Lecture Notes: BCS theory of superconductivity

Prof. Rafael M. Fernandes

Here we will discuss a new ground state of the interacting electron gas: the superconducting state. In this macroscopic quantum state, the electrons form coherent bound states called Cooper pairs, which dramatically change the macroscopic properties of the system, giving rise to perfect conductivity and perfect diamagnetism. We will mostly focus on conventional superconductors, where the Cooper pairs originate from a small attractive electron-electron interaction mediated by phonons. However, in the so-called unconventional superconductors - a topic of intense research in current solid state physics - the pairing can originate even from purely repulsive interactions.

1 Phenomenology

Superconductivity was discovered by Kamerlingh-Onnes in 1911, when he was studying the transport properties of Hg (mercury) at low temperatures. He found that below the liquifying temperature of helium, at around 4.2 K, the resistivity of Hg would suddenly drop to zero. Although at the time there was not a well established model for the low-temperature behavior of transport in metals, the result was quite surprising, as the expectations were that the resistivity would either go to zero or diverge at $T = 0$, but not vanish at a finite temperature.

In a metal the resistivity at low temperatures has a constant contribution from impurity scattering, a T^2 contribution from electron-electron scattering, and a T^5 contribution from phonon scattering. Thus, the vanishing of the resistivity at low temperatures is a clear indication of a new ground state.

Another key property of the superconductor was discovered in 1933 by Meissner. He found that the magnetic flux density $\mathbf{B}$ is expelled below the superconducting transition temperature T_c, i.e. $\mathbf{B} = 0$ inside a superconductor material - the so-called *Meissner effect*. This means that the superconductor is a perfect diamagnet. Recall that the relationship between $\mathbf{B}$, the magnetic field $\mathbf{H}$, and the magnetization $\mathbf{M}$ is given by:

$$\mathbf{B} = \mathbf{H} + 4\pi\mathbf{M} \tag{1}$$

Therefore, since $\mathbf{B} = 0$ for a superconductor, the magnetic susceptibility $\chi = \partial\mathbf{M}/\partial\mathbf{H}$ is given by:

$$\chi = -\frac{1}{4\pi} \tag{2}$$

If one increases the magnetic field applied to a superconductor, it eventually destroys the superconducting state, driving the system back to the normal state. In type I superconductors, there is no intermediate state separating the transition from the superconducting to the normal state upon increasing field. In type II superconductors, on the other hand, there is an intermediate state, called mixed state, which appears before the transition to the normal state. In the mixed state, the magnetic field partially penetrates the material via the formation of an array of flux tubes carrying a multiple of the magnetic flux quantum $\Phi_0 = \frac{hc}{2|e|}$.

The first question we want to address is: which one of these two properties is "more fundamental", perfect conductivity or perfect diamagnetism? Let us study the implications of perfect conductivity using Maxwell equations. If a material is a perfect conductor, application of a an electric field freely accelerates the electric charge:

$$m\ddot{\mathbf{r}} = -e\mathbf{E} \tag{3}$$

But since the current density is given by $\mathbf{J} = -en_s\dot{\mathbf{r}}$, where n_s is the number of "superconducting electrons", we have:

$$\dot{\mathbf{J}} = \frac{n_s e^2}{m}\mathbf{E} \tag{4}$$

From Faraday law, we have:

$$\boldsymbol{\nabla} \times \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{B}}{\partial t} \tag{5}$$

which implies:

$$\boldsymbol{\nabla} \times \frac{\partial \mathbf{J}}{\partial t} = -\frac{n_s e^2}{cm}\frac{\partial \mathbf{B}}{\partial t} \tag{6}$$

But Ampere law gives

$$\boldsymbol{\nabla} \times \mathbf{B} = \frac{4\pi}{c}\mathbf{J} \tag{7}$$

and we obtain:

$$\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \frac{\partial \mathbf{B}}{\partial t} = -\frac{4\pi n_s e^2}{mc^2}\frac{\partial \mathbf{B}}{\partial t} \tag{8}$$

Using the identity $\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \mathbf{C} = \boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \mathbf{C}) - \nabla^2\mathbf{C}$ and Maxwell equation $\boldsymbol{\nabla} \cdot \mathbf{B} = 0$, we obtain the equation:

$$\nabla^2\left(\frac{\partial \mathbf{B}}{\partial t}\right) = \lambda^{-2}\left(\frac{\partial \mathbf{B}}{\partial t}\right) \tag{9}$$

where we defined the penetration depth:

$$\lambda = \sqrt{\frac{mc^2}{4\pi n_s e^2}} \tag{10}$$

What is the meaning of Eq. (9)? Consider a one-dimensional system that is a perfect conductor for $x > 0$. Solving the differential equation for x, and taking into account the boundary conditions, we obtain that the derivative $\partial \mathbf{B}/\partial t$ decays exponentially with x, i.e.

$$\frac{\partial \mathbf{B}}{\partial t} = \left(\frac{\partial \mathbf{B}}{\partial t}\right)_{x=0} \mathrm{e}^{-x/\lambda} \tag{11}$$

This means that the magnetic field inside a perfect conductor is constant over time. However, this is not the Meissner effect, which implies that the magnetic field is *zero* - not a constant - inside the superconductor. For instance, consider that a magnetic field $\mathbf{B}_0$ is applied to the material above T_c, when it is not yet a superconductor. If we cool down the system below T_c, the Meissner effect says that $\mathbf{B}_0$ has to be expelled from the material, since $\mathbf{B} = 0$ inside it. However, for a perfect conductor the field would remain $\mathbf{B}_0$ inside the material. This exercise tells us that a *superconductor is not just a perfect conductor*!

Based on this fact, the London brothers proposed a phenomenological model to describe the superconductors that arbitrarily eliminates the time derivatives from Eq. (9):

$$\nabla^2 \mathbf{B} = \lambda^{-2} \mathbf{B} \tag{12}$$

This equation correctly captures the Meissner effect, as we discussed above, emphasizing the perfect diamagnetic properties of the superconductor. Combined with Ampere law, this equation implies the following relationship between $\mathbf{J}$ and $\mathbf{B}$:

$$\boldsymbol{\nabla} \times \mathbf{J} = -\frac{n_s e^2}{mc} \mathbf{B} \tag{13}$$

Since $\mathbf{B} = \boldsymbol{\nabla} \times \mathbf{A}$, where $\mathbf{A}$ is the magnetic vector potential, the equation above becomes the London equation

$$\mathbf{J} = -\frac{n_s e^2}{mc} \mathbf{A} \tag{14}$$

in the so-called Coulomb gauge $\boldsymbol{\nabla} \cdot \mathbf{A} = 0$, i.e. in the gauge where the vector potential has only a non-zero transverse component. This gauge must be chosen because, from the continuity equation, the identity $\boldsymbol{\nabla} \cdot \mathbf{J} = 0$ must be satisfied.

How can we justify London equation? From a phenomenological point of view, it follows from the *rigidity* of the wave-function in the superconducting state. For instance, according to Bloch theorem, the total momentum of the system in its ground state (i.e. in the absence of any applied field) has a zero average value, $\langle \Psi | \mathbf{p} | \Psi \rangle = 0$. Now, let us assume that the wave-function Ψ is rigid, i.e. that this relationship holds even in the presence of an external field. Then, since the canonical momentum is given by $\mathbf{p} = m\mathbf{v} - e\mathbf{A}/c$, we obtain:

$$\langle \mathbf{v} \rangle = \frac{e\mathbf{A}}{mc} \tag{15}$$

Since $\mathbf{J} = -en_s \langle \mathbf{v} \rangle$, we recover London equation (14).

Of course, the main question is about the microscopic mechanism that gives rise to this wave-function rigidity and, ultimately, to the superconducting state. Several of the most brilliant physicists of the last century tried to address this question - such as Bohr, Einstein, Feynman, Born, Heisenberg - but the answer only came in 1957 with the famous theory of Bardeen, Cooper, and Schrieffer (BCS) - almost 50 years after the experimental discovery by Kamerlingh-Onnes!

Key experimental contributions made the main properties of the superconductors more transparent before the BCS theory appeared in 1957. The observation of an exponential decay of the specific heat at low temperatures showed that the energy spectrum of a superconductor is gapped. This is in contrast to the spectrum of a regular metal, which is gapless - recall that exciting an electron-hole pair near the Fermi surface costs very little energy to the metal.

Another key experiment was the observation of the isotope effect. By studying the superconducting transition temperature T_c of materials containing a different element isotope, it was shown that T_c decays with $M^{-1/2}$, where M is the mass of the isotope. Since this mass is related only to the ions forming the lattice, this experimental observation indicated that the lattice - and therefore the phonons - must play a key role in the formation of the superconducting state.

The main point of the BCS theory is that the attractive electron-electron interaction mediated by the phonons gives rise to Cooper pairs, i.e. bound states formed by two electrons of opposite spins and momenta. These Cooper pairs form then a coherent macroscopic ground state, which displays a gapped spectrum and perfect diamagnetism. Key to the formation of Cooper pairs is the existence of a well-defined Fermi surface, as we will discuss below.

2 One Cooper pair

Much of the physics involved in the BCS theory can be discussed in the context of a simple quantum mechanics problem. Consider two electrons that interact with each other via an attractive potential $V(\mathbf{r}_1 - \mathbf{r}_2)$. The Schrödinger equation is given by:

$$\left[-\frac{\hbar^2 \nabla^2_{\mathbf{r}_1}}{2m} - \frac{\hbar^2 \nabla^2_{\mathbf{r}_2}}{2m} + V(\mathbf{r}_1 - \mathbf{r}_2) \right] \Psi(\mathbf{r}_1, \mathbf{r}_2) = E \Psi(\mathbf{r}_1, \mathbf{r}_2) \tag{16}$$

where $\Psi(\mathbf{r}_1, \mathbf{r}_2)$ is the wave-function and E, the energy. As usual, we change variables to the relative displacement $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$ and to the position of the center of mass $\mathbf{R} = \frac{1}{2}(\mathbf{r}_1 + \mathbf{r}_2)$. In terms of these new variables, the Schrödinger equation becomes:

$$\left[-\frac{\hbar^2\nabla^2_{\mathbf{R}}}{2m^*}-\frac{\hbar^2\nabla^2_{\mathbf{r}}}{2\mu}+V\left(\mathbf{r}\right)\right]\Psi\left(\mathbf{r},\mathbf{R}\right)=E\Psi\left(\mathbf{r},\mathbf{R}\right) \tag{17}$$

where $m^*=2m$ is the total mass and $\mu=m/2$ is the reduced mass. Since the potential does not depend on the center of mass coordinate $\mathbf{R}$, we look for the solution:

$$\Psi\left(\mathbf{r},\mathbf{R}\right)=\psi\left(\mathbf{r}\right)\mathrm{e}^{i\mathbf{K}\cdot\mathbf{R}} \tag{18}$$

which gives:

$$\left[-\frac{\hbar^2\nabla^2_{\mathbf{r}}}{2\mu}+V\left(\mathbf{r}\right)\right]\psi\left(\mathbf{r}\right)=\tilde{E}\psi\left(\mathbf{r}\right) \tag{19}$$

where we defined $\tilde{E}=E-\frac{\hbar^2K^2}{2m^*}$. For a given eigenvalue $\tilde{E}$, the lowest energy E is the one for which $\mathbf{K}=0$, i.e. for which the momentum of the center of mass vanishes. Thus, for now we consider $E=\tilde{E}$. In this case, the two electrons have opposite momenta. Depending on the symmetry of the spatial part of the wave-function, even $\psi\left(\mathbf{r}\right)=\psi\left(-\mathbf{r}\right)$ or odd $\psi\left(\mathbf{r}\right)=-\psi\left(-\mathbf{r}\right)$, the spins of the electrons will form either a singlet or a triplet state, respectively, in order to ensure the anti-symmetry of the total wave-function.

To proceed, we take the Fourier transform of the Schrödinger equation, by introducing:

$$\psi\left(\mathbf{k}\right)=\int d^3r\,\psi\left(\mathbf{r}\right)\mathrm{e}^{-i\mathbf{k}\cdot\mathbf{r}} \tag{20}$$

It follows that:

$$\begin{aligned}
\frac{\hbar^2k^2}{2\mu}\psi\left(\mathbf{k}\right)+\int d^3r\,V\left(\mathbf{r}\right)\psi\left(\mathbf{r}\right)\mathrm{e}^{-i\mathbf{k}\cdot\mathbf{r}} &= E\psi\left(\mathbf{k}\right)\\
\int\frac{d^3q}{\left(2\pi\right)^3}V\left(\mathbf{q}\right)\int d^3r\psi\left(\mathbf{r}\right)\mathrm{e}^{-i\left(\mathbf{k}-\mathbf{q}\right)\cdot\mathbf{r}} &= \left(E-\frac{\hbar^2k^2}{m}\right)\psi\left(\mathbf{k}\right)\\
\int\frac{d^3k'}{\left(2\pi\right)^3}V\left(\mathbf{k}-\mathbf{k}'\right)\psi\left(\mathbf{k}'\right) &= \left(E-2\varepsilon_{\mathbf{k}}\right)\psi\left(\mathbf{k}\right)
\end{aligned} \tag{21}$$

In the last line, we changed variables to $\mathbf{q}=\mathbf{k}-\mathbf{k}'$ and defined the free electron energy $\varepsilon_{\mathbf{k}}=\frac{\hbar^2k^2}{2m}$. A bound state between the electrons has $E<2\varepsilon_{\mathbf{k}}$, i.e. the total energy is smaller than the energy of two independent free electrons. Therefore, we define the modified wave-function

$$\Delta\left(\mathbf{k}\right)=\left(E-2\varepsilon_{\mathbf{k}}\right)\psi\left(\mathbf{k}\right) \tag{22}$$

which gives:

$$\Delta(\mathbf{k}) = -\int \frac{d^3k'}{(2\pi)^3} \frac{V(\mathbf{k}-\mathbf{k}')}{2\varepsilon_{\mathbf{k}'} - E} \Delta(\mathbf{k}') \tag{23}$$

Notice that the previous equation is nothing but the Schrödinger equation written in a different form. Inspired on our results for the phonon-mediated electron-electron interaction, let us consider a potential that is attractive $V(\mathbf{k}-\mathbf{k}') = -V_0$ for $\varepsilon_{\mathbf{k}'}, \varepsilon_{\mathbf{k}} < \hbar\omega_D$ and zero otherwise. Recall that ω_D is the Debye frequency. We look for a solution with constant $\Delta(\mathbf{k}) = \Delta$. Since this implies an even spatial wave-function $\psi(\mathbf{r}) = \psi(-\mathbf{r})$, the spins of the two electrons must be anti-parallel (a singlet).

Defining the "density of states per spin" (recall that we are considering only a two-electron system):

$$\rho(\varepsilon) = \frac{m^{3/2}}{\sqrt{2}\hbar^3\pi^2}\sqrt{\varepsilon} \tag{24}$$

we obtain:

$$\begin{aligned}
\Delta &= \frac{V_0 \Delta m^{3/2}}{\sqrt{2}\hbar^3\pi^2} \int_0^{\omega_D} \frac{d\varepsilon\sqrt{\varepsilon}}{2\varepsilon - E} \\
1 &= \frac{V_0 m^{3/2}}{\sqrt{2}\hbar^3\pi^2} \left[\sqrt{\omega_D} - \sqrt{\frac{-E}{2}} \arctan\left(\sqrt{\frac{2\omega_D}{-E}} \right) \right]
\end{aligned} \tag{25}$$

This equation determines the value of the bound state energy $E < 0$ as function of the attractive potential V_0. In order to have a bound state, we set $E \to 0^-$ to obtain the minimum value of V_0:

$$V_{0,\text{min}} = \frac{\sqrt{2}\hbar^3\pi^2}{m^{3/2}\sqrt{\omega_D}} \tag{26}$$

Therefore, we find that there will be a bound state only if the attractive interaction is strong enough.

However, in this exercise we overlooked an important feature: in the actual many-body system, only the electrons near the Fermi level will be affected by the attractive interaction. To mimic this property, we consider an attractive potential $V(\mathbf{k}-\mathbf{k}') = -V_0$ for the unoccupied electronic states above the Fermi energy ε_F, $\varepsilon_{\mathbf{k}'} - \varepsilon_F, \varepsilon_{\mathbf{k}} - \varepsilon_F < \hbar\omega_D$. Since $\hbar\omega_D \ll \varepsilon_F$, we can approximate the density of states for its value at ε_F. Then Eq. (23) becomes, for $\Delta(\mathbf{k}) = \Delta$:

$$\begin{aligned}
\Delta &= V_0\rho(\varepsilon_F)\Delta \int_{\varepsilon_F}^{\varepsilon_F+\omega_D} \frac{d\varepsilon}{2\varepsilon - E} \\
\frac{2}{V_0\rho(\varepsilon_F)} &= \ln\left(\frac{2\varepsilon_F - E + 2\omega_D}{2\varepsilon_F - E} \right)
\end{aligned} \tag{27}$$

In the limit of small $V_0\rho(\varepsilon_F) \ll 1$, E is close to $2\varepsilon_F$, and we can approximate $2\varepsilon_F - E + 2\omega_D \approx 2\omega_D$. Defining the binding energy $E_b \equiv 2\varepsilon_F - E$, we obtain:

$$E_b = 2\omega_D \mathrm{e}^{-\frac{2}{V_0\rho(\varepsilon_F)}} \tag{28}$$

This shows that a bound state will be formed regardless of how small the attractive interaction V_0 is. Such a bound state is called a Cooper pair. This is fundamentally different from the free electron case we considered before, where the attractive interaction has to overcome a threshold to create a bound state. The key property responsible for this different behavior is the existence of a well-defined Fermi surface, separating states that are occupied from states that are unoccupied.

To finish this section, let us recall that the total energy in the case where the center of mass has a finite momentum $\mathbf{K}$ is given by:

$$\begin{aligned} E &= E_{\mathbf{K}=0} + \frac{\hbar^2 K^2}{4m} \\ E &= 2\varepsilon_F - E_b + \frac{\hbar^2 K^2}{4m} \end{aligned}$$

Thus, in the limit where $E \to 2\varepsilon_F$, we can still obtain a bound state with a finite center-of-mass momentum:

$$K = \frac{2}{\hbar}\sqrt{mE_b} \tag{29}$$

This gives rise to a finite current density:

$$J = n_s e \frac{\hbar K}{m} = 2 n_s e \sqrt{\frac{E_b}{m}} \tag{30}$$

3 Many Cooper pairs: BCS state

In the previous section we saw that two electrons near the Fermi level are unstable towards the formation of a Cooper pair for an arbitrarily small attractive interaction. Thus, we expect that the many-body electronic system will be unstable towards the formation of a new ground state, where these Cooper pairs proliferate. In this section, we will study this BCS state using mean-field theory.

3.1 Effective Hamiltonian and the BCS wave-function

To investigate the onset of superconductivity, we consider the following effective Hamiltonian:

$$H = \sum_{\mathbf{k}\sigma} \xi_{\mathbf{k}} c^{\dagger}_{\mathbf{k}\sigma} c_{\mathbf{k}\sigma} + \frac{1}{N} \sum_{\mathbf{k}\mathbf{k}'} V_{\mathbf{k}\mathbf{k}'} c^{\dagger}_{\mathbf{k}\uparrow} c^{\dagger}_{-\mathbf{k}\downarrow} c_{-\mathbf{k}'\downarrow} c_{\mathbf{k}'\uparrow} \tag{31}$$

Here, $c^{\dagger}_{\mathbf{k}\sigma}$ creates an electron with momentum $\mathbf{k}$ and spin σ, and we already included the chemical potential by defining $\xi_{\mathbf{k}} = \varepsilon_{\mathbf{k}} - \mu$. The second term describes the destruction of a Cooper pair (two electrons with opposite momenta and spin) and the subsequent creation of another Cooper pair.

To proceed, we perform the usual mean-field decoupling of the quartic term:

$$\left\langle c^\dagger_{\mathbf{k}\uparrow} c^\dagger_{-\mathbf{k}\downarrow} c_{-\mathbf{k}'\downarrow} c_{\mathbf{k}'\uparrow} \right\rangle \approx \left\langle c^\dagger_{\mathbf{k}\uparrow} c^\dagger_{-\mathbf{k}\downarrow} \right\rangle c_{-\mathbf{k}'\downarrow} c_{\mathbf{k}'\uparrow} + c^\dagger_{\mathbf{k}\uparrow} c^\dagger_{-\mathbf{k}\downarrow} \left\langle c_{-\mathbf{k}'\downarrow} c_{\mathbf{k}'\uparrow} \right\rangle - \left\langle c^\dagger_{\mathbf{k}\uparrow} c^\dagger_{-\mathbf{k}\downarrow} \right\rangle \left\langle c_{-\mathbf{k}'\downarrow} c_{\mathbf{k}'\uparrow} \right\rangle \tag{32}$$

Differently than the previous mean-field calculations we did, the mean value $\left\langle c^\dagger_{\mathbf{k}\uparrow} c^\dagger_{-\mathbf{k}\downarrow} \right\rangle$ is not zero, since it corresponds to one Cooper pair in the superconducting state. Thus, we define the *gap function*:

$$\Delta_{\mathbf{k}} = -\frac{1}{N} \sum_{\mathbf{k}'} V_{\mathbf{k}\mathbf{k}'} \left\langle c_{-\mathbf{k}'\downarrow} c_{\mathbf{k}'\uparrow} \right\rangle \tag{33}$$

For now, there is no reason to call it a gap, but we will discuss its meaning very soon. The effective Hamiltonian becomes:

$$H = \sum_{\mathbf{k}\sigma} \xi_{\mathbf{k}} c^\dagger_{\mathbf{k}\sigma} c_{\mathbf{k}\sigma} - \sum_{\mathbf{k}} \left(\Delta_{\mathbf{k}} c^\dagger_{\mathbf{k}\uparrow} c^\dagger_{-\mathbf{k}\downarrow} + \Delta^*_{\mathbf{k}} c_{-\mathbf{k}\downarrow} c_{\mathbf{k}\uparrow} \right) + \sum_{\mathbf{k}} \Delta_{\mathbf{k}} \left\langle c^\dagger_{\mathbf{k}\uparrow} c^\dagger_{-\mathbf{k}\downarrow} \right\rangle \tag{34}$$

To solve it, we employ the so-called Bogoliubov transformation. In particular, we define new fermionic operators $\gamma_{\mathbf{k}\sigma}$ and coefficients $u_{\mathbf{k}}$, $v_{\mathbf{k}}$:

$$\begin{aligned} c_{\mathbf{k}\uparrow} &= u^*_{\mathbf{k}} \gamma_{\mathbf{k}\uparrow} + v_{\mathbf{k}} \gamma^\dagger_{-\mathbf{k}\downarrow} \\ c^\dagger_{-\mathbf{k}\downarrow} &= u_{\mathbf{k}} \gamma^\dagger_{-\mathbf{k}\downarrow} - v^*_{\mathbf{k}} \gamma_{\mathbf{k}\uparrow} \end{aligned} \tag{35}$$

In order for the fermionic commutation relations to be satisfied, the normalization condition:

$$|u_{\mathbf{k}}|^2 + |v_{\mathbf{k}}|^2 = 1 \tag{36}$$

must be satisfied. Substituting in the effective Hamiltonian yields:

$$\begin{aligned} \sum_{\mathbf{k}\sigma} \xi_{\mathbf{k}} c^\dagger_{\mathbf{k}\sigma} c_{\mathbf{k}\sigma} &= \sum_{\mathbf{k}} \xi_{\mathbf{k}} \left[c^\dagger_{\mathbf{k}\uparrow} c_{\mathbf{k}\uparrow} + c^\dagger_{-\mathbf{k}\downarrow} c_{-\mathbf{k}\downarrow} \right] \\ &= \sum_{\mathbf{k}} \xi_{\mathbf{k}} \left[\left(|u_{\mathbf{k}}|^2 - |v_{\mathbf{k}}|^2 \right) \left(\gamma^\dagger_{\mathbf{k}\uparrow} \gamma_{\mathbf{k}\uparrow} + \gamma^\dagger_{-\mathbf{k}\downarrow} \gamma_{-\mathbf{k}\downarrow} \right) + 2 |v_{\mathbf{k}}|^2 + 2 u_{\mathbf{k}} v_{\mathbf{k}} \gamma^\dagger_{\mathbf{k}\uparrow} \gamma^\dagger_{-\mathbf{k}\downarrow} + 2 u^*_{\mathbf{k}} v^*_{\mathbf{k}} \gamma_{-\mathbf{k}\downarrow} \gamma_{\mathbf{k}\uparrow} \right] \end{aligned}$$

as well as:

$$\begin{aligned} -\sum_{\mathbf{k}} \left(\Delta_{\mathbf{k}} c^\dagger_{\mathbf{k}\uparrow} c^\dagger_{-\mathbf{k}\downarrow} + \Delta^*_{\mathbf{k}} c_{-\mathbf{k}\downarrow} c_{\mathbf{k}\uparrow} \right) &= \sum_{\mathbf{k}} \left[\left(\Delta_{\mathbf{k}} u_{\mathbf{k}} v^*_{\mathbf{k}} + \Delta^*_{\mathbf{k}} u^*_{\mathbf{k}} v_{\mathbf{k}} \right) \left(\gamma^\dagger_{\mathbf{k}\uparrow} \gamma_{\mathbf{k}\uparrow} + \gamma^\dagger_{-\mathbf{k}\downarrow} \gamma_{-\mathbf{k}\downarrow} \right) - \left(\Delta_{\mathbf{k}} u_{\mathbf{k}} v^*_{\mathbf{k}} + \Delta^*_{\mathbf{k}} u^*_{\mathbf{k}} v_{\mathbf{k}} \right) \right] \\ &\quad - \sum_{\mathbf{k}} \left[\left(\Delta_{\mathbf{k}} u^2_{\mathbf{k}} - \Delta^*_{\mathbf{k}} v^2_{\mathbf{k}} \right) \gamma^\dagger_{\mathbf{k}\uparrow} \gamma^\dagger_{-\mathbf{k}\downarrow} + \left(\Delta^*_{\mathbf{k}} \left(u^*_{\mathbf{k}} \right)^2 - \Delta_{\mathbf{k}} \left(v^*_{\mathbf{k}} \right)^2 \right) \gamma_{-\mathbf{k}\downarrow} \gamma_{\mathbf{k}\uparrow} \right] \end{aligned}$$

Therefore, the effective Hamiltonian becomes:

$$H = H_0 + H_1 + H_2 \tag{37}$$

with:

$$\begin{aligned}
H_0 &= \sum_{\mathbf{k}} \left[2\xi_{\mathbf{k}} \left|v_{\mathbf{k}}\right|^2 - \Delta_{\mathbf{k}} u_{\mathbf{k}} v_{\mathbf{k}}^* - \Delta_{\mathbf{k}}^* u_{\mathbf{k}}^* v_{\mathbf{k}} + \Delta_{\mathbf{k}} \left\langle c_{\mathbf{k}\uparrow}^\dagger c_{-\mathbf{k}\downarrow}^\dagger \right\rangle\right] \\
H_1 &= \sum_{\mathbf{k}} \left[\xi_{\mathbf{k}} \left(\left|u_{\mathbf{k}}\right|^2 - \left|v_{\mathbf{k}}\right|^2\right) + \Delta_{\mathbf{k}} u_{\mathbf{k}} v_{\mathbf{k}}^* + \Delta_{\mathbf{k}}^* u_{\mathbf{k}}^* v_{\mathbf{k}}\right] \left(\gamma_{\mathbf{k}\uparrow}^\dagger \gamma_{\mathbf{k}\uparrow} + \gamma_{-\mathbf{k}\downarrow}^\dagger \gamma_{-\mathbf{k}\downarrow}\right) \\
H_2 &= \sum_{\mathbf{k}} \left[\left(2\xi_{\mathbf{k}} u_{\mathbf{k}} v_{\mathbf{k}} - \Delta_{\mathbf{k}} u_{\mathbf{k}}^2 + \Delta_{\mathbf{k}}^* v_{\mathbf{k}}^2\right)\right] \left(\gamma_{\mathbf{k}\uparrow}^\dagger \gamma_{-\mathbf{k}\downarrow}^\dagger\right) + \text{h.c.}
\end{aligned} \tag{38}$$

where h.c. denotes the hermitian conjugate. To diagonalize the Hamiltonian, we must find the coefficients $u_{\mathbf{k}}$, $v_{\mathbf{k}}$ that make the undesired term H_2 vanish. Hence, we obtain the quadratic equation:

$$2\xi_{\mathbf{k}} u_{\mathbf{k}} v_{\mathbf{k}} - \Delta_{\mathbf{k}} u_{\mathbf{k}}^2 + \Delta_{\mathbf{k}}^* v_{\mathbf{k}}^2 = 0 \tag{39}$$

Solving for the ratio $v_{\mathbf{k}}/u_{\mathbf{k}}$ yields:

$$\frac{v_{\mathbf{k}}}{u_{\mathbf{k}}} = \frac{\sqrt{\xi_{\mathbf{k}}^2 + \left|\Delta_{\mathbf{k}}\right|^2} - \xi_{\mathbf{k}}}{\Delta_{\mathbf{k}}^*} \tag{40}$$

where we picked only the positive root to ensure that the energy of the BCS state is a minimum and not a maximum. Notice that because the numerator is real, the phase of the complex gap function $\Delta_{\mathbf{k}}$ must be the same as the relative phase between $v_{\mathbf{k}}$ and $u_{\mathbf{k}}$. Since we can set the phase of $u_{\mathbf{k}}$ to be zero without loss of generality, it follows that the phases of $v_{\mathbf{k}}$ and $\Delta_{\mathbf{k}}$ are the same.

Using the normalization condition $\left|u_{\mathbf{k}}\right|^2 + \left|v_{\mathbf{k}}\right|^2 = 1$, we obtain:

$$\begin{aligned}
\left|u_{\mathbf{k}}\right|^2 &= \frac{1}{1 + \left|\frac{v_{\mathbf{k}}}{u_{\mathbf{k}}}\right|^2} = \frac{1}{2} \frac{\left|\Delta_{\mathbf{k}}\right|^2}{\xi_{\mathbf{k}}^2 + \left|\Delta_{\mathbf{k}}\right|^2 - \xi_{\mathbf{k}} \sqrt{\xi_{\mathbf{k}}^2 + \left|\Delta_{\mathbf{k}}\right|^2}} \\
\left|u_{\mathbf{k}}\right|^2 &= \frac{1}{2} \frac{\left|\Delta_{\mathbf{k}}\right|^2 \left(\sqrt{\xi_{\mathbf{k}}^2 + \left|\Delta_{\mathbf{k}}\right|^2} + \xi_{\mathbf{k}}\right)}{\left(\sqrt{\xi_{\mathbf{k}}^2 + \left|\Delta_{\mathbf{k}}\right|^2}\right) \left(\xi_{\mathbf{k}}^2 + \left|\Delta_{\mathbf{k}}\right|^2 - \xi_{\mathbf{k}}^2\right)} \\
\left|u_{\mathbf{k}}\right|^2 &= \frac{1}{2} \left(1 + \frac{\xi_{\mathbf{k}}}{\sqrt{\xi_{\mathbf{k}}^2 + \left|\Delta_{\mathbf{k}}\right|^2}}\right)
\end{aligned} \tag{41}$$

from which follows:

$$|v_{\mathbf{k}}|^2 = \frac{1}{2}\left(1 - \frac{\xi_{\mathbf{k}}}{\sqrt{\xi_{\mathbf{k}}^2 + |\Delta_{\mathbf{k}}|^2}}\right) \tag{42}$$

It is convenient to define the excitation energy:

$$E_{\mathbf{k}} = \sqrt{\xi_{\mathbf{k}}^2 + |\Delta_{\mathbf{k}}|^2} \tag{43}$$

Using the above relations, we obtain:

$$\begin{aligned} H_1 &= \sum_{\mathbf{k}} \left[\xi_{\mathbf{k}}\left(|u_{\mathbf{k}}|^2 - |v_{\mathbf{k}}|^2\right) + \Delta_{\mathbf{k}} u_{\mathbf{k}} v_{\mathbf{k}}^* + \Delta_{\mathbf{k}}^* u_{\mathbf{k}}^* v_{\mathbf{k}}\right]\left(\gamma_{\mathbf{k}\uparrow}^\dagger \gamma_{\mathbf{k}\uparrow} + \gamma_{-\mathbf{k}\downarrow}^\dagger \gamma_{-\mathbf{k}\downarrow}\right) \\ &= \sum_{\mathbf{k}} \left[\frac{\xi_{\mathbf{k}}^2}{\sqrt{\xi_{\mathbf{k}}^2 + |\Delta_{\mathbf{k}}|^2}} + \left(1 + \frac{\xi_{\mathbf{k}}}{\sqrt{\xi_{\mathbf{k}}^2 + |\Delta_{\mathbf{k}}|^2}}\right)\left(\sqrt{\xi_{\mathbf{k}}^2 + |\Delta_{\mathbf{k}}|^2} - \xi_{\mathbf{k}}\right)\right]\left(\gamma_{\mathbf{k}\uparrow}^\dagger \gamma_{\mathbf{k}\uparrow} + \gamma_{-\mathbf{k}\downarrow}^\dagger \gamma_{-\mathbf{k}\downarrow}\right) \\ &= \sum_{\mathbf{k}} \sqrt{\xi_{\mathbf{k}}^2 + |\Delta_{\mathbf{k}}|^2}\left(\gamma_{\mathbf{k}\uparrow}^\dagger \gamma_{\mathbf{k}\uparrow} + \gamma_{-\mathbf{k}\downarrow}^\dagger \gamma_{-\mathbf{k}\downarrow}\right) \end{aligned} \tag{44}$$

as well as:

$$\begin{aligned} H_0 &= \sum_{\mathbf{k}} \left[2\xi_{\mathbf{k}} |v_{\mathbf{k}}|^2 - \Delta_{\mathbf{k}} u_{\mathbf{k}} v_{\mathbf{k}}^* - \Delta_{\mathbf{k}}^* u_{\mathbf{k}}^* v_{\mathbf{k}} + \Delta_{\mathbf{k}}\left\langle c_{\mathbf{k}\uparrow}^\dagger c_{-\mathbf{k}\downarrow}^\dagger\right\rangle\right] \\ H_0 &= \sum_{\mathbf{k}} \left[\left(\xi_{\mathbf{k}} - \frac{\xi_{\mathbf{k}}^2}{\sqrt{\xi_{\mathbf{k}}^2 + |\Delta_{\mathbf{k}}|^2}}\right) - \left(1 + \frac{\xi_{\mathbf{k}}}{\sqrt{\xi_{\mathbf{k}}^2 + |\Delta_{\mathbf{k}}|^2}}\right)\left(\sqrt{\xi_{\mathbf{k}}^2 + |\Delta_{\mathbf{k}}|^2} - \xi_{\mathbf{k}}\right) + \Delta_{\mathbf{k}}\left\langle c_{\mathbf{k}\uparrow}^\dagger c_{-\mathbf{k}\downarrow}^\dagger\right\rangle\right] \\ H_0 &= \sum_{\mathbf{k}} \left(\xi_{\mathbf{k}} - \sqrt{\xi_{\mathbf{k}}^2 + |\Delta_{\mathbf{k}}|^2} + \Delta_{\mathbf{k}}\left\langle c_{\mathbf{k}\uparrow}^\dagger c_{-\mathbf{k}\downarrow}^\dagger\right\rangle\right) \end{aligned} \tag{45}$$

Therefore, the effective Hamiltonian is:

$$H = \sum_{\mathbf{k}\sigma} E_{\mathbf{k}} \gamma_{\mathbf{k}\sigma}^\dagger \gamma_{\mathbf{k}\sigma} + E_0 \tag{46}$$

where E_0 is just the ground-state energy:

$$E_0 = \sum_{\mathbf{k}} \left(\xi_{\mathbf{k}} - E_{\mathbf{k}} + \Delta_{\mathbf{k}}\left\langle c_{\mathbf{k}\uparrow}^\dagger c_{-\mathbf{k}\downarrow}^\dagger\right\rangle\right) \tag{47}$$

It becomes clear from Eq. (46) why we called $\Delta_{\mathbf{k}}$ the gap function: even at the Fermi level, where $\xi_{\mathbf{k}} = 0$, the energy spectrum of the superconductor has a gap of size $|\Delta_{\mathbf{k}}|$. Thus, we need to give the minimum energy of $2|\Delta_{\mathbf{k}}|$ to the system to excite its quasi-particles, which are described by the operators $\gamma_{\mathbf{k}\sigma}^\dagger$ and are usually called *Bogoliubons*.

Note from Eq. (35) that a Bogoliubon is a mixture of electrons and holes:

$$\begin{aligned}\gamma_{\mathbf{k}\uparrow} &= u_{\mathbf{k}} c_{\mathbf{k}\uparrow} - v_{\mathbf{k}} c^{\dagger}_{-\mathbf{k}\downarrow} \\ \gamma^{\dagger}_{-\mathbf{k}\downarrow} &= u^{*}_{\mathbf{k}} c^{\dagger}_{-\mathbf{k}\downarrow} + v^{*}_{\mathbf{k}} c_{\mathbf{k}\uparrow}\end{aligned}$$

From Eqs. (41) and (42) describing the behavior of $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$, we have that as $\Delta_{\mathbf{k}} \to 0$, $|u_{\mathbf{k}}|^2 \to 1$ for $\xi_{\mathbf{k}} > 0$ and $|u_{\mathbf{k}}|^2 \to 0$ for $\xi_{\mathbf{k}} < 0$ whereas $|v_{\mathbf{k}}|^2 \to 1$ for $\xi_{\mathbf{k}} < 0$ and $|v_{\mathbf{k}}|^2 \to 0$ for $\xi_{\mathbf{k}} > 0$. Thus, at the normal state, creating a Bogoliubon excitation corresponds to creating an electron for energies above the Fermi level and creating a hole (destroying an electron) of opposite momentum and spin for energies below the Fermi level. At the superconducting state, a Bogoliubon becomes a superposition of both an electron and a hole state.

The BCS ground state wave-function, therefore, corresponds to the vacuum of Bogoliubons:

$$\gamma_{\mathbf{k}\sigma} |\Psi_{\text{BCS}}\rangle = 0 \tag{48}$$

How can this wave-function be written in terms of the original vacuum of electrons $|0\rangle$? To find this out, it is sufficient to consider only one spin species of Bogoliubons. Written in terms of the electron operators, the condition above becomes:

$$u_{\mathbf{k}} c_{\mathbf{k}\uparrow} |\Psi_{\text{BCS}}\rangle = v_{\mathbf{k}} c^{\dagger}_{-\mathbf{k}\downarrow} |\Psi_{\text{BCS}}\rangle \tag{49}$$

We now write the BCS wave-function as an arbitrary combination of Cooper pairs:

$$|\Psi_{\text{BCS}}\rangle = \mathcal{N} \prod_{\mathbf{q}} \mathrm{e}^{\alpha_{\mathbf{q}} c^{\dagger}_{\mathbf{q}\uparrow} c^{\dagger}_{-\mathbf{q}\downarrow}} |0\rangle \tag{50}$$

where $\mathcal{N}$ is a normalization constant and $\alpha_{\mathbf{q}}$ is a function to be determined. Clearly, when $c_{\mathbf{k}\uparrow}$ acts on the wave-function above, the only term inside the product that does not commute with $c_{\mathbf{k}\uparrow}$ is the one for which $\mathbf{q} = \mathbf{k}$. Let us focus on this term. Defining $\theta_{\mathbf{k}} = \alpha_{\mathbf{k}} c^{\dagger}_{\mathbf{k}\uparrow} c^{\dagger}_{-\mathbf{k}\downarrow}$ to simplify the notation, we have:

$$c_{\mathbf{k}\uparrow} \mathrm{e}^{\alpha_{\mathbf{k}} c^{\dagger}_{\mathbf{k}\uparrow} c^{\dagger}_{-\mathbf{k}\downarrow}} |0\rangle = c_{\mathbf{k}\uparrow} \mathrm{e}^{\theta_{\mathbf{k}}} |0\rangle = \sum_{n=1}^{\infty} \frac{c_{\mathbf{k}\uparrow} \theta^n_{\mathbf{k}}}{n!} |0\rangle \tag{51}$$

Now, we have the commutation relation:

$$\left[c_{\mathbf{k}\uparrow}, \theta_{\mathbf{k}}\right] = \alpha_{\mathbf{k}} \left\{c_{\mathbf{k}\uparrow}, c^{\dagger}_{\mathbf{k}\uparrow}\right\} c^{\dagger}_{-\mathbf{k}\downarrow} = \alpha_{\mathbf{k}} c^{\dagger}_{-\mathbf{k}\downarrow} \tag{52}$$

where we used the fact that $[A, BC] = \{A, B\} C - B \{A, C\}$. Hence, since $c_{\mathbf{k}\uparrow} |0\rangle = 0$, it follows that:

$$
\begin{aligned}
c_{\mathbf{k}\uparrow}\theta_{\mathbf{k}}\left|0\right\rangle &= \alpha_{\mathbf{k}}c^{\dagger}_{-\mathbf{k}\downarrow}\left|0\right\rangle \\
c_{\mathbf{k}\uparrow}\theta^{2}_{\mathbf{k}}\left|0\right\rangle &= \left(\left[c_{\mathbf{k}\uparrow}\theta_{\mathbf{k}},\theta_{\mathbf{k}}\right]+\theta_{\mathbf{k}}c_{\mathbf{k}\uparrow}\theta_{\mathbf{k}}\right)\left|0\right\rangle = \theta_{\mathbf{k}}\left(\left[c_{\mathbf{k}\uparrow},\theta_{\mathbf{k}}\right]+c_{\mathbf{k}\uparrow}\theta_{\mathbf{k}}\right)\left|0\right\rangle = 2\theta_{\mathbf{k}}\alpha_{\mathbf{k}}c^{\dagger}_{-\mathbf{k}\downarrow}\left|0\right\rangle
\end{aligned}
$$

and, in general,

$$
c_{\mathbf{k}\uparrow}\theta^{n}_{\mathbf{k}}\left|0\right\rangle = n\theta^{n-1}_{\mathbf{k}}\alpha_{\mathbf{k}}c^{\dagger}_{-\mathbf{k}\downarrow}\left|0\right\rangle \tag{53}
$$

Therefore, we obtain:

$$
c_{\mathbf{k}\uparrow}\mathrm{e}^{\alpha_{\mathbf{k}}c^{\dagger}_{\mathbf{k}\uparrow}c^{\dagger}_{-\mathbf{k}\downarrow}}\left|0\right\rangle = \alpha_{\mathbf{k}}\sum_{n=1}^{\infty}\frac{\theta^{n-1}_{\mathbf{k}}}{(n-1)!}c^{\dagger}_{-\mathbf{k}\downarrow}\left|0\right\rangle \tag{54}
$$

Now, since:

$$
\left[\theta_{\mathbf{k}},c^{\dagger}_{-\mathbf{k}\downarrow}\right] = \alpha_{\mathbf{k}}\left[c^{\dagger}_{\mathbf{k}\uparrow}c^{\dagger}_{-\mathbf{k}\downarrow},c^{\dagger}_{-\mathbf{k}\downarrow}\right] = 0 \tag{55}
$$

we arrive at the result:

$$
c_{\mathbf{k}\uparrow}\left(\mathrm{e}^{\alpha_{\mathbf{k}}c^{\dagger}_{\mathbf{k}\uparrow}c^{\dagger}_{-\mathbf{k}\downarrow}}\left|0\right\rangle\right) = \alpha_{\mathbf{k}}c^{\dagger}_{-\mathbf{k}\downarrow}\sum_{n'=0}^{\infty}\frac{\theta^{n'}_{\mathbf{k}}}{(n')!}\left|0\right\rangle = \alpha_{\mathbf{k}}c^{\dagger}_{-\mathbf{k}\downarrow}\left(\mathrm{e}^{\alpha_{\mathbf{k}}c^{\dagger}_{\mathbf{k}\uparrow}c^{\dagger}_{-\mathbf{k}\downarrow}}\left|0\right\rangle\right) \tag{56}
$$

Substituting in Eq. (49) then gives:

$$
u_{\mathbf{k}}c_{\mathbf{k}\uparrow}\left|\Psi_{\mathrm{BCS}}\right\rangle = u_{\mathbf{k}}\alpha_{\mathbf{k}}c^{\dagger}_{-\mathbf{k}\downarrow}\left|\Psi_{\mathrm{BCS}}\right\rangle = v_{\mathbf{k}}c^{\dagger}_{-\mathbf{k}\downarrow}\left|\Psi_{\mathrm{BCS}}\right\rangle \tag{57}
$$

implying that the function $\alpha_{\mathbf{k}}$ is given by:

$$
\alpha_{\mathbf{k}} = \frac{v_{\mathbf{k}}}{u_{\mathbf{k}}} \tag{58}
$$

Hence, the BCS wave-function is:

$$
\begin{aligned}
\left|\Psi_{\mathrm{BCS}}\right\rangle &= \mathcal{N}\prod_{\mathbf{k}}\mathrm{e}^{\frac{v_{\mathbf{k}}}{u_{\mathbf{k}}}c^{\dagger}_{\mathbf{k}\uparrow}c^{\dagger}_{-\mathbf{k}\downarrow}}\left|0\right\rangle \\
\left|\Psi_{\mathrm{BCS}}\right\rangle &= \mathcal{N}\prod_{\mathbf{k}}\left(1+\frac{v_{\mathbf{k}}}{u_{\mathbf{k}}}c^{\dagger}_{\mathbf{k}\uparrow}c^{\dagger}_{-\mathbf{k}\downarrow}\right)\left|0\right\rangle
\end{aligned}
$$

where we used the fact that, due to Pauli's principle, $\left(c^\dagger_{\mathbf{k}\uparrow}c^\dagger_{-\mathbf{k}\downarrow}\right)^n = 0$ for $n > 1$. To normalize this wave-function, we notice that:

$$\begin{aligned}
\langle 0|\left(u^*_{\mathbf{k}} + v^*_{\mathbf{k}} c_{\mathbf{k}\uparrow}c_{-\mathbf{k}\downarrow}\right)\left(u_{\mathbf{k}} + v_{\mathbf{k}} c^\dagger_{\mathbf{k}\uparrow}c^\dagger_{-\mathbf{k}\downarrow}\right)|0\rangle &= \langle 0|\left(|u_{\mathbf{k}}|^2 + |v_{\mathbf{k}}|^2 c_{\mathbf{k}\uparrow}c^\dagger_{\mathbf{k}\uparrow}c_{-\mathbf{k}\downarrow}c^\dagger_{-\mathbf{k}\downarrow}\right)|0\rangle \\
&= \langle 0|\left(|u_{\mathbf{k}}|^2 + |v_{\mathbf{k}}|^2 \left(1 - c^\dagger_{\mathbf{k}\uparrow}c_{\mathbf{k}\uparrow}\right)\left(1 - c^\dagger_{\mathbf{k}\uparrow}c_{-\mathbf{k}\downarrow}\right)\right)|0\rangle \\
&= \langle 0|\left(|u_{\mathbf{k}}|^2 + |v_{\mathbf{k}}|^2\right)|0\rangle
\end{aligned}$$

Therefore, the normalized BCS wave-function is given by:

$$|\Psi_{\mathrm{BCS}}\rangle = \prod_{\mathbf{k}}\left(u_{\mathbf{k}} + v_{\mathbf{k}} c^\dagger_{\mathbf{k}\uparrow}c^\dagger_{-\mathbf{k}\downarrow}\right)|0\rangle \tag{59}$$

Recall that the phase of the Cooper pairs is determined solely by the coefficient $v_{\mathbf{k}}$, and this phase coincides with the phase of the gap function $\Delta_{\mathbf{k}}$.

3.2 The gap equation

We still need to determine the gap function $\Delta_{\mathbf{k}}$, given self-consistently by Eq. (33). Using the Bogoliubov transformation (35), we have:

$$\begin{aligned}
\Delta_{\mathbf{k}} &= -\frac{1}{N}\sum_{\mathbf{k}'} V_{\mathbf{k}\mathbf{k}'}\left\langle\left(u^*_{\mathbf{k}'}\gamma_{-\mathbf{k}'\downarrow} - v_{\mathbf{k}'}\gamma^\dagger_{\mathbf{k}'\uparrow}\right)\left(u^*_{\mathbf{k}'}\gamma_{\mathbf{k}'\uparrow} + \gamma^\dagger_{-\mathbf{k}'\downarrow}\right)\right\rangle \\
\Delta_{\mathbf{k}} &= -\frac{1}{N}\sum_{\mathbf{k}'} V_{\mathbf{k}\mathbf{k}'} u^*_{\mathbf{k}'} v_{\mathbf{k}'}\left(\left\langle\gamma_{-\mathbf{k}'\downarrow}\gamma^\dagger_{-\mathbf{k}'\downarrow}\right\rangle - \left\langle\gamma^\dagger_{\mathbf{k}'\uparrow}\gamma_{\mathbf{k}'\uparrow}\right\rangle\right)
\end{aligned}$$

The Bogoliubons follow the Fermi-Dirac distribution and have an energy dispersion $E_{\mathbf{k}}$. Thus:

$$\left\langle\gamma^\dagger_{\mathbf{k}'\uparrow}\gamma_{\mathbf{k}'\uparrow}\right\rangle = \left\langle\gamma^\dagger_{-\mathbf{k}'\downarrow}\gamma_{-\mathbf{k}'\downarrow}\right\rangle = \frac{1}{\mathrm{e}^{\beta E_{\mathbf{k}'}} + 1} \tag{60}$$

yielding:

$$\left\langle\gamma_{-\mathbf{k}'\downarrow}\gamma^\dagger_{-\mathbf{k}'\downarrow}\right\rangle - \left\langle\gamma^\dagger_{\mathbf{k}'\uparrow}\gamma_{\mathbf{k}'\uparrow}\right\rangle = 1 - \frac{2}{\mathrm{e}^{\beta E_{\mathbf{k}'}} + 1} = \frac{\mathrm{e}^{\beta E_{\mathbf{k}'}} - 1}{\mathrm{e}^{\beta E_{\mathbf{k}'}} + 1} = \tanh\left(\frac{E_{\mathbf{k}'}}{2k_B T}\right) \tag{61}$$

From Eqs. (40) and (41), we also have:

$$u^*_{\mathbf{k}'} v_{\mathbf{k}'} = |u_{\mathbf{k}}|^2 \frac{v_{\mathbf{k}'}}{u_{\mathbf{k}'}} = \frac{1}{2}\left(1 + \frac{\xi_{\mathbf{k}'}}{\sqrt{\xi^2_{\mathbf{k}'} + |\Delta_{\mathbf{k}'}|^2}}\right)\left(\frac{\sqrt{\xi^2_{\mathbf{k}'} + |\Delta_{\mathbf{k}'}|^2} - \xi_{\mathbf{k}'}}{\Delta^*_{\mathbf{k}'}}\right) = \frac{\Delta_{\mathbf{k}'}}{2\sqrt{\xi^2_{\mathbf{k}'} + |\Delta_{\mathbf{k}'}|^2}} \tag{62}$$

yielding the gap equation:

$$\Delta_{\mathbf{k}} = -\frac{1}{N}\sum_{\mathbf{k}'}\frac{V_{\mathbf{k}\mathbf{k}'}\Delta_{\mathbf{k}'}}{2E_{\mathbf{k}'}}\tanh\left(\frac{E_{\mathbf{k}'}}{2k_BT}\right) \tag{63}$$

We can now study for which values of the potential $V_{\mathbf{k}\mathbf{k}'}$ and of the temperature T we obtain a non-zero gap, and therefore the BCS solution discussed in the previous section.

To proceed, we need to discuss the form of the potential. Based on our results for the phonon-mediated electronic interaction, we consider a constant attractive potential $V_{\mathbf{k}\mathbf{k}'} = -V_0$ for a shell of thickness $\hbar\omega_D$ around the Fermi energy, $|\xi_{\mathbf{k}}|, |\xi_{\mathbf{k}'}| < \hbar\omega_D$ (recall that $\xi_{\mathbf{k}} = \varepsilon_{\mathbf{k}} - \mu$). Since the potential does not depend on $\mathbf{k}, \mathbf{k}'$, we look for a gap function that is also $\mathbf{k}$ independent and real, $\Delta_{\mathbf{k}} = \Delta$. This type of gap function is called an *s-wave* gap, since its angular dependence is that of the Y_{00} spherical harmonic.

Therefore, we obtain:

$$1 = \frac{V_0}{N}\sum_{k<k_D}\frac{1}{2E_{\mathbf{k}}}\tanh\left(\frac{E_{\mathbf{k}}}{2k_BT}\right) \tag{64}$$

Introducing the density of states *per spin* $\rho(\varepsilon)$ (notice that it has half the value of the density of states we considered so far), we obtain:

$$\begin{aligned} 1 &= V_0\int_{-\hbar\omega_D}^{\hbar\omega_D}\frac{\rho(\varepsilon)\,d\varepsilon}{2\sqrt{\varepsilon^2+\Delta^2}}\tanh\left(\frac{\sqrt{\varepsilon^2+\Delta^2}}{2k_BT}\right) \\ 1 &= V_0\rho_F\int_0^{\hbar\omega_D}\frac{d\varepsilon}{\sqrt{\varepsilon^2+\Delta^2}}\tanh\left(\frac{\sqrt{\varepsilon^2+\Delta^2}}{2k_BT}\right) \end{aligned} \tag{65}$$

In the last line, we used the fact that $\hbar\omega_D \ll \mu$ to approximate the density of states by its value at the Fermi level.

This self-consistent equation gives the gap function for an arbitrary temperature $\Delta(T)$. Let us study the limiting behaviors. At $T = 0$, since $\tanh(x \to \infty) \to 1$, we have:

$$1 = V_0\rho_F\int_0^{\hbar\omega_D}\frac{d\varepsilon}{\sqrt{\varepsilon^2+\Delta_0^2}} \tag{66}$$

where we denoted $\Delta_0 \equiv \Delta(T=0)$. Evaluation of the integral is straightforward and gives:

$$\frac{1}{V_0\rho_F} = \text{arcsinh}\left(\frac{\hbar\omega_D}{\Delta_0}\right) \tag{67}$$

In most cases, Δ_0 is of the order of a few meV, much smaller than $\hbar\omega_D$, which is of the order of a few hundreds of meV. Hence, we can expand the $\text{arcsinh}(x)$ for large x to obtain:

$$\frac{1}{V_0\rho_F} = \ln\left(\frac{2\hbar\omega_D}{\Delta_0}\right)$$

$$\Delta_0 \quad = \quad 2\hbar\omega_D\, \mathrm{e}^{-\frac{1}{V_0\rho_F}} \tag{68}$$

Thus, we recover a result similar to our simplified analysis of the Schrödinger equation: an arbitrarily small attractive interaction V_0 gives rise to a finite gap at zero temperature, showing that the Fermi liquid state is unstable towards the formation of the BCS superconducting state. We also see that superconductivity is a non-perturbative effect, given the dependence on $\mathrm{e}^{-\frac{1}{V_0\rho_F}}$.

What is the critical temperature T_c for which a non-zero gap first appears? To determine it, we go back to Eq. (65) and send $\Delta \to 0$, yielding:

$$\frac{1}{V_0\rho_F} = \int_0^{\hbar\omega_D} \frac{d\varepsilon}{\varepsilon} \tanh\left(\frac{\varepsilon}{2k_B T_c}\right) = \int_0^{\frac{\hbar\omega_D}{2k_B T_c}} dx\, \frac{\tanh x}{x} \tag{69}$$

To evaluate the integral, we perform it by parts and use the fact that $\hbar\omega_D \gg k_B T_c$:

$$\begin{aligned}
\int_0^{\frac{\hbar\omega_D}{2k_B T_c}} dx\, \frac{\tanh x}{x} \quad &\approx \quad (\tanh x\, \ln x)_0^{\frac{\hbar\omega_D}{2k_B T_c}} - \int_0^{\infty} dx\, \frac{\ln x}{\cosh^2 x} \\
&\approx \quad \ln\left(\frac{\hbar\omega_D}{2k_B T_c}\right) - \ln\left(\frac{\pi}{4\mathrm{e}^{\gamma_E}}\right) = \ln\left(\frac{2\mathrm{e}^{\gamma_E}\hbar\omega_D}{\pi k_B T_c}\right)
\end{aligned}$$

where $\gamma_E \approx 0.577$ is the Euler constant. The superconducting transition temperature is then given by:

$$T_c = \frac{2\mathrm{e}^{\gamma_E}}{\pi} \frac{\hbar\omega_D}{k_B}\, \mathrm{e}^{-\frac{1}{V_0\rho_F}} \tag{70}$$

which again depends on $\mathrm{e}^{-\frac{1}{V_0\rho_F}}$, being non-zero for any arbitrarily small V_0. Combining Eqs. (68) and (70) gives the universal ratio between the zero-temperature gap and the critical temperature:

$$\frac{\Delta_0}{k_B T_c} \approx 1.76 \tag{71}$$

One of the early successes of BCS theory was the verification that this relationship is approximately satisfied in most of the known superconductors at the time. The BCS theory also addresses the isotope effect we discussed earlier: T_c in Eq. (70) depends linearly on the Debye frequency ω_D, which in turn varies as the inverse square root of the ionic mass M, i.e. $T_c \propto \omega_D \propto M^{-1/2}$, in agreement with the experimental observations.

3.3 Thermodynamic properties: specific heat

A key feature of the BCS theory is the presence of an energy gap Δ in the spectrum. Such a gap is manifest in several thermodynamic quantities, such as the low-temperature specific heat and the density of states $\rho(\varepsilon)$. The latter can be measured experimentally via tunneling. In the superconducting state, we have, for positive energies $\varepsilon > 0$ (once again, we focus on the density of states per spin):

$$
\begin{aligned}
\rho(\varepsilon) &= \int \frac{d^3k}{(2\pi)^3}\,\delta\left(\varepsilon - \sqrt{\Delta^2 + \xi_{\mathbf{k}}^2}\right) \\
\rho(\varepsilon) &= \int d\xi\,\rho_0(\xi)\,\delta\left(\varepsilon - \sqrt{\Delta^2 + \xi^2}\right) \\
\rho(\varepsilon) &= \rho_F \int d\xi\,\delta\left(\varepsilon - \sqrt{\Delta^2 + \xi^2}\right)
\end{aligned} \tag{72}
$$

where $\rho_0(\xi)$ is the density of states of the normal phase, which has been approximated by its value at the Fermi level, since these are the energies we are interested in. Clearly, the argument of the delta function can only be zero if $\varepsilon > \Delta$, i.e. there are no states inside the gap - as expected. Using:

$$
\delta\left(\varepsilon - \sqrt{\Delta^2 + \xi^2}\right) = \sum_{\pm} \frac{\delta\left(\xi \pm \sqrt{\varepsilon^2 - \Delta^2}\right)}{\left|\frac{\xi}{\sqrt{\Delta^2+\xi^2}}\right|} \tag{73}
$$

we obtain:

$$
\rho(\varepsilon) = \frac{2\rho_F \varepsilon}{\sqrt{\varepsilon^2 - \Delta^2}}\,\theta(\varepsilon - \Delta) \tag{74}
$$

where $\theta(x)$ is the usual step function. The factor of 2 here (absent in Tinkham's book) is a consequence of the fact that as $\Delta \to 0$, $E \to |\xi|$, i.e. it contains two branches of particle-hole excitations, doubling the density of states.

Using this expression for the density of states inside the superconducting state, it is straightforward to show that the specific heat at low temperatures displays activated behavior, i.e. $C \sim \mathrm{e}^{-\Delta/k_B T}$. The superconducting transition also affects the specific heat at T_c. To investigate it, we could in principle calculate the total internal energy due to the quasi-particle excitations:

$$
E_{\text{int}} = E_0 + \sum_{\mathbf{k}\sigma} E_{\mathbf{k}} \left\langle \gamma_{\mathbf{k}\sigma}^{\dagger} \gamma_{\mathbf{k}\sigma} \right\rangle \tag{75}
$$

and evaluate the derivative $\partial E_{\text{int}}/\partial T$. The issue is that the ground-state energy E_0, given by Eq. (47), also depends on temperature. To avoid this issue, it is easier to compute the entropy of the free fermionic gas formed by the Bogoliubon excitations:

$$
S = -k_B \sum_{\mathbf{k}\sigma} \left[(1 - f_{\mathbf{k}}) \ln(1 - f_{\mathbf{k}}) + f_{\mathbf{k}} \ln f_{\mathbf{k}}\right] \tag{76}
$$

where $f_{\mathbf{k}} \equiv \left\langle \gamma_{\mathbf{k}\sigma}^{\dagger} \gamma_{\mathbf{k}\sigma} \right\rangle = 1/\left(\mathrm{e}^{\beta E_{\mathbf{k}}} + 1\right)$ is the Fermi-Dirac function. The specific heat density is given by:

$$C=\frac{T}{V}\frac{dS}{dT}=\frac{T}{V}\frac{d\beta}{dT}\frac{dS}{d\beta}=-\frac{\beta}{V}\frac{dS}{d\beta} \tag{77}$$

We then have:

$$\begin{aligned}
C &= \frac{k_B\beta}{V}\sum_{\mathbf{k}\sigma}\frac{df_{\mathbf{k}}}{d\beta}\left[-\ln\left(1-f_{\mathbf{k}}\right)-1+\ln f_{\mathbf{k}}+1\right] \\
C &= \frac{k_B\beta}{V}\sum_{\mathbf{k}\sigma}\frac{df_{\mathbf{k}}}{d\beta}\ln\left[\frac{1}{\mathrm{e}^{\beta E_{\mathbf{k}}}+1}\frac{\mathrm{e}^{\beta E_{\mathbf{k}}}+1}{\mathrm{e}^{\beta E_{\mathbf{k}}}}\right] \\
C &= -\frac{2k_B\beta^2}{V}\sum_{\mathbf{k}}\frac{df_{\mathbf{k}}}{d\beta}E_{\mathbf{k}} \qquad (78)
\end{aligned}$$

The total derivative gives:

$$\frac{df_{\mathbf{k}}}{d\beta}=\frac{\partial f_{\mathbf{k}}}{\partial\beta}+\frac{\partial f_{\mathbf{k}}}{\partial E_{\mathbf{k}}}\frac{\partial E_{\mathbf{k}}}{\partial\beta}=\frac{E_{\mathbf{k}}}{\beta}\frac{\partial f_{\mathbf{k}}}{\partial E_{\mathbf{k}}}+\frac{\partial f_{\mathbf{k}}}{\partial E_{\mathbf{k}}}\frac{1}{2E_{\mathbf{k}}}\frac{\partial\Delta^2}{\partial\beta} \tag{79}$$

where we used the fact that $E_{\mathbf{k}}=\sqrt{\xi_{\mathbf{k}}^2+\Delta^2}$. Therefore, we obtain:

$$C=\frac{2k_B\beta}{V}\sum_{\mathbf{k}}\left(-\frac{\partial f_{\mathbf{k}}}{\partial E_{\mathbf{k}}}\right)\left(E_{\mathbf{k}}^2+\frac{\beta}{2}\frac{\partial\Delta^2}{\partial\beta}\right) \tag{80}$$

Let us analyze this expression close to T_c. Above T_c, $\Delta^2=0$ and $E_{\mathbf{k}}\to|\xi_{\mathbf{k}}|$. Since $\partial f_{\mathbf{k}}/\partial\xi_{\mathbf{k}}$ is an even function of $\xi_{\mathbf{k}}$, we have $\frac{\partial f_{\mathbf{k}}}{\partial|\xi_{\mathbf{k}}|}=\frac{\partial f_{\mathbf{k}}}{\partial\xi_{\mathbf{k}}}$. Using the Sommerfeld expansion:

$$-\frac{\partial f_{\mathbf{k}}}{\partial\xi_{\mathbf{k}}}\approx\delta\left(\xi\right)+\frac{\pi^2}{6\beta^2}\delta''\left(\xi\right) \tag{81}$$

we obtain, in the normal state:

$$\begin{aligned}
C\left(T_c+0^+\right) &= \frac{\pi^2k_B}{3\beta}\int d\xi\,\xi^2\rho\left(\xi\right)\delta''\left(\xi\right) \\
C\left(T_c+0^+\right) &= \frac{\pi^2k_B}{3\beta}\frac{\partial^2}{\partial\xi^2}\left[\xi^2\rho\left(\xi\right)\right]_{\xi=0} \\
C\left(T_c+0^+\right) &= \left(\frac{2\pi^2k_B^2\rho_F}{3}\right)T_c\equiv\gamma T_c \qquad (82)
\end{aligned}$$

As expected, we recover the result from the free Fermi gas (recall that ρ_F here is the density of states per spin). Immediately below T_c, we can again make $E_{\mathbf{k}}\to|\xi_{\mathbf{k}}|$ and $\frac{\partial f_{\mathbf{k}}}{\partial E_{\mathbf{k}}}=\frac{\partial f_{\mathbf{k}}}{\partial\xi_{\mathbf{k}}}$, but now Δ^2 is non-zero. Therefore, we obtain:

$$
\begin{aligned}
C\left(T_c+0^-\right) &= C\left(T_c+0^+\right)+k_B\beta^2\left(\frac{\partial\Delta^2}{\partial\beta}\right)_{T_c}\int d\xi\left(-\frac{\partial f_{\mathbf{k}}}{\partial\xi}\right)\rho\left(\xi\right) \\
C\left(T_c+0^-\right) &= C\left(T_c+0^+\right)+\rho_F\left(-\frac{\partial\Delta^2}{\partial T}\right)_{T_c}
\end{aligned}
\tag{83}
$$

i.e. at T_c the specific heat is discontinuous, displaying a jump $\Delta C\equiv C\left(T_c+0^-\right)-C\left(T_c+0^+\right)$:

$$
\Delta C=\rho_F\left(-\frac{\partial\Delta^2}{\partial T}\right)_{T_c} \tag{84}
$$

Close to T_c, the gap function behaves as (Homework):

$$
\Delta^2\approx\frac{8\pi^2}{7\zeta\left(3\right)}k_B^2T_c\left(T_c-T\right) \tag{85}
$$

where $\zeta(x)$ is the zeta Riemann function. Then, we obtain the following universal ratio between the specific heat jump and its value in the normal state (given by Eq. (82)):

$$
\frac{\Delta C}{\gamma T_c}=\frac{12}{7\zeta\left(3\right)}\approx 1.43 \tag{86}
$$

The experimental observation of this universal ratio in several superconducting materials is another success of the BCS theory.

3.4 London equation and the Meissner effect

As we discussed, the fundamental property of a superconductor is perfect diamagnetism, i.e. the Meissner effect. Here we show that the BCS theory naturally addresses the Meissner effect, justifying microscopically the phenomenological London equation (14).

Let us consider the kinetic term of the Hamiltonian in the presence of a magnetic field. The momentum is given by $\mathbf{p}+\frac{e}{c}\mathbf{A}$, where $\mathbf{A}$ is the magnetic vector potential, $\mathbf{B}=\nabla\times\mathbf{A}$. In second-quantization language, introducing the field operator $\hat{\psi}_\sigma\left(\mathbf{r}\right)$, we have:

$$
H=\sum_\sigma\int d^3r\,\hat{\psi}_\sigma^\dagger\left(\mathbf{r}\right)\frac{1}{2m}\left(\mathbf{p}+\frac{e}{c}\mathbf{A}\right)^2\hat{\psi}_\sigma\left(\mathbf{r}\right) \tag{87}
$$

We work in the Coulomb gauge, where $\mathbf{p}\cdot\mathbf{A}\propto\nabla\cdot\mathbf{A}=0$. Then, to lowest order in perturbation theory in $\mathbf{A}$, we have $H=H_0+H_1$, where H_0 is the kinetic Hamiltonian in the absence of any external fields and H_1 is given by:

$$
H_1=\frac{e}{mc}\sum_\sigma\int d^3r\,\hat{\psi}_\sigma^\dagger\left(\mathbf{r}\right)\left(\mathbf{A}\cdot\mathbf{p}\right)\hat{\psi}_\sigma\left(\mathbf{r}\right) \tag{88}
$$

Now, the total current operator is given by:

$$\begin{aligned}\hat{\mathbf{J}} &= -\frac{e}{v}\sum_{\sigma}\int d^3r\,\hat{\psi}^{\dagger}_{\sigma}(\mathbf{r})\,\frac{1}{m}\left(\mathbf{p}+\frac{e}{c}\mathbf{A}\right)\hat{\psi}_{\sigma}(\mathbf{r})\\ \hat{\mathbf{J}} &= -\frac{e^2}{mc}\left(\frac{1}{v}\sum_{\sigma}\int d^3r\,\hat{\psi}^{\dagger}_{\sigma}(\mathbf{r})\,\hat{\psi}_{\sigma}(\mathbf{r})\right)\mathbf{A}-\frac{e}{mv}\sum_{\sigma}\int d^3r\,\hat{\psi}^{\dagger}_{\sigma}(\mathbf{r})\,\mathbf{p}\hat{\psi}_{\sigma}(\mathbf{r})\end{aligned} \tag{89}$$

Evaluating the mean value in the ground state (i.e. at zero temperature), we obtain $\mathbf{J}=\mathbf{J}_d+\mathbf{J}_p$ with the so-called diamagnetic current:

$$\mathbf{J}_d = -\frac{ne^2}{mc}\mathbf{A} \tag{90}$$

and the paramagnetic current $\mathbf{J}_p=\left\langle \hat{\mathbf{J}}_p\right\rangle$ with:

$$\hat{\mathbf{J}}_p = -\frac{e}{mv}\sum_{\sigma}\int d^3r\,\hat{\psi}^{\dagger}_{\sigma}(\mathbf{r})\,\mathbf{p}\hat{\psi}_{\sigma}(\mathbf{r}) \tag{91}$$

If $\mathbf{J}_p=0$, we would recover the London equation (14) with all electrons being part of the superconducting condensate, $n_s=n$. However, the ground state in the presence of a field is not the BCS ground state - which we denote here by $|0\rangle$ - because of the contribution (88) to the kinetic energy. Since this term is linear in $\mathbf{A}$, in principle $\mathbf{J}_p$ can also have a term linear in $\mathbf{A}$ that could cancel the diamagnetic contribution $\mathbf{J}_d$. This is exactly what happens in the normal state, where no Meissner effect is observed.

In the superconducting state, however, the situation is different. Using first-order perturbation theory, the ground state is changed to:

$$|0\rangle \rightarrow |0\rangle + \sum_{l\neq 0}|l\rangle\,\frac{\langle l\,|H_1|0\,\rangle}{E_0-E_l} \tag{92}$$

where $|l\rangle$ is an excited state. Then, since $\langle 0\,|\hat{\mathbf{J}}_p|0\,\rangle=0$, we have:

$$\mathbf{J}_p=\sum_{l\neq 0}\frac{\langle 0\,|\hat{\mathbf{J}}_p|l\,\rangle\langle l\,|H_1|0\,\rangle}{E_0-E_l}+\sum_{l\neq 0}\frac{\langle 0\,|H_1|l\,\rangle\langle l\,|\hat{\mathbf{J}}_p|0\,\rangle}{E_0-E_l} \tag{93}$$

Let us analyze the matrix element $\langle l\,|H_1|0\,\rangle$, which is linearly proportional to $\mathbf{A}$, see Eq. (88). Changing basis from the coordinate to the momentum representation, $\hat{\psi}_{\sigma}(\mathbf{r})=\frac{1}{\sqrt{v}}\sum_{\mathbf{k}}c_{\mathbf{k}\sigma}\mathrm{e}^{i\mathbf{k}\cdot\mathbf{r}}$, and considering the Fourier transformation $\mathbf{A}=\sum_{\mathbf{q}}\mathbf{A}_{\mathbf{q}}\mathrm{e}^{i\mathbf{q}\cdot\mathbf{r}}$, we have:

$$\begin{aligned}H_1 &= \frac{\hbar e}{mc}\sum_{\sigma}\sum_{\mathbf{k}\mathbf{k}'\mathbf{q}}\left(\frac{1}{v}\int d^3r\,\mathrm{e}^{i(\mathbf{k}-\mathbf{k}'+\mathbf{q})\cdot\mathbf{r}}\right)c^{\dagger}_{\mathbf{k}'\sigma}\left(\mathbf{A}_{\mathbf{q}}\cdot\mathbf{k}\right)c_{\mathbf{k}\sigma}\\ H_1 &= \frac{\hbar e}{mc}\sum_{\sigma}\sum_{\mathbf{k}\mathbf{q}}\left(\mathbf{k}\cdot\mathbf{A}_{\mathbf{q}}\right)c^{\dagger}_{\mathbf{k}+\mathbf{q}\sigma}c_{\mathbf{k}\sigma}\end{aligned} \tag{94}$$

To make contact with the BCS theory, we rewrite this term as follows:

$$
\begin{aligned}
H_1 &= \frac{\hbar e}{mc}\left[\sum_{\mathbf{kq}} \mathbf{k}\cdot\mathbf{A_q} c^{\dagger}_{\mathbf{k+q}\uparrow} c_{\mathbf{k}\uparrow} + \sum_{\mathbf{kq}} \mathbf{k}\cdot\mathbf{A_q} c^{\dagger}_{\mathbf{k+q}\downarrow} c_{\mathbf{k}\downarrow}\right] \\
H_1 &= \frac{\hbar e}{mc}\left[\sum_{\mathbf{kq}} \mathbf{k}\cdot\mathbf{A_q} c^{\dagger}_{\mathbf{k+q}\uparrow} c_{\mathbf{k}\uparrow} - \sum_{\mathbf{k'q}} \left(\mathbf{k'}+\mathbf{q}\right)\cdot\mathbf{A_q} c^{\dagger}_{-\mathbf{k'}\downarrow} c_{-\mathbf{k'}-\mathbf{q}\downarrow}\right] \\
H_1 &= \frac{\hbar e}{mc}\sum_{\mathbf{kq}} \mathbf{k}\cdot\mathbf{A_q}\left(c^{\dagger}_{\mathbf{k+q}\uparrow} c_{\mathbf{k}\uparrow} - c^{\dagger}_{-\mathbf{k}\downarrow} c_{-\mathbf{k}-\mathbf{q}\downarrow}\right) \qquad (95)
\end{aligned}
$$

where we used the fact that, in the Coulomb gauge, $\mathbf{q}\cdot\mathbf{A_q}=0$. Using the Bogoliubov transformation (35) and the fact that $\gamma_{\mathbf{k}\sigma}\left|0\right\rangle = 0$, we obtain:

$$
\begin{aligned}
\left\langle l\right| c^{\dagger}_{\mathbf{k+q}\uparrow} c_{\mathbf{k}\uparrow}\left|0\right\rangle &= \left\langle l\right|\left(u_{\mathbf{k+q}}\gamma^{\dagger}_{\mathbf{k+q}\uparrow} + v^{*}_{\mathbf{k+q}}\gamma_{-\mathbf{k}-\mathbf{q}\downarrow}\right)\left(u^{*}_{\mathbf{k}}\gamma_{\mathbf{k}\uparrow} + v_{\mathbf{k}}\gamma^{\dagger}_{-\mathbf{k}\downarrow}\right)\left|0\right\rangle \\
&= u_{\mathbf{k+q}} v_{\mathbf{k}} \left\langle l\right| \gamma^{\dagger}_{\mathbf{k+q}\uparrow}\gamma^{\dagger}_{-\mathbf{k}\downarrow}\left|0\right\rangle \qquad (96)
\end{aligned}
$$

as well as:

$$
\begin{aligned}
\left\langle l\right| c^{\dagger}_{-\mathbf{k}\downarrow} c_{-\mathbf{k}-\mathbf{q}\downarrow}\left|0\right\rangle &= \left\langle l\right|\left(u_{\mathbf{k}}\gamma^{\dagger}_{-\mathbf{k}\downarrow} - v^{*}_{\mathbf{k}}\gamma_{\mathbf{k}\uparrow}\right)\left(u^{*}_{\mathbf{k+q}}\gamma_{-\mathbf{k}-\mathbf{q}\downarrow} - v_{\mathbf{k+q}}\gamma^{\dagger}_{\mathbf{k+q}\uparrow}\right)\left|0\right\rangle \\
&= -u_{\mathbf{k}} v_{\mathbf{k+q}} \left\langle l\right| \gamma^{\dagger}_{-\mathbf{k}\downarrow}\gamma^{\dagger}_{\mathbf{k+q}\uparrow}\left|0\right\rangle \qquad (97)
\end{aligned}
$$

Using the anticommutation relations of the Bogoliubon operators, we obtain:

$$
\left\langle l\right| H_1\left|0\right\rangle = \frac{\hbar e}{mc}\sum_{\mathbf{kq}} \mathbf{k}\cdot\mathbf{A_q}\left(u_{\mathbf{k+q}}v_{\mathbf{k}} - u_{\mathbf{k}}v_{\mathbf{k+q}}\right)\left\langle l\right| \gamma^{\dagger}_{\mathbf{k+q}\uparrow}\gamma^{\dagger}_{-\mathbf{k}\downarrow}\left|0\right\rangle \qquad (98)
$$

To obtain the conductivity, we must take the limit $\mathbf{q}\rightarrow 0$ for a uniform field. From the previous equation, it is clear that $\left\langle l\right| H_1\left|0\right\rangle \rightarrow 0$ in this limit. Furthermore, since the energy spectrum is gapped, $|E_0 - E_l| > 2\Delta$ in Eq. (93) - this is the rigidity of the superconducting state. Then, it follows that $\mathbf{J}_p = 0$, and we obtain:

$$
\mathbf{J} = \mathbf{J}_p + \mathbf{J}_d = -\frac{ne^2}{mc}\mathbf{A} \qquad (99)
$$

i.e. we recover London equation (90) and, consequently, the Meissner effect. By comparing to Eq. (14), we notice that in the ground state (zero temperature) all the electrons participate in the superconducting condensate, i.e. $n_s = n$, and not only the electrons near the Fermi level. At finite temperatures, the number of superconducting electrons decreases and eventually vanishes at T_c. Experimentally, the superfluid density n_s can be indirectly measured via the penetration depth, as shown by Eq. (10).

4 Ginzburg-Landau model

We finish these notes by discussing briefly another approach to understand the rigidity of the superconducting state and its relationship to persistent currents. It is based on the Ginzburg-Landau model, originally conceived as a phenomenological model to describe superconductivity and later shown by Gor'kov to be derived from the BCS theory.

The main quantity in the Ginzburg-Landau model is the complex order parameter $\Psi(\mathbf{r})$, which can be interpreted as the superconducting wave-function. The idea is that, below T_c, the average value of the superconducting wave-function is non-zero, i.e. $\langle\Psi\rangle \neq 0$, while above T_c it remains zero. Let $F[\Psi(\mathbf{r})]$ be the functional that gives the difference between the free energy of the superconducting state and the normal state, $\mathcal{F} = \int d\mathbf{r}\, F[\Psi(\mathbf{r})]$. It follows that the equilibrium value of F must be positive above T_c (so that the free energy of the normal state is smaller than the free energy of the superconducting state) and negative below T_c (so that the free energy of the normal state is larger than the free energy of the superconducting state). Therefore, it must vanish at T_c. Near T_c, one can then expand the free energy $F[\Psi(\mathbf{r})]$ in powers of Ψ. Symmetry and analyticity requirements impose that the only possible terms in the expansion are those involving even powers of $|\Psi|$. Thus, in the case where $\Psi(\mathbf{r})$ does not depend on the position $\mathbf{r}$, we obtain:

$$F(\Psi, \Psi^*) = \alpha|\Psi|^2 + \frac{\beta}{2}|\Psi|^4 \tag{100}$$

This is the so-called Landau free energy expansion. Recall that $|\Psi|^2 = \Psi\Psi^*$ since Ψ is a complex function. The quartic coefficient β must be positive, otherwise the free energy is not bounded. To understand the meaning of the quadratic coefficient α, we minimize the free energy function by taking its derivative with respect to Ψ^* (the same result is obtained if one takes the derivative with respect to Ψ instead), since we know that in equilibrium the free energy takes its minimum value:

$$\begin{aligned}\frac{\partial F}{\partial \Psi^*} &= \alpha\Psi + \beta\Psi|\Psi|^2 = 0 \\ &\Psi\left(a + \beta|\Psi|^2\right) = 0\end{aligned} \tag{101}$$

Therefore, there are two possible solutions:

$$|\Psi| = 0 \quad \text{or} \quad |\Psi| = \sqrt{-\frac{\alpha}{\beta}} \tag{102}$$

corresponding to the normal state ($\Psi = 0$) and to the superconducting state ($\Psi \neq 0$). The free energy of each solution is given by:

$$F = 0 \quad \text{or} \quad F = -\frac{\alpha^2}{2\beta} \tag{103}$$

respectively. Thus, if the superconducting solution exists (i.e. the one with $\Psi \neq 0$), it gives the global minimum of the free energy. Clearly, because $\beta > 0$, this solution can only be physical if $\alpha < 0$. Consequently, the normal state is the global minimum and $\Psi = 0$ for $\alpha > 0$, whereas the superconducting state is the global minimum and $\Psi \neq 0$ for $\alpha < 0$. Thids analysis allows us to conclude that α must vanish and change sign across T_c:

$$\alpha = a\left(T - T_c\right) \tag{104}$$

Substituting back into the solution, we find:

$$|\Psi| \propto \sqrt{T_c - T} \tag{105}$$

Thus, the superconducting wave function vanishes as the system approaches T_c from below with a square-root dependence.

We now consider the most general case, in which the function $\Psi\left(\mathbf{r}\right)$ is no longer constant. Symmetry and analiticity requirements impose that only second-order derivatives can appear in the free energy expansion, i.e. terms of the form $\left|\boldsymbol{\nabla}\Psi\right|^2$. The coefficient of this term must be positive, since the system pays energy if the wave-function is not uniform – this is linked to the concept of rigidity. Because the Cooper-pair is charged, it must couple to the eletromagnetic field via the usual minimal coupling $\frac{\hbar}{i}\boldsymbol{\nabla} + \frac{2e}{c}\mathbf{A}$, where $\mathbf{A}$ is the magnetic vector potential. The factor $2e$ is because the Cooper pair has charge $-2e$. Therefore, the free energy functional becomes:

$$F\left[\Psi\left(\mathbf{r}\right), \Psi^*\left(\mathbf{r}\right), \mathbf{A}\right] = \alpha\left|\Psi\left(\mathbf{r}\right)\right|^2 + \frac{\beta}{2}\left|\Psi\left(\mathbf{r}\right)\right|^4 + \frac{1}{4m}\left|\left(\frac{\hbar}{i}\boldsymbol{\nabla} + \frac{2e}{c}\mathbf{A}\right)\Psi\right|^2 + \frac{B^2}{8\pi} \tag{106}$$

The last term is just the energy of the electromagnetic field. The fact that we have $4m$ instead of the usual $2m$ is because the Cooper pair has two electrons. This is the so-called Ginzburg-Landau free energy expansion. It was first proposed by Ginzburg and Landau on phenomenological grounds before the BCS theory. It was later shown by Gor'kov that this free energy can be directly derived from the microscopic BCS theory.

Let us derive the equilibrium equations. Note that we need to minimize the free energy with respect to both Ψ and $\mathbf{A}$. To do that, it is convenient to write the gradient term explicitly:

$$\begin{aligned} \frac{1}{4m}\left(\frac{\hbar}{i}\boldsymbol{\nabla}\Psi + \frac{2e}{c}\mathbf{A}\Psi\right)\cdot\left(-\frac{\hbar}{i}\boldsymbol{\nabla}\Psi^* + \frac{2e}{c}\mathbf{A}\Psi^*\right) &= \\ \frac{\hbar^2}{4m}\left(\boldsymbol{\nabla}\Psi\right)\cdot\left(\boldsymbol{\nabla}\Psi^*\right) - \frac{i\hbar e}{2mc}\left(\Psi^*\boldsymbol{\nabla}\Psi - \Psi\boldsymbol{\nabla}\Psi^*\right)\cdot\mathbf{A} + \frac{e^2A^2}{mc^2}\left|\Psi\right|^2 &= \\ \sum_i \frac{\hbar^2}{4m}\partial_i\Psi\partial_i\Psi^* - \sum_i \frac{i\hbar e}{2mc}\left(\Psi^*\partial_i\Psi - \Psi\partial_i\Psi^*\right)A_i + \sum_i \frac{e^2A_i^2}{mc^2}\left|\Psi\right|^2 & \end{aligned} \tag{107}$$

where, in the last line, we expressed the equation in terms of the vector components of the nabla operator and of the vector potential. Minimizing the functional with respect to Ψ^* gives the Euler-Lagrange equation:

$$\begin{aligned}
\frac{\partial F}{\partial \Psi^*} - \sum_j \partial_j \frac{\partial F}{\partial\left(\partial_j \Psi^*\right)} &= 0 \\
\alpha\Psi + \beta\Psi\left|\Psi\right|^2 + \frac{e^2 A^2}{mc^2}\Psi - \sum_i \frac{i\hbar e}{2mc}\left(\partial_i \Psi\right) A_i - \sum_j \partial_j \left[\frac{\hbar^2}{4m}\partial_j \Psi + \frac{i\hbar e}{2mc}\Psi A_j\right] &= 0 \\
\alpha\Psi + \beta\Psi\left|\Psi\right|^2 + \frac{1}{4m}\left(\frac{\hbar}{i}\boldsymbol{\nabla} + \frac{2e}{c}\mathbf{A}\right)^2 \Psi &= 0
\end{aligned} \tag{108}$$

Notice its similarity with the Schroedinger equation. To minimize the free energy functional with respect to $\mathbf{A}$, it is convenient to rewrite the magnetic contribution to the free energy as:

$$\frac{B^2}{8\pi} = \frac{\left|\boldsymbol{\nabla}\times\mathbf{A}\right|^2}{8\pi} = \frac{1}{8\pi}\sum_{i,j,k,l,m} \varepsilon_{ijk}\varepsilon_{ilm}\partial_j A_k \partial_l A_m \tag{109}$$

where we used the Levi-Civita symbol ε_{ijk}. Therefore, the corresponding Euler-Lagrange equation becomes:

$$\begin{aligned}
\frac{\partial F}{\partial A_p} - \sum_q \partial_q \frac{\partial F}{\partial\left(\partial_q A_p\right)} &= 0 \\
-\frac{i\hbar e}{2mc}\left(\Psi^* \partial_p \Psi - \Psi \partial_p \Psi^*\right) + \frac{2e^2 A_p}{mc^2}\left|\Psi\right|^2 &= \frac{1}{4\pi}\sum_{q,i,l,m} \varepsilon_{iqp}\varepsilon_{ilm}\partial_q \partial_l A_m \\
-\frac{i\hbar e}{2mc}\left(\Psi^* \boldsymbol{\nabla}\Psi - \Psi\boldsymbol{\nabla}\Psi^*\right) + \frac{2e^2\mathbf{A}}{mc^2}\left|\Psi\right|^2 &= -\frac{1}{4\pi}\boldsymbol{\nabla}\times\left(\boldsymbol{\nabla}\times\mathbf{A}\right)
\end{aligned} \tag{110}$$

Using the fourth Maxwell equation, $\boldsymbol{\nabla}\times\mathbf{B} = 4\pi\mathbf{J}/c$, we obtain an equation for the superfluid current:

$$\mathbf{J} = -\frac{e\hbar}{2mi}\left(\Psi^* \boldsymbol{\nabla}\Psi - \Psi\boldsymbol{\nabla}\Psi^*\right) - \frac{2e^2\mathbf{A}}{mc}\left|\Psi\right|^2 \tag{111}$$

Additional analysis that will not be discussed here reveals that the amplitude of the superconducting wave-function $\left|\Psi\left(\mathbf{r}\right)\right|^2$ must be equal to half the superfluid density $n_s/2$. Therefore, in general we can write:

$$\Psi\left(\mathbf{r}\right) = \frac{1}{\sqrt{2}}\sqrt{n_s\left(\mathbf{r}\right)}\,\mathrm{e}^{i\theta(\mathbf{r})} \tag{112}$$

with $\theta\left(\mathbf{r}\right)$ denoting the phase of the superconducting condensate. The factor of 1/2 accounts for the fact that the charge associated with the wave-function is the Cooper pair's charge $-2e$. In the case where the superfluid density is homogeneous, only the phase of the superconducting wave-function depends on the position, yielding the superfluid current:

$$\mathbf{J} = -\left(\frac{e\hbar n_s}{2m}\right)\boldsymbol{\nabla}\theta - \left(\frac{n_s e^2}{mc}\right)\mathbf{A} \tag{113}$$

The second term shows that for a uniform superconducting phase $\boldsymbol{\nabla}\theta = 0$, we recover the London equation. The first term shows that when $\mathbf{A} = 0$ a non-uniform phase gives rise to a current flow in the superconducting state, and vice-versa. In most quantum mechanical systems, macroscopic changes in the global phase do not change the properties of the system. Here, however, the entire superconducting state has the same phase, and macroscopic changes in θ lead to changes in macroscopic properties of the system due to this global phase coherence. In the BCS language, the phase coherence comes from the factor $v_{\mathbf{k}}$ in the wave-function (59), which endows every Cooper pair with the same phase. If we enforce a slow variation in the phase on the macroscopic scale, resulting in a small non-zero $\boldsymbol{\nabla}\theta$, the superconducting condensate responds by developing a current $\mathbf{J}$. Now, because this current is a result of minimizing the Ginzburg-Landau free energy, it must be an equilibrium property, and cannot dissipate energy. This allows the system to behave as a perfect conductor.

The expression (113) has other important consequences. First, notice that if we put two superconductors next to each other, separated by a thin insulating barrier, the difference in the phase of the two superconducting wave-functions will give rise to a current flowing in the junction. This is known as the *Josephson effect*.

Second, consider the situation where a hole is made inside a superconductor. Inside this hole, the system is in the normal state. If we consider a closed path surrounding this hole, deep inside the superconducting state, the current across this loop is zero. Then, integrating Eq. (113) across this loop yields:

$$\oint \mathbf{A}\cdot d\mathbf{l} = -\frac{\hbar c}{2e}\oint \boldsymbol{\nabla}\theta\cdot d\mathbf{l} \tag{114}$$

Using Stokes theorem:

$$\oint \mathbf{A}\cdot d\mathbf{l} = \int_S (\boldsymbol{\nabla}\times\mathbf{A})\cdot d\mathbf{S} = \int_S \mathbf{B}\cdot d\mathbf{S} = \Phi \tag{115}$$

where Φ is the magnetic flux. Since the phase θ can only change by multiples of 2π from initial point to the final point of the loop, we obtain:

$$\Phi = \frac{hc}{2\,|e|}\,n \tag{116}$$

where n is an arbitrary integer. Thus, the magnetic flux of a normal region inside a superconductor has to be a multiple of the flux quantum $\Phi_0 = \frac{hc}{2|e|}$.

Let us comment more on the phase of the superconductor. Note that the free energy functional (106) is invariant under a gauge transformation, i.e. gauge changes in the vector potential $\mathbf{A} \to \mathbf{A} + \boldsymbol{\nabla}\chi$ are cancelled by a local change in the phase $\theta \to \theta - \frac{2e}{\hbar c}\chi$. In the superconducting state, however, because the phase is fixed, the system actually breaks gauge invariance. Therefore, the symmetry broken by the superconducting state is the $U(1)$ gauge symmetry. One would expect that breaking this continuous symmetry would give rise to a Goldstone mode. However, this is not true because this is a local (i.e. gauge) - not a global - symmetry that couples to the electromagnetic vector potential. This is the main difference from a neutral superfluid, which does have a Goldstone mode associated with the phase.

In fact, it can be shown that the breaking of gauge invariance gives rise to an effective mass for the electromagnetic field. This is the celebrated *Anderson-Higgs mechanism.* Consider, for instance, the free energy associated with changes in the phase of a superconductor (i.e. the superfluid density is assumed to be constant). From Eq. (106), the free energy becomes:

$$\mathcal{F} = \frac{n_s}{4m} \int d^3r \left(\hbar \boldsymbol{\nabla} \theta + \frac{2e}{c} \mathbf{A} \right)^2 \tag{117}$$

One can add to this free energy the electromagnetic energy, which is proportional to $q^2 A_{\perp}^2$, where q is the wave-vector of the field and $\mathbf{A}_{\perp}$ is the transverse component of the field. Without going into details, we mention that if the phase fluctuations are integrated out from the free energy, we obtain an effective free energy for the electromagnetic field of the form:

$$\mathcal{F}_{\text{eff}} \propto \sum_{\mathbf{q}} \left(\lambda^{-2} + q^2 \right) \mathbf{A}_{\perp} \left(\mathbf{q} \right) \cdot \mathbf{A}_{\perp} \left(-\mathbf{q} \right) \tag{118}$$

The term $\lambda^{-2} \propto n_s$ is the inverse squared penetration depth and acts as an effective mass for the electromagnetic field. This is not surprising: the Meissner effect implies that the magnetic field is "massive" inside a superconductor, since it decays as it propagates from the interface to the interior of the superconductor. The agent responsible for giving mass to the superconductor - i.e. the "Higgs boson" - is the superconducting condensate - more specifically, its rigidity n_s. Thus, the rigidity is the key property responsible for the Meissner effect, and not the gap function Δ - in fact, one can find gapless superconductors that still display Meissner effect and persistent currents.