frac{L-x}{L^2} \sin \Delta\phi \right] \\ &= -2 \frac{e\hbar n_s}{mL} \sin \Delta\phi. \end{aligned} \quad (11.11)$$

The current is obviously obtained by integrating over the cross-sectional area,

$$I_s = -\frac{2e\hbar n_s}{m} \frac{A}{L} \sin \Delta\phi, \quad (11.12)$$

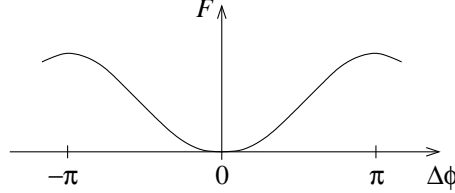
so that we get

$$I_c = -\frac{2e\hbar n_s}{m} \frac{A}{L}. \quad (11.13)$$

The negative sign is due to the negative charge $-2e$ of the Cooper pairs. The amplitude of the current-phase relation is clearly $|I_c|$.

Ginzburg-Landau theory also gives us the free energy of the wire. Since we have neglected the α and β terms when solving the Ginzburg-Landau equation, we must for consistency do the same here,

$$\begin{aligned}
 F &= A \int_0^L dx \frac{\hbar^2}{4m} |\psi'(x)|^2 = A \int_0^L dx \frac{\hbar^2}{4m} \left(-\frac{\alpha}{\beta} \right) \left| -\frac{1}{L} + \frac{1}{L} e^{i\Delta\phi} \right|^2 = A \frac{\hbar^2}{4m} \frac{2n_s}{L^2} \int_0^L dx 2(1 - \cos \Delta\phi) \\
 &= \frac{A}{L} \frac{\hbar^2 n_s}{m} (1 - \cos \Delta\phi).
 \end{aligned}
 \tag{11.14}$$



The free energy is minimal when the phases of the two superconductors coincide. Thus if there existed any mechanism by which the phases could relax, they would approach a state with uniform phase across the junction, a highly plausible result.

We can now also derive the AC Josephson effect. Assuming that the free energy of the junction is only changed by the supercurrent, we have

$$\frac{d}{dt} F = I_s V,
 \tag{11.15}$$

i.e., the electrical power. This relation implies that

$$\frac{\partial F}{\partial \Delta\phi} \frac{d}{dt} \Delta\phi = I_s V
 \tag{11.16}$$

$$\Rightarrow \frac{A}{L} \frac{\hbar^2 n_s}{m} \sin \Delta\phi \frac{d}{dt} \Delta\phi = -\frac{2e\hbar n_s}{m} \frac{A}{L} \sin \Delta\phi V
 \tag{11.17}$$

$$\Rightarrow \frac{d}{dt} \Delta\phi = -\frac{2e}{\hbar} V,
 \tag{11.18}$$

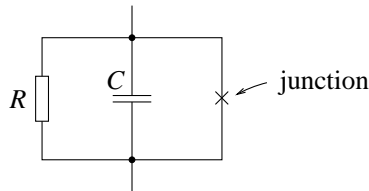
as stated above. Physically, if a supercurrent is flowing in the presence of a bias voltage, it generates power. Since energy is conserved, this power must equal the change of (free) energy per unit time of the junction.

11.2 Dynamics of Josephson junctions

For a discussion of the dynamical current-voltage characteristics of a Josephson junction, it is crucial to realize that a real junction also

1. permits single-particle tunneling (see Sec. 10.4), which we model by an ohmic resistivity R in parallel to the junction,
2. has a non-zero capacitance C .

This leads to the *resistively and capacitively shunted junction (RCSJ) model* represented by the following circuit diagram:



The current through the device is the sum of currents through the three branches,

$$I = \frac{V}{R} + C \frac{dV}{dt} - I_c \sin \Delta\phi, \quad (11.19)$$

where we take $I_c > 0$ and have made the sign explicit. With

$$\frac{d}{dt} \Delta\phi = -\frac{2e}{\hbar} V \quad (11.20)$$

we obtain

$$I = -\frac{\hbar}{2eR} \frac{d}{dt} \Delta\phi - \frac{\hbar C}{2e} \frac{d^2}{dt^2} \Delta\phi - I_c \sin \Delta\phi. \quad (11.21)$$

We introduce the *plasma frequency*

$$\omega_p := \sqrt{\frac{2eI_c}{\hbar C}} \quad (11.22)$$

and the *quality factor*

$$Q := \omega_p RC \quad (11.23)$$

of the junction. This leads to

$$\frac{I}{I_c} = -\frac{1}{\omega_p^2} \frac{d^2}{dt^2} \Delta\phi - \frac{1}{Q\omega_p} \frac{d}{dt} \Delta\phi - \sin \Delta\phi \quad (11.24)$$

and with $\tau := \omega_p t$ finally to

$$\frac{d^2}{d\tau^2} \Delta\phi + \frac{1}{Q} \frac{d}{d\tau} \Delta\phi + \sin \Delta\phi = -\frac{I}{I_c}. \quad (11.25)$$

Compare this equation to the Newton equation for a particle moving in one dimension in a potential $V_{\text{pot}}(x)$ with Stokes friction,

$$m\ddot{x} + \alpha\dot{x} = -\frac{dV_{\text{pot}}}{dx} \quad (11.26)$$

$$\Rightarrow \ddot{x} + \frac{\alpha}{m}\dot{x} = -\frac{1}{m} \frac{dV_{\text{pot}}}{dx}. \quad (11.27)$$

This Newton equation has the same form as the equation of motion of $\Delta\phi$ if we identify

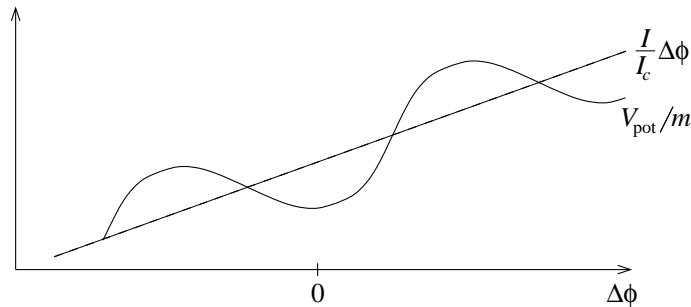
$$t \leftrightarrow \tau, \quad (11.28)$$

$$x \leftrightarrow \Delta\phi, \quad (11.29)$$

$$\frac{\alpha}{m} \leftrightarrow \frac{1}{Q}, \quad (11.30)$$

$$\frac{1}{m} V_{\text{pot}}(x) \leftrightarrow \frac{I}{I_c} \Delta\phi - \cos \Delta\phi. \quad (11.31)$$

Thus the time dependence of $\Delta\phi(\tau)$ corresponds to the damped motion of a particle in a “tilted-washboard” potential



Equation (11.25) can be used to study a Josephson junction in various regimes. First, note that a stationary solution exists as long as $|I| \leq I_c$. Then

$$\sin \Delta\phi = \text{const} = -\frac{I}{I_c} \quad \text{and} \quad V \equiv 0. \quad (11.32)$$

This solution does not exist for $|I| > I_c$. What happens if we impose a time-independent current that is larger than the critical current? We first consider a strongly damped junction, $Q \ll 1$. Then we can neglect the acceleration term and write

$$\frac{1}{Q} \frac{d}{d\tau} \Delta\phi + \sin \Delta\phi = -\frac{I}{I_c} \quad (11.33)$$

$$\Rightarrow \frac{1}{Q} \frac{d}{d\tau} \Delta\phi = -\frac{I}{I_c} - \sin \Delta\phi \quad (11.34)$$

$$\Rightarrow -\frac{d\Delta\phi}{\frac{I}{I_c} + \sin \Delta\phi} = Q d\tau \quad (11.35)$$

$$\Rightarrow Q(\tau - \tau_0) = -\int_0^{\Delta\phi} \frac{d\Delta\phi'}{\frac{I}{I_c} + \sin \Delta\phi'} \stackrel{I \geq I_c}{=} -\frac{2}{\sqrt{\left(\frac{I}{I_c}\right)^2 - 1}} \arctan \frac{1 + \frac{I}{I_c} \tan \frac{\Delta\phi'}{2}}{\sqrt{\left(\frac{I}{I_c}\right)^2 - 1}} \Big|_0^{\Delta\phi}. \quad (11.36)$$

We are interested in periodic solutions for $e^{i\Delta\phi}$ or $\Delta\phi \bmod 2\pi$. One period T is the time it takes for $\Delta\phi$ to change from 0 to -2π (note that $d\Delta\phi/d\tau < 0$). Thus

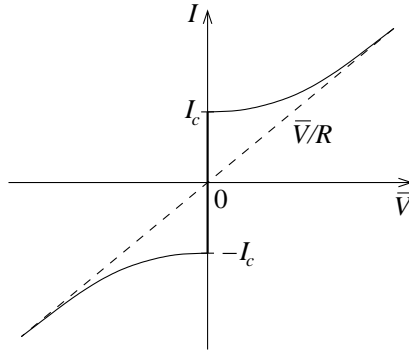
$$Q\omega_p T = -\int_0^{-2\pi} \frac{d\Delta\phi'}{\frac{I}{I_c} + \sin \Delta\phi'} \stackrel{I \geq I_c}{=} -\frac{2\pi}{\sqrt{\left(\frac{I}{I_c}\right)^2 - 1}} \quad (11.37)$$

$$\Rightarrow T = \frac{2\pi}{Q\omega_p} \frac{1}{\sqrt{\left(\frac{I}{I_c}\right)^2 - 1}} = 2\pi \frac{\hbar}{2eI_c R} \frac{1}{\sqrt{\left(\frac{I}{I_c}\right)^2 - 1}} = \frac{\pi\hbar}{eR} \frac{1}{\sqrt{I^2 - I_c^2}}. \quad (11.38)$$

The voltage $V \propto d\Delta\phi/dt$ is of course time-dependent but the *time-averaged voltage* is simply

$$\bar{V} = \frac{1}{T} \int_0^T dt V(t) = -\frac{\hbar}{2e} \frac{1}{T} \int_0^T dt \frac{d}{dt} \Delta\phi = -\frac{\hbar}{2e} \frac{1}{T} \underbrace{[\Delta\phi(T) - \Delta\phi(0)]}_{=-2\pi} = \frac{\pi\hbar}{e} \frac{1}{T} = R \sqrt{I^2 - I_c^2} \quad (11.39)$$

for $I > I_c$. By symmetry, $\bar{V} = -R \sqrt{I^2 - I_c^2}$ for $I < -I_c$. The current-voltage characteristics for given direct current thus look like this:



For $|I| \leq I_c$, the current flows without resistance. At I_c , non-zero DC and AC voltages set in gradually. For $|I| \gg I_c$, the DC voltage approaches the ohmic result for a normal contact.

The solution for general Q requires numerical calculation but we can analyze the opposite case of weak damping, $Q \gg 1$. The stationary solution $\Delta\phi = \text{const}$, $V = 0$ still exists for $I \leq I_c$. The mechanical analogy suggests that the time-dependent solution with periodic $e^{i\Delta\phi}$ will be a very rapid slide down the washboard, overlaid by a small-amplitude oscillation,

$$\Delta\phi \cong -\omega t + \delta\phi, \quad (11.40)$$

where $\omega \gg \omega_p$ and $\delta\phi$ is periodic in time and small. Inserting this ansatz into the equation of motion we find

$$\omega = Q \frac{I}{I_c} \omega_p \gg \omega_p, \quad (11.41)$$

$$\delta\phi \cong - \underbrace{\frac{\omega_p^2}{\omega^2}}_{\ll 1} \sin \omega t. \quad (11.42)$$

Thus

$$\Delta\phi \cong -\frac{\omega}{\omega_p} \tau - \frac{\omega_p^2}{\omega^2} \sin \frac{\omega}{\omega_p} \tau. \quad (11.43)$$

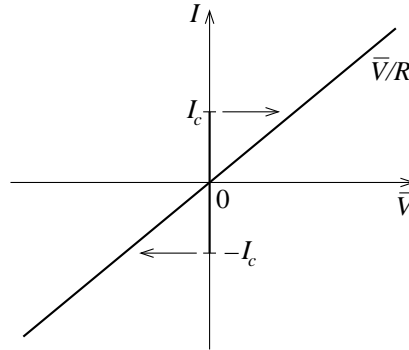
We convince ourselves that this is a good solution for $Q \gg 1$: Inserting it into Eq. (11.25), we obtain for the left-hand side

$$\begin{aligned} & \sin \frac{\omega}{\omega_p} \tau - \frac{1}{Q} \frac{\omega}{\omega_p} - \frac{1}{Q} \frac{\omega_p}{\omega} \cos \frac{\omega}{\omega_p} \tau - \sin \left(\frac{\omega}{\omega_p} \tau + \frac{\omega_p^2}{\omega^2} \sin \frac{\omega}{\omega_p} \tau \right) \\ & \cong \cancel{\sin \frac{\omega}{\omega_p} \tau} - \frac{I}{I_c} - \frac{1}{Q^2} \frac{I_c}{I} \cos \frac{\omega}{\omega_p} \tau - \cancel{\sin \frac{\omega}{\omega_p} \tau} - \left(\cos \frac{\omega}{\omega_p} \tau \right) \frac{1}{Q^2} \left(\frac{I_c}{I} \right)^2 \cos \frac{\omega}{\omega_p} \tau + \mathcal{O} \left(\frac{1}{Q^4} \right). \end{aligned} \quad (11.44)$$

To leading order in $1/Q$ this is just $-I/I_c$, which agrees with the right-hand side. We thus find an averaged voltage of

$$\bar{V} = \frac{1}{T} \int_0^T dt \left(-\frac{\hbar}{2e} \frac{d}{dt} \Delta\phi \right) = \frac{\hbar\omega}{2e} = \frac{\hbar}{2e} Q \frac{I}{I_c} \omega_p = \frac{\hbar}{2e} \frac{2eI_c}{\hbar C} RC \frac{I}{I_c} = RI, \quad (11.45)$$

i.e., the ohmic behavior of the normal junction. Note that the time-dependent solution exists for *all* currents, not just for $|I| > I_c$. Thus for $|I| \leq I_c$ there are now *two* solutions, with $V \equiv 0$ and with $\bar{V} = RI$. If we would change the imposed current we could expect hysteretic behavior. This is indeed observed.



If we instead impose a constant voltage we obtain, for any Q ,

$$\frac{d}{dt} \Delta\phi = -\frac{2e}{\hbar} V = \text{const} \quad \Rightarrow \quad \frac{d^2}{dt^2} \Delta\phi = 0 \quad (11.46)$$

and thus

$$-\frac{1}{Q\omega_p} \frac{2e}{\hbar} V + \sin \left(-\frac{2e}{\hbar} Vt + \Delta\phi_0 \right) = -\frac{I(t)}{I_c} \quad (11.47)$$

$$\Rightarrow I(t) = \frac{I_c}{Q\omega_p} \frac{2e}{\hbar} V + I_c \sin \left(\frac{2e}{\hbar} Vt - \Delta\phi_0 \right). \quad (11.48)$$

The averaged current is just

$$\bar{I} = \frac{I_c}{Q\omega_p} \frac{2e}{\hbar} V = \frac{\hbar C}{2eI_c} \frac{I_c}{RC} \frac{2e}{\hbar} V = \frac{V}{R}. \quad (11.49)$$

Note that this result holds for any damping. It is evidently important to carefully specify whether a constant current or a constant voltage is imposed.

11.3 Bogoliubov-de Gennes Hamiltonian

It is often necessary to describe inhomogeneous systems, Josephson junctions are typical examples. So far, the only theory we know that is able to treat inhomogeneity is the Ginzburg-Landau theory, which has the disadvantage that the quasiparticles are not explicitly included. It is in this sense not a microscopic theory. We will now discuss a microscopic description that allows us to treat inhomogeneous systems. The essential idea is to make the BCS mean-field Hamiltonian spatially dependent. This leads to the *Bogoliubov-de Gennes Hamiltonian*. It is useful to revert to a first-quantized description. To this end, we introduce the condensate state $|\psi_{\text{BCS}}\rangle$ as the ground state of the BCS Hamiltonian

$$H_{\text{BCS}} = \sum_{\mathbf{k}\sigma} \xi_{\mathbf{k}} c_{\mathbf{k}\sigma}^\dagger c_{\mathbf{k}\sigma} - \sum_{\mathbf{k}} \Delta_{\mathbf{k}}^* c_{-\mathbf{k},\downarrow} c_{\mathbf{k}\uparrow} - \sum_{\mathbf{k}} \Delta_{\mathbf{k}} c_{\mathbf{k}\uparrow}^\dagger c_{-\mathbf{k},\downarrow}^\dagger + \text{const.} \quad (11.50)$$

$|\psi_{\text{BCS}}\rangle$ agrees with the BCS ground state defined in Sec. 9.2 in the limit $T \rightarrow 0$ (recall that $\Delta_{\mathbf{k}}$ and thus H_{BCS} is temperature-dependent). We have

$$H_{\text{BCS}} |\psi_{\text{BCS}}\rangle = E_{\text{BCS}} |\psi_{\text{BCS}}\rangle, \quad (11.51)$$

where E_{BCS} is the temperature-dependent energy of the condensate.

Since H_{BCS} is bilinear, it is sufficient to consider single-particle excitations. Many-particle excitations are simply product states, or more precisely Slater determinants, of single-particle excitations. We first define a two-component spinor

$$|\Psi_{\mathbf{k}}\rangle \equiv \begin{pmatrix} |\Psi_{\mathbf{k}1}\rangle \\ |\Psi_{\mathbf{k}2}\rangle \end{pmatrix} := \begin{pmatrix} c_{\mathbf{k}\uparrow}^\dagger \\ c_{-\mathbf{k},\downarrow} \end{pmatrix} |\psi_{\text{BCS}}\rangle. \quad (11.52)$$

It is easy to show that

$$[H_{\text{BCS}}, c_{\mathbf{k}\uparrow}^\dagger] = \xi_{\mathbf{k}} c_{\mathbf{k}\uparrow}^\dagger - \Delta_{\mathbf{k}}^* c_{-\mathbf{k},\downarrow}, \quad (11.53)$$

$$[H_{\text{BCS}}, c_{-\mathbf{k},\downarrow}] = -\xi_{\mathbf{k}} c_{-\mathbf{k},\downarrow} - \Delta_{\mathbf{k}} c_{\mathbf{k}\uparrow}^\dagger. \quad (11.54)$$

With these relations we obtain

$$\begin{aligned} H_{\text{BCS}} |\Psi_{\mathbf{k}1}\rangle &= H_{\text{BCS}} c_{\mathbf{k}\uparrow}^\dagger |\psi_{\text{BCS}}\rangle = \left(\xi_{\mathbf{k}} c_{\mathbf{k}\uparrow}^\dagger - \Delta_{\mathbf{k}}^* c_{-\mathbf{k},\downarrow} + c_{\mathbf{k}\uparrow}^\dagger H_{\text{BCS}} \right) |\psi_{\text{BCS}}\rangle \\ &= (E_{\text{BCS}} + \xi_{\mathbf{k}}) |\Psi_{\mathbf{k}1}\rangle - \Delta_{\mathbf{k}}^* |\Psi_{\mathbf{k}2}\rangle \end{aligned} \quad (11.55)$$

and

$$\begin{aligned} H_{\text{BCS}} |\Psi_{\mathbf{k}2}\rangle &= H_{\text{BCS}} c_{-\mathbf{k},\downarrow} |\psi_{\text{BCS}}\rangle = \left(-\xi_{\mathbf{k}} c_{-\mathbf{k},\downarrow} - \Delta_{\mathbf{k}} c_{\mathbf{k}\uparrow}^\dagger + c_{-\mathbf{k},\downarrow} H_{\text{BCS}} \right) |\psi_{\text{BCS}}\rangle \\ &= (E_{\text{BCS}} - \xi_{\mathbf{k}}) |\Psi_{\mathbf{k}2}\rangle - \Delta_{\mathbf{k}} |\Psi_{\mathbf{k}1}\rangle. \end{aligned} \quad (11.56)$$

Thus for the basis $\{|\Psi_{\mathbf{k}1}\rangle, |\Psi_{\mathbf{k}2}\rangle\}$ the Hamiltonian has the matrix form

$$\begin{pmatrix} E_{\text{BCS}} + \xi_{\mathbf{k}} & -\Delta_{\mathbf{k}} \\ -\Delta_{\mathbf{k}}^* & E_{\text{BCS}} - \xi_{\mathbf{k}} \end{pmatrix}. \quad (11.57)$$

This is the desired Hamiltonian in first-quantized form, except that we want to measure excitation energies relative to the condensate energy. Thus we write as the first-quantized Hamiltonian in \mathbf{k} space

$$\mathcal{H}_{\text{BdG}}(\mathbf{k}) = \begin{pmatrix} \xi_{\mathbf{k}} & -\Delta_{\mathbf{k}} \\ -\Delta_{\mathbf{k}}^* & -\xi_{\mathbf{k}} \end{pmatrix}. \quad (11.58)$$

This is the Bogoliubov-de Gennes Hamiltonian for non-magnetic superconductors. Its eigenvalues are

$$\pm \sqrt{\xi_{\mathbf{k}}^2 + |\Delta_{\mathbf{k}}|^2} = \pm E_{\mathbf{k}} \quad (11.59)$$

with corresponding eigenstates

$$u_{\mathbf{k}} |\Psi_{\mathbf{k}1}\rangle - v_{\mathbf{k}}^* |\Psi_{\mathbf{k}2}\rangle = \left(u_{\mathbf{k}} c_{\mathbf{k}\uparrow}^\dagger - v_{\mathbf{k}}^* c_{-\mathbf{k},\downarrow} \right) |\psi_{\text{BCS}}\rangle = \gamma_{\mathbf{k}\uparrow}^\dagger |\psi_{\text{BCS}}\rangle \quad (11.60)$$

and

$$v_{\mathbf{k}} |\Psi_{\mathbf{k}1}\rangle + u_{\mathbf{k}}^* |\Psi_{\mathbf{k}2}\rangle = \left(v_{\mathbf{k}} c_{\mathbf{k}\uparrow}^\dagger + u_{\mathbf{k}}^* c_{-\mathbf{k},\downarrow} \right) |\psi_{\text{BCS}}\rangle = \gamma_{-\mathbf{k},\downarrow} |\psi_{\text{BCS}}\rangle \quad (11.61)$$

with $u_{\mathbf{k}}, v_{\mathbf{k}}$ defined as above. (A lengthy but straightforward calculation has been omitted.) We can now understand why the second eigenvalue comes out negative: The corresponding eigenstate contains a quasiparticle *annihilation* operator, not a *creation* operator. Hence, $\mathcal{H}_{\text{BdG}}(\mathbf{k})$ reproduces the excitation energies we already know.

The next step is to Fourier-transform the Hamiltonian to obtain its real-space representation, which we write as

$$\mathcal{H}_{\text{BdG}}(\mathbf{r}) := \frac{1}{N} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}} \mathcal{H}_{\text{BdG}}(\mathbf{k}) = \begin{pmatrix} H_0(\mathbf{r}) & -\Delta(\mathbf{r}) \\ -\Delta^*(\mathbf{r}) & -H_0(\mathbf{r}) \end{pmatrix}, \quad (11.62)$$

where we expect

$$H_0(\mathbf{r}) = -\frac{\hbar^2}{2m} \nabla^2 - \mu + V(\mathbf{r}) \quad (11.63)$$

as the free-electron Hamiltonian. But in this form it becomes easy to include spatially inhomogeneous situations: Both $V(\mathbf{r})$ and $\Delta(\mathbf{r})$ can be chosen spatially dependent (and not simply lattice-periodic). The corresponding Schrödinger equation

$$\mathcal{H}_{\text{BdG}}(\mathbf{r}) \Psi(\mathbf{r}) = E \Psi(\mathbf{r}) \quad (11.64)$$

with

$$\Psi(\mathbf{r}) = \begin{pmatrix} \Psi_1(\mathbf{r}) \\ \Psi_2(\mathbf{r}) \end{pmatrix} \quad (11.65)$$

is called the *Bogoliubov-de Gennes equation*. Note that in this context the gap $\Delta(\mathbf{r})$ is usually defined with the opposite sign, which is just a phase change, so that the explicit minus signs in the off-diagonal components of \mathcal{H}_{BdG} are removed. Furthermore, in Bogoliubov-de Gennes theory, the gap function is typically not evaluated selfconsistently from the averages $\langle c_{-\mathbf{k},\downarrow} c_{\mathbf{k}\uparrow} \rangle$. Rather, $\Delta(\mathbf{r})$ is treated as a given function characterizing the tendency of superconducting pairing.

11.4 Andreev reflection

As an application of the Bogoliubov-de Gennes approach, we study what happens to an electron that impinges on a normal-superconducting interface from the normal side. We model this situation by the Bogoliubov-de Gennes Hamiltonian

$$\mathcal{H}_{\text{BdG}} = \begin{pmatrix} -\frac{\hbar^2}{2m} \nabla^2 - \mu & \Delta_0 \Theta(x) \\ \Delta_0 \Theta(x) & \frac{\hbar^2}{2m} \nabla^2 + \mu \end{pmatrix} \quad (11.66)$$

(note the changed sign of $\Delta(\mathbf{r})$) so that

$$\mathcal{H}_{\text{BdG}} \Psi(\mathbf{r}) = E \Psi(\mathbf{r}). \quad (11.67)$$

In the normal region, $x < 0$, the two components $\Psi_1(\mathbf{r}), \Psi_2(\mathbf{r})$ are just superpositions of plane waves with wave vectors $\mathbf{k}_1, \mathbf{k}_2$ that must satisfy

$$k_1^2 = 2m(\mu + E) = k_F^2 + 2mE, \quad (11.68)$$

$$k_2^2 = 2m(\mu - E) = k_F^2 - 2mE, \quad (11.69)$$

where $\hbar = 1$. In the superconductor, $x > 0$, we have

$$\left(-\frac{1}{2m}\nabla^2 - \mu\right)\Psi_1(\mathbf{r}) + \Delta_0\Psi_2(\mathbf{r}) = E\Psi_1(\mathbf{r}), \quad (11.70)$$

$$\left(\frac{1}{2m}\nabla^2 + \mu\right)\Psi_2(\mathbf{r}) + \Delta_0\Psi_1(\mathbf{r}) = E\Psi_2(\mathbf{r}) \quad (11.71)$$

$$\Rightarrow \Psi_2(\mathbf{r}) = \frac{E + \frac{1}{2m}\nabla^2 + \mu}{\Delta_0}\Psi_1(\mathbf{r}) \quad (11.72)$$

$$\Rightarrow \left(\frac{1}{2m}\nabla^2 + \mu\right)^2\Psi_1(\mathbf{r}) = (E^2 - \Delta_0^2)\Psi_1(\mathbf{r}) \quad (11.73)$$

and analogously

$$\left(\frac{1}{2m}\nabla^2 + \mu\right)^2\Psi_2(\mathbf{r}) = (E^2 - \Delta_0^2)\Psi_2(\mathbf{r}). \quad (11.74)$$

If the energy is above the gap, $|E| > \Delta_0$, the solutions are again plain wave vectors $\mathbf{q}_1, \mathbf{q}_2$, where now

$$\left(\frac{q_{1,2}^2}{2m} - \mu\right)^2 = E^2 - \Delta_0^2 > 0 \quad (11.75)$$

$$\Rightarrow q_{1,2}^2 = 2m\left(\mu + \sqrt{E^2 - \Delta_0^2}\right), \quad (11.76)$$

and amplitudes coupled by Eq. (11.72).

We are here interested in the more surprising case $|E| < \Delta_0$. Since the solution must be continuous across the interface and is plane-wave-like in the normal region, we make the ansatz

$$\Psi_1(\mathbf{r}) = e^{i(k_{1y}y + k_{1z}z)}\Phi_1(x), \quad (11.77)$$

from which

$$\left(-\frac{k_{1\parallel}^2}{2m} + \frac{1}{2m}\frac{d^2}{dx^2} + \mu\right)^2\Phi_1(x) = (E^2 - \Delta_0^2)\Phi_1(x) \quad (11.78)$$

with $\mathbf{k}_{1\parallel} := (k_{1y}, k_{1z})$. Since this equation is linear with constant coefficients, we make an exponential ansatz

$$\Phi_1(x) = e^{\kappa x + iqx} \quad (11.79)$$

with $\kappa, q \in \mathbb{R}$. This leads to

$$\left(-\frac{k_{1\parallel}^2}{2m} + \frac{(\kappa + iq)^2}{2m} + \mu\right)^2 = E^2 - \Delta_0^2 < 0 \quad (11.80)$$

$$\begin{aligned} \Rightarrow \left(-\frac{k_{1\parallel}^2}{2m} + q^2 + \mu + \frac{i\kappa q}{m} + \frac{\kappa^2}{2m}\right)^2 &= \left(\frac{k_{1\parallel}^2}{2m} - \mu - \frac{\kappa^2}{2m} - \frac{i\kappa q}{m}\right)^2 \\ &= \left(\frac{k_{1\parallel}^2}{2m} + q^2 - \mu - \frac{\kappa^2}{2m}\right)^2 - \frac{\kappa^2 q^2}{m^2} - 2\frac{i\kappa q}{m}\left(\frac{k_{1\parallel}^2}{2m} - \mu - \frac{\kappa^2}{2m}\right) = E^2 - \Delta_0^2. \end{aligned} \quad (11.81)$$

Since the right-hand side is real, we require

$$\frac{k_{1\parallel}^2}{2m} + q^2 - \mu - \frac{\kappa^2}{2m} = 0 \quad (11.82)$$

$$\Rightarrow \kappa^2 = k_{1\parallel}^2 + q^2 - 2m\mu. \quad (11.83)$$

For the real part it follows that

$$-\frac{\kappa^2 q^2}{m^2} = E^2 - \Delta_0^2 \quad (11.84)$$

$$\Rightarrow \frac{\kappa^2 q^2}{m^2} = \frac{(k_{1\parallel}^2 + q^2 - 2m\mu) q^2}{m^2} = \Delta_0^2 - E^2 > 0 \quad (11.85)$$

$$\Rightarrow q^4 + (k_{1\parallel}^2 - 2m\mu) q^2 - m^2(\Delta_0^2 - E^2) = 0 \quad (11.86)$$

$$\begin{aligned} \Rightarrow q^2 &= \frac{2m\mu - k_{1\parallel}^2}{2} \pm \sqrt{\left(\frac{2m\mu - k_{1\parallel}^2}{2}\right)^2 + m^2(\Delta_0^2 - E^2)} \\ &= \frac{1}{2} \left[k_F^2 - k_{1\parallel}^2 \pm \sqrt{(k_F^2 - k_{1\parallel}^2)^2 + 4m^2(\Delta_0^2 - E^2)} \right]. \end{aligned} \quad (11.87)$$

Both solutions are clearly real but the one with the minus sign is negative so that q would be imaginary, contrary to our assumption. Thus the relevant solutions are

$$q = \pm q_1 := \pm \frac{1}{\sqrt{2}} \sqrt{k_F^2 - k_{1\parallel}^2 + \sqrt{(k_F^2 - k_{1\parallel}^2)^2 + 4m^2(\Delta_0^2 - E^2)}}. \quad (11.88)$$

From this we get

$$\begin{aligned} \kappa^2 &= k_{1\parallel}^2 - k_F^2 + \frac{1}{2} \left[k_F^2 - k_{1\parallel}^2 + \sqrt{(k_F^2 - k_{1\parallel}^2)^2 + 4m^2(\Delta_0^2 - E^2)} \right] \\ &= \frac{1}{2} \left[k_{1\parallel}^2 - k_F^2 + \sqrt{(k_F^2 - k_{1\parallel}^2)^2 + 4m^2(\Delta_0^2 - E^2)} \right] \end{aligned} \quad (11.89)$$

and

$$\kappa = -\kappa_1 := -\frac{1}{\sqrt{2}} \sqrt{k_{1\parallel}^2 - k_F^2 + \sqrt{(k_F^2 - k_{1\parallel}^2)^2 + 4m^2(\Delta_0^2 - E^2)}}. \quad (11.90)$$

The positive root exists but would lead to a solution that grows exponentially for $x \rightarrow \infty$. For $\Psi_2(\mathbf{r})$ the derivation is completely analogous. However, $\Psi_1(\mathbf{r})$ and $\Psi_2(\mathbf{r})$ are related by Eq. (11.72), which for exponential functions becomes a simple proportionality. Therefore, we must have $\mathbf{k}_{1\parallel} = \mathbf{k}_{2\parallel}$, which already implies $q_1 = q_2$ and $\kappa_1 = \kappa_2$. Then we have

$$\begin{aligned} \Psi_2(\mathbf{r}) &= \frac{E + \frac{1}{2m}\nabla^2 + \mu}{\Delta_0} \Psi_1(\mathbf{r}) = \frac{1}{\Delta_0} \left[E + \frac{(-\kappa_1 \pm iq_1)^2}{2m} - \frac{k_{1\parallel}^2}{2m} + \mu \right] \Psi_1(\mathbf{r}) \\ &= \frac{1}{\Delta_0} \left[E - \underbrace{\frac{k_{1\parallel}^2 + q^2}{2m}}_{=0} + \mu + \frac{\kappa_1^2}{2m} \mp i \frac{\kappa_1 q_1}{m} \right] \Psi_1(\mathbf{r}). \end{aligned} \quad (11.91)$$

Since we already know that

$$\frac{\kappa_1^2 q_1^2}{m^2} = \Delta_0^2 - E^2 \quad (11.92)$$

and κ_1, q_1 have been defined as positive, we get

$$\Psi_2(\mathbf{r}) = \frac{E \mp i\sqrt{\Delta_0^2 - E^2}}{\Delta_0} \Psi_1(\mathbf{r}). \quad (11.93)$$

We now write down an ansatz and show that it satisfies the Bogoliubov-de Gennes equation and the continuity conditions at the interface. The ansatz reads

$$\Psi(\mathbf{r}) = \begin{pmatrix} e^{i\mathbf{k}_1 \cdot \mathbf{r}} + r e^{i\tilde{\mathbf{k}}_1 \cdot \mathbf{r}} \\ a e^{i\mathbf{k}_2 \cdot \mathbf{r}} \end{pmatrix} \quad \text{for } x \leq 0, \quad (11.94)$$

with $\tilde{\mathbf{k}}_1 := (-k_{1x}, k_{1y}, k_{1z})$ and $\mathbf{k}_2 = (k_{2x}, k_{1y}, k_{1z})$ with $k_{2x} > 0$, where $k_{1x}^2 + k_{1\parallel}^2 = k_F^2 + 2mE$ and $k_{2x}^2 + k_{1\parallel}^2 = k_F^2 - 2mE$, and

$$\Psi(\mathbf{r}) = e^{-\kappa_1 x} \begin{pmatrix} \alpha_+ e^{i(q_1 x + k_{1y} y + k_{1z} z)} + \alpha_- e^{i(-q_1 x + k_{1y} y + k_{1z} z)} \\ \beta_+ e^{i(q_1 x + k_{1y} y + k_{1z} z)} + \beta_- e^{i(-q_1 x + k_{1y} y + k_{1z} z)} \end{pmatrix} \quad \text{for } x \geq 0. \quad (11.95)$$

Note that $\tilde{\mathbf{k}}_1$ is the wave vector of a specularly reflected electron. From Eq. (11.93) we get

$$\beta_{\pm} = \frac{E \mp i\sqrt{\Delta_0^2 - E^2}}{\Delta_0} \alpha_{\pm}. \quad (11.96)$$

From the continuity of Ψ_1, Ψ_2 , and their x -derivatives we obtain

$$1 + r = \alpha_+ + \alpha_-, \quad (11.97)$$

$$a = \beta_+ + \beta_-, \quad (11.98)$$

$$ik_{1x} - r ik_{1x} = \alpha_+(-\kappa_1 + iq_1) + \alpha_-(-\kappa_1 - iq_1), \quad (11.99)$$

$$a ik_{2x} = \beta_+(-\kappa_1 + iq_1) + \beta_-(-\kappa_1 - iq_1). \quad (11.100)$$

We thus have six coupled linear equations for the six unknown coefficients $r, a, \alpha_+, \alpha_-, \beta_+, \beta_-$. The equations are linearly independent so that they have a unique solution, which we can obtain by standard methods. The six coefficients are generally non-zero and complex. We do not give the lengthy expressions here but discuss the results physically.

- The solution in the superconductors decays exponentially, which is reasonable since the energy lies in the superconducting gap.
- In the normal region there is a specularly reflected electron wave (coefficient r), which is also expected. So far, the same results would be obtained for a simple potential step. However, explicit evaluation shows that in general $|r|^2 < 1$, i.e., not all electrons are reflected.
- There is also a term

$$\Psi_2(\mathbf{r}) = a e^{i\mathbf{k}_2 \cdot \mathbf{r}} \quad \text{for } x \leq 0. \quad (11.101)$$

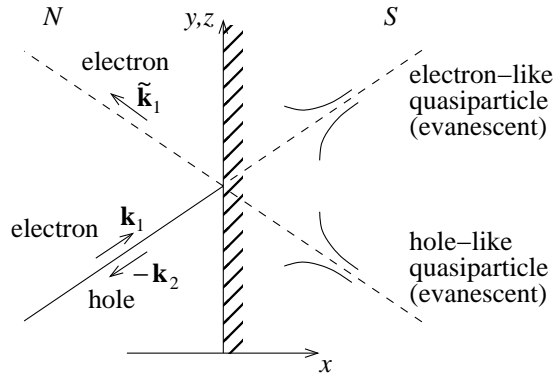
Recall that the second spinor component was defined by

$$|\Psi_{\mathbf{k}_2}\rangle = c_{-\mathbf{k}, \downarrow} |\psi_{\text{BCS}}\rangle. \quad (11.102)$$

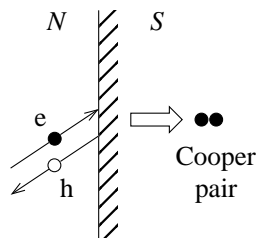
Hence, the above term represents a spin-down hole with wave vector $-\mathbf{k}_2$. Now $\mathbf{k}_{1\parallel} = \mathbf{k}_{2\parallel}$ and $k_{1x}^2 = k_F^2 - k_{1\parallel}^2 + 2mE$ and $k_{2x}^2 = k_F^2 - k_{1\parallel}^2 - 2mE$. But the last terms $\pm 2mE$ are small since

$$|E| < \Delta_0 \ll \mu = \frac{k_F^2}{2m} \quad (11.103)$$

in conventional superconductors. Thus $|k_{2x} - k_{1x}|$ is small and the hole is traveling nearly in the opposite direction compared to the incoming electron wave. This phenomenon is called *Andreev reflection*.



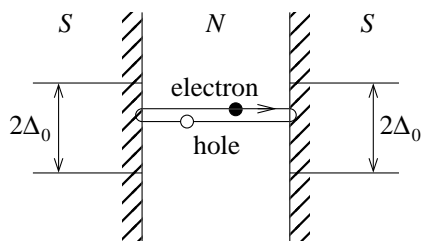
Since not all electrons are reflected and in addition some holes are generated, where does the missing charge go? The quasiparticle states in the superconductor are evanescent and thus cannot accommodate the missing charge. The only possible explanation is that the charge is added to the superconducting condensate, i.e., that additional Cooper pairs are formed. (The whole process can also run backwards, in which case Cooper pairs are removed.) Recall that the condensate does not have a sharp electron number and can therefore absorb or emit electrons without changing the state. But it can only absorb or emit electrons in pairs. The emerging picture is that if an incoming electron is not specularly reflected, a Cooper pair is created, which requires a second electron. This second electron is taken from the normal region, creating a hole, which, as we have seen, travels in the direction the original electron was coming from.



Andreev bound states

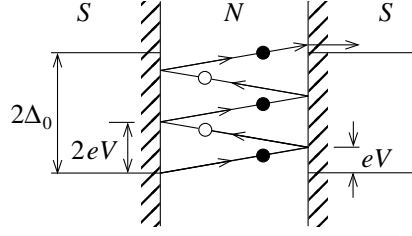
An interesting situation arises if a normal region is delimited by superconductors on two sides. We here only qualitatively consider a superconductor-normal-superconductor (SNS) hetero structure. Similar effects can also occur for example in the normal core of a vortex.

If no voltage is applied between the two superconductors, an electron in the normal region, with energy within the gap, is Andreev reflected as a hole at one interface. It is then Andreev reflected as an electron at the other interface. It is plausible that multiple reflections can lead to the formation of bound states. The real physics is somewhat more complicated since the electron is also partially specularly reflected as an *electron*. It is conceptually clear, though, how to describe Andreev bound states within the Bogoliubov-de Gennes formalism: We just have to satisfy continuity conditions for both interfaces.



If Andreev reflection dominates, as assumed for the sketch above, a Cooper pair is emitted into the right superconductor for every reflection at the right interface. Conversely, a Cooper pair is absorbed from the left superconductor for every reflection at the left interface. This corresponds to a supercurrent through the device. Andreev bound states thus offer a microscopic description of the Josephson effect in superconductor-normal-superconductor junctions.

If we apply a voltage V , the situation changes dramatically: If an electron moving, say, to the right, increases its kinetic energy by eV due to the bias voltage, an Andreev reflected hole traveling to the left *also increases* its kinetic energy by eV since it carries the opposite charge. An electron/hole Andreev-reflected multiple times can thus gain arbitrarily high energies for any non-vanishing bias voltage.



In particular, an electron-like quasiparticle from an occupied state below the gap in, say, the left superconductor can after multiple reflections emerge in a previously unoccupied state above the gap in the right superconductor. A new transport channel becomes available whenever the full gap $2\Delta_0$ is an odd integer multiple of eV :

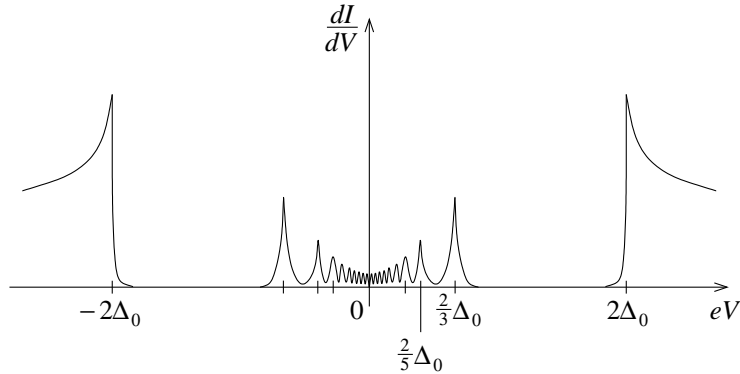
$$2\Delta_0 = (2n + 1) eV, \quad n = 0, 1, \dots \quad (11.104)$$

$$\Rightarrow eV = \frac{\Delta_0}{n + \frac{1}{2}}, \quad n = 0, 1, \dots \quad (11.105)$$

The case $n = 0$ corresponds to direct quasiparticle transfer from one superconductor to the other, similar to quasiparticle tunneling in a superconductor-insulator-superconductor junction. The opening of new transport channels for $n = 0, 1, \dots$, i.e., at

$$eV = \frac{2}{3} \Delta_0, \frac{2}{5} \Delta_0, \frac{2}{7} \Delta_0, \dots \quad (11.106)$$

leads to structures in the current-voltage characteristics *below* the gap, specifically to peaks in the differential conductance dI/dV .



12

Unconventional pairing

In this chapter we first discuss why interactions different from the phonon-mediated one might lead to unconventional pairing, that is to a gap function $\Delta_{\mathbf{k}}$ with non-trivial \mathbf{k} dependence. Then we will briefly consider the origin of such interactions.

12.1 The gap equation for unconventional pairing

We will still use the BCS gap equation even when discussing unconventional superconductors. While the BCS mean-field theory is not quantitatively correct in such cases, it will give a clear understanding of why the gap $\Delta_{\mathbf{k}}$ can have non-trivial symmetry. To get started, we briefly review results from BCS theory. The screened effective interaction was derived in Sec. 8.4,

$$V_{\text{eff}}^{\text{RPA}}(\mathbf{q}, i\nu_n) = V_C^{\text{RPA}}(\mathbf{q}, i\nu_n) \frac{(i\nu_n)^2}{(i\nu_n)^2 - \omega_{\mathbf{q}}^2(i\nu_n)}, \quad (12.1)$$

where V_C^{RPA} is the screened Coulomb interaction and $\omega_{\mathbf{q}}$ is the renormalized phonon dispersion. The retarded interaction at small frequencies is

$$V_{\text{eff}}^R(\mathbf{q}, i\nu_n) \cong 4\pi \frac{e^2}{q^2 + \kappa_s^2} \frac{\nu^2}{\nu^2 - \omega_{\mathbf{q}}^R(\nu)^2 + i0^+ \text{sgn } \nu}, \quad (12.2)$$

where κ_s is the inverse screening length. The behavior at small distances r and small but non-zero frequency ν is determined by V_{eff}^R at large q , where $\omega_{\mathbf{q}}$ can be approximated by the Debye frequency. The interaction is thus attractive and decays like $1/r$ for small r . The interaction is strongest at the same site in a tight-binding model. In order to understand the physics, it makes sense to replace the interaction by a simplified one that is completely local (attractive Hubbard model) or, equivalently, constant in \mathbf{k} space, as we have done above. However, the BCS gap equation

$$\Delta_{\mathbf{k}} = -\frac{1}{N} \sum_{\mathbf{k}'} V_{\mathbf{k}\mathbf{k}'} \frac{\Delta_{\mathbf{k}'}}{2E_{\mathbf{k}'}} [1 - n_F(E_{\mathbf{k}'})] \quad (12.3)$$

is in fact much more general. In the gap equation, $V_{\mathbf{k}\mathbf{k}'}$ describes the amplitude for scattering of two electrons with momenta \mathbf{k}' and $-\mathbf{k}'$ and opposite spins into states with momenta \mathbf{k} and $-\mathbf{k}$.

Let us first consider the case that the interaction is local in real space (flat in \mathbf{k} space) but repulsive. This would apply if the phonons were for some reason ineffective in overscreening the Coulomb interaction. Then we obtain

$$\Delta_{\mathbf{k}} = -\frac{1}{N} \sum_{\mathbf{k}'} V_0 \frac{\Delta_{\mathbf{k}'}}{2E_{\mathbf{k}'}} [1 - n_F(E_{\mathbf{k}'})] \quad (12.4)$$

with $V_0 > 0$. The right-hand side is clearly independent of \mathbf{k} so that we have $\Delta_{\mathbf{k}} = \Delta_0$ and can cancel a factor of Δ_0 if it is non-zero:

$$1 = -\frac{V_0}{N} \sum_{\mathbf{k}'} \frac{1 - n_F(E_{\mathbf{k}'})}{2E_{\mathbf{k}'}}. \quad (12.5)$$

But now the right-hand side is always negative. Consequently, there is no non-trivial solution and thus no superconductivity for a \mathbf{k} -independent repulsion.

Now let us look at a strong interaction between nearest-neighbor sites. We consider a two-dimensional square lattice for simplicity and since it is thought to be a good model for the cuprates. In momentum space, a nearest-neighbor interaction is written as

$$V_{\mathbf{k}\mathbf{k}'} = 2V_1 [\cos(k_x - k'_x)a + \cos(k_y - k'_y)a], \quad (12.6)$$

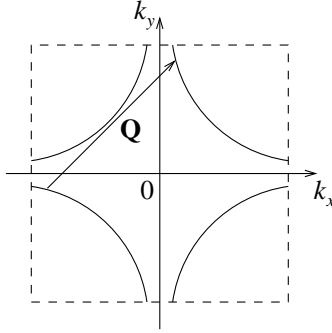
where $V_1 > 0$ ($V_1 < 0$) for a repulsive (attractive) interaction.

The electronic properties of a superconductor are most strongly affected by the states close to the normal-state Fermi surface since elsewhere $|\Delta_{\mathbf{k}}|$ is small compared to the normal-state energy $|\zeta_{\mathbf{k}}|$. We therefore concentrate on these states. Also note that the coupling to states away from the Fermi surface is suppressed in the gap equation by the factor $\Delta_{\mathbf{k}'}/E_{\mathbf{k}'}$. This is a relatively weak, power-law suppression.

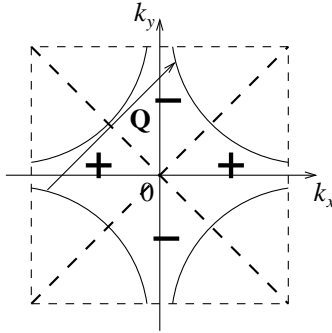
Let us first consider the repulsive case $V_1 > 0$. Then $V_{\mathbf{k}\mathbf{k}'}$ is most strongly repulsive (positive) for momentum transfer $\mathbf{k} - \mathbf{k}' \rightarrow 0$. But $\Delta_{\mathbf{k}}$ should be a smooth function of \mathbf{k} , thus for \mathbf{k} and \mathbf{k}' close together, $\Delta_{\mathbf{k}}$ and $\Delta_{\mathbf{k}'}$ are also similar. In particular, $\Delta_{\mathbf{k}}$ will rarely change its sign between \mathbf{k} and \mathbf{k}' . Consequently, the right-hand side of the gap equation always contains a large contribution with sign *opposite* to that of $\Delta_{\mathbf{k}}$, coming from the sum over \mathbf{k}' close to \mathbf{k} . Hence, a repulsive nearest-neighbor interaction is unlikely to lead to superconductivity.

For the attractive case, $V_1 < 0$, $V_{\mathbf{k}\mathbf{k}'}$ is most strongly attractive for $\mathbf{k} - \mathbf{k}' \rightarrow 0$ and most strongly repulsive for $\mathbf{k} - \mathbf{k}' \rightarrow (\pi/a, \pi/a)$ and equivalent points in the Brillouin zone. The attraction at small $\mathbf{q} = \mathbf{k} - \mathbf{k}'$ is always favorable for superconductivity. However, we also have an equally strong repulsion around $\mathbf{q} \approx (\pi/a, \pi/a)$. A critical situation thus arises if both \mathbf{k} and \mathbf{k}' lie close to the Fermi surface and their difference is close to $(\pi/a, \pi/a)$. The central insight is that this can still help superconductivity if the gaps $\Delta_{\mathbf{k}}$ and $\Delta_{\mathbf{k}'}$ at \mathbf{k} and \mathbf{k}' , respectively, have *opposite sign*. In this case the contribution to the right-hand side of the gap equation from such \mathbf{k}' has the same sign as $\Delta_{\mathbf{k}}$ since $V_{\mathbf{k}\mathbf{k}'} > 0$ and there is an explicit minus sign.

This effect is crucial in the cuprates, which do have an effective attractive nearest-neighbor interaction and have a large normal-state Fermi surface shown here for a two-dimensional model:



The vector \mathbf{Q} in the sketch is $\mathbf{Q} = (\pi/a, \pi/a)$. Following the previous discussion, $\Delta_{\mathbf{k}}$ close to the Fermi surface should have different sign between points separated by \mathbf{Q} . On the other hand, the small- \mathbf{q} attraction favors gaps $\Delta_{\mathbf{k}}$ that change sign “as little as possible.” By inspection, these conditions are met by a gap changing sign on the diagonals:



This type of gap is called a $d_{x^2-y^2}$ -wave (or just d -wave) gap since it has the symmetry of a $d_{x^2-y^2}$ -orbital (though in \mathbf{k} -space, not in real space). The simplest gap function with this symmetry and consistent with the lattice structure is

$$\Delta_{\mathbf{k}} = \Delta_0 (\cos k_x a - \cos k_y a). \quad (12.7)$$

Recall that the gap function away from the Fermi surface is of limited importance.

The d -wave gap $\Delta_{\mathbf{k}}$ is distinct from the conventional, approximately constant s -wave gap in that it has zeroes on the Fermi surface. These zeroes are called *gap nodes*. In the present case they appear in the (11) and equivalent directions. The quasiparticle dispersion in the vicinity of such a node is

$$E_{\mathbf{k}} = \sqrt{\xi_{\mathbf{k}}^2 + |\Delta_{\mathbf{k}}|^2} \cong \sqrt{(\epsilon_{\mathbf{k}} - \mu)^2 + \Delta_0^2 (\cos k_x a - \cos k_y a)^2}. \quad (12.8)$$

One node is at

$$\mathbf{k}_0 = \frac{k_F}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \quad (12.9)$$

Writing $\mathbf{k} = \mathbf{k}_0 + \mathbf{q}$ and expanding for small \mathbf{q} , we obtain

$$E_{\mathbf{k}_0+\mathbf{q}} \cong \sqrt{(\mathbf{v}_F \cdot \mathbf{q})^2 + \Delta_0^2 [-(\sin k_x^0 a)q_x a + (\sin k_y^0 a)q_y a]^2}, \quad (12.10)$$

where

$$\mathbf{v}_F := \left. \frac{\partial \epsilon_{\mathbf{k}}}{\partial \mathbf{k}} \right|_{\mathbf{k}=\mathbf{k}_0} = \frac{v_F}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad (12.11)$$

is the normal-state Fermi velocity at the node. Thus

$$E_{\mathbf{k}_0+\mathbf{q}} \cong \sqrt{(\mathbf{v}_F \cdot \mathbf{q})^2 + \Delta_0^2 \left[\left(-a \sin \frac{k_F a}{\sqrt{2}}, a \sin \frac{k_F a}{\sqrt{2}} \right) \cdot \mathbf{q} \right]^2} = \sqrt{(\mathbf{v}_F \cdot \mathbf{q})^2 + (\mathbf{v}_{\text{qp}} \cdot \mathbf{q})^2}, \quad (12.12)$$

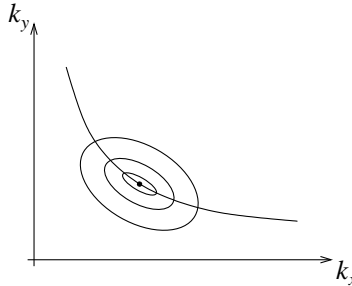
where

$$\mathbf{v}_{\text{qp}} := \Delta_0 a \sin \frac{k_F a}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix} \perp \mathbf{v}_F. \quad (12.13)$$

Thus the quasiparticle dispersion close to the node is a *cone* like for massless relativistic particles, but with different velocities in the directions normal and tangential to the Fermi surface. Usually one finds

$$v_F > v_{\text{qp}}. \quad (12.14)$$

The sketch shows equipotential lines of $E_{\mathbf{k}}$.

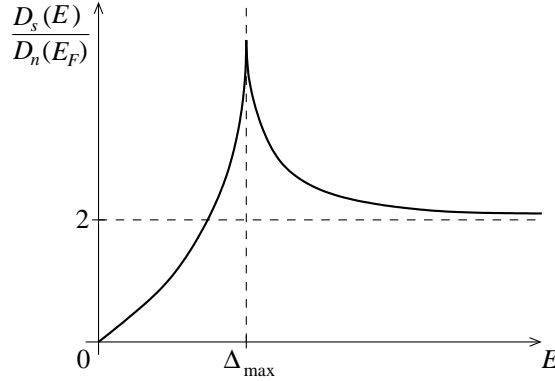


The fact that the gap closes at some \mathbf{k} points implies that the quasiparticle density of states does not have a gap.

At low energies we can estimate it from our expansion of the quasiparticle energy,

$$\begin{aligned}
D_s(E) &= \frac{1}{N} \sum_{\mathbf{k}} \delta(E - E_{\mathbf{k}}) \\
&\cong \frac{4}{N} \sum_{\mathbf{q}} \delta\left(E - \sqrt{(\mathbf{v}_F \cdot \mathbf{q})^2 + (\mathbf{v}_{\text{qp}} \cdot \mathbf{q})^2}\right) \\
&\cong 4a_{\text{uc}} \int \frac{d^2q}{(2\pi)^2} \delta\left(E - \sqrt{(\mathbf{v}_F \cdot \mathbf{q})^2 + (\mathbf{v}_{\text{qp}} \cdot \mathbf{q})^2}\right) \\
&= \frac{4a_{\text{uc}}}{v_F v_{\text{qp}}} \int \frac{d^2u}{(2\pi)^2} \delta\left(E - \sqrt{u_x^2 + u_y^2}\right) \\
&= \frac{2a_{\text{uc}}}{\pi v_F v_{\text{qp}}} \int_0^\infty du u \delta(E - u) = \frac{2a_{\text{uc}}}{\pi v_F v_{\text{qp}}} E,
\end{aligned} \tag{12.15}$$

where a_{uc} is the area of the two-dimensional unit cell. We see that the density of states starts *linearly* at small energies. The full dependence is sketched here:



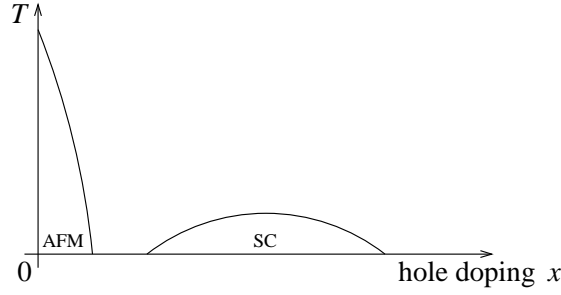
An additional nice feature of the d -wave gap is the following: The interaction considered above is presumably not of BCS (Coulomb + phonons) type. However, there should also be a strong short-range Coulomb repulsion, which is not overscreened by phonon exchange. This repulsion can again be modeled by a constant $V_0 > 0$ in \mathbf{k} -space. This additional interaction adds the term

$$-\frac{1}{N} \sum_{\mathbf{k}'} V_0 \frac{1 - n_F(E_{\mathbf{k}'})}{2E_{\mathbf{k}'}} \Delta_{\mathbf{k}'} \tag{12.16}$$

to the gap equation. But since $E_{\mathbf{k}'}$ does not change sign under rotation of \mathbf{k}' by $\pi/2$ (i.e., 90°), while $\Delta_{\mathbf{k}'}$ does change sign under this rotation, the sum over \mathbf{k}' vanishes. d -wave pairing is thus robust against on-site Coulomb repulsion.

12.2 Cuprates

Estimates of T_c based on phonon-exchange and using experimentally known values of the Debye frequency, the electron-phonon coupling, and the normal-state density of states are much lower than the observed critical temperatures. Also, as we have seen, such an interaction is flat in \mathbf{k} -space, which favors an s -wave gap. An s -wave gap is inconsistent with nearly all experiments on the cuprates that are sensitive to the gap. The last section has shown that $d_{x^2-y^2}$ -wave pairing in the cuprates is plausible *if* there is an attractive interaction for momentum transfers $\mathbf{q} \approx (\pi/a, \pi/a)$. We will now discuss where this attraction could be coming from.



A glance at typical phase diagrams shows that the undoped cuprates tend to be antiferromagnetic. Weak hole doping or slightly stronger electron doping destroy the antiferromagnetic order, and at larger doping, superconductivity emerges. At even larger doping (the “overdoped” regime), superconductivity is again suppressed. Also in many other unconventional superconductors superconductivity is found in the vicinity of but rarely coexisting with magnetic order. This is true for most pnictide and heavy-fermion superconductors. The vicinity of a magnetically ordered phase makes itself felt by strong magnetic fluctuations and strong, but short-range, spin correlations. These are seen as an enhanced spin susceptibility.

At a magnetic second-order phase transition, the static spin susceptibility $\chi_{\mathbf{q}}$ diverges at $\mathbf{q} = \mathbf{Q}$, where \mathbf{Q} is the ordering vector. It is $\mathbf{Q} = 0$ for ferromagnetic order and $\mathbf{Q} = (\pi/a, \pi/a)$ for checkerboard (Néel) order on a square lattice. Even some distance from the transition or at non-zero frequencies ν , the susceptibility $\chi_{\mathbf{q}}(\nu)$ tends to have a maximum close to \mathbf{Q} . Far away from the magnetic phase or at high frequencies this remnant of magnetic order becomes small. This discussion suggests that the exchange of spin fluctuations, which are strong close to \mathbf{Q} , could provide the attractive interaction needed for Cooper pairing.

The Hubbard model

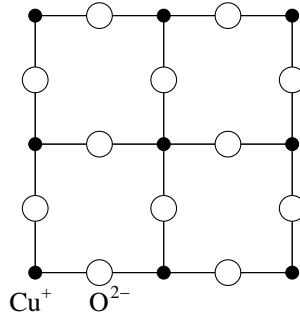
The two-dimensional, single-band, repulsive Hubbard model is thought (by many experts, not by everyone) to be the simplest model that captures the main physics of the cuprates. The Hamiltonian reads, in real space,

$$H = - \sum_{ij\sigma} t_{ij} c_{i\sigma}^\dagger c_{j\sigma} + U \sum_i c_{i\uparrow}^\dagger c_{i\uparrow} c_{i\downarrow}^\dagger c_{i\downarrow} \quad (12.17)$$

with $U > 0$ and, in momentum space,

$$H = \sum_{\mathbf{k}\sigma} \epsilon_{\mathbf{k}} c_{\mathbf{k}\sigma}^\dagger c_{\mathbf{k}\sigma} + \frac{U}{N} \sum_{\mathbf{k}\mathbf{k}'\mathbf{q}} c_{\mathbf{k}+\mathbf{q},\uparrow}^\dagger c_{\mathbf{k}'-\mathbf{q},\downarrow}^\dagger c_{\mathbf{k}'\downarrow} c_{\mathbf{k}\uparrow}. \quad (12.18)$$

Also, the undoped cuprate parent compounds have, from simple counting, an odd number of electrons per unit cell. Thus for them the single band must be half-filled, while for doped cuprates it is still close to half filling. The underlying lattice in real space is a two-dimensional square lattice with each site i corresponding to a Cu^+ ion.



The transverse spin susceptibility is defined by

$$\chi^{+-}(\mathbf{q}, \tau) = - \langle T_\tau S^+(\mathbf{q}, \tau) S^-(\mathbf{-q}, 0) \rangle, \quad (12.19)$$

where τ is the imaginary time, T_τ is the time-ordering directive, and

$$S^\pm(\mathbf{q}, \tau) := S^x(\mathbf{q}, \tau) \pm iS^y(\mathbf{q}, \tau) \quad (12.20)$$

with

$$S^\alpha(\mathbf{q}, \tau) := \frac{1}{\sqrt{N}} \sum_{\mathbf{k}\sigma\sigma'} c_{\mathbf{k}+\mathbf{q},\sigma}^\dagger(\tau) \frac{\sigma_{\sigma\sigma'}^\alpha}{2} c_{\mathbf{k}\sigma'}(\tau) \quad (12.21)$$

are electron-spin operators. $\boldsymbol{\sigma} = (\sigma^x, \sigma^y, \sigma^z)$ is the vector of Pauli matrices. The susceptibility can be rewritten as

$$\chi^{+-}(\mathbf{q}, \tau) = -\frac{1}{N} \sum_{\mathbf{k}\mathbf{k}'} \left\langle T_\tau \left(c_{\mathbf{k}+\mathbf{q},\uparrow}^\dagger c_{\mathbf{k}\downarrow} \right) (\tau) \left(c_{\mathbf{k}'-\mathbf{q},\downarrow}^\dagger c_{\mathbf{k}'\uparrow} \right) (0) \right\rangle. \quad (12.22)$$

In the non-interacting limit of $U \rightarrow 0$, the average of four fermionic operators can be written in terms of products of averages of two operators (Wick's theorem). The resulting *bare* susceptibility reads

$$\begin{aligned} \chi_0^{+-}(\mathbf{q}, \tau) &= -\frac{1}{N} \sum_{\mathbf{k}} \left\langle T_\tau c_{\mathbf{k}+\mathbf{q},\uparrow}^\dagger(\tau) c_{\mathbf{k}+\mathbf{q},\uparrow}(0) \right\rangle_0 \left\langle T_\tau c_{\mathbf{k}\downarrow}(\tau) c_{\mathbf{k}\downarrow}^\dagger(0) \right\rangle_0 \\ &= \frac{1}{N} \sum_{\mathbf{k}} \left\langle T_\tau c_{\mathbf{k}+\mathbf{q},\uparrow}(0) c_{\mathbf{k}+\mathbf{q},\uparrow}^\dagger(\tau) \right\rangle_0 \left\langle T_\tau c_{\mathbf{k}\downarrow}(\tau) c_{\mathbf{k}\downarrow}^\dagger(0) \right\rangle_0 \\ &= \frac{1}{N} \sum_{\mathbf{k}} \mathcal{G}_{\mathbf{k}+\mathbf{q},\uparrow}^0(-\tau) \mathcal{G}_{\mathbf{k}\downarrow}^0(\tau). \end{aligned} \quad (12.23)$$

The Fourier transform as a function of the bosonic Matsubara frequency $i\nu_n$ is

$$\begin{aligned} \chi_0^{+-}(\mathbf{q}, i\nu_n) &= \int_0^\beta d\tau e^{i\nu_n\tau} \chi_0^{+-}(\mathbf{q}, \tau) \\ &= \frac{1}{N} \int_0^\beta d\tau e^{i\nu_n\tau} \sum_{\mathbf{k}} \frac{1}{\beta} \sum_{i\omega'_n} e^{-i\omega'_n(-\tau)} \mathcal{G}_{\mathbf{k}+\mathbf{q},\uparrow}^0(i\omega'_n) \frac{1}{\beta} \sum_{i\omega_n} e^{-i\omega_n\tau} \mathcal{G}_{\mathbf{k}\downarrow}^0(i\omega_n) \\ &= \frac{1}{N} \sum_{\mathbf{k}} \frac{1}{\beta^2} \sum_{i\omega_n, i\omega'_n} \beta \delta_{\nu_n + \omega'_n, \omega_n} \mathcal{G}_{\mathbf{k}+\mathbf{q},\uparrow}^0(i\omega'_n) \mathcal{G}_{\mathbf{k}\downarrow}^0(i\omega_n) = \frac{1}{N} \sum_{\mathbf{k}} \frac{1}{\beta} \sum_{i\omega_n} \mathcal{G}_{\mathbf{k}+\mathbf{q},\uparrow}^0(i\omega_n - i\nu_n) \mathcal{G}_{\mathbf{k}\downarrow}^0(i\omega_n) \\ &= \frac{1}{N} \sum_{\mathbf{k}} \frac{1}{\beta} \sum_{i\omega_n} \frac{1}{i\omega_n - i\nu_n - \xi_{\mathbf{k}+\mathbf{q}}} \frac{1}{i\omega_n - \xi_{\mathbf{k}}}. \end{aligned} \quad (12.24)$$

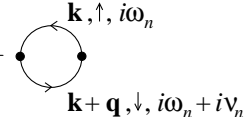
This expression can be written in a more symmetric form by making use of the identity $\xi_{\mathbf{k}} = \xi_{-\mathbf{k}}$ and replacing the summation variables \mathbf{k} by $-\mathbf{k} - \mathbf{q}$ and $i\omega_n$ by $i\omega_n + i\nu_n$. The result is

$$\chi_0^{+-}(\mathbf{q}, i\nu_n) = \frac{1}{N} \sum_{\mathbf{k}} \frac{1}{\beta} \sum_{i\omega_n} \frac{1}{i\omega_n - \xi_{\mathbf{k}}} \frac{1}{i\omega_n + i\nu_n - \xi_{\mathbf{k}+\mathbf{q}}}. \quad (12.25)$$

The Matsubara frequency sum can be evaluated using methods from complex analysis. We here give the result without proof,

$$\chi_0^{+-}(\mathbf{q}, i\nu_n) = \frac{1}{N} \sum_{\mathbf{k}} \frac{n_F(\xi_{\mathbf{k}}) - n_F(\xi_{\mathbf{k}+\mathbf{q}})}{i\nu_n + \xi_{\mathbf{k}} - \xi_{\mathbf{k}+\mathbf{q}}}. \quad (12.26)$$

Note that the same result is found for the bare charge susceptibility except for a spin factor of 2. Diagrammatically, the result can be represented by the bubble diagram

$$\chi_0^{+-}(\mathbf{q}, i\nu_n) = -\frac{1}{2} \Pi_0(\mathbf{q}, i\nu_n) = - \text{bubble diagram} \quad (12.27)$$


What changes when we switch on the Hubbard interaction U ? In analogy with the RPA theory for the screened Coulomb interaction we might guess that the RPA spin susceptibility is given by

$$-\chi_{\text{RPA}}^{+-} \stackrel{?}{=} \begin{array}{c} \uparrow \\ \circlearrowleft \\ \downarrow \end{array} + \begin{array}{c} \uparrow \\ \circlearrowright \\ \downarrow \end{array} - U \begin{array}{c} \uparrow \\ \circlearrowleft \\ \downarrow \end{array} + \dots \quad (12.28)$$

but the second and higher terms vanish since they contain vertices at which the Hubbard interaction supposedly flips the spin,

$$\begin{array}{c} \downarrow \\ \swarrow \\ \bullet \\ \nearrow \\ \uparrow \end{array} - U \dots, \quad (12.29)$$

which it cannot do. On the other hand, the following ladder diagrams do not vanish and represent the RPA susceptibility:

$$-\chi_{\text{RPA}}^{+-} = \begin{array}{c} \uparrow \\ \circlearrowleft \\ \downarrow \end{array} + \begin{array}{c} \downarrow \\ \circlearrowleft \\ \uparrow \\ U \\ \downarrow \\ \circlearrowright \\ \uparrow \end{array} + \begin{array}{c} \downarrow \\ \circlearrowleft \\ \uparrow \\ U \\ \downarrow \\ \circlearrowleft \\ \uparrow \\ U \\ \downarrow \\ \circlearrowright \\ \uparrow \end{array} + \dots \quad (12.30)$$

Since the Hubbard interaction does not depend on momentum, this series has a rather simple mathematical form,

$$\begin{aligned} \chi_{\text{RPA}}^{+-}(\mathbf{q}, i\nu_n) &= \chi_0^{+-}(\mathbf{q}, i\nu_n) + \chi_0^{+-}(\mathbf{q}, i\nu_n) U \chi_0^{+-}(\mathbf{q}, i\nu_n) + \chi_0^{+-}(\mathbf{q}, i\nu_n) U \chi_0^{+-}(\mathbf{q}, i\nu_n) U \chi_0^{+-}(\mathbf{q}, i\nu_n) + \dots \\ &= \chi_0^{+-}(\mathbf{q}, i\nu_n) [1 + U \chi_0^{+-}(\mathbf{q}, i\nu_n) + U \chi_0^{+-}(\mathbf{q}, i\nu_n) U \chi_0^{+-}(\mathbf{q}, i\nu_n) + \dots] \\ &= \frac{\chi_0^{+-}(\mathbf{q}, i\nu_n)}{1 - U \chi_0^{+-}(\mathbf{q}, i\nu_n)} \end{aligned} \quad (12.31)$$

[the signs in the first line follow from the Feynman rules, in particular each term contains a single fermionic loop, which gives a minus sign, which cancels the explicit one in Eq. (12.30)]. The RPA spin susceptibility can be evaluated numerically for given dispersion $\xi_{\mathbf{k}}$. It is clear that it predicts an instability of the Fermi liquid if the static RPA spin susceptibility

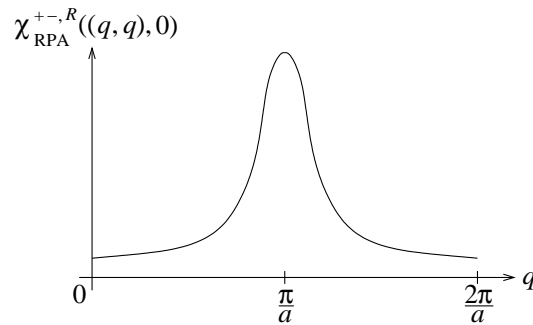
$$\chi_{\text{RPA}}^{+-,R}(\mathbf{q}, 0) = \chi_{\text{RPA}}^{+-}(\mathbf{q}, i\nu_n \rightarrow \nu + i0^+) |_{\nu \rightarrow 0} \equiv \frac{\chi_0^{+-,R}(\mathbf{q}, 0)}{1 - U \chi_0^{+-,R}(\mathbf{q}, 0)} \quad (12.32)$$

diverges at some $\mathbf{q} = \mathbf{Q}$, i.e., if

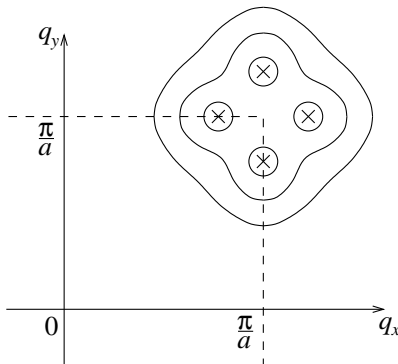
$$U \chi_0^{+-,R}(\mathbf{Q}, 0) = 1. \quad (12.33)$$

Since the *spin* susceptibility diverges, this would be a magnetic ordering transition with ordering vector \mathbf{Q} . Note that the RPA is not a good theory for the antiferromagnetic transition of the cuprates since it is a resummation of a perturbative series in U/t , which is not small in cuprates. t is the typical hopping amplitude. Nevertheless it gives qualitatively reasonable results in the paramagnetic phase, which is of interest for superconductivity.

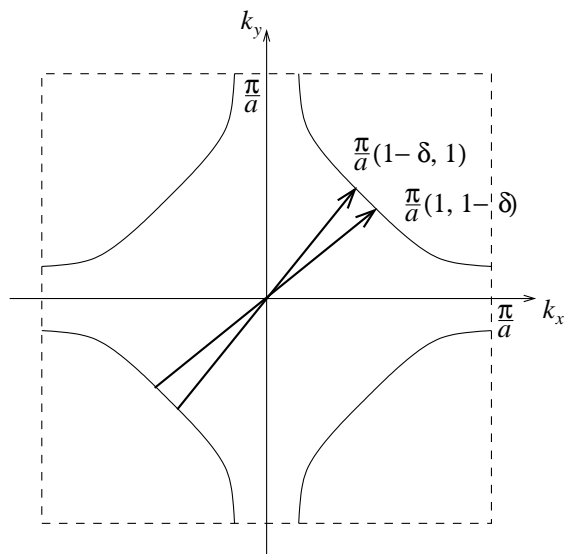
The numerical evaluation at intermediate doping and high temperatures gives a broad and high peak in $\chi_{\text{RPA}}^{+-,R}(\mathbf{q}, 0)$ centered at $\mathbf{Q} = (\pi/a, \pi/a)$. This is consistent with the ordering at \mathbf{Q} observed at weak doping.



At lower temperatures, details become resolved that are obscured by thermal broadening at high T . The RPA and also more advanced approaches are very sensitive to the electronic bands close to the Fermi energy; states with $|\xi_{\mathbf{k}}| \gg k_B T$ have exponentially small effect on the susceptibility. Therefore, the detailed susceptibility at low T strongly depends on details of the model Hamiltonian. Choosing nearest-neighbor and next-nearest-neighbor hopping in such a way that a realistic Fermi surface emerges, one obtains a spin susceptibility with *incommensurate* peaks at $\frac{\pi}{a}(1, 1 \pm \delta)$ and $\frac{\pi}{a}(1 \pm \delta, 1)$.



These peaks are due to *nesting*: Scattering is enhanced between parallel portions of the Fermi surface, which in turn enhances the susceptibility [see M. Norman, Phys. Rev. B **75**, 184514 (2007)].



The results for the spin susceptibility are in qualitative agreement with neutron-scattering experiments. However, the RPA overestimates the tendency toward magnetic order, which is reduced by more advanced approaches.

Spin-fluctuation exchange

The next step is to construct an effective electron-electron interaction mediated by the exchange of spin fluctuations. The following diagrammatic series represents the simplest way of doing this, though certainly not the only one:

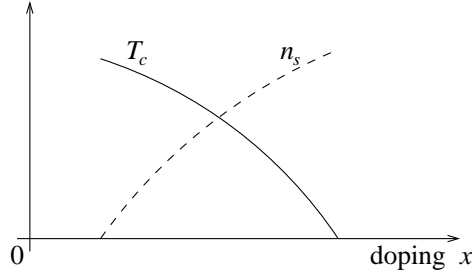
$$V_{\text{eff}} := \begin{array}{c} \uparrow \downarrow \\ \bullet \bullet \\ \downarrow \uparrow \end{array} + \begin{array}{c} \uparrow \downarrow \\ \bullet \bullet \\ \downarrow \uparrow \end{array} + \begin{array}{c} \uparrow \downarrow \\ \bullet \bullet \\ \downarrow \uparrow \end{array} + \begin{array}{c} \uparrow \downarrow \\ \bullet \bullet \\ \downarrow \uparrow \end{array} + \dots \quad (12.34)$$

Note that the external legs do not represent electronic Green functions but only indicate the states of incoming and outgoing electrons. The series is very similar to the one for the RPA susceptibility. Indeed, the effective interaction is

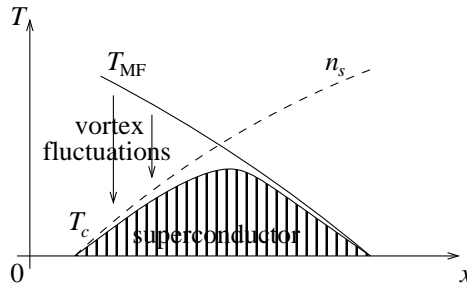
$$\begin{aligned}
V_{\text{eff}}(\mathbf{q}, i\nu_n) &= U + U\chi_0^{+-}(\mathbf{q}, i\nu_n)U + U\chi_0^{+-}(\mathbf{q}, i\nu_n)U\chi_0^{+-}(\mathbf{q}, i\nu_n) + \dots \\
&= U + U^2[\chi_0^{+-}(\mathbf{q}, i\nu_n) + \chi_0^{+-}(\mathbf{q}, i\nu_n)U\chi_0^{+-}(\mathbf{q}, i\nu_n) + \dots] \\
&= U + U^2\chi_{\text{RPA}}^{+-}(\mathbf{q}, i\nu_n).
\end{aligned}
\tag{12.35}$$

Typically one goes beyond the RPA at this point by including additional diagrams. In particular, also charge fluctuations are included through the charge susceptibility and the bare Green function \mathcal{G}^0 is replaced by a selfconsistent one incorporating the effect of spin and charge fluctuations on the electronic self-energy. This leads to the *fluctuation-exchange approximation (FLEX)*. One could now obtain the Cooper instability due to the FLEX effective interaction in analogy to Sec. 9.1 and use a BCS mean-field theory to describe the superconducting state. However, since the system is not in the weak-coupling limit—the typical interaction times the electronic density of states is not small—one usually employs a strong-coupling generalization of BCS theory known as Eliashberg theory. Since the effective interaction is, like the spin susceptibility, strongly peaked close to $(\pi/a, \pi/a)$, it favors $d_{x^2-y^2}$ -wave pairing, as we have seen.

The numerical result of the FLEX for T_c and for the superfluid density n_s are sketched here:



The curve for T_c vs. doping does not yet look like the experimentally observed dome. There are several aspects that make the region of weak doping (underdoping) difficult to treat theoretically. One is indicated in the sketch: n_s is strongly reduced, which indicates that superconductivity may be in some sense fragile in this regime. Furthermore, the cuprates are nearly two-dimensional solids. In fact we have used a two-dimensional model so far. If we take this seriously, we know from chapter 7 that any mean-field theory, which Eliashberg theory with FLEX effective interaction still is, fails miserably. Instead, we expect a BKT transition at a strongly reduced critical temperature. We reinterpret the FLEX critical temperature as the mean-field temperature T_{MF} and the FLEX superfluid density as the unrenormalized superfluid density n_s^0 . We have seen in chapter 7 that the bare stiffness K_0 is proportional to n_s^0 . The BKT transition temperature T_c is defined by $K(l \rightarrow \infty) = 2/\pi$. It is thus reduced by small n_s^0 , corresponding to a small initial value $K(0) \equiv K_0$. This is physically clear: Small stiffness makes it easy to create vortex-antivortex pairs. T_c is of course also reduced by T_{MF} and can never be larger than T_{MF} . A BKT theory on top of the FLEX gives the following phase diagram, which is in qualitative agreement with experiments:

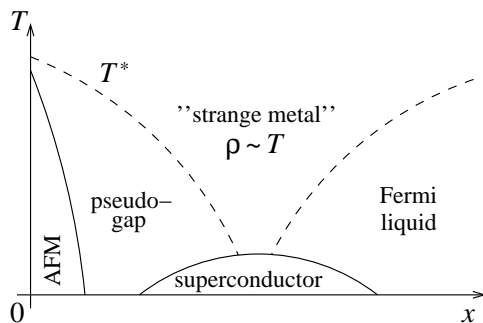


This scenario is consistent with the Nernst effect (an electric field measured normal to both an applied magnetic field and a temperature gradient) in underdoped cuprates, which is interpreted in terms of free vortices in a broad temperature range, which we would understand as the range from T_c to T_{MF} .

Also note that spin fluctuations strongly affect the electronic properties in underdoped cuprates up to a temperature T^* significantly higher than T_{MF} . For example, below T^* the electronic density of states close to the Fermi energy is suppressed compared to the result of band-structure calculations. The FLEX describes this effect qualitatively correctly. This suppression is the well-known *pseudogap*.

Quantum critical point

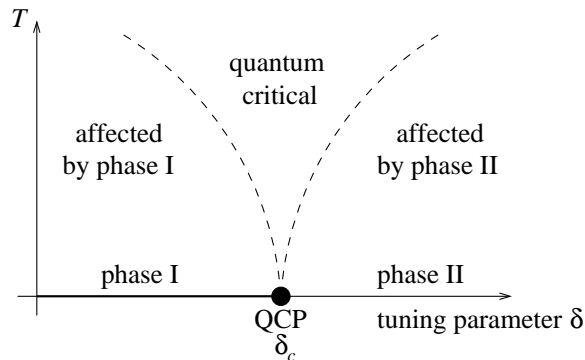
The previously discussed approach relies on a resummation of a perturbative series in U/t . This is questionable for cuprates, where U/t is on the order of 3. Many different approaches have been put forward that supposedly work in this strong-coupling regime. They emphasize different aspects of the cuprates, showing that it is not even clear which ingredients are the most important for understanding the phase diagram. Here we will review a line of thought represented by Chandra Varma and Subir Sachdev, among others. Its starting point is an analysis of the *normal* region of the phase diagram.



Roughly speaking, there are three regimes in the normal-conducting state:

- a pseudo-gap regime below T^* at underdoping, in which the electronic density of states at low energies is suppressed,
- a “strange-metal” regime above the superconducting dome, without a clear suppression of the density of states but with unusual temperature and energy dependencies of various observables, for example a resistivity linear in temperature, $\rho \propto T$,
- an apparently ordinary normal-metal (Fermi-liquid) regime at overdoping, with standard $\rho \propto \text{const} + T^2$ dependence.

The two crossover lines look very much like what one expects to find for a *quantum critical point* (QCP), i.e., a phase transition at zero temperature.

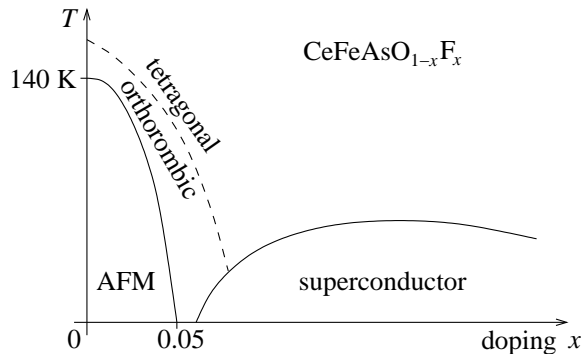


The regions to the left and right have a characteristic energy scale $\epsilon(\delta)$ inherited from the ground state ($T = 0$), which dominates the thermal fluctuations. The energy scales go to zero at the QCP, i.e., for $\delta \rightarrow \delta_c$ from both sides. Right at the QCP there then is no energy scale and energy or temperature-dependent quantities have to be power laws. The emerging idea is that the superconducting dome hides a QCP between antiferromagnetic (spin-density-wave) and paramagnetic order at $T = 0$. The spin fluctuations associated with this QCP become stronger

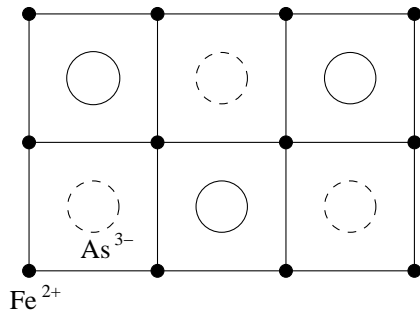
as it is approached. Consequently, the superconductivity caused by them is strongest and has the highest T_c right above the QCP. Compare the previous argument: There, the spin fluctuations are assumed to be strongest close to the (finite-temperature) antiferromagnetic phase and one needs to invoke small n_s and vortex fluctuations to argue why the maximum T_c is not close to the antiferromagnetic phase.

12.3 Pnictides

The iron pnictide superconductors are a more heterogeneous group than the cuprates. However, for most of them the phase diagram is roughly similar to the one of a typical cuprate in that superconductivity emerges at finite doping in the vicinity of an antiferromagnetic phase. This antiferromagnetic phase is metallic, not a Mott insulator, though, suggesting that interactions are generally weaker in the pnictides.

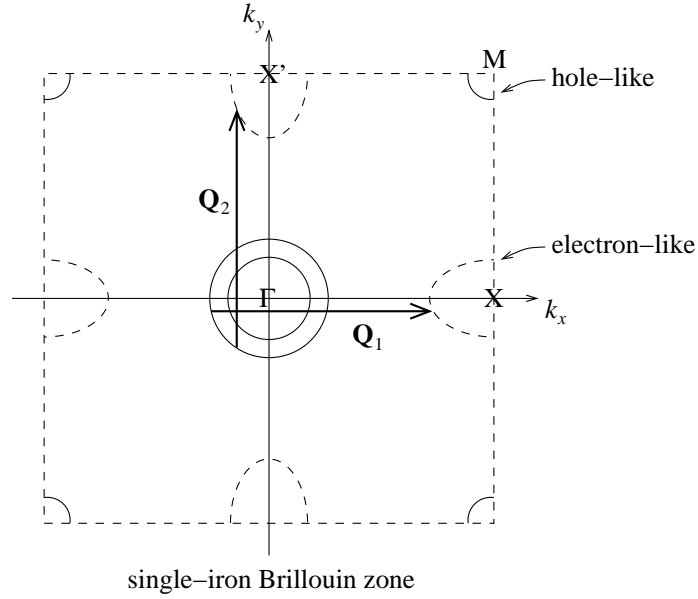


The crystal structure is quasi-two-dimensional, though probably less so than in the cuprates. The common structural motif is an iron-pnictogen, in particular $\text{Fe}^{2+}\text{As}^{3-}$, layer with Fe^{2+} forming a square lattice and As^{3-} sitting alternately above and below the Fe^{2+} plaquettes.



While the correct unit cell contains two As and two Fe ions, the glide-mirror symmetry with respect to the Fe plane allows to formulate two-dimensional models using a *single-iron* unit cell. The fact that different unit cells are used leads to some confusion in the field. Note that since the single-iron unit cell is *half* as large as the two-iron unit cell, the corresponding single-iron Brillouin zone is *twice* as large as the two-iron Brillouin zone.

A look at band-structure calculations or angular resolved photoemission data shows that the Fermi surface of pnictides is much more complicated than the one of cuprates. The $k_z = 0$ cut typically shows five Fermi pockets.

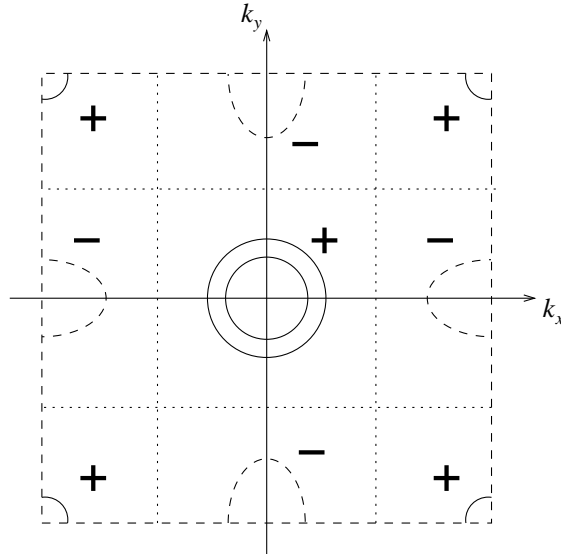


The (probably outer) hole pocket at Γ and the electron pockets at X and X' are well nested with nesting vectors $\mathbf{Q}_1 = (\pi/a, 0)$ and $\mathbf{Q}_2 = (0, \pi/a)$, respectively. Not surprisingly, the spin susceptibility is peaked at \mathbf{Q}_1 and \mathbf{Q}_2 in the paramagnetic phase and in the antiferromagnetic phase the system orders antiferromagnetically at either of three vectors. Incidentally, the antiferromagnet emerges through the formation and condensation of electron-hole pairs (*excitons*), described by a BCS-type theory. The same excitonic instability is for example responsible for the magnetism of chromium.

Assuming that the exchange of spin fluctuations in the paramagnetic phase is the main pairing interaction, the gap equation

$$\Delta_{\mathbf{k}} = -\frac{1}{N} \sum_{\mathbf{k}'} V_{\mathbf{k}\mathbf{k}'} \frac{1 - n_F(E_{\mathbf{k}'})}{2E_{\mathbf{k}'}} \Delta_{\mathbf{k}'} \quad (12.36)$$

with $V_{\mathbf{k}\mathbf{k}'}$ large and positive for $\mathbf{k} - \mathbf{k}' = \pm\mathbf{Q}_1$ or $\pm\mathbf{Q}_2$, favors a gap function $\Delta_{\mathbf{k}}$ changing sign between \mathbf{k} and $\mathbf{k} + \mathbf{Q}_1$ and between \mathbf{k} and $\mathbf{k} + \mathbf{Q}_2$, see Sec. 12.1. This is most easily accommodated by a nodeless gap changing sign between the electron and hole pockets:



This gap is said to have s -wave symmetry in that it does not have lower symmetry than the lattice, unlike d -wave. To emphasize the sign change, it is often called an s_{\pm} -wave gap. The simplest realization would be

$$\Delta_{\mathbf{k}} = \Delta_0 \cos k_x a \cos k_y a. \quad (12.37)$$

12.4 Triplet superconductors and He-3

So far, we have assumed that Cooper pairs are formed by two electrons with opposite spin so that the total spin of pair vanishes (spin-singlet pairing). This assumption becomes questionable in the presence of strong *ferromagnetic* interactions, which favor parallel spin alignment. If superconductivity is possible at all in such a situation, we could expect to find spin-1 Cooper pairs. Since they would be spin triplets, one is talking of *triplet superconductors*. This scenario is very likely realized in Sr_2RuO_4 (which is, interestingly, isostructural to the prototypical cuprate La_2CuO_4), a few organic salts, and some heavy-fermion compounds. It is even more certain to be responsible for the superfluidity of He-3, where neutral He-3 atoms instead of charged electrons form Cooper pairs, see Sec. 2.2.

Formally, we restrict ourselves to a BCS-type mean-field theory. We generalize the effective interaction to allow for an arbitrary spin dependence,

$$H = \sum_{\mathbf{k}\sigma} \xi_{\mathbf{k}} c_{\mathbf{k}\sigma}^\dagger c_{\mathbf{k}\sigma} + \frac{1}{N} \sum_{\mathbf{k}\mathbf{k}'} \sum_{\sigma\tau\sigma'\tau'} V_{\sigma\tau\sigma'\tau'}(\mathbf{k}, \mathbf{k}') c_{\mathbf{k}\sigma}^\dagger c_{-\mathbf{k},\tau}^\dagger c_{-\mathbf{k}',\tau'} c_{\mathbf{k}'\sigma'}, \quad (12.38)$$

where $\sigma, \tau, \sigma', \tau' = \uparrow, \downarrow$ are spin indices. In decomposing the interaction, we now allow the averages $\langle c_{-\mathbf{k},\tau} c_{\mathbf{k}\sigma} \rangle$ to be non-zero for all τ, σ . Thus the mean-field Hamiltonian reads

$$H_{\text{MF}} = \sum_{\mathbf{k}\sigma} \xi_{\mathbf{k}} c_{\mathbf{k}\sigma}^\dagger c_{\mathbf{k}\sigma} + \frac{1}{N} \sum_{\mathbf{k}\mathbf{k}'} \sum_{\sigma\tau\sigma'\tau'} V_{\sigma\tau\sigma'\tau'}(\mathbf{k}, \mathbf{k}') \left(\langle c_{\mathbf{k}\sigma}^\dagger c_{-\mathbf{k},\tau}^\dagger \rangle c_{-\mathbf{k}',\tau'} c_{\mathbf{k}'\sigma'} + c_{\mathbf{k}\sigma}^\dagger c_{-\mathbf{k},\tau}^\dagger \langle c_{-\mathbf{k}',\tau'} c_{\mathbf{k}'\sigma'} \rangle \right) + \text{const.} \quad (12.39)$$

We define

$$\Delta_{\sigma\tau}(\mathbf{k}) := -\frac{1}{N} \sum_{\mathbf{k}'} \sum_{\sigma'\tau'} V_{\sigma\tau\sigma'\tau'}(\mathbf{k}, \mathbf{k}') \langle c_{-\mathbf{k}',\tau'} c_{\mathbf{k}'\sigma'} \rangle \quad (12.40)$$

so that

$$\Delta_{\sigma\tau}^*(\mathbf{k}) = -\frac{1}{N} \sum_{\mathbf{k}'} \sum_{\sigma'\tau'} V_{\sigma'\tau'\sigma\tau}(\mathbf{k}', \mathbf{k}) \langle c_{\mathbf{k}'\sigma'}^\dagger c_{-\mathbf{k}',\tau'}^\dagger \rangle. \quad (12.41)$$

Here, we have used that

$$[V_{\sigma\tau\sigma'\tau'}(\mathbf{k}, \mathbf{k}')]^* = V_{\sigma'\tau'\sigma\tau}(\mathbf{k}', \mathbf{k}), \quad (12.42)$$

which follows from hermiticity of the Hamiltonian H . Then

$$H_{\text{MF}} = \sum_{\mathbf{k}\sigma} \xi_{\mathbf{k}} c_{\mathbf{k}\sigma}^\dagger c_{\mathbf{k}\sigma} - \sum_{\mathbf{k}\sigma\tau} \Delta_{\sigma\tau}^*(\mathbf{k}) c_{-\mathbf{k},\tau} c_{\mathbf{k}\sigma} - \sum_{\mathbf{k}\sigma\tau} \Delta_{\sigma\tau}(\mathbf{k}) c_{\mathbf{k}\sigma}^\dagger c_{-\mathbf{k},\tau}^\dagger + \text{const.} \quad (12.43)$$

The gap function $\Delta_{\sigma\tau}(\mathbf{k})$ is now a matrix in spin space,

$$\hat{\Delta}(\mathbf{k}) = \begin{pmatrix} \Delta_{\uparrow\uparrow}(\mathbf{k}) & \Delta_{\uparrow\downarrow}(\mathbf{k}) \\ \Delta_{\downarrow\uparrow}(\mathbf{k}) & \Delta_{\downarrow\downarrow}(\mathbf{k}) \end{pmatrix}. \quad (12.44)$$

The function $\hat{\Delta}(\mathbf{k})$ has an important symmetry property that follows from the symmetry of averages $\langle c_{-\mathbf{k},\tau} c_{\mathbf{k}\sigma} \rangle$: It is clear that

$$\langle c_{\mathbf{k}\sigma} c_{-\mathbf{k},\tau} \rangle = -\langle c_{-\mathbf{k},\tau} c_{\mathbf{k}\sigma} \rangle. \quad (12.45)$$

Furthermore, by relabeling $\sigma \leftrightarrow \tau$, $\sigma' \leftrightarrow \tau'$, $\mathbf{k} \rightarrow -\mathbf{k}$, $\mathbf{k}' \rightarrow -\mathbf{k}'$ in the interaction term of H , we see that the interaction strength must satisfy the relation

$$V_{\sigma\tau\sigma'\tau'}(\mathbf{k}, \mathbf{k}') = V_{\tau\sigma\tau'\sigma'}(-\mathbf{k}, -\mathbf{k}'). \quad (12.46)$$

Thus

$$\Delta_{\tau\sigma}(-\mathbf{k}) = -\frac{1}{N} \sum_{\mathbf{k}'} \sum_{\sigma'\tau'} V_{\tau\sigma\tau'\sigma'}(-\mathbf{k}, -\mathbf{k}') \langle c_{\mathbf{k}'\sigma'} c_{-\mathbf{k}'\tau'} \rangle = +\frac{1}{N} \sum_{\mathbf{k}'} \sum_{\sigma'\tau'} V_{\sigma\tau\sigma'\tau'}(\mathbf{k}, \mathbf{k}') \langle c_{-\mathbf{k}'\tau'} c_{\mathbf{k}'\sigma'} \rangle = -\Delta_{\sigma\tau}(\mathbf{k}) \quad (12.47)$$

or, equivalently,

$$\hat{\Delta}(-\mathbf{k}) = -\hat{\Delta}^T(\mathbf{k}). \quad (12.48)$$

We now want to write $\hat{\Delta}(\mathbf{k})$ in terms of singlet and triplet components. From elementary quantum theory, a spin-singlet pair is created by

$$s_{\mathbf{k}}^{\dagger} := \frac{c_{\mathbf{k}\uparrow}^{\dagger} c_{-\mathbf{k}\downarrow}^{\dagger} - c_{\mathbf{k}\downarrow}^{\dagger} c_{-\mathbf{k}\uparrow}^{\dagger}}{\sqrt{2}}, \quad (12.49)$$

while the $m = 1, 0, -1$ components of a spin-triplet pair are created by

$$t_{\mathbf{k}1}^{\dagger} := c_{\mathbf{k}\uparrow}^{\dagger} c_{-\mathbf{k}\uparrow}^{\dagger}, \quad (12.50)$$

$$t_{\mathbf{k}0}^{\dagger} := \frac{c_{\mathbf{k}\uparrow}^{\dagger} c_{-\mathbf{k}\downarrow}^{\dagger} + c_{\mathbf{k}\downarrow}^{\dagger} c_{-\mathbf{k}\uparrow}^{\dagger}}{\sqrt{2}}, \quad (12.51)$$

$$t_{\mathbf{k},-1}^{\dagger} := c_{\mathbf{k}\downarrow}^{\dagger} c_{-\mathbf{k}\downarrow}^{\dagger}, \quad (12.52)$$

respectively. Alternatively, we can transform onto states with maximum spin along the x , y , and z axis. This is analogous to the mapping from ($l = 1, m$) eigenstates onto p_x , p_y , and p_z orbitals for the hydrogen atom. The new components are created by

$$t_{\mathbf{k}x}^{\dagger} := \frac{-c_{\mathbf{k}\uparrow}^{\dagger} c_{-\mathbf{k}\uparrow}^{\dagger} + c_{\mathbf{k}\downarrow}^{\dagger} c_{-\mathbf{k}\downarrow}^{\dagger}}{\sqrt{2}}, \quad (12.53)$$

$$t_{\mathbf{k}y}^{\dagger} := i \frac{c_{\mathbf{k}\uparrow}^{\dagger} c_{-\mathbf{k}\uparrow}^{\dagger} + c_{\mathbf{k}\downarrow}^{\dagger} c_{-\mathbf{k}\downarrow}^{\dagger}}{\sqrt{2}}, \quad (12.54)$$

$$t_{\mathbf{k}z}^{\dagger} := \frac{c_{\mathbf{k}\uparrow}^{\dagger} c_{-\mathbf{k}\downarrow}^{\dagger} + c_{\mathbf{k}\downarrow}^{\dagger} c_{-\mathbf{k}\uparrow}^{\dagger}}{\sqrt{2}} = t_{\mathbf{k}0}^{\dagger}, \quad (12.55)$$

which form a vector $\mathbf{t}_{\mathbf{k}}^{\dagger}$. The term in H_{MF} involving $\Delta_{\sigma\tau}(\mathbf{k})$ can now be expressed in terms of the new operators,

$$\begin{aligned} -\sum_{\mathbf{k}\sigma\tau} \Delta_{\sigma\tau}(\mathbf{k}) c_{\mathbf{k}\sigma}^{\dagger} c_{-\mathbf{k}\tau}^{\dagger} &= -\sum_{\mathbf{k}} \left[\Delta_{\uparrow\uparrow}(\mathbf{k}) \frac{-t_{\mathbf{k}x}^{\dagger} - it_{\mathbf{k}y}^{\dagger}}{\sqrt{2}} + \Delta_{\uparrow\downarrow}(\mathbf{k}) \frac{s_{\mathbf{k}}^{\dagger} + t_{\mathbf{k}z}^{\dagger}}{\sqrt{2}} + \Delta_{\downarrow\uparrow}(\mathbf{k}) \frac{-s_{\mathbf{k}}^{\dagger} + t_{\mathbf{k}z}^{\dagger}}{\sqrt{2}} + \Delta_{\downarrow\downarrow}(\mathbf{k}) \frac{t_{\mathbf{k}x}^{\dagger} - it_{\mathbf{k}y}^{\dagger}}{\sqrt{2}} \right] \\ &\stackrel{!}{=} -\sum_{\mathbf{k}} \sqrt{2} \left[\Delta_{\mathbf{k}} + \mathbf{d}(\mathbf{k}) \cdot \mathbf{t}_{\mathbf{k}}^{\dagger} \right] \end{aligned} \quad (12.56)$$

(the factor $\sqrt{2}$ is conventional), which requires

$$\Delta_{\mathbf{k}} = \frac{\Delta_{\uparrow\downarrow}(\mathbf{k}) - \Delta_{\downarrow\uparrow}(\mathbf{k})}{2}, \quad (12.57)$$

$$d_x(\mathbf{k}) = \frac{-\Delta_{\uparrow\uparrow}(\mathbf{k}) + \Delta_{\downarrow\downarrow}(\mathbf{k})}{2}, \quad (12.58)$$

$$d_y(\mathbf{k}) = -i \frac{\Delta_{\uparrow\uparrow}(\mathbf{k}) + \Delta_{\downarrow\downarrow}(\mathbf{k})}{2}, \quad (12.59)$$

$$d_z(\mathbf{k}) = \frac{\Delta_{\uparrow\downarrow}(\mathbf{k}) + \Delta_{\downarrow\uparrow}(\mathbf{k})}{2}. \quad (12.60)$$

The term in H_{MF} involving $\Delta_{\sigma\tau}^*(\mathbf{k})$ is just the hermitian conjugate of the one considered. We have now identified the singlet component of the gap, $\Delta_{\mathbf{k}}$, and the triplet components, $\mathbf{d}(\mathbf{k})$. Since

$$\hat{\Delta}(\mathbf{k}) = \begin{pmatrix} -d_x(\mathbf{k}) + id_y(\mathbf{k}) & \Delta_{\mathbf{k}} + d_z(\mathbf{k}) \\ -\Delta_{\mathbf{k}} + d_z(\mathbf{k}) & d_x(\mathbf{k}) + id_y(\mathbf{k}) \end{pmatrix}, \quad (12.61)$$

we can write the gap matrix in a compact form as

$$\hat{\Delta}(\mathbf{k}) = (\Delta_{\mathbf{k}}\mathbb{1} + \mathbf{d}(\mathbf{k}) \cdot \boldsymbol{\sigma}) i\sigma^y, \quad (12.62)$$

where $\boldsymbol{\sigma}$ is the vector of Pauli matrices. The mean-field Hamiltonian H_{MF} is diagonalized by a Bogoliubov transformation and the gap function $\hat{\Delta}(\mathbf{k})$ is obtained selfconsistently from a gap equation in complete analogy to the singlet case discussed in Sec. 10.1, except that $\hat{\Delta}(\mathbf{k})$ is now a matrix and that we require four coefficients $u_{\mathbf{k}\uparrow}, u_{\mathbf{k}\downarrow}, v_{\mathbf{k}\uparrow}, v_{\mathbf{k}\downarrow}$. We do not show this here explicitly.

A few remarks on the physics are in order, though. The symmetry $\hat{\Delta}(-\mathbf{k}) = -\hat{\Delta}^T(\mathbf{k})$ implies

$$\begin{aligned} (\Delta_{-\mathbf{k}}\mathbb{1} + \mathbf{d}(-\mathbf{k}) \cdot \boldsymbol{\sigma}) i\sigma^y &= \underbrace{-i(\sigma^y)^T}_{i\sigma^y} (\Delta_{\mathbf{k}}\mathbb{1} + \mathbf{d}(\mathbf{k}) \cdot \boldsymbol{\sigma}^T) \\ \Rightarrow \quad \Delta_{-\mathbf{k}}\mathbb{1} + \mathbf{d}(-\mathbf{k}) \cdot \boldsymbol{\sigma} &= \sigma^y (\Delta_{\mathbf{k}}\mathbb{1} + \mathbf{d}(\mathbf{k}) \cdot \boldsymbol{\sigma}^T) \sigma^y \\ &= \Delta_{\mathbf{k}}\mathbb{1} + \mathbf{d}(\mathbf{k}) \cdot \begin{pmatrix} \sigma^y(\sigma^x)^T\sigma^y \\ \sigma^y(\sigma^y)^T\sigma^y \\ \sigma^y(\sigma^z)^T\sigma^y \end{pmatrix} \\ &= \Delta_{\mathbf{k}}\mathbb{1} + \mathbf{d}(\mathbf{k}) \cdot \begin{pmatrix} \sigma^y\sigma^x\sigma^y \\ -\sigma^y\sigma^y\sigma^y \\ \sigma^y\sigma^z\sigma^y \end{pmatrix} \\ &= \Delta_{\mathbf{k}}\mathbb{1} - \mathbf{d}(\mathbf{k}) \cdot \boldsymbol{\sigma}. \end{aligned} \quad (12.63)$$

Thus we conclude that $\Delta_{\mathbf{k}}$ is even,

$$\Delta_{-\mathbf{k}} = \Delta_{\mathbf{k}}, \quad (12.65)$$

whereas $\mathbf{d}(\mathbf{k})$ is odd,

$$\mathbf{d}(-\mathbf{k}) = -\mathbf{d}(\mathbf{k}). \quad (12.66)$$

Hence, the d -vector can never be constant, unlike the single gap in Sec. 10.1. Furthermore, we can expand $\Delta_{\mathbf{k}}$ into even basis functions of rotations in real (and \mathbf{k}) space,

$$\Delta_{\mathbf{k}} = \Delta_s\psi_s(\mathbf{k}) + \Delta_{d_{x^2-y^2}}\psi_{d_{x^2-y^2}}(\mathbf{k}) + \Delta_{d_{3z^2-r^2}}\psi_{d_{3z^2-r^2}}(\mathbf{k}) + \Delta_{d_{xy}}\psi_{d_{xy}}(\mathbf{k}) + \Delta_{d_{yz}}\psi_{d_{yz}}(\mathbf{k}) + \Delta_{d_{zx}}\psi_{d_{zx}}(\mathbf{k}) + \dots, \quad (12.67)$$

and expand (the components of) $\mathbf{d}(\mathbf{k})$ into odd basis functions,

$$\mathbf{d}(\mathbf{k}) = \mathbf{d}_{p_x}\psi_{p_x}(\mathbf{k}) + \mathbf{d}_{p_y}\psi_{p_y}(\mathbf{k}) + \mathbf{d}_{p_z}\psi_{p_z}(\mathbf{k}) + \dots \quad (12.68)$$

For the singlet case we have already considered the basis functions $\psi_s(\mathbf{k}) = 1$ for conventional superconductors, $\psi_s(\mathbf{k}) = \cos k_x a \cos k_y a$ for the pnictides, and $\psi_{d_{x^2-y^2}}(\mathbf{k}) = \cos k_x a - \cos k_y a$ for the cuprates. A typical basis function for a triplet superconductor would be $\psi_{p_x} = \sin k_x a$. He-3 in the so-called B phase realized at now too high pressures (see Sec. 2.2) has the d -vector

$$\mathbf{d}(\mathbf{k}) = \hat{\mathbf{x}}\psi_{p_x}(\mathbf{k}) + \hat{\mathbf{y}}\psi_{p_y}(\mathbf{k}) + \hat{\mathbf{z}}\psi_{p_z}(\mathbf{k}), \quad (12.69)$$

where the basis functions are that standard expressions for p_x , p_y , and p_z orbitals (note that there is no Brillouin zone since He-3 is a liquid),

$$\psi_{p_x}(\mathbf{k}) = \sqrt{\frac{3}{4\pi}} \sin \theta_{\mathbf{k}} \cos \phi_{\mathbf{k}}, \quad (12.70)$$

$$\psi_{p_y}(\mathbf{k}) = \sqrt{\frac{3}{4\pi}} \sin \theta_{\mathbf{k}} \sin \phi_{\mathbf{k}}, \quad (12.71)$$

$$\psi_{p_z}(\mathbf{k}) = \sqrt{\frac{3}{4\pi}} \cos \theta_{\mathbf{k}}. \quad (12.72)$$

This is the so-called *Balian-Werthamer state*.

It is plausible that in a crystal the simultaneous presence of a non-vanishing \mathbf{k} -even order parameter $\Delta_{\mathbf{k}}$ and a non-vanishing \mathbf{k} -odd order parameter $\mathbf{d}(\mathbf{k})$ would break spatial inversion symmetry. Thus in an inversion-symmetric crystal singlet and triplet superconductivity do not coexist. However, in crystals lacking inversion symmetry (*noncentrosymmetric* crystals), singlet and triplet pairing can be realized at the same time. Moreover, one can show that spin-orbit coupling in a noncentrosymmetric superconductor mixes singlet and triplet pairing. In this case, $\Delta_{\mathbf{k}}$ and $\mathbf{d}(\mathbf{k})$ *must* both become non-zero simultaneously below T_c . Examples for such superconductors are CePt_3Si , CeRhSi_3 , and Y_2C_3 .