# NOTES ON
# SUPERCONDUCTIVITY

ALESSANDRO GORI
github/nepero27178

Last compiled: September 1, 2024

# Contents

# 1 SUPERCONDUCTIVITY AND THERMODYNAMICS



Superconductivity is a special state of matter, in which a material exhibits zero resistance and magnetic flux is expelled. It turns out some critical parameters exist, $\mathbf{H}_c$ (critical field) and $T_c$ (critical temperature) separating the normal phase to the superconductive one. Thus, transition to the state of superconductivity is described in the the context of critical phase transition.

Superconductivity involves a wide number of theoretical interesting aspects, from continuous symmetry breaking to condensation of electron pairs, to surprising technological implementations. All in all, it represents an exciting field of physics where thermodynamics, statistical mechanics, quantum mechanics, fluidodynamics and topology collide.

> **M. Tinkham**: ***Introduction to Superconductivity*** ([5] @ 1.1)
>
> What Kamerlingh Onnes observed was that the electrical resistance of various metals such as mercury, lead, and tin disappeared completely in a small temperature range at a critical temperature $T_c$, which is characteristic of the material. The complete disappearance of resistance is most sensitively demonstrated by experiments with persistent currents in superconducting rings [...]. Once set up, such currents have been observed to flow without measurable decrease for a year, and a lower bound of some $10^5$ years for theirs characteristic decay time has been established by using nuclear magnetic resonance to detect any slight decrease in the field produced by the circulating current. In fact [...] under many circumstances we expect absolutely no change in field or current to occur in times less than $10^{10^{10}}$ years!

This first chapter is devoted to phenomenology. We start by the experimental observations, trying to understand what actually a superconductor *is*. Then we move to the realm of thermodynamics, and study the general phase diagram of ideal superconductors. What presented here is pretty much general and many corrections must be understood in order to describe real superconductors.

## 1.1 MACROSCOPIC THEORY OF SUPERCONDUCTIVITY

In 1935, the brothers Fritz and Heinz London published their macroscopic theory of superconductivity [8]. The key idea is the following: electrons can be in the normal state (subscript $n$) or in the superconductive one (subscript $s$). Then the density is given by

$$n = n_n + n_s$$

A critical temperature $T_c$ exists, such that $n_s \neq 0$ if $T \leq T_c$. We aim to find some equations to describe the motion of superconducting currents,

$$\mathbf{J}_s = e n_s \mathbf{v}_s$$

where we consider the electron charge $e = -|e|$ in order to keep the results valid for general charges. Our equations need to exhibit **perfect conductivity** and **perfect diamagnetism**. We shall start by the perfect conductivity.

### 1.1.1 The land where electrons do not collide

Take the Drude theory for the electron motion in crystals,

$$m\frac{d\mathbf{v}}{dt} = e\mathbf{E} - \frac{m}{\tau}\mathbf{v}$$

with $m$ the electron mass and $\tau$ the mean scattering time. The superconducting electrons move dissipation-less, meaning $\tau_s \to +\infty$. Then

$$men_s\frac{d\mathbf{v}}{dt} = e^2 n_s \mathbf{E} \quad \Longrightarrow \quad \mathbf{E} = \Lambda\frac{\partial \mathbf{J}_s}{\partial t} \quad \text{with} \quad \Lambda = \frac{m}{e^2 n_s}$$

Obviously this equation describes a free system accelerated by a uniform field. This is the **first London equation**, describing perfect conductivity.

We momentarily drop the subscript $s$, and take for granted the Drude theory of transport (for reference, check [2]); also, every material parameter can be replaced by its effective version in crystals ($m \to m^\star, \varepsilon_0 \to \varepsilon^\star$ and so on). By classical electrodynamics, metals are those material capable of perfectly cancel the electric field inside. By the means of Drude theory, dynamic conductivity for large $\tau$ is given by

$$\sigma(\omega) = \frac{\sigma_0}{1 - i\omega\tau} \simeq i\frac{\sigma_0}{\omega\tau} \quad \text{with} \quad \sigma_0 = \frac{ne^2\tau}{m}$$

thus perfectly imaginary, $\sigma \simeq iy$ with $y \geq 0$. Then, being the dielectric function related to the conductivity by

$$\varepsilon(\omega) = 1 + \frac{4\pi i}{\omega}\sigma(\omega) \simeq 1 - \frac{\omega_p^2}{\omega^2} \quad \text{with} \quad \omega_p^2 = \frac{ne^2}{\varepsilon_0 m}$$

we have $\varepsilon(\omega) < 0$ for $\omega < \omega_p$. Finally, being the dielectric function related to the complex refraction index by

$$N^2(\omega) = \varepsilon(\omega) \quad \text{with} \quad N(\omega) = n(\omega) + ik(\omega)$$

for $\omega \leq \omega_p$ we expect $N \in i\mathbb{R}$, which means the electromagnetic wave is evanescent in the crystal. In first approximation, in dynamic regime, metals are those materials capable of reflecting low-energy waves.

We now justify this conclusion. Take Maxwell's equations,

$$\boldsymbol{\nabla} \times \mathbf{B} - \frac{1}{c^2}\frac{\partial \mathbf{E}}{\partial t} = \mu_0 \mathbf{J}$$
$$\boldsymbol{\nabla} \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$$

By taking the time derivative of the first we get

$$\boldsymbol{\nabla} \times \left(-\frac{\partial \mathbf{B}}{\partial t}\right) + \frac{1}{c^2}\frac{\partial^2 \mathbf{E}}{\partial t^2} = -\mu_0 \frac{\partial \mathbf{J}}{\partial t}$$

and inserting both the second Maxwell equation and the first London equation, we get

$$\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \mathbf{E}) + \frac{1}{c^2}\frac{\partial^2 \mathbf{E}}{\partial t^2} = -\frac{\mu_0}{\Lambda}\mathbf{E}$$

We define $\lambda^2 \equiv \Lambda/\mu_0$, and use the identity

$$\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \mathbf{A}) = \boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \mathbf{A}) - \nabla^2 \mathbf{A}$$

then

$$-\boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \mathbf{E}) + \nabla^2 \mathbf{E} - \frac{1}{c^2}\frac{\partial^2 \mathbf{E}}{\partial t^2} = \frac{\mathbf{E}}{\lambda^2} \tag{1.1}$$

Now, consider a purely transversal wave, $\mathbf{E}(\mathbf{x}, t) = E_z(x, t)\bar{\mathbf{z}}$. Then, Fourier-transforming in frequency domain, we get

$$\frac{\partial^2}{\partial x^2} E_z(x, \omega) - \left(\frac{1}{\lambda^2} - \frac{\omega^2}{c^2}\right) E_z(x, \omega) = 0$$

the solution to the above equation is a wave

$$E_z(x, \omega) = e^{-x/\ell(\omega)} E_z(0, \omega) \quad \text{with} \quad \frac{1}{\ell^2(\omega)} \equiv \frac{1}{\lambda^2} - \frac{\omega^2}{c^2}$$

If

$$\frac{1}{\lambda} - \frac{\omega}{c} > 0 \quad \Longrightarrow \quad \omega < \frac{c}{\lambda} = \sqrt{\frac{\mu_0}{\Lambda}} c = \sqrt{\frac{ne^2}{m\varepsilon_0}} = \omega_p$$

then $\ell(\omega) \in \mathbb{R}$, meaning the wave is evanescent and suppressed on a length scale $\ell(\omega)$, as expected. This makes sense because the ability of exciting plasmonic modes is independent of the presence of dissipation modes. For $\omega \leq \omega_p$, we shall approximate $\ell(\omega) \simeq \lambda$: thus $\lambda$ is the characteristic length scale over which a static field penetrates the superconductor (**penetration length**):

$$\lambda^2 = \frac{m}{\mu_0 e^2 n_s}$$

Evidently, for vanishing superconducting densities $n_s \to 0$, $\lambda$ is divergent.

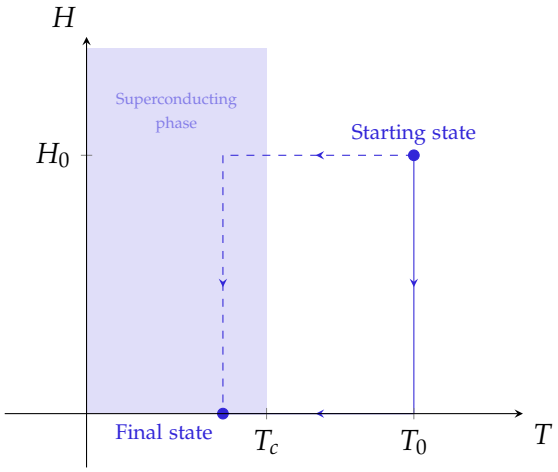



**Figure 1.1:** The sketch of the two processes described in text. The starting state is at temperature $T_0$ and the applied magnetic field is $H_0$. The final state is at temperature below the critical temperature and at zero applied field. In this simple context we are neglecting the fact that for high fields the system escapes the superconducting region.

What about magnetic fields? The same procedure yields

$$\left[\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \frac{1}{\lambda^2}\right]\frac{\partial \mathbf{B}}{\partial t} = 0 \quad \Rightarrow \quad \left[\nabla^2 - \left(\frac{1}{\lambda^2} - \frac{\omega}{c^2}\right)\right](-i\omega)\mathbf{B}(\omega) = 0$$

The same limit, $\omega \to 0$, now poses a difference: any magnetic flux density **B** (we shall call it that way not to confuse it with the external magnetic field **H**) now solves this equation. Then, **a perfect conductor perfectly cancels any static electric field and traps any static magnetic flux density**. Also, perfect conductors do not exist.

The reason is of thermodynamic stability. Consider the a system at temperature $T_0 > T_c$, under the influence of an external field $H_0$, as sketched in Fig. 1.1. Suppose we work at suitably low fields to neglect the $H$ dependence of the superconducting region (as it will turn out, such a region will exhibit a critical boundary $H_c(T)$). Suppose, also, to change the parameters adiabatically: at any given instant $H$ (and thus $B$) is static. If we follow the solid path, entering the superconducting region at $H = 0$: then we expect the final state to show $B = 0$. If, instead, we follow the dashed path, we enter the superconducting region at $H \neq 0$; then the magnetic flux density $B \neq 0$ is trapped in the following. Then the final state is ill-defined, being its thermodynamic quantities ambivalent. It follows that a perfect conductor is **not** a stable state, thus it doesn't exist.

### 1.1.2 The land where magnetic fields are not welcome

Evidently, our first attempt to describe superconductivity needs some refinement, and superconductors are something more than very good conductors.

> **M. Tinkham**: ***Introduction to Superconductivity*** ([5] @ 1.1)
>
> Thus, *perfect conductivity* is the first traditional hallmark of superconduc-

> tivity. It is also the prerequisite for the most potential applications, such as high current transmission lines or high-field magnets.
>
> The next hallmark to be discovered was *perfect diamagnetism*, found in 1933 by Meissner and Ochsenfeld. They found that not only a magnetic field is *excluded* from entering a superconductor [...], as might appear to be explained by perfect conductivity, but also that a field in an originally normal sample is *expelled* as it is cooled through $T_c$. This certainly could *not* be explained by perfect conductivity, which would tend to trap the flux *in*.

We take the curl of the first London equation using Faraday's law

$$\boldsymbol{\nabla} \times \mathbf{E} = \boldsymbol{\nabla} \times \Lambda \frac{\partial \mathbf{J}_s}{\partial t} \quad \Longrightarrow \quad \boldsymbol{\nabla} \times \Lambda \frac{\partial \mathbf{J}_s}{\partial t} + \frac{\partial \mathbf{B}}{\partial t} = 0$$

Then we postulate this equation holds without the time derivative, from a purely phenomenological perspective. Thus, we got the **London equations**

$$\mathbf{E} = \Lambda \dot{\mathbf{J}}_s \tag{1.2}$$
$$\mathbf{B} = \boldsymbol{\nabla} \times (-\Lambda \mathbf{J}_s) \tag{1.3}$$

Notice that the second one, at this stage, is not well-justified. Evidently, it is equivalent to the equation

$$\mathbf{A}_s(\mathbf{x}, t) = -\Lambda \mathbf{J}_s(\mathbf{x}, t) + \boldsymbol{\nabla} F(\mathbf{x}, t)$$

with $\mathbf{A}_s$ the electromagnetic vector potential and $F$ some scalar function. Now we can impose physical constraints, to fix the gauge. First, the supercurrent must be zero in the bulk and only flow on the sample boundary $\partial\Omega$. Also, we impose such a boundary supercurrent to flow parallel to the surface $\partial\Omega$ (a perpendicular component would imply charge cumulating on the superconductor surface). Finally, we impose no superconductive charge to cumulate inside the bulk. These conditions read

$$\boldsymbol{\nabla} \cdot \mathbf{J}_s = 0 \qquad \mathbf{J}_s(\mathbf{x}, t)\big|_{\mathbf{x} \notin \partial\Omega} = 0 \qquad \mathbf{J}_s(\mathbf{x}, t)\big|_{\mathbf{x} \in \partial\Omega} \perp \partial\Omega \bar{\mathbf{n}} \tag{1.4}$$

with $\bar{\mathbf{n}}$ the surface element versor. Those conditions ensure **perfect diamagnetism**, thanks to the second London equation (1.3). In fact, from Maxwell's equation for static fields

$$\boldsymbol{\nabla} \times \mathbf{B} = \mu_0 \mathbf{J}_s \;\Rightarrow\; -\nabla^2 \mathbf{B} = -\frac{\mu_0}{\Lambda} \boldsymbol{\nabla} \times (-\Lambda \mathbf{J}_s) \;\Rightarrow\; \left[\nabla^2 - \frac{1}{\lambda^2}\right] \mathbf{B} = 0$$

which is solved by an exponentially suppressed field,

$$\mathbf{B}(\mathbf{x}) = e^{-|\mathbf{x}|/\lambda} \mathbf{B}(\mathbf{0})$$

We can impose the same conditions as Eq. (1.4) on the vector potential,

$$\boldsymbol{\nabla} \cdot \mathbf{A}_s = 0 \qquad \mathbf{A}_s(\mathbf{x}, t)\big|_{\mathbf{x} \notin \partial\Omega} = 0 \qquad \mathbf{A}_s(\mathbf{x}, t)\big|_{\mathbf{x} \in \partial\Omega} \perp \partial\Omega \bar{\mathbf{n}} \tag{1.5}$$

thus defining the so-called **London gauge**. In this gauge we completely identify the vector potential and the superconducting current up to a factor $-\Lambda$. Notice that, in this gauge

$$-\frac{\partial \mathbf{A}_s}{\partial t} = \Lambda \frac{\partial \mathbf{J}_s}{\partial t} = \mathbf{E}$$

which is the correct potential equation for the electric field, provided a null electric potential gradient. Then Eq. (1.2) and Eq. (1.3) can be summarized in

$$\mathbf{A}_s = -\Lambda \mathbf{J}_s \tag{1.6}$$

which is coherent because gauge was fixed. London's gauge, essentially, expresses the conservation of the particles number.

## 1.2 THERMODYNAMICS

We shall now concentrate on some thermodynamic aspects of the theory. In the absence of electromagnetic contributions, the free energy is given by

$$F \equiv E - TS$$

In order to include the electromagnetic contributions over a volume $\Omega$, we need to correct its infinitesimal part by

$$dF_m \equiv dF + \mu_0 \int_\Omega d^D\mathbf{r}\, \mathbf{H} \cdot d\mathbf{M}$$

Suppose now no field is applied. Thus we may drop the subscript $m$, being the free energy non magnetic. We indicate with the superscript $(\mathrm{n})$ the normal state and with $(\mathrm{s})$ the superconducting state. The **condensation energy** $\Delta f$ is given by

$$\Delta f(T) = f^{(\mathrm{n})}(T) - f^{(\mathrm{s})}(T)$$

Evidently, if $\Delta f_m > 0$, the superconducting phase is favorite. Pay attention to the fact that we are at zero magnetic field.

### 1.2.1 The critical field

$F_m$, the **magnetic free energy**, has the magnetization as its natural variable. We prefer to work with $\mathbf{H}$ as a natural variable, therefore we define the **magnetic Gibbs free energy** as its Legendre transform

$$G_m = F_m - \mu_0 \int_\Omega d^D\mathbf{r}\, \mathbf{H} \cdot \mathbf{M} \quad \Longrightarrow \quad dG_m = dF - \mu_0 \int_\Omega d^D\mathbf{r}\, \mathbf{M} \cdot d\mathbf{H}$$

We define the energy densities per unit volume

$$f = \frac{F}{\Omega} \qquad g_m = \frac{G_m}{\Omega}$$

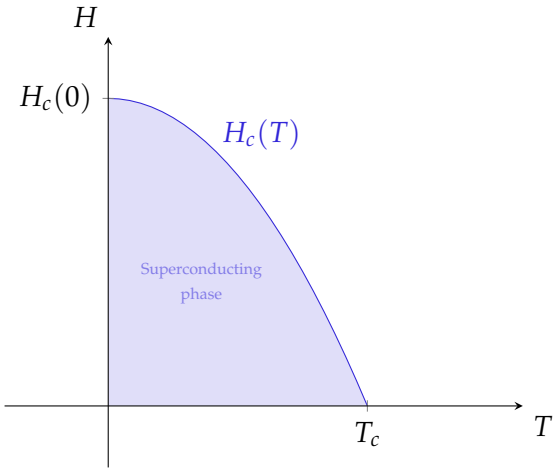



**Figure 1.2:** Sketch of the general critical field dependence on temperature, as depicted in the text.

Thus at some fixed temperature $T$

$$g_m(T,\mathbf{H}) = \int_{(T,\mathbf{0})}^{(T,\mathbf{H})} dg_m(T,\mathbf{H}') = f(T) - \mu_0 \int_{(T,\mathbf{0})}^{(T,\mathbf{H})} d\mathbf{H}' \cdot \langle \mathbf{M}(\mathbf{H}') \rangle$$

where

$$\langle \mathbf{M} \rangle = \frac{1}{\Omega} \int_\Omega d^D\mathbf{r}\, \mathbf{M}$$

Let us comment briefly the above equations. The integrals are intended as adiabatic along the path connecting the states $(T,\mathbf{0})$ to $(T,\mathbf{H})$. We start from a zero applied field because $g_m$ is to be interpreted as the free energy for establishing a $\mathbf{H}$ field starting from a rest condition. The physical picture could be one of a solenoid embedding a superconductor, with a slowly varying current applying a locally static field $\mathbf{H}'$. Notice that the free enegy part, $f$, is non-magnetic and thus depends only on the temperature.

Inside a superconductor all the magnetic flux density, $\mathbf{B}$, is expelled. Thus we have $\mathbf{M} = \langle \mathbf{M} \rangle = -\mathbf{H}$. Substituting this result in the equation for $g_m$ we get

$$g_m^{(\mathrm{s})}(T,\mathbf{H}) = f^{(\mathrm{s})}(T) + \mu_0 \int_{(T,\mathbf{0})}^{(T,\mathbf{H})} d\mathbf{H}' \cdot \mathbf{H}' = f^{(\mathrm{s})}(T) + \left.\frac{\mu_0 H^2}{2}\right|_{@T} \tag{1.7}$$

We expect a critical field $\mathbf{H}_c(T)$ to exist: it is unphysical to think that any material could expel an arbitrarily large applied magnetic field. This is proved by experiments. Then, we understand that the phase transition occurs whenever

$$g_m^{(\mathrm{s})}(T,\mathbf{H}) = g_m^{(\mathrm{n})}(T,\mathbf{H})$$

We can neglect the magnetic contribution to $g_m^{(\mathrm{n})}$, considering non-magnetic materials. Then we define the **critical field** $H_c(T)$ as

$$H_c^2(T) = \frac{2}{\mu_0} \Delta f(T) \tag{1.8}$$

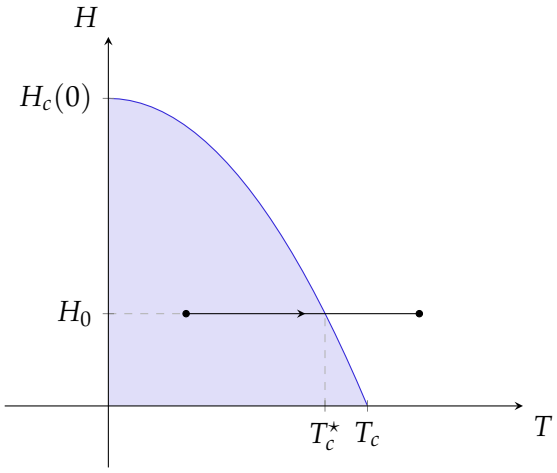



**Figure 1.3:** The phase transition described in Sec. 1.2.2. $H_0$ is the uniform applied field, kept constant and non-zero. Phase transition occurs at temperature $T_c^\star < T_c$.

delimiting a superconducting region. Experimentally it turns out

$$H_c(T) \simeq H_c(0)\left(1 - \frac{T^2}{T_c^2}\right)$$

as plotted in Fig. 1.2.

In the following, we will always use the variable $H$ instead of $\mathbf{H}$, assuming isotropy and therefore dependence only on the field intensity. We now make use of thermodynamics to analyze the order of the phase transition in the two cases $H \neq 0$ and $H = 0$.

### 1.2.2 First order transition at non-zero field

First, we consider the transition at finite applied field $H_0 \neq 0$, varying the temperature $T$ as sketched in Fig. 1.3. We may look at entropy per unit volume, defined as

$$s(T, H_0) \equiv \frac{S(T, H_0)}{\Omega} = -\left[\frac{\partial}{\partial T} g_m(T, H)\right]_{H=H_0}$$

Notice that, due to Eq. (1.7),

$$\begin{aligned} g_m^{(\mathrm{s})}(T, H) - g_m^{(\mathrm{s})}(T, 0) &= \left[g_m^{(\mathrm{n})}(T, H) + \Delta f(T) + \frac{\mu_0 H^2}{2}\right] \\ &\quad - \left[g_m^{(\mathrm{n})}(T, 0) + \Delta f(T)\right] \\ &= \left[g_m^{(\mathrm{n})}(T, H) - g_m^{(\mathrm{n})}(T, 0)\right] + \frac{\mu_0 H^2}{2} = \frac{\mu_0 H^2}{2} \end{aligned}$$

where in the last passage we used that the normal state is non-magnetic, therefore its free energy is independent of the field applied. It follows

$$g_m^{(\mathrm{s})}(T, H) = g_m^{(\mathrm{s})}(T, 0) + \frac{\mu_0 H^2}{2}$$

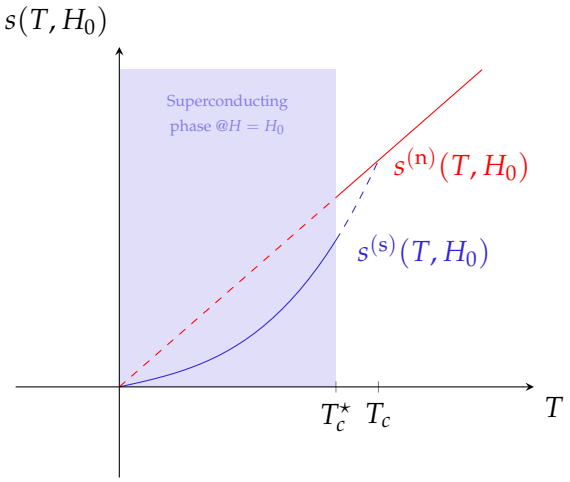



**Figure 1.4:** For an external field $H_0 \neq 0$, transition occurs at temperature $T_c^\star < T_c$. Here is sketched entropy per unit volume for the normal phase (n) and the superconducting one (s). The dashed line represents analytic continuation of the function, the solid line represents the physical entropy in the two regimes. Evidently for any $T_c^\star < T_c$, i.e. for any $H_0 \neq 0$, the physical entropy undergoes a discontinuity at the transition.

From Eq. (1.7) it follows

$$g_m^{(s)}(T,0) = g_m^{(s)}(T, H_c(T)) - \frac{\mu_0 H_c^2(T)}{2}$$

since both sides are equal to $f^{(s)}(T)$, unvaried by the presence of the field. Moreover, $g_m^{(s)}(T, H_c(T)) = g_m^{(n)}(T, H_c(T))$ by the definition of the critical field, Eq. (1.8). Then

$$\begin{aligned} g_m^{(s)}(T,H) &= g_m^{(n)}(T, H_c(T)) + \frac{\mu_0}{2}\left[H^2 - H_c^2(T)\right] \\ &= g_m^{(n)}(T,H) + \frac{\mu_0}{2}\left[H^2 - H_c^2(T)\right] \end{aligned}$$

being, once again, the normal state Gibbs energy independent of the field. Then the entropy difference in following the path in Fig. 1.3 is a function of temperature

$$\begin{aligned} \Delta s(T, H_0) &= s^{(n)}(T, H_0) - s^{(s)}(T, H_0) \\ &= -\left[\frac{\partial}{\partial T} g_m^{(n)}(T,H) - \frac{\partial}{\partial T} g_m^{(s)}(T,H)\right]_{H=H_0} \\ &= \frac{\mu_0}{2}\frac{\partial}{\partial T}\left[H_c^2(T) - H_0^2\right] \\ &= \mu_0 H_c(T)\frac{\partial}{\partial T} H_c(T) \\ &\simeq \mu_0 H_c^2(0)\left(1 - \frac{T^2}{T_c^2}\right)\left(-\frac{2T}{T_c^2}\right) \quad \text{for} \quad T < T_c \end{aligned}$$

since for $T \geq T_c$, evidently, $H_c(T) = 0$ and thus $\Delta s = 0$.

Two things are worth of noticing. As depicted in Fig. 1.4,

$$\Delta s(T_c^\star, H_0 \neq 0) < 0$$

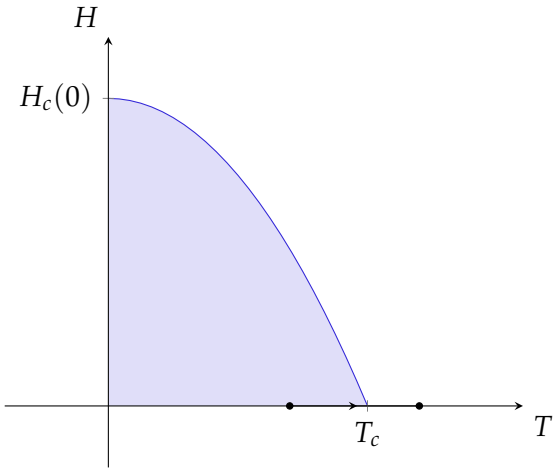



**Figure 1.5:** The transformation described in Sec. 1.2.3, at zero applied field.

where $T_c^\star$ is the temperature at which transition occurs, as represented in Fig. 1.3. First, entropy - a first order derivative of a thermodynamic function - suffers a finite discontinuity, making the superconducting transition a one of the **first order**. Secondly, being $\Delta s < 0$, the transition from normal state to superconducting state increases the global order of the system. This is coherent with what we will find: superconductivity is an essentially collective phenomenon, requiring some notion of long-range order.

### 1.2.3 Second order transition at zero field

Now we turn to the other case, i.e. $H_0 = 0$. The transition is depicted in Fig. 1.5: at zero applied field, we adiabatically increase the temperature above $T_c$. Looking to Fig. 1.4, evidently in this case $T_c^\star = T_c$, meaning that entropy does **not** undergo a discontinuity, but so does its derivative. Then we expect this transition to be one of the second order. The relevant physical quantity involved is the **specific heat** (per unit volume)

$$c(T,H) \equiv T\frac{\partial}{\partial T}s(T,H)$$

then

$$\Delta c(T,H) = T\frac{\partial}{\partial T}\Delta s(T,H)$$

where, as always,

$$\Delta c(T,H) = c^{(\mathrm{n})}(T,H) - c^{(\mathrm{s})}(T,H)$$

Recovering the precedent expression for $\Delta s$,

$$\begin{aligned}\Delta c(T,H) &= T\frac{\partial}{\partial T}\left[\mu_0 H_c(T)\frac{\partial}{\partial T}H_c(T)\right] \\ &= \mu_0 T\left[\left(\frac{\partial}{\partial T}H_c(T)\right)^2 + H_c(T)\frac{\partial^2}{\partial T^2}H_c(T)\right] \\ &\simeq \mu_0 T H_c^4(0)\left[\left(-\frac{2}{T_c^2}+\frac{6T^2}{T_c^4}\right)^2 + \left(1-\frac{T^2}{T_c^2}\right)\left(-\frac{2T}{T_c^2}\right)\frac{12T}{T_c^4}\right]\end{aligned}$$

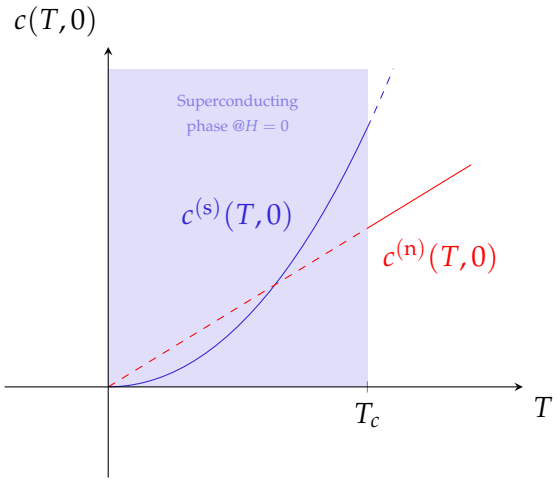



**Figure 1.6:** Specific heat of the normal phase (n) and superconducting phase (s), with zero applied field. The finite discontinuity of $\Delta c$ is the hallmark of the second order transition.

as can be seen by direct inspection. We now evaluate $\Delta c$ at $T = T_c$ and $H = 0$. The second term drops, and we get immediately

$$\Delta c(T_c, 0) > 0$$

As expected, the transition is one of the **second order**, being the specific heat essentially a second derivative of free energy. In Fig. 1.6 is represented the specific heat for the normal and superconducting phases. Similar calculations can be carried out for other quantities, leading to analogous conclusions.

We have discovered, on purely statistical grounds, that superconductivity exhibits many interesting phenomena. The next chapter is devoted to write a formal statistical theory for the phase transition.

# 2 GINZBURG-LANDAU THEORY OF SUPERCONDUCTIVITY



to do: intro

## 2.1 GINZBURG–LANDAU THEORY FOR CLASSICAL SPINS

We start with a simple set of $N$ classical spins $s_i = \pm 1$, rigidly fixed to a lattice in position $i$, evolving through the hamiltonian

$$\mathcal{H}\left[s_1, \cdots, s_N\right]$$

The partition function is

$$\mathcal{Z} = \mathrm{Tr}\left[\sum_{(\forall i)\ s_i} e^{-\beta \mathcal{H}[s_1, \cdots, s_N]}\right]$$

We group the lattice sites; each group $\Omega_\alpha$ contains $g \ll N$ spins and is denoted by its position $\mathbf{x}_\alpha$, as represented schematically in Fig. 2.1. An average magnetization per unit volume can be defined,

$$m(\mathbf{x}_\alpha) \equiv \frac{1}{g} \sum_{j \in \Omega_\alpha} s_j$$

Evidently, in the thermodynamic limit we can take $\mathbf{x}_\alpha$ as a continuous variable, thus considering the average magnetization per unit volume $m(\mathbf{x})$ as a field. As a consequence, the partition function takes the form of a functional integral over the field configurations

$$\mathcal{Z} = \int \mathcal{D}_m e^{-\beta F[m]}$$

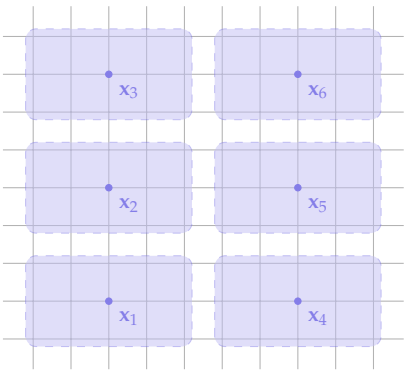



**Figure 2.1:** Coarse graining of a bidimensional square lattice. At each side of the grid, a spin $s_i = \pm 1$ is located. The colored regions represent the grouping of spins.

Such a system, in absence of fields, evidently exhibits a $\mathbb{Z}_2$ symmetry. Moreover, we suppose a critical temperature $\beta_c$ exists such that

$$(\beta < \beta_c) \qquad \langle m(\mathbf{x}) \rangle = \frac{1}{\mathcal{Z}} \int \mathcal{D}_m m(\mathbf{x}) e^{-\beta F[m]} = 0$$

Evidently, coupling a uniaxial field $h(\mathbf{x})$ to the system

$$F[m,h] = F[m] - \int d\mathbf{x}\, h(\mathbf{x}) m(\mathbf{x})$$

the favoured configuration will be the one where, at each point, the product $h(\mathbf{x})m(\mathbf{x})$ is positive (meaning: the local magnetization is aligned with the field). This is a clear symmetry breaking. We suppose the system exhibits such spontaneous symmetry breaking for low temperatures, described by

$$(\beta > \beta_c) \qquad \langle m(\mathbf{x}) \rangle = \lim_{h \to 0^{\pm}} \frac{1}{\mathcal{Z}} \int \mathcal{D}_m m(\mathbf{x}) e^{-\beta F[m,h]} = \pm m_0(\mathbf{x})$$

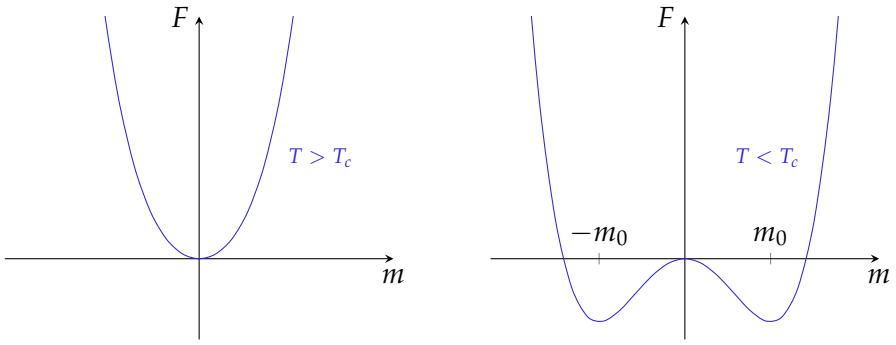
with $m_0(\mathbf{x}) \neq 0$. In essence, for subcritical temperatures we suppose the system to start spontaneously magnetizing, in a manner similar to the application of an infinitesimal field. Then the local magnetization (a physical, therefore continuous quantity) is a order parameter signaling a phase transition. It changes from zero to non-zero undergoing the transition. We are interested in the point $\beta_c$ itself, then we can expand the free energy in even powers (odd powers of expansion are suppressed due to $\mathbb{Z}_2$ symmetry):

$$F[m] = F_0 + \int d\mathbf{x} \left[ a m^2(\mathbf{x}) + \frac{b}{2} m^4(\mathbf{x}) + c |\boldsymbol{\nabla} m(\mathbf{x})|^2 + \cdots \right]$$

where $a, b$ and $c$ are the expansion coefficients, in general depending on $\beta$. We suppose $b > 0$ for any temperature, in order to keep the free energy positive-defined (for $\beta < 0$ an absolute minimum in this expansion could not be defined). We also require $c > 0$: in this system, to have local magnetization oscillations should increase the energy, not decreasing it.

Since the free energy enters the partition function and the configurational probability as an exponential, we perform a stationary phase approximation minimizing the free energy

$$\frac{\delta F}{\delta m} = 0$$




**(a)** Supercritical free energy.

**(b)** Subcritical free energy.

**Figure 2.2:** The free energy in Ginzburg-Landau theory for the magnetizable array of classical spins in $\mathbb{Z}_2$ symmetry. The supercritical ($T > T_c$) free energy 2.2a only displays a clear minimum for $m = 0$. The subcritical ($T < T_c$) free energy 2.2b has two minima, $m = \pm m_0$.

### 2.1.1 Homogeneous magnetization in absence of fields

The special case of an homogeneous order parameter,

$$m(\mathbf{x}) = m_0$$

gives immediately $\boldsymbol{\nabla} m(\mathbf{x}) = 0$. Therefore, minimizing the free energy functional,

$$\frac{\delta}{\delta m} \int d\mathbf{x} \left[ a m^2(\mathbf{x}) + \frac{b}{2} m^4(\mathbf{x}) \right]_{m(\mathbf{x})=m_0} = 2\left(a m_0 + b m_0^3\right) \stackrel{!}{=} 0$$

which allows for the solutions

$$m_0 = 0 \qquad \text{and} \qquad m_0^2 = -\frac{a}{b}$$

Since $b > 0$, and we assume $m_0 = 0$ to be the solution for $T > T_c$, then $a$ needs to change sign across the transition. In particular, it must be

$$a > 0 \quad \text{for } T > T_c \qquad\qquad a < 0 \quad \text{for } T < T_c$$

In order to make the second solution acceptable only in the low-temperatures regime. Finally, $a$ is assumed to be a continuous function of the temperature, since free energy can't be gapped. Then, expanding $a$ around the critical temperature, its expansion must be odd, and the simplest term is the linear one

$$a \simeq a_0 \left(T - T_c\right)$$

as confirmed by the general theory of $\mathbb{Z}_2$ phase transitions. The free energy $F$ takes the form sketched in Fig. 2.2.

### 2.1.2 Breaking of symmetry through field coupling

We now consider the coupling of an inhomogeneous field $h(\mathbf{x})$. Of course, now we cannot neglect space variations of the magnetization, $\boldsymbol{\nabla} m(\mathbf{x}) \neq 0$. The free energy functional is given by

$$F\left[m, h\right] = F_0 + \int d\mathbf{x} \left[ a m^2(\mathbf{x}) + \frac{b}{2} m^4(\mathbf{x}) + c|\boldsymbol{\nabla} m(\mathbf{x})|^2 - h(\mathbf{x}) m(\mathbf{x}) + \cdots \right]$$

With some calculations, it turns out

$$\frac{\delta F}{\delta m} = 0 \quad \Longrightarrow \quad 2\left[a m_h(\mathbf{x}) + b m_h^3(\mathbf{x}) - c\nabla^2 m_h(\mathbf{x})\right] = h(\mathbf{x}) \tag{2.1}$$

where $m_h(\mathbf{x})$ is the magnetization in presence of the field. The great advantage in introducing such field is the following: the magnetization in presence of the field is simply

$$m_h(\mathbf{x}) = \frac{1}{\mathcal{Z}[h]} \int \mathcal{D}_m \, m(\mathbf{x}) e^{-\beta F[m] + \beta \int d\mathbf{x}' \, h(\mathbf{x}') m(\mathbf{x}')}$$

with the new partition function

$$\mathcal{Z}[h] \equiv \int \mathcal{D}_m \, e^{-\beta F[m] + \beta \int d\mathbf{x}' \, h(\mathbf{x}') m(\mathbf{x}')}$$

Evidently

$$\begin{aligned}\frac{\delta m_h(\mathbf{x})}{\delta h(\mathbf{x}')} = \frac{\delta}{\delta h(\mathbf{x}')}\left[\frac{1}{\mathcal{Z}[h]}\right] \int \mathcal{D}_m \, m(\mathbf{x}) e^{-\beta F[m] + \beta \int d\mathbf{x}' \, h(\mathbf{x}') m(\mathbf{x}')} \\ + \frac{1}{\mathcal{Z}[h]} \frac{\delta}{\delta h(\mathbf{x}')} \left[\int \mathcal{D}_m \, m(\mathbf{x}) e^{-\beta F[m] + \beta \int d\mathbf{x}' \, h(\mathbf{x}') m(\mathbf{x}')}\right]\end{aligned}$$

Performing the calculation, and taking the $h \to 0$ limit, it turns out

$$\frac{1}{\beta} \frac{\delta m_h(\mathbf{x})}{\delta h(\mathbf{x}')} = \left\langle m(\mathbf{x}) m(\mathbf{x}') \right\rangle - \left\langle m(\mathbf{x}) \right\rangle \left\langle m(\mathbf{x}') \right\rangle \Big|_{@h \to 0}$$

The above expression is exactly the **spatial correlation function** for the order parameter,

$$\mathcal{C}\left(\mathbf{x} - \mathbf{x}'\right) \equiv \left\langle m(\mathbf{x}) m(\mathbf{x}') \right\rangle - \left\langle m(\mathbf{x}) \right\rangle \left\langle m(\mathbf{x}') \right\rangle \Big|_{@h \to 0}$$

which is a measure of how much magnetization is correlated taking two points $\mathbf{x}$, $\mathbf{x}'$. The dependence on the difference of position was included in order to implement assumed translational symmetry in the thermodynamic limit. We consider small fluctuations around the rest magnetization $m_0$,

$$m_h(\mathbf{x}) = m_0 + \delta m_h(\mathbf{x})$$

induced by a small field $\delta h(\mathbf{x})$. Then, linearizing Eq. (2.1),

$$\left(a + 3b m_0^2\right) \delta m_h(\mathbf{x}) - c\nabla^2 \delta m_h(\mathbf{x}) = \frac{\delta h(\mathbf{x})}{2}$$

which translates in Fourier transform as

$$\left(a + 3b m_0^2\right) \delta m_h(\mathbf{q}) + c|\mathbf{q}|^2 \delta m_h(\mathbf{q}) = \frac{\delta h(\mathbf{q})}{2}$$

Thus

$$\beta \mathcal{C}(\mathbf{q}) = \frac{\delta m_h(\mathbf{q})}{\delta h(\mathbf{q})} = \frac{1}{2c} \frac{1}{\xi^{-2} + |\mathbf{q}|^2} \quad \text{with} \quad \xi^{-2} \equiv \frac{a + 3b m_0^2}{c}$$

which is the Fourier transform of the Yukawa potential, with screening $\xi$. Finally, we have an expression for the correlator

$$\mathcal{C}(\mathbf{x}-\mathbf{x}') = \int_\star d\mathbf{q}\, \mathcal{C}(\mathbf{q}) e^{-i\mathbf{q}\cdot(\mathbf{x}-\mathbf{x}')}$$

The subscript $\star$ indicates the momentum region over which we integrate. Since we are coarse-graining the model (see Fig. 2.1) we must impose a cutoff $\Lambda$ to $|\mathbf{q}|$, since for extremely high momenta (which means, extremely small distances) the magnetization field $m(\mathbf{x})$ wouldn't even be properly defined. Then

$$\mathcal{C}(\mathbf{x}-\mathbf{x}') = \frac{k_B T}{2c} \int_{|\mathbf{q}|<\Lambda} d\mathbf{q}\, \frac{e^{-i\mathbf{q}\cdot(\mathbf{x}-\mathbf{x}')}}{\xi^{-2}+|\mathbf{q}|^2} = (\text{factors}) \frac{e^{-|\mathbf{x}-\mathbf{x}'|/\xi}}{|\mathbf{x}-\mathbf{x}'|}$$

The interpretation of $\xi$ becomes clear: over a distance $\xi$, the order parameter becomes approximately uncorrelated. For such a distance we expect the magnetization (and in the following the superconducting order parameter) to fluctuate significantly. This length will play a fundamental role in understanding the behavior of different classes of superconductors. Most importantly, it can be shown that for low symmetry-breaking fields the transition is a second-order one.

## 2.2 SYMMETRY BREKING IN SUPERCONDUCTORS

We intend to apply all this machinery to superconductors. We know from Chap. 1 that for zero field the superconducting transition is a second order one, just like for the Ising model. The key idea of Landau was the following: the order parameter arising from zero to non-zero in the transition must be somehow connected to $n_s(\mathbf{x})$. From a quantum perspective, it's intuitive to define a pseudo-wavefunction $\Psi(\mathbf{x})$ such that

$$|\Psi(\mathbf{x})|^2 = n_s(\mathbf{x})$$

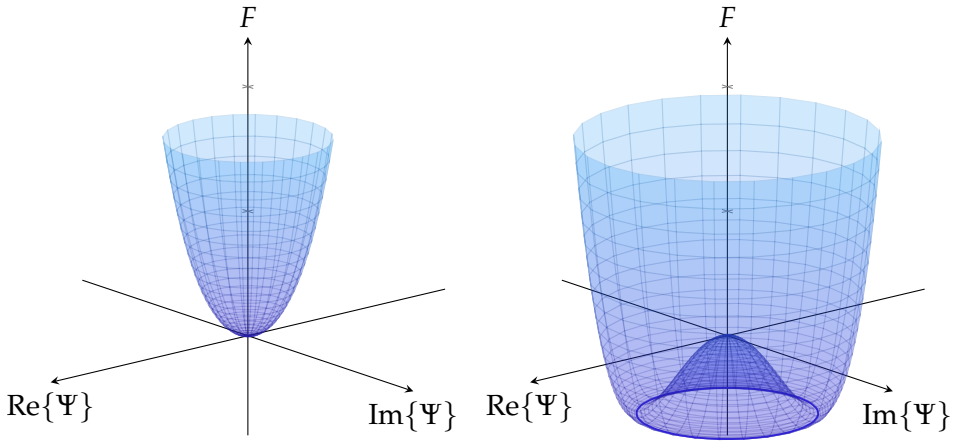
Superconductivity is an inherently quantum phenomenon. Then it's most reasonable to assume $\Psi$ itself as the order parameter. Since no measurable quantity can depend on a global sign change of the pseudo-wavefunction (which is e $e^{i\pi}$ phase change), a similar expansion as before must hold for $\Psi$

$$F[\Psi,\Psi^*] = F_0 + \int d\mathbf{x} \left[ a|\Psi(\mathbf{x})|^2 + \frac{b}{2}|\Psi(\mathbf{x})|^4 + c|\boldsymbol{\nabla}\Psi(\mathbf{x})|^2 + \cdots \right]$$

The parameter arguments are $\Psi$, $\Psi^*$, equivalent to $\mathrm{Re}\{\Psi\}$ and $\mathrm{Im}\{\Psi\}$. In the homogeneous case the parameter is specified up to a phase, thus giving the free energy sketched in Fig. 2.3. The equilibrium state is now degenerate over a non-countable set of solutions, one for each choice of the phase, lying on the circular minimum of the free energy.

The underlying symmetry of Quantum Mechanics is the simple $U(1)$ symmetry of phases, so it is that symmetry that gets broken in the phase transition. We consider a symmetry breaking field $\eta(\mathbf{x})$, conjugated to the order parameter via minimum coupling as follows

$$F[\Psi,\Psi^*,\eta,\eta^*] = F[\Psi,\Psi^*] - \int d\mathbf{x}\, [\eta^*(\mathbf{x})\Psi(\mathbf{x}) + \eta(\mathbf{x})\Psi^*(\mathbf{x})]$$




**(a)** Free energy for $T > T_c$ **(b)** Free energy for $T < T_c$

**Figure 2.3:** In Fig. 2.3a is represented the supercritical Ginzburg-Landau free-energy expansion for the $U(1)$ order parameter of Quantum Mechanics. As can be seen, the equilibrium condition is given by a single point. In Fig. 2.3b the subcritical situation is represented. The minimum here is degenerate into a ring. The topology of such equilibrium manifold is the reason of the spontaneous symmetry breaking mechanism in superconductors.

Proceeding as in the previous section, we need to minimize the free energy with respect to the order parameters $\Psi$, $\Psi^*$:

$$\frac{\delta F}{\delta \Psi^*} \stackrel{!}{=} 0 \qquad \Longrightarrow \qquad a\Psi(\mathbf{x}) + b|\Psi(\mathbf{x})|^2\Psi(\mathbf{x}) - c\nabla^2\Psi(\mathbf{x}) \;= \eta(\mathbf{x}) \tag{2.2}$$

$$\frac{\delta F}{\delta \Psi} \stackrel{!}{=} 0 \qquad \Longrightarrow \qquad a\Psi^*(\mathbf{x}) + b|\Psi(\mathbf{x})|^2\Psi^*(\mathbf{x}) - c\nabla^2\Psi^*(\mathbf{x}) \;= \eta^*(\mathbf{x}) \tag{2.3}$$

If we choose a homogeneous symmetry breaking field with specified phase,

$$\eta(\mathbf{x}) = \eta_0 e^{i\varphi}$$

and then take the limit $\eta_0 \to 0$, a specific phase for the order parameter gets selected. In fact, the order parameter minimizes the free energy in the homogeneous configuration $\Psi(\mathbf{x}) = \Psi_0 e^{i\varphi}$, with $\Psi_0 \in \mathbb{R}$. This gives rise to the equation

$$\left[a + b|\Psi_0|^2\right] \Psi_0 e^{i\varphi} = \eta_0 e^{i\varphi}$$

Evidently, for $\eta_0 \to 0$, the transition occurs with a homogeneous order parameter

$$|\Psi_0|^2 = -\frac{a}{b} \qquad \text{with} \qquad a < 0$$

As in the case of the Ising model, we may suppose a linear dependence of $a$ on temperature. In Fig. 2.3b, of all the equivalent states in the *sombrero* minimum (the ring), the one at angle $\varphi$ becomes the physical state.

In the general case, the symmetry breaking field has a complex spatial dependence

$$\eta(\mathbf{x}) = |\eta(\mathbf{x})|e^{i\varphi(\mathbf{x})}$$

which incorporates both a magnitude and a phase part. From physical intuition, we can postulate that our complex Ginzburg-Landau system will

behave like the classical Ising system of Sec. 2.1 for fluctuations of the $\eta$ field at a fixed phase. In other terms, the free energy of Fig. 2.2b looks exactly like the one in Fig. 2.3b, as seen by a specific intersecting plane at a given angle. $\mathbb{Z}_2$ symmetry is incorporated in $U(1)$ symmetry, if we limit ourselves to antipodal phases. In the next section we discuss how the order parameter fluctuates around the rest position.

## 2.3 FLUCTUATIONS OF THE COMPLEX ORDER PARAMETER

How to deal with fluctuations in presence of a complex order parameter? We need to extend the paradigm developed in Sec. 2.1. The starting point is the partition function,

$$\begin{aligned}\mathcal{Z}\left[\Psi,\Psi^*,\eta,\eta^*\right] &\equiv \int \mathcal{D}_{\Psi,\Psi^*}\exp\{-\beta F\left[\Psi,\Psi^*,\eta,\eta^*\right]\}\\ &= \int \mathcal{D}_{\Psi,\Psi^*}\exp\Big\{-\beta F\left[\Psi,\Psi^*\right]\\ &\qquad +\beta\int d\mathbf{x}\,\left[\eta^*(\mathbf{x})\Psi(\mathbf{x})+\eta(\mathbf{x})\Psi^*(\mathbf{x})\right]\Big\}\end{aligned}$$

With some fixed field $\eta$, we have

$$\Psi_\eta(\mathbf{x}) = \frac{1}{\mathcal{Z}\left[\Psi,\Psi^*,\eta,\eta^*\right]}\int \mathcal{D}_{\Psi,\Psi^*}\Psi(\mathbf{x})\exp\{-\beta F\left[\Psi,\Psi^*,\eta,\eta^*\right]\}$$

then, we see that taking the functional derivative

$$\frac{\delta\Psi_\eta(\mathbf{x})}{\delta\eta(\mathbf{x}')}$$

this must act both on the partition function and the free energy inside the integral. Skipping some steps, it turns out

$$\begin{aligned}\frac{\delta\Psi_\eta(\mathbf{x})}{\delta\eta(\mathbf{x}')} &= \frac{1}{\mathcal{Z}}\int \mathcal{D}_{\Psi,\Psi^*}\Psi(\mathbf{x})\Psi^*(\mathbf{x}')\exp\{-\beta F\}\\ &\quad -\frac{1}{\mathcal{Z}}\int \mathcal{D}_{\Psi,\Psi^*}\Psi(\mathbf{x})\exp\{-\beta F\}\frac{1}{\mathcal{Z}}\int \mathcal{D}_{\Psi,\Psi^*}\Psi^*(\mathbf{x}')\exp\{-\beta F\}\end{aligned}$$

then, being the correlator a measure of *how much* $\Psi$ changes at point $\mathbf{x}$ if a perturbation $\eta$ acts at point $\mathbf{x}'$, we see that

$$\mathcal{C}(\mathbf{x}-\mathbf{x}') = \left\langle\Psi(\mathbf{x})\Psi^*(\mathbf{x}')\right\rangle - \left\langle\Psi(\mathbf{x})\right\rangle\left\langle\Psi^*(\mathbf{x}')\right\rangle = \left\langle\delta\Psi(\mathbf{x})\delta\Psi^*(\mathbf{x}')\right\rangle$$

as can be seen by direct inspection. This will be important in the following.

We are now interested in analyzing how correlation acts on a broken-symmetry situation. Without loss of generality, we suppose the system exhibits $U(1)$ symmetry breaking at $\varphi = 0$, due to a homogeneous field. Thus a precise state on the degenerate ring in Fig 2.3b is selected. We want to analyze small fluctuations of the order parameter around the homogeneous mean-field $\Psi(\mathbf{x}) = \Psi_0$ condition. The symmetry breaking field is complex, thus fluctuations can occur both in the magnitude and in the phase of the field.

### 2.3.1 Magnitude fluctuations

We now suppose to let the magnitude of the real symmetry breaking field fluctuate space-wise as follows

$$\eta_0 \to \eta_0 e^{\delta k(\mathbf{x})}$$

The exponential form is the common one for magnitude transformations, like contractions. We suppose the fluctuations $\delta k(\mathbf{x}) \in \mathbb{R}$ to be small, i.e.

$$\eta_0 \to \eta_0 \left(1 + \delta k(\mathbf{x})\right)$$

For small perturbations, we expect the order parameter to accommodate accordingly

$$\Psi_0 \to \Psi_0 e^{\delta k(\mathbf{x})} \simeq \Psi_0 \left(1 + \delta k(\mathbf{x})\right)$$

Consider Eq. (2.2). Under a small magnitude perturbation,

$$a\Psi(\mathbf{x}) + b|\Psi(\mathbf{x})|^2\Psi(\mathbf{x}) - c\nabla^2\Psi(\mathbf{x}) = \eta(\mathbf{x}) \tag{2.4}$$

$$a\Psi_0 e^{\delta k(\mathbf{x})} + b\left[\Psi_0 e^{\delta k(\mathbf{x})}\right]^2 \Psi_0 e^{\delta k(\mathbf{x})} - c\nabla^2\Psi_0 e^{\delta k(\mathbf{x})} = \eta_0 e^{\delta k(\mathbf{x})} \tag{2.5}$$

$$\left[a + b\Psi_0^2\right]\Psi_0 + \left[a + 3b\Psi_0^2\right]\delta\Psi(\mathbf{x}) - c\nabla^2\delta\Psi(\mathbf{x}) \simeq \eta_0 + \delta\eta(\mathbf{x}) \tag{2.6}$$

where we took the linear order in $\delta k$, and

$$\delta\Psi(\mathbf{x}) = \Psi_0\delta k(\mathbf{x}) \qquad \text{and} \qquad \delta\eta(\mathbf{x}) = \eta_0\delta k(\mathbf{x})$$

The first terms of Eq. (2.6) both sides cancel,

$$\left[a + b\Psi_0^2\right]\Psi_0 = \eta_0$$

representing the rest homogeneous solution. The remaining terms,

$$\left[a + 3b\Psi_0^2\right]\delta\Psi(\mathbf{x}) - c\nabla^2\delta\Psi(\mathbf{x}) = \delta\eta(\mathbf{x})$$

give rise to a solution perfectly analogous to the classical Ising model analyzed in Sec. 2.1, with the mapping

$$\delta\Psi(\mathbf{x}) \to \delta m(\mathbf{x}) \qquad \text{and} \qquad \delta\eta(\mathbf{x}) \to \delta h(\mathbf{x})$$

which can be done flawlessly because everything is real by assumption. This is rather intuitive, because for any fixed phase $\varphi$, the system is physically equivalent to the above discussed classical system, so no more physics is expected. For such a fluctuation we expect a Yukawa correlation function on a scale $\xi$, as defined above:

$$\mathcal{C}\left(\mathbf{x} - \mathbf{x}'\right) \sim \frac{e^{-|\mathbf{x}-\mathbf{x}'|/\xi}}{|\mathbf{x} - \mathbf{x}'|} \qquad \text{with} \qquad \xi^{-2} = \frac{a + 3b\Psi_0^2}{c}$$

We highlight something important to understand the huge difference arising in the next section. Pay attention to the factor 3 present in the definition of $\xi$ in the above equation. It emerges from the fact that we are dealing with a **magnitude** fluctuation: passing from Eq. (2.4) to Eq. (2.5) we made the following substitution

$$|\Psi(\mathbf{x})|^2 \to \left[\Psi_0 e^{\delta k(\mathbf{x})}\right]^2 = \Psi_0^2\, e^{2\delta k(\mathbf{x})}$$

which of course cannot be done if $\delta k \in \mathbb{C}$, as in the next section. Linearizing, the factor 3 emerges, and prevents $\xi$ from diverging in the subcritical $\eta_0 \to 0$ case: in such a case, in fact, we know that the subcritical rest solution is the real homogeneous one $\Psi(\mathbf{x}) = \Psi_0$ with

$$\Psi_0^2 = -\frac{a}{b} \qquad \text{for} \qquad T < T_c$$

Then, a small magnitude-perturbing field at fixed phase and temperature does **not** make correlations *go crazy*.

### 2.3.2 Phase fluctuations

Suppose now the fluctuation occurs in the phase,

$$\eta_0 \to \eta_0 e^{i\delta\varphi(\mathbf{x})} \simeq \eta_0 + i\eta_0\delta\varphi(\mathbf{x})$$

then a specular fluctuation in the phase of $\Psi$ emerges,

$$\Psi(\mathbf{x}) = \Psi_0 + i\Psi_0\delta\varphi(\mathbf{x})$$

We need to be a bit more cautious here. We start from Eq. (2.2) at linear order,

$$a\Psi(\mathbf{x}) + b|\Psi(\mathbf{x})|^2\Psi(\mathbf{x}) - c\nabla^2\Psi(\mathbf{x}) = \eta(\mathbf{x}) \tag{2.7}$$

$$a\Psi_0 e^{i\delta\varphi(\mathbf{x})} + b\left[\Psi_0^2\right]\Psi_0 e^{i\delta\varphi(\mathbf{x})} - c\nabla^2\Psi_0 e^{i\delta\varphi(\mathbf{x})} = \eta_0 e^{i\delta\varphi(\mathbf{x})} \tag{2.8}$$

$$\left[a + b\Psi_0^2\right]\Psi_0 + \left[a + b\Psi_0^2\right]\delta\Psi(\mathbf{x}) - c\nabla^2\delta\Psi(\mathbf{x}) \simeq \eta_0 + \delta\eta(\mathbf{x}) \tag{2.9}$$

where $\delta\Psi(\mathbf{x}) = i\Psi_0\delta\varphi(\mathbf{x})$ and $\delta\eta(\mathbf{x}) = i\eta_0\delta\varphi(\mathbf{x})$. Eq. (2.9) differs from Eq. (2.6) in the absence of the factor 3. As we'll see, this will have a great impact. From Eq. (2.9), neglecting the first term both sides (they cancel out) and taking the Fourier transform it's easy to see

$$\delta\Psi(\mathbf{q}) = \frac{1}{c}\frac{\delta\eta(\mathbf{q})}{\kappa^2 + |\mathbf{q}|^2} \quad \text{with} \quad \kappa^2 = \frac{a + b\Psi_0^2}{c}$$

We intentionally referred to $\kappa$ with a different notation with respect to $\xi$. Notice that, **for vanishing symmetry breaking fields** $\eta_0 \to 0$ we recover the situation in Fig. 2.3b, and $\Psi_0^2 \to -a/b$, which means $\kappa \to 0$. As before,

$$\mathcal{C}(\mathbf{q}) = \frac{1}{\beta}\frac{\delta\Psi(\mathbf{q})}{\delta\eta(\mathbf{q})} = \frac{k_B T}{c}\frac{1}{\kappa^2 + |\mathbf{q}|^2}$$

We can already see the problem here. If $\kappa \to 0$, this is the Fourier transform of the $3D$ Coulomb potential, which presents infrared divergence in low dimensionality. Note that, instead, for non-vanishing symmetry breaking fields ($\eta_0 \neq 0$) no problem arises at all.

What consequences has this? Take the **local phase fluctuations** of the phase parameter,

$$\langle\delta\varphi(\mathbf{x})\delta\varphi(\mathbf{x})\rangle = \langle\delta\varphi(\mathbf{0})\delta\varphi(\mathbf{0})\rangle$$

where we used the assumption of translational invariance. We will refer to this quantity simply as $\langle\delta\varphi^2\rangle$. It's not needed to subtract the disconnected

component of the fluctuation, $\langle \delta\varphi(\mathbf{x})\rangle \langle \delta\varphi(\mathbf{x})\rangle$, since it is assumed to have zero mean. Evidently

$$\langle \delta\varphi^2 \rangle = \frac{\langle (i\Psi_0 \delta\varphi(\mathbf{0}))\,(-i\Psi_0 \delta\varphi(\mathbf{0}))\rangle}{\Psi_0^2} = \frac{\langle \delta\Psi(\mathbf{0})\delta\Psi^*(\mathbf{0})\rangle}{\Psi_0^2}$$

We will use the brief notation $\langle \delta\Psi(\mathbf{0})\delta\Psi^*(\mathbf{0})\rangle = \langle \delta\Psi^2\rangle$. Now, since

$$\langle \delta\Psi(\mathbf{x})\delta\Psi^*(\mathbf{x}')\rangle = \mathcal{C}\left(\mathbf{x}-\mathbf{x}'\right)$$

it turns out

$$\langle \delta\varphi^2 \rangle = \frac{\langle \delta\Psi^2\rangle}{\Psi_0^2} = \frac{\mathcal{C}(\mathbf{0})}{\Psi_0^2} = \frac{k_B T}{c\Psi_0^2}\int_\star \mathcal{C}(\mathbf{q})$$

Again, the subscript $\star$ indicates we impose an integration cutoff $|\mathbf{q}| < \Lambda$. Skipping some steps, we see

$$\langle \delta\varphi^2 \rangle \sim \kappa^{D-2}\int_0^{\Lambda/\kappa} dq \frac{q^{D-1}}{1+q^2}$$

and going under the critical temperature

$$(@T < T_c) \qquad \langle \delta\varphi^2 \rangle \sim \lim_{\kappa\to 0}\kappa^{D-2}\int_0^{\Lambda/\kappa} dq \frac{q^{D-1}}{1+q^2}$$

Then

1. For $D = 1$,
$$\lim_{\kappa\to 0}\frac{1}{\kappa}\int_0^{\Lambda/\kappa} dq \frac{1}{1+q^2} = +\infty$$
since the integral is finite, but the $\kappa^{-1}$ prefactor diverges;

2. For $D = 2$,
$$\lim_{\kappa\to 0}\int_0^{\Lambda/\kappa} dq \frac{q}{1+q^2} = +\infty$$
since the integral is logarithmic for large $q$.

3. For $D > 2$,
$$\lim_{\kappa\to 0}\kappa^{D-2}\int_0^{\Lambda/\kappa} dq \frac{q^{D-1}}{1+q^2} < \lim_{\kappa\to 0}\kappa^{D-2}\int_0^{\Lambda/\kappa} dq\, q^{D-3} = \frac{\Lambda^{D-2}}{D-2}$$
which remains finite.

Then, at low dimensionality ($D \leq 2$) phase fluctuations (**in absence of fields**) are completely disruptive! The symmetry cannot be properly broken, since phase will start fluctuate due to heavy correlations happening in low dimensions. Then we expect, for low-dimensional system, that the system does **not** exhibit long-range order. This result can be inferred by Mermin-Wagner theorem on spontaneous continuous symmetry breaking.

## 2.4 SUPERCONDUCTIVITY

Let's turn now to proper superconductors. All the mechanism we developed lies upon a thermodynamic free energy expansion. We shall do so.

### 2.4.1 Free energy expansion and Ginzburg-Landau equations

We need an expression for the free energy. This should be something of the form

$$F = F^{(\mathrm{n})} + F^{(\mathrm{s})} + F^{(\mathrm{m})}$$

where $(\mathrm{n})$ represents the normal state, $(\mathrm{s})$ the superconducting one and $(\mathrm{m})$ the bare magnetic contribution to free energy. In fact, to find an equilibrium position by minimization of the free energy is to find the equilibrium configuration of the order parameter *and* the fields. In the following we assume $F^{(\mathrm{n})}$ only depends on the temperature. The magnetic energy stored in the field is

$$E^{(\mathrm{m})} = \int d\mathbf{x}\, \frac{|\mathbf{B}|^2}{2\mu_0}$$

then the free energy is given by

$$F^{(\mathrm{m})} = \int d\mathbf{x}\, \left[\frac{|\mathbf{B}|^2}{2\mu_0} - \mathbf{B}\cdot\mathbf{H}\right] = \int d\mathbf{x}\, \left[\frac{|\boldsymbol{\nabla}\times\mathbf{A}|^2}{2\mu_0} - (\boldsymbol{\nabla}\times\mathbf{A})\cdot\mathbf{H}\right]$$

Note that here $\mathbf{A}$ is a gauge field, no gauge choice has been made.

Let's move to the superconducting part. The key (highly non-trivial) assumption, here, is that the pseudo-wavefunction $\Psi(\mathbf{x})$ actually *is* a proper wavefunction. Of course in this case any gauge choice on $\mathbf{A}$ translates into a gauge transformation of the wavefunction. Then, recovering the expansion of the last section

$$F\left[\Psi,\Psi^*\right] = F_0 + \int d\mathbf{x}\, \left[a|\Psi(\mathbf{x})|^2 + \frac{b}{2}|\Psi(\mathbf{x})|^4 + c|\boldsymbol{\nabla}\Psi(\mathbf{x})|^2\right]$$

we may interpret the gradient operator as the momentum operator of Quantum Mechanics,

$$c|\boldsymbol{\nabla}\Psi(\mathbf{x})|^2 = \frac{1}{2m}|-i\hbar\boldsymbol{\nabla}\Psi(\mathbf{x})|^2 = \frac{1}{2m}|\mathbf{p}\Psi(\mathbf{x})|^2$$

where $m$, *a priori*, is not the electron mass but a just the quantity $m = \hbar^2/2c$. A charge $q$ is coupled to the electromagnetic field via the Peierls substitution,

$$\mathbf{p} \to \mathbf{p} - q\mathbf{A}$$

then the free energy turns out to be

$$\begin{aligned} F\left[\Psi,\Psi^*,\mathbf{A};\mathbf{H},T\right] = F^{(\mathrm{n})}[T] &+ \int d\mathbf{x}\, \left[\frac{|\boldsymbol{\nabla}\times\mathbf{A}|^2}{2\mu_0} - (\boldsymbol{\nabla}\times\mathbf{A})\cdot\mathbf{H}\right] \\ &+ \int d\mathbf{x}\, \left[a|\Psi(\mathbf{x})|^2 + \frac{b}{2}|\Psi(\mathbf{x})|^4 + \frac{1}{2m}|-i\hbar\boldsymbol{\nabla}\Psi(\mathbf{x}) - q\mathbf{A}\Psi(\mathbf{x})|^2\right] \end{aligned}$$

The externally fixed parameters, $\mathbf{H}$ and $T$, are separated from the physical parameters (with respect to which we must differentiate) by a semicolon. Minimal coupling of the field parameter $\Psi$ is implemented, as it is easy to see by the means of a second-order expansion (in powers of the vector potential). Functional derivation (be kind, I can't find the will to perform it) leads to the **Ginzburg-Landau equations for superconductors**

$$a\Psi + b|\Psi|^2\Psi + \frac{1}{2m}\left[-i\hbar\boldsymbol{\nabla} - q\mathbf{A}\right]^2\Psi = 0 \tag{2.10}$$

$$\frac{q}{2im}\left[\hbar\Psi^*\boldsymbol{\nabla}\Psi - \hbar\Psi\boldsymbol{\nabla}\Psi^* - 2qi|\Psi|^2\mathbf{A}\right] = \mathbf{J} \tag{2.11}$$

where spatial dependence is intended. Eq. (2.10) is obtained differentiating with respect to $\Psi^*$, while Eq. (2.11) with respect to $\mathbf{A}$.

### 2.4.2 The critical field

We want to determine the critical field $H_c = |\mathbf{H}_c|$ defined as in Sec. 1.2.1. For the normal state ($T > T_c$), inside the non-magnetic material

$$\mathbf{B} = \mu_0\mathbf{H} \quad , \quad \Psi = 0 \quad \Longrightarrow \quad f^{(\mathrm{m})} = -\frac{\mu_0}{2}|\mathbf{H}|^2 \quad , \quad f^{(\mathrm{s})} = 0$$

while in the superconducting state ($T < T_c$)

$$\mathbf{B} = 0 \quad , \quad |\Psi|^2 = -\frac{a}{b} \quad \Longrightarrow \quad f^{(\mathrm{m})} = 0 \quad , \quad f^{(\mathrm{s})} = a|\Psi|^2 + \frac{b}{2}|\Psi|^4 = -\frac{a^2}{2b}$$

Of course perfect screening of the magnetic flux density $\mathbf{B}$ is possible only for low external fields $\mathbf{H}$; although superconductivity exists in a large superconducting region, it is not true that *everywhere* the field is perfectly cancelled for large fields, as will become clear in the following sections. Then, requesting at the transition the free energy to be moved from the fields to the superconducting state,

$$\frac{\mu_0}{2}|\mathbf{H}_c|^2 = \frac{a^2}{2b}$$

Similarly to the analogous parameter $a$ of Sec. 2.1, we assume $a$ to vary linearly with $T - T_c$. This leads to

$$H_c(T) \sim |T - T_c|$$

This linear dependence seems in contrast with what derived in Sec. 1.2.1. However, is to be noted that all this argument holds for small external fields, i.e. only in the region $[T_c - \Delta T, T_c + \Delta T] \times [0, 0 + \Delta H]$, where the quadratic dependence of $H_c(T)$ is approximately linear. The reason for that is that we neglected, in $f^{(\mathrm{s})}$, the contribution due to the vector potential $\mathbf{A}$. This is reasonable if $\mathbf{B} \simeq 0$ everywhere, which is true if the superconducting currents expelling the flux occupy a small portion of the sample. This last is true if the field to be expelled is low enough.

## 2.5 SPONTANEOUS SYMMETRY BREAKING IN SUPERCONDUCTORS

One of the main features of superconductors, as understood in Ginzburg-Landau theory, is the spontaneous breaking of the natural $U(1)$ symmetry

of the complex order parameter. But what *is* spontaneous symmetry breaking? Briefly, it's what happens when the lagrangian of a system exhibits a certain number of symmetries, and the lowest energy state (the equilibrium state) does not. In the case of superconductors, a perfectly homogeneous order parameter is phase-definite, but the free-energy minimum manifold (see Fig. 2.3) is phase-symmetric!

> **A. Altland and B. D. Simons**: ***Condensed Matter Field Theory*** ([7] @ 6.3)
>
> The mechanism encountered here is one of spontaneous symmetry breaking. To understand the general principle, consider an action $S[\psi]$ with a global continuous symmetry under some transformation $g$ (not to be confused with the aforementioned coupling constant of the Bose gas): specifically, the action remains invariant under a global transformation of the fields such that $\forall i \in M\colon \psi_i \to g\psi_i$ , where $M$ is the base manifold, i.e. $S[\psi] = S[g\psi]$. The transformation is "continuous" in the sense that g takes values in some manifold, typically a group $G$.
>
> Examples: The action of a Heisenberg ferromagnet is invariant under rotation of all spins simultaneously by the same amount, $\mathbf{S}_i \to g\mathbf{S}_i$ . In this case, $g \in G = O(3)$, the three-dimensional group of rotations ($g$ not to be confused with the coupling constant of the interaction). The action of the displacement fields $\mathbf{u}$ describing elastic deformations of a solid (phonons) is invariant under simultaneous translation of all displacements $\mathbf{u}_i \to \mathbf{u}_i + \mathbf{a}$, i.e. the symmetry manifold is the d-dimensional translation group $G \sim \mathbb{R}^d$ . In the example above, we encountered a $U(1)$ symmetry under phase multiplication $\psi_0 \to \psi_0 e^{i\phi}$. This phase freedom expresses the global gauge symmetry of quantum mechanics under transformation by a phase, a point we discuss in more detail below.
>
> Now, given a theory with globally $G$ invariant action, two scenarios are conceivable: either the ground states share the invariance properties of the action or they do not [...].
>
> In spite of the undeniable existence of solids, magnets, and Bose-Einstein condensates of definite phase, the notion of a ground state that does not share the full symmetry of the theory may appear paradoxical, or at least "unnatural." For example, even if any particular ground state of the "Mexican hat" potential shown in the figure above "breaks" the rotational symmetry, should not all these states enter the partition sum with equal statistical weight, such that the net outcome of the theory is again fully symmetric?

### 2.5.1 The Goldstone boson of superconductivity

We go back to the symmetric free energy in absence of fields,

$$F\left[\Psi,\Psi^*\right] = F_0 + \int d\mathbf{x}\left[a|\Psi(\mathbf{x})|^2 + \frac{b}{2}|\Psi(\mathbf{x})|^4 + c|\boldsymbol{\nabla}\Psi(\mathbf{x})|^2\right]$$

and consider a general fluctuation of the order parameter (both in its magnitude and phase)

$$\Psi(\mathbf{x}) = \Psi_0 e^{\delta k(\mathbf{x}) + i\delta\varphi(\mathbf{x})} \quad \text{with} \quad \delta k(\mathbf{r}) \ll 1\ ,\ \delta\varphi(\mathbf{r}) \ll 1\ ,\ \Psi_0 \in \mathbb{R}$$

The reality of $\Psi_0$ can be assumed without loss of generality. We substitute in the free energy expansion, neglecting terms above the second order

$$F\left[\Psi+\delta\Psi,\Psi^*+\delta\Psi^*\right] \simeq F_0 + \int d\mathbf{x}\left[a\left|\Psi_0 e^{\delta k(\mathbf{x})+i\delta\varphi(\mathbf{x})}\right|^2 + c\left|\boldsymbol{\nabla}\Psi_0 e^{\delta k(\mathbf{x})+i\delta\varphi(\mathbf{x})}\right|^2\right]$$
$$= F_0 + \int d\mathbf{x}\Psi_0^2\left[a\left|e^{\delta k(\mathbf{x})}\right|^2 + c\left|\boldsymbol{\nabla}e^{\delta k(\mathbf{x})+i\delta\varphi(\mathbf{x})}\right|^2\right]$$

We define for simplicity $\delta\lambda(\mathbf{x}) = \delta k(\mathbf{x}) + i\delta\varphi(\mathbf{x})$. Expansion up to second order gives

$$\left|e^{\delta k(\mathbf{x})}\right|^2 \simeq \left|1+\delta k(\mathbf{x})+\frac{1}{2}\delta k^2(\mathbf{x})\right|^2 \simeq 1+2\delta k(\mathbf{x})+2\delta k^2(\mathbf{x})$$

and

$$\left|\boldsymbol{\nabla}e^{\delta\lambda(\mathbf{x})}\right|^2 \simeq |\boldsymbol{\nabla}\delta\lambda(\mathbf{x})|^2 = [\boldsymbol{\nabla}\delta k(\mathbf{x})]^2 + [\boldsymbol{\nabla}\delta\varphi(\mathbf{x})]^2$$

then, since the term with the gradient does not contribute to the rest free energy

$$F\left[\Psi+\delta\Psi,\Psi^*+\delta\Psi^*\right] \simeq F_0 + F\left[\Psi,\Psi^*\right] + 2\Psi_0^2\int d\mathbf{x}\, a\delta k(\mathbf{x})$$
$$+\int d\mathbf{x}\,\Psi_0^2\left[2a\delta k^2(\mathbf{x}) + c|\boldsymbol{\nabla}\delta k(\mathbf{x})|^2 + c|\boldsymbol{\nabla}\delta\varphi(\mathbf{x})|^2\right]$$

The linear term can be safely neglected, since normal thermal fluctuations are at zero average over the volume. Passing in Fourier transform, and collecting in $F^\star$ all the terms in the first line,

$$F\left[\Psi+\delta\Psi,\Psi^*+\delta\Psi^*\right] \simeq F^\star\left[\Psi,\Psi^*\right]$$
$$+\Psi_0^2\sum_{\mathbf{q}}\left[c|\mathbf{q}|^2-2a\right]\delta k(\mathbf{q})\delta k(-\mathbf{q}) + \Psi_0^2\sum_{\mathbf{q}} c|\mathbf{q}|^2\delta\varphi(\mathbf{q})\delta\varphi(-\mathbf{q})$$

We use the notation

$$\delta K_{\mathbf{q}} = \Psi_0\delta k(\mathbf{q}) \qquad\qquad \delta\Phi_{\mathbf{q}} = \Psi_0\delta\varphi(\mathbf{q})$$

Then the above expansion reads

$$\delta F \simeq \sum_{\mathbf{q}}\left[c|\mathbf{q}|^2-2a\right]\delta K_{\mathbf{q}}\delta K_{-\mathbf{q}} + \sum_{\mathbf{q}} c|\mathbf{q}|^2\delta\Phi_{\mathbf{q}}\delta\Phi_{-\mathbf{q}}$$

This quadratic expression for the free energy allows for interpretation of $K$ and $\Phi$ as normal modes of the system. For a normal particle, energy contributions are of the form

$$\delta F = \sum_{\mathbf{q}}\varepsilon_{\mathbf{q}}n_{\mathbf{q}} = \sum_{\mathbf{q}}\left(\kappa_{\mathbf{q}}+m\right)n_{\mathbf{q}}$$

with $\kappa_{\mathbf{q}}$ the kinetic contribution, vanishing for $\mathbf{q}\to\mathbf{0}$. Looking at the expression for $\delta F$, we see that of the $K$ mode is "massive" since its coefficient does not vanish in the long wavelength limit (moreover: it's positive, since $a<0$ for $T<T_c$); the $\Phi$ mode is "massless", since it does the opposite. This mode is strongly associated with the continuous $U(1)$ symmetry breaking, and is the bosonic manifestation of its existence: it is the Nambu-Goldstone boson for the $U(1)$ symmetry of superconductivity.

**K. Huang**: ***Statistical Mechanics, 2nd Ed*** ([6] @ 16.6)

We must comment on that most remarkable fact, which underlies Landau's original conception of the order parameter, that macroscopic systems generally have a lesser degree of symmetry manifested at low temperatures than at high temperatures. The symmetry manifested at high temperatures is usually a property if the system's microscopic Hamiltonian. As such, it cannot cease to exist, even when it appears to be violated. The question is, where does it go?

For example, the microscopic Hamiltonian of a ferromagnet is rotationally invariant. Lowering the temperature of the system surely does not change that. What is changed, however, is the mode in which the symmetry is expressed. One might find it natural to assume (in the best tradition of Aristotelian logic) that in the most perfect expression of symmetry, the ground state should be invariant under the symmetry operation. Nature, alas, is not perfect. In most instances, the system possesses many equivalent ground states that transform into one another under the symmetry operation. But, since the system can actually exists in only one of these states, the symmetry appears to be broken. This phenomenon, that the ground state of the system does not possess the symmetry of the Hamiltonian, is called "spontaneous symmetry breaking".

We have encountered the simplest manifestation of spontaneous symmetry breaking in the two-dimensional Ising model. The total energy is invariant under a simultaneous sign change of all the spins. Yet, the lowest-energy configuration is that in which all spins are aligned. In this case the symmetry is expressed through the fact that the spins could also have aligned themselves in the opposite direction [...].

When a continuous symmetry is spontaneously broken in a quantum mechanical system, interesting consequences follow. In this case, there is non-countable infinity of ground-states (with energy to be taken as zero) orthogonal to one another. The system must choose one of these. As a consequence of the degeneracy, there emerges a type of excited state in which the local ground states changes very gradually over space, so as to form a "wave" of very long wavelength. Such a state is orthogonal to any one ground state, with an energy approaching zero when the wavelength approaches infinity. This is called a "Goldstone excitation". The underlying symmetry is said to be realized in the "Goldstone mode" [...].

A new twist occurs when the system is coupled to to a zero-mass vector field such as the photon field. In this case the Goldstone mode gives way to the "Higgs mode", wherein the "photon" acquires a mass dynamically, and the would-be Goldstone excitation becomes the longitudinal degree of freedom of the massive "photon". An example of this is the Meissner effect in superconductivity, in which the photon mass corresponds to the inverse penetration depth.

We now pass to the most important section of this chapter, which treats one of the most beautiful mechanisms in Physics (on a *very* introductory level) and makes superconductivity a really unique phenomenon.

### 2.5.2 The Anderson-Higgs mechanism

Suppose to switch on the field. On a purely euristic level, even with a homogeneous phase-definite field we are now changing actively the lagrangian of the problem, *explicitly* breaking the symmetry. We know that for a given phase-definite symmetry breaking the field $\Psi$ accommodates accordingly. Thus we expect the minimum manifold to collapse into a single point, since there are no more symmetries to be broken. Then the Nambu-Goldstone mode should not arise anymore. This is exactly what happens: the Goldstone mode of superconductivity is "dressed" by interactions, and becomes non-vanishing in the long-wavelength limit. It gets *massive*... just like matter does thanks to the Higgs field. turns The mechanism works as follows. Suppose to turn on the electromagnetic field $\mathbf{A}$,

$$F\left[\Psi,\Psi^*,\mathbf{A}\right] = F_0 + \int d\mathbf{x}\left[a|\Psi(\mathbf{x})|^2 + \frac{b}{2}|\Psi(\mathbf{x})|^4 + \frac{c}{\hbar^2}|-i\hbar\boldsymbol{\nabla}\Psi(\mathbf{x}) - q\mathbf{A}\Psi(\mathbf{x})|^2\right]$$

We will proceed exactly like in the previous section, performing a fluctuations analysis, collecting into $F^\star$ everything below the second order and neglecting everything above. Once again we consider a small fluctuation over the mean-field solution

$$\Psi(\mathbf{x}) = \Psi_0 e^{\delta\lambda(\mathbf{x})}$$

As always $\Psi_0 \in \mathbb{R}$. Nothing changes for the $a$ term, while for the $c$ term

$$\begin{aligned}\left|\left[-i\hbar\boldsymbol{\nabla} - q\mathbf{A}\right]\Psi_0 e^{\delta\lambda(\mathbf{x})}\right|^2 &= \Psi_0^2\left|e^{\delta\lambda(\mathbf{x})}\right|^2\left|-i\hbar\boldsymbol{\nabla}\delta\lambda(\mathbf{x}) - q\mathbf{A}\right|^2 \\ &= \Psi_0^2 e^{2\delta k(\mathbf{x})}\left[|\hbar\boldsymbol{\nabla}\delta\varphi(\mathbf{x}) - q\mathbf{A}|^2 + |\hbar\boldsymbol{\nabla}\delta k(\mathbf{x})|^2\right]\end{aligned}$$

The astonishing result, here, is that a gauge transformation can be made upon the vector potential without changing any physical result, and making the phase fluctuations field **disappear**. Notice that, being the gauge-transformation of wavefunctions essentially a multiplication by a phase factor, and since here the order parameter enters everywhere only through its modulus, we can neglect such gauge transformation. The said gauge choice (indicated by the $\star$ superscript) is the following

$$\mathbf{A}^\star(\mathbf{x}) = \mathbf{A}(\mathbf{x}) - \frac{q}{\hbar}\boldsymbol{\nabla}\varphi(\mathbf{x})$$

which finally gives

$$F\left[\Psi+\delta\Psi,\Psi^*+\delta\Psi^*,\mathbf{A}^\star\right] \simeq F^\star\left[\Psi,\Psi^*\right] + \int d\mathbf{x}\Psi_0^2\left\{2a\delta k^2(\mathbf{x}) + c\left[|\boldsymbol{\nabla}\delta k(\mathbf{x})|^2 + |\mathbf{A}^\star|^2\right]\right\}$$

Notice that this expansion is somehow incomplete, since we hid inside $F^\star$ a term linear in $\delta k(\mathbf{x})$ and multiplied by $|\mathbf{A}^\star|^2$, which is in general non-trivial and represents interactions between the gauge field and the massive mode

of the theory. However, the main feature of this formula is that, if we take its Fourier transform,

$$\delta F \simeq \sum_{\mathbf{q}} \left[ c|\mathbf{q}|^2 - 2a \right] \delta K_{\mathbf{q}} \delta K_{-\mathbf{q}} + \sum_{\mathbf{q}} c\Psi_0^2 |\mathbf{A}^\star|_{\mathbf{q}} |\mathbf{A}^\star|_{-\mathbf{q}}$$

now **none** of these modes is massive! In fact, taking the long-wavelength limit, all their dispersions remain finite. This is remarkable, *since one of these modes is a photon*. Quite impressive, isn't it? The final majestic detail to notice before moving on to the next chapter, is that the photon "mass" is proportional to $\Psi_0^2 \sim n_s$, and we know that $\lambda^{-2} \sim n_s$ as well. Thus the inverse penetration depth is connected to the photon mass: the smallest the path electromagnetic fields travel through the superconductor, the greater the photon mass.

# 3 CONVENTIONAL SUPERCONDUCTORS



We now have two length scales: one is the penetration depth $\lambda$, arising from general London theory of superconductivity: it measures the length scale over which electromagnetic field are suppressed in a superconductor. The other is the coherence length $\xi$, which arose from Ginzburg-Landau theory: it measures the length over which correlations of the order field at different points are suppressed. Two situations are possible, both $\xi \gg \lambda$, or $\xi \ll \lambda$ (neglect the one in which they are comparable). On this difference "type I" and "type II" superconductors are distinguished. These two classes form **conventional superconductors**.

## 3.1 SUPERCONDUCTORS OF TYPE I AND TYPE II

Consider a superconducting sample, homogeneous in the $y$ and $z$ directions, with the interface with vacuum placed at $x = 0$, as represented in Fig. 3.1. The $x > 0$ is in superconducting state, the $x < 0$ region is in normal state. A homogeneous magnetic field $\mathbf{H} = H\bar{\mathbf{z}}$ is applied everywhere.

Evidently the order parameter $\Psi(\mathbf{x})$ can only depend on $x$, and the magnetic flux density $\mathbf{B}(\mathbf{x})$ only has the $\bar{\mathbf{z}}$ component and only depends on $x$. In

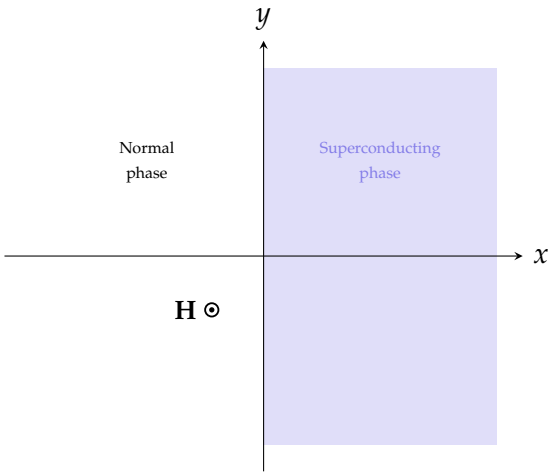



**Figure 3.1:** A superconducting sample has the interface with vacuum placed at $x = 0$, as described in Sec. 3.1. A uniform magnetic field $\mathbf{H} = H\bar{\mathbf{z}}$ is applied everywhere.

the region $x \gtrsim 0$, the behavior of the superconductor depends on its type. The general boundary conditions hold anyways

$$\begin{array}{lll} x \leq 0 & \Psi(x) = 0 & B(x) = \mu_0 H \\ x \to +\infty & \Psi(x) = \Psi_0 & B(x) = 0 \end{array}$$

where $\Psi_0^2 = -a/b$. A good gauge choice for the vector potential (also, only dependent on $x$ due to symmetry) is

$$\mathbf{A}(x) = A(x)\bar{\mathbf{y}} \quad \Longrightarrow \quad B(x) = \bar{\mathbf{z}} \cdot (\boldsymbol{\nabla} \times \mathbf{A}) = \left( \frac{\partial}{\partial x} \bar{\mathbf{y}} \cdot \mathbf{A}(x) \right) = \frac{\partial A(x)}{\partial x}$$

Such a choice is coherent with the Coulomb-London gauge $\boldsymbol{\nabla} \cdot \mathbf{A} = 0$. Also, since $\Psi(x)$ changes along the $x$ direction,

$$\mathbf{A} \cdot \boldsymbol{\nabla} \Psi = 0$$

Take now the first Ginzburg-Landau equation, Eq. (2.10). Expanding the term $\left(-i\hbar\boldsymbol{\nabla} - q\mathbf{A}\right)^2 \Psi(x)$ it can be recast in the form

$$\begin{aligned} a\Psi(x) + b|\Psi(x)|^2\Psi(x) + \frac{c}{\hbar^2}\left(-\hbar^2\nabla^2 + q^2A^2(x)\right)\Psi(x) = 0 \\ -c\nabla^2\Psi(x) + \left[a + \frac{c}{\hbar^2/q^2}A^2(x)\right]\Psi(x) + b|\Psi(x)|^2\Psi(x) = 0 \end{aligned}$$

We now define the magnetic flux unity

$$\Phi_0 = \frac{h}{q} = 2\pi\frac{\hbar}{q}$$

and use the known result for the correlation length,

$$\xi^{-2} = \frac{a + 3b\Psi_0^2}{c} = -\frac{2a}{c} \quad \Longrightarrow \quad -c = 2a\xi^2$$

then, dividing everything by $a$ and substituting the laplacian with $\partial_x^2$ being the order field only dependent on $x$

$$2\xi^2\frac{\partial^2}{\partial x^2}\Psi(x)+\left[1-2\xi^2\left(\frac{2\pi}{\Phi_0}\right)^2A^2(x)\right]\Psi(x)-\frac{|\Psi(x)|^2}{\Psi_0^2}\Psi(x)=0$$

Finally, defining the **reduced order parameter** $f(x)\equiv\Psi(x)/\Psi_0$, and assuming it real without loss of generality we get

$$2\xi^2\frac{\partial^2}{\partial x^2}f(x)+\left[1-2\xi^2\left(\frac{2\pi}{\Phi_0}\right)^2A^2(x)\right]f(x)-f^3(x)=0 \tag{3.1}$$

A **very important specification on notation** is here needed, in order to avoid the terrible confusion the author has gone through when writing these notes. In standard notation, the symbol $\xi$ is used to indicate a quantity closely related to our $\xi$, in particular

$$\xi^{\text{Ⓢ}}=\sqrt{2}\xi$$

where $\xi^{\text{Ⓢ}}$ is the "standard notation xi". The reason is that our $\xi$ comes from statistical grounds and indicates the correlation length of the Ising model (extended to a complex order parameter, as seen in Chap. 2); instead, $\xi^{\text{Ⓢ}}$ is defined as the typical length of the differential equation (3.1), thus substituting $2\xi^2=\xi^{\text{Ⓢ}2}$. Of course we could have adapted to the standard notation, however the correlation definition seems just more physically fundamental than the standard definition. This specification is of great importance: in order to avoid useless entropy increasing, we will come back on this later when needed.

For a real order parameter the second Ginzburg-Landau equation, Eq. (2.11) reads

$$-\frac{q^2}{m}|\Psi(x)|^2\mathbf{A}(x)=\mathbf{J}(x)=\frac{1}{\mu_0}\boldsymbol{\nabla}\times\mathbf{B}(x)=-\frac{1}{\mu_0}\nabla^2\mathbf{A}(x)$$

again using $\boldsymbol{\nabla}\cdot\mathbf{A}=0$. Then, since

$$\lambda^{-2}=\frac{\mu_0}{\Lambda}=\frac{\mu_0q^2n_s}{m}$$

and $n_s=\Psi_0^2$

$$\left[\frac{\partial^2}{\partial x^2}-\lambda^{-2}f^2(x)\right]A(x)=0 \tag{3.2}$$

To solve these Eq. (3.1) and Eq. (3.2) we need boundary conditions. In the bulk, $f=1$ and $A=0$.

### 3.1.1 Type I

Superconductors of the first type are characterized by $\xi\gg\lambda$. Then we select the region $x\sim\xi$, where $f=1$ and $A=0$. The vector potential is given by

$$A(x)\simeq -Hx\theta(-x)$$

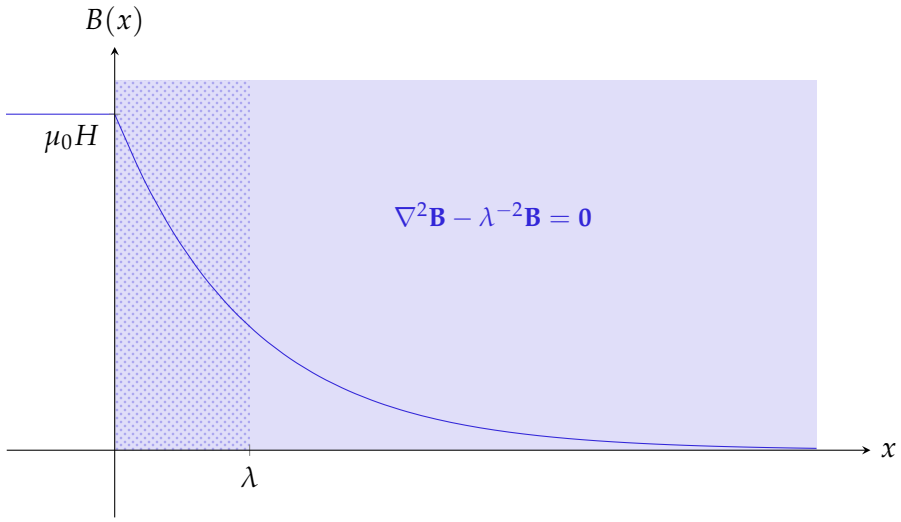



**Figure 3.2:** Inside a superconducting sample the magnetic flux density decays on a distance $\lambda$. The dotted region represents the interface in which the field is able to penetrate. The ratio $\xi/\lambda$ defines the class of the conventional superconductor.

By Eq. (3.1), nearly everywhere in the sample (see Fig. 3.3a)

$$\left[2\xi^2\frac{\partial^2}{\partial x^2}+1\right]f(x)-f^3(x)=0$$

Of course in any case $A$ vanishes for $x\to+\infty$ leading to the above equation for $f$, however in this situation such equation is (approximately) true in the whole sample. By the powerful means of Wolfram Alpha®, it turns out

$$f(x)=\tanh\left(\frac{x}{2\xi}\right)$$

We see here that $\xi$ is the length over which the order parameter rises from $\Psi=0$ to $\Psi=\Psi_0$, the mean-field solution.

### 3.1.2 Type II

Superconductors of the second type are a little more interesting. They are characterized by $\xi\ll\lambda$, meaning that in the interface (the dotted region of Fig. 3.2) the order parameter is already saturated, $\Psi=\Psi_0$, but the fields are still non-vanishing

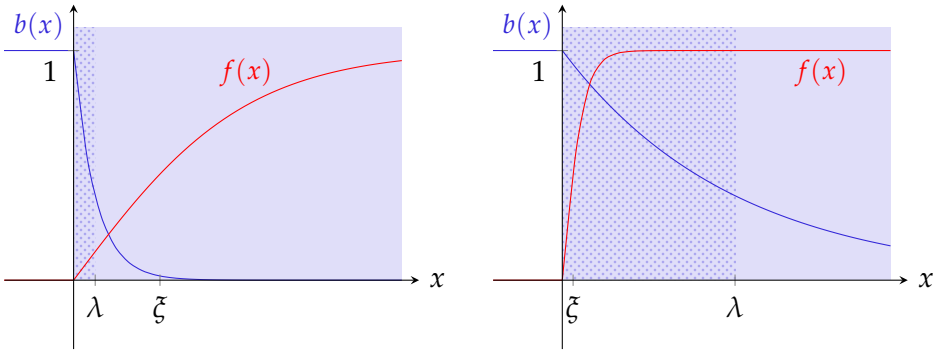
$$\frac{\partial^2 A}{\partial x^2}-\frac{A}{\lambda^2}=0\qquad(\text{for }\xi<x<\lambda)$$

then the spatial dependence of $A$ is smoother

$$A(x)=\mu_0 H\left[(x-\lambda)\theta(-x)-\lambda e^{-x/\lambda}\theta(x)\right]$$

(the $\lambda$ term was included for continuity) and the flux density is

$$B(x)=\mu_0 H\left[\theta(-x)+e^{-x/\lambda}\theta(x)\right]$$




**(a)** Type I superconductor.

**(b)** Type II superconductor.

**Figure 3.3:** The two classes of conventional superconductors are distinguished by the value of the ratio $\xi/\lambda$. In Fig. 3.3a a type I superconductor is represented, with $b(x) \equiv B(x)/\mu_0 H$ and $f(x) = \Psi(x)/\Psi_0$ the reduced order field. In this case $\lambda \ll \xi$ and the order field saturates in a field-free region. In Fig. 3.3b a type II superconductor is represented. In this case, being $\lambda \gg \xi$, the order field saturates in a region where fields penetrate. In both regions, the interface is dotted for clarity.

as represented in Fig. 3.3b. In this case the solution for the order parameter $f(x)$ is a little more complex, but being the fields decreasing on a length scale much larger, locally they can be neglected and a step-like solution for $f$ works as a suitable approximation.

## 3.2 MAGNETIC PROPERTIES OF SUPERCONDUCTORS

The process of creating a "hole" inside a superconducting sample, meaning a region of space where the material is in normal state or physically a hole, is of great importance. It is characteristic of type II superconductors to have a particular nucleation behavior, with the formation of a triangular lattice of "holes". We start here by analyzing the properties of holes in superconducting materials.

### 3.2.1 Flux quantization

Consider the bulk of a conventional superconductor, subject to an external magnetic field $\mathbf{H} = H\bar{\mathbf{z}}$. At point $(0,0)$ of the $xy$ plane perpendicular to the field, a hole is present. For the sake of simplicity we consider a round hole of radius $r_0$, as in Fig. 3.4. Due to $z$ symmetry, we may work in cylindrical coordinates $(r,\theta)$. Then the order parameter has spatial dependence

$$\Psi(r,\theta) = a(r,\theta)e^{ip(r,\theta)}$$

with $a \in \mathbb{R}$ the amplitude and $p \in \mathbb{R}$ the phase. Being $n_s = |\Psi|^2 = a^2$ cylindrical symmetric, $a$ cannot depend on $\theta$.

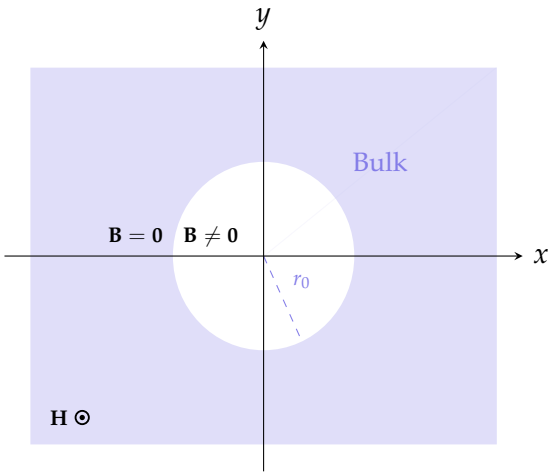



**Figure 3.4:** The hole of radius $r_0$ in the superconductor bulk described in Sec. 3.2.1. The uniform magnetic field $\mathbf{H} = H\bar{\mathbf{z}}$ is applied everywhere, and in the bulk is perfectly screened. Magnetic flux density $\mathbf{B}$ is able to penetrate in the hole.

Consider the flux density in the hole: $B$ cannot depend on $\theta$ due to symmetry. Having only the $z$ component, this means that $\mathbf{A}$ only has the $\bar{\boldsymbol{\theta}}$ component and depends only on $r$. Therefore, since $\boldsymbol{\nabla} \times \mathbf{B} = \mu_0 \mathbf{J}$,

$$\mu_0 \mathbf{J} = \boldsymbol{\nabla} \times \bar{\mathbf{z}} B(r) = -\bar{\boldsymbol{\theta}} \frac{\partial}{\partial z} B(r) = \bar{\boldsymbol{\theta}} J(r)$$

Rather intuitively, current flows in the $\theta$ direction to screen the field in the superconductor. We expect $J(r)$ to decrease rapidly for $r > r_0$. Now, take the second Ginzburg-Landau equation, Eq. (2.11),

$$\begin{aligned}
\mathbf{J}(r) &= \frac{\hbar q}{2im} \left[\Psi^*(r,\theta)\boldsymbol{\nabla}\Psi(r,\theta) - \Psi(r,\theta)\boldsymbol{\nabla}\Psi^*(r,\theta)\right] - \frac{q^2}{m}|\Psi(r,\theta)|^2 \mathbf{A}(r) \\
&= \frac{\hbar q}{m} a^2(r) \boldsymbol{\nabla} p(r,\theta) - \frac{q^2}{m} a^2(r) \mathbf{A}(r) \\
&= \frac{q}{m} n_s(r) \left[\hbar \boldsymbol{\nabla} p(r,\theta) - q\mathbf{A}(r)\right]
\end{aligned}$$

because the gradients of the amplitude cancel out. Since both $\mathbf{J}$ and $\mathbf{A}$ only have the angular component, it must be $\partial_r p = 0$, which means that **the phase only depends on the angle $\boldsymbol{\theta}$**, $p(r,\theta) = p(\theta)$. We must ensure monodromy

$$\Psi(r,\theta) = a(r) e^{ip(\theta)} \overset{!}{=} a(r) e^{ip(\theta+2\pi)} \quad \Longrightarrow \quad p(\theta+2\pi) - p(\theta) = 2n\pi$$

with $n \in \mathbb{R}$.

We now wish to integrate the current on the path $\partial\Sigma$ represented in Fig. 3.5, in order to extract the magnetic flux through the hole and its vicinity. For $r \gg r_0$, deep in the bulk, we expect $\Psi(r,\theta) = \Psi_0$ (the mean-field solution): the region $\Sigma$, delimited by the path $\partial\Sigma$, is chosen large enough to

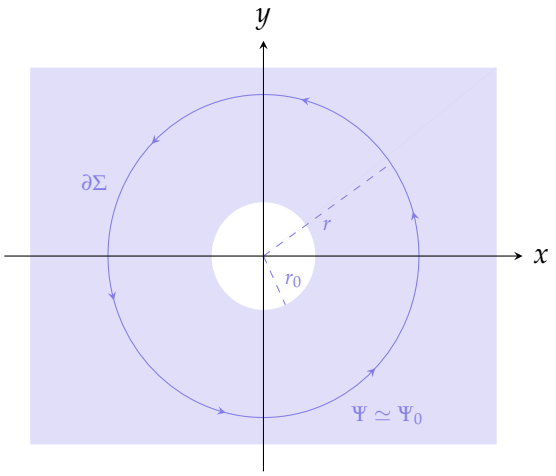



**Figure 3.5:** The deep-bulk integration path $\partial\Sigma$ around the hole described in Sec. 3.2.1.

enclose **all** magnetic flux penetrating in the sample (we expect the magnetic flux density not to drop to zero for $r \simeq r_0^+$). It turns out

$$\begin{aligned}\oint_{\partial\Sigma} d\boldsymbol{\ell}\cdot\mathbf{J} &= \oint_{\partial\Sigma} d\boldsymbol{\ell}\cdot\left[\frac{q}{m}n_s(r)\left[\hbar\boldsymbol{\nabla}p(\theta)-q\mathbf{A}(r)\right]\right]\\ &= \frac{\hbar q}{m}\Psi_0^2\oint_{\partial\Sigma} d\boldsymbol{\ell}\cdot\boldsymbol{\nabla}p(\theta)-\frac{q^2}{m}\Psi_0^2\oint_{\Gamma} d\boldsymbol{\ell}\cdot\mathbf{A}(r)\\ &= \frac{q}{m}\Psi_0^2\int_0^{2\pi} d\theta\frac{\partial p}{\partial\theta}-\frac{q^2}{m}\Psi_0^2\int_{\Sigma} d\boldsymbol{\Sigma}\cdot\mathbf{B}(r)\end{aligned}$$

where we used the Stokes theorem. Now, for $r$ deep in the bulk the current is expected to vanish (supercurrents flow near interfaces). Then

$$\Phi_\Sigma(B)=\frac{\hbar}{q}\left[p(2\pi)-p(0)\right]=n\Phi_0 \qquad \text{with } n\in\mathbb{R} \text{ and } \Phi_0=\frac{h}{q}$$

We discover that **when a hole is pierced in a superconductor, the amount of magnetic flux passing through it is quantized in units of $\Phi_0$**. The same results holds true for different shapes of the hole, since the spatial dependencies of $p$ and $\mathbf{J}$ far from the hole are kind of independent of the specific local structure of the impurity.

### 3.2.2 Nucleation field

Given some material in the normal state ($\Psi = 0$), on an empirical and thermodynamic level we know that for low enough fields superconducting transition can occur. We now want to find the largest field for which a finite order field $\Psi \neq 0$ can arise. Such a process is called **nucleation**.

Take the first Ginzburg-Landau equation, Eq. (2.10). We want to study the transition, so we suppose $\Psi$ to be low enough to ignore the $b$ term, which is cubic. We get

$$\begin{aligned} a\Psi(\mathbf{x}) + \frac{c}{\hbar^2}\left[-i\hbar\boldsymbol{\nabla} - q\mathbf{A}\right]^2\Psi(\mathbf{x}) &\simeq 0 \\ \left[-i\boldsymbol{\nabla} - \frac{q}{\hbar}\mathbf{A}\right]^2\Psi(\mathbf{x}) &\simeq -\frac{a}{c}\Psi(\mathbf{x}) \\ \left[-i\boldsymbol{\nabla} - \frac{2\pi}{\Phi_0}\mathbf{A}\right]^2\Psi(\mathbf{x}) &\simeq \frac{1}{2\xi^2}\Psi(\mathbf{x}) \end{aligned}$$

since $-c/a = 2\xi^2$. This is a Schrödinger equation for a particle in a magnetic field, and gives rise to **Landau levels**. The energy levels of such free system are

$$E_n(k_z) = \hbar\omega_B\left(n + \frac{1}{2}\right) + \frac{\hbar^2 k_z^2}{2m} \quad \text{with} \quad \omega_B = \frac{qB}{m}$$

Since we divided everything by $c = \hbar^2/2m$, we have

$$\begin{aligned} \frac{1}{2\xi^2} &= \frac{2m}{\hbar^2}\hbar\frac{qB}{m}\left(n + \frac{1}{2}\right) + k_z^2 \\ \frac{2\pi}{\Phi_0}B\left(n + \frac{1}{2}\right) &= \frac{1}{2}\left(\frac{1}{2\xi^2} - k_z^2\right) \end{aligned}$$

Evidently the right side is independent of the field. The largest field satisfying this equation is given by $n = 0$, $k_z = 0$,

$$(@n = 0, k_z = 0) \qquad B^\star = \frac{\Phi_0}{\pi\left(2\xi\right)^2}$$

Lowering the external field, symmetry breaking occurs **when a circle of radius $2\xi$ encloses exactly one quantum of flux**. Since for slightly more intense fields the sample is in the normal state, such field $B^\star$ defines a critical external field, usually indicated as $H_{c2}$ and called **upper critical field**

$$H_{c2} = \frac{B^\star}{\mu_0}$$

Recall now:

$$\xi^2 = -\frac{c}{2a} = -\frac{\hbar^2}{4ma} \qquad \lambda^2 = \frac{\Lambda}{\mu_0} = \frac{m}{\mu_0 q^2 \Psi_0^2} \qquad \Psi_0 = \sqrt{\frac{-a}{b}}$$

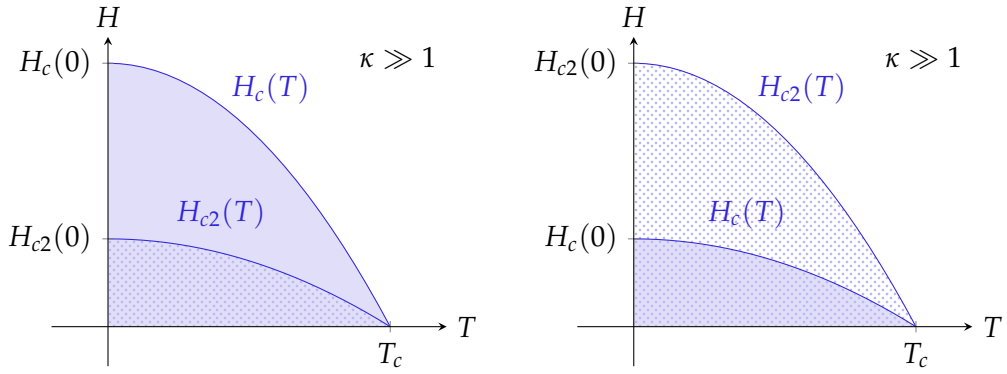
Then

$$\lambda^2\xi^2 = \frac{m}{\mu_0 q^2}\frac{b}{-a}\frac{\hbar^2}{4m(-a)} = \frac{1}{4\mu_0}\frac{\hbar^2}{q^2}\frac{b}{a^2}$$

From thermodynamic arguments we found out $\mu_0 H_c^2 = a^2/b$. Then

$$\mu_0 H_c = \frac{1}{\lambda\xi}\frac{\Phi_0}{4\pi} = \frac{\xi}{\lambda}\frac{\Phi_0}{\pi\left(2\xi\right)^2} \quad \Longrightarrow \quad H_{c2} = \kappa H_c \quad \text{with} \quad \kappa \equiv \frac{\lambda}{\xi}$$

Remember: we are not using the standard notation, in which $\kappa^{\text{Ⓢ}} = \lambda/\xi^{\text{Ⓢ}}$. Again, the superscript Ⓢ here indicates how the symbol "$\kappa$" is intended in standard notation. With such a choice of symbols, clearly $H_{c2} = \sqrt{2}\kappa^{\text{Ⓢ}} H_c$. Now:




**(a)** Type I superconductor.

**(b)** Type II superconductor.

**Figure 3.6:** Phase diagrams for the two types of conventional superconductors described in Sec. 3.2.2. Inside the shaded region the **uniform** superconducting phase presents a energy gain; inside the dotted region symmetry can be broken and superconductivity can emerge. Regions which are both shaded and dotted are those in which the sample presents the mean-field uniform superconducting solution.

1. For type I superconductors $\kappa \ll 1$. Then the upper critical field lies below the conventional critical field. The phase diagram is the one represented in Fig. 3.6a. Inside the shaded region in figure, thermodynamics indicates a free energy gain in forming a uniform $\Psi(x) = \Psi_0$ superconducting phase. However nucleation can occur only when one quantum of flux is enclosed by a circle of radius $2\xi$. Then, in the region

$$H_{c2}(T) \leq H \leq H_c(T)$$

there is not enough magnetix flux to break symmetry, and the sample remains in a **uniform, normal and supercooled state**. Such region is the one shaded, but not dotted, in Fig. 3.6a. For $H < H_{c2}$ nucleation is possible, and since $H < H_c$ the solution is the uniform one. For this reason the transition is a one of the first order: $\Psi$ jumps from 0 to $\Psi_0$ in a discontinuous manner. In standard notation, type I superconductors are characterized by $\kappa^{\text{Ⓢ}} \ll 1/\sqrt{2}$.

2. For type II superconductors $\kappa \gg 1$. Symmetry breaking is **locally** possible in the region

$$H_c(T) \leq H \leq H_{c2}(T)$$

which is the one dotted but not shaded in Fig. 3.6b. This means that even if the uniform solution is not thermodynamically favourite everywhere in the sample, local nucleation can in principle occur. As we show in the next section, it actually happens. We expect such a region to be in a **mixed normal-superconducting phase**. For $H < H_c$ the mean-field solution arises. Then we expect the transition to be one of the second order, since by physical intuition $\Psi$ is expected to vary from 0 to $\Psi_0$ continuously when lowering the external field. In standard notation, type II superconductors are those where $\kappa^{\text{Ⓢ}} \gg 1/\sqrt{2}$.

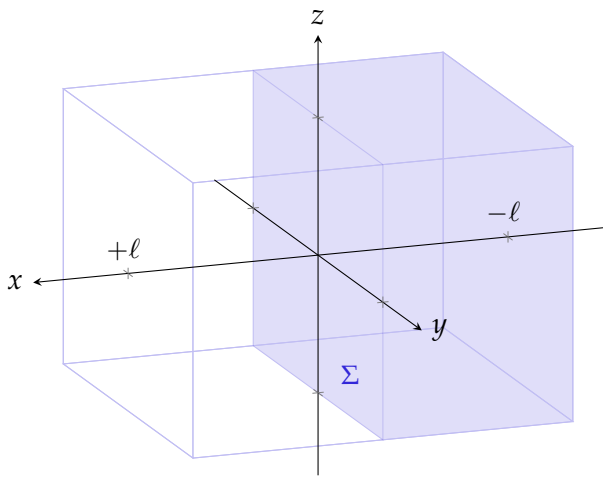



**Figure 3.7:** The volume element described in Sec. 3.2.3. The surface $\Sigma$ is assumed to be part of the interface, and the volume is constructed by extrusion of of a quantity $\pm\ell$. In the mixed phase, the shaded volume is in the superconducting phase.

Notice that in the mixed phase of type II superconductors it is to be expected an incomplete magnetic flux expulsion. The next section is devoted to justifying this expectation.

### 3.2.3 The mixed phase

A physical state at equilibrium is realized if it minimizes the free energy. We now show that in superconductors of type II in the mixed phase described in the above section, $H_c(T) < H < H_{c2}(T)$ at some fixed sub-critical temperature $T$, forming an interface between a normal and a superconducting region is thermodynamically favourite. Such phase is also known as **Shubnikov phase**.

Since $H < H_{c2}$ nucleation is possible. It can occur homogeneously or it can form normal regions as well as superconducting regions, thus interfaces. We consider one of those "potential" interfaces, which is a continuous $2D$ manifold in $3D$ space (not necessarily closed) and focus on a small cubic volume element of surface $\Sigma$ and depth $2\ell$ on it, as the one depicted in Fig. 3.7. The reference axis are chosen in order to have half of the volume "before" the interface and the other half "after". Of course, we are assuming here the interface to be locally approximated by a flat surface at sufficiently high length scales in order to neglect the atomic structure of the material. We assume $\ell \gg \lambda, \xi$.

We need to compute the free energy difference for creating a superconducting configuration with respect to the completely normal phase. For the homogeneous mean-field solution such quantity is simply

$$\Delta F^{(\mathrm{h})}(T,H) = -2\ell\Sigma\Delta f(T) \qquad \text{with} \qquad \Delta f(T) = \frac{\mu_0 H_c^2(T)}{2}$$

where $\Delta f(T)$ is the condensation energy, as derived in Chap. 1. Instead, considering a normal phase for $0 < x < \ell$ and a superconducting phase for

$-\ell < x < 0$ (see Fig. 3.7), we need to replace half of the above free energy with a purely magnetic contribution

$$\Delta F^{(\mathrm{i})}(T,H) = -\mu_0 \ell \Sigma \frac{H_c^2(T)}{2} - \mu_0 \ell \Sigma \frac{H^2}{2} + \delta F$$

plus a correction $\delta F$ accounting for the fact that $\Psi$ and $B$ vary rapidly but somehow continuously across the interface. It can be estimated roughly thinking that fields penetrate the interface on a length $\lambda$ and the order parameter $\Psi$ is different from $\Psi_0$ on a length $\xi$. Then, approximating both behaviors as step-like,

$$\delta F \simeq \mu_0 \xi \Sigma \frac{H_c^2(T)}{2} - \mu_0 \lambda \Sigma \frac{H^2}{2}$$

The positive sign of the first term is due to the fact that we need to **remove** a volume $\xi\Sigma$ from the superconducting region. Thus

$$\Delta F^{(\mathrm{i})}(T,H) = -\mu_0 \left(\ell - \xi\right) \Sigma \frac{H_c^2(T)}{2} - \mu_0 \left(\ell + \lambda\right) \Sigma \frac{H^2}{2}$$

The interface is produced if $\Delta F^{(\mathrm{h})} - \Delta F^{(\mathrm{i})} > 0$. Such difference is given by

$$\Delta F^{(\mathrm{h})} - \Delta F^{(\mathrm{i})} = -\mu_0 \left(\ell + \xi\right) \Sigma \frac{H_c^2}{2} + \mu_0 \left(\ell + \lambda\right) \Sigma \frac{H^2}{2}$$

Now we consider $H \simeq H_c(T)$: the first terms in parenthesis cancel out, while we're left with

$$\Delta F^{(\mathrm{h})} - \Delta F^{(\mathrm{i})} \simeq \mu_0 \left(\lambda - \xi\right) \Sigma \frac{H_c^2}{2} > 0$$

since $\lambda \gg \xi$ in type II superconductors. Then we conclude that when $\kappa \gg 1$, we must expect the system to produce a mixed phase with interfaces between normal and superconducting phase, being **locally** thermodynamically convenient. At this stage we cannot infer anything about the structure of such interfaces. Notice that the same exact argument can be used for type I superconductors, where $\kappa \ll 1$: as expected the homogeneous solution is globally and locally convenient and no interfaces are expected.

> **M. Tinkham**: ***Introduction to Superconductivity*** ([5] @ 4.3)
>
> Special consideration is required to show that the exact crossover from positive to negative surface energy occurs for $\kappa = 1/\sqrt{2}$. This was found by numerical integration by Ginzburg and Landau in their original paper, and they already anticipated that a conventional laminar intermediate state would only occur for lower values of $\kappa$. But until Abrikosov's path-breaking paper, no one fully anticipated the radically different behavior that resulted from the negative surface energy at higher values of $\kappa$. In one stroke, his paper created the study of type II superconductivity, the name he gave to materials with $\kappa > 1/\sqrt{2}$. Since this is the subject of the next chapter, for the present we shall simply remark that the negative surface energy causes the flux-bearing (normal) regions to subdivide until a quantum limit is reached in which each quantum of flux $\Phi_0 = hc/2e$

> passes through the sample as a distinct flux tube. These flux tubes form a regular array, and $\Psi \to 0$ along the axis of each one.

Notice, in the above citation, that standard notation is used ($\kappa^{\textcircled{S}}$) and the charge unity is $q = 2e$, which anticipates the fact that the elementary charged element in superconductors is **an electron pair**.

## 3.3 THE ABRIKOSOV VORTEX IMPURITY

We now turn to one of the most fascinating topics in the physics of conventional superconductivity. In 1957 the breakthrough article of the Soviet Physicist Alexei Abrikosov "On the Magnetic properties of superconductors of the second group" [9] laid the foundation for physically interpreting a very special phase of matter in certain temperature-field configurations of type II superconductors. As it turns out, in such phase magnetic flux can penetrate the sample creating a topological excitation, a robust configuration of localized flux density (which consists of exactly one quantum of flux $\Phi_0$), which defines a small core where the sample is in the normal state, and superconducting currents flowing in a vortex around it. This configuration is a quasiparticle excitation of the system.

Alexei Abrikosov was awarded the Nobel Prize in Physics in 2003, altogether with Vitaly Ginzburg and Anthony James Leggett. His theory explained how a sample could maintain a superconducting behavior for fields higher than the critical field $H_c(T)$, and today is of fundamental importance in a broad range of science and technology, from particle accelerators and fusion reactors to topological materials as well as medical instrumentation using nuclear magnetic resonance.

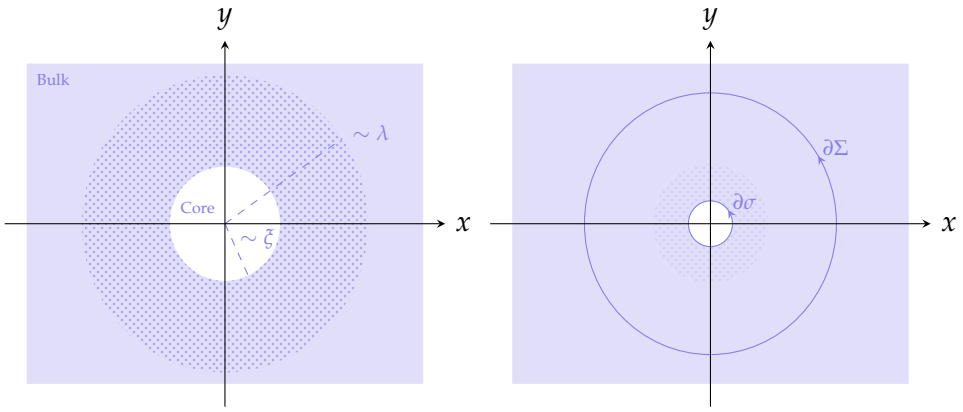
### 3.3.1 Impurities: the Abrikosov vortex

Consider an infinite sample, with an external field $\mathbf{H} = H\bar{\mathbf{z}}$ applied in the $z$ direction. Suppose the sample to be in the homogeneous superconducting state, with exclusion of a certain finite region where the Ginzburg-Landau equations (2.10) and (2.11) admit an inhomogeneous solution. Such a region is an impurity with some geometrical structure. Being non-superconducting, the impurity is expected not to completely exclude magnetic flux density (i.e. non exhibiting a local Meissner effect).

Now, the geometry of the impurity, which is supposed a stable state of the system, must obey the general symmetries of the problem: thus, it must exhibit cylindrical symmetry around the $z$ axis. We then select the center of this symmetric structure as the coordinates origin. The order parameter will obey the same dependence,

$$\Psi(\mathbf{x}) = \Psi(r)$$

Inside the impurity the sample is in normal state, $\Psi = 0$; outside it is in superconducting state, $\Psi \neq 0$. The basic cylindrical structure with such behavior has a single cylindrical core of radius $\xi$, since $\mathcal{O}(\xi)$ is the length scale




**(a)** Abrikosov vortex structure. **(b)** Integration paths.

**Figure 3.8:** The vortex structure and integration paths described in Sec. 3.3.1. The superconducting region extends approximately for $r > \xi$ and is shaded in figure. The dotted region, of radius $\sim \lambda$, is a non-Meissner phase with incomplete expulsion of the magnetic flux density. In Fig. 3.8b two integration paths are specified: $\partial\sigma$ delimits the inner core, $\sigma$, while $\partial\Sigma$ is a circular path deep in the bulk ($r \gg \lambda$) delimiting the surface $\Sigma$.

needed for $\Psi$ to break symmetry. We are dealing with a type II superconductor, so $\lambda > \xi$: we expect the field to penetrate in the sample in a region of radius $\mathcal{O}(\lambda)$. Then

| | | | |
|---|---|---|---|
| $r < \xi$ | core of the impurity | $\Psi = 0$ | $B \neq 0$ |
| $\xi < r < \lambda$ | field penetration region | $\Psi \neq 0$ | $B \neq 0$ |
| $r > \lambda$ | bulk of the superconductor | $\Psi \neq 0$ | $B = 0$ |

The structure is represented in Fig. 3.8a (neglecting the $z$ direction). On the $xy$ plane, we denote the surface occupied by the inner core by $\sigma$. We expect currents to circulate in the intermediate region in order to sustain the $B$ flux penetrating inside the impurity inner core. This basic structure is called the Abrikosov vortex.

Deep in the bulk, currents are null. Then, considering an integration path $\partial\Sigma$ as in Fig. 3.8b, we easily get

$$0 = \oint_{\partial\Sigma} d\boldsymbol{\ell} \cdot \mathbf{J} = \int_{\Sigma} d\Sigma\, \boldsymbol{\nabla} \times \mathbf{J} = \int_{\Sigma\setminus\sigma} d\Sigma\, \boldsymbol{\nabla} \times \mathbf{J}^{(\mathrm{s})} + \int_{\sigma} d\sigma\, \boldsymbol{\nabla} \times \mathbf{J}^{(\mathrm{c})}$$

In the last passage we decomposed the surface integral in two parts, being simply $\Sigma = (\Sigma \setminus \sigma) \cup \sigma$. Moreover, we defined two contributions to the current: the **core current** $\mathbf{J}^{(\mathrm{c})}$ flowing in the core region $\sigma$ and the **superconducting current** $\mathbf{J}^{(\mathrm{s})}$ flowing in the remaining region $\Sigma \setminus \sigma$. Clearly only the latter obeys London equation (1.3),

$$0 = \oint_{\partial\Sigma} d\boldsymbol{\ell} \cdot \mathbf{J} = -\frac{1}{\Lambda} \int_{\Sigma\setminus\sigma} d\Sigma\, \boldsymbol{\nabla} \times \mathbf{A} + \int_{\sigma} d\sigma\, \boldsymbol{\nabla} \times \mathbf{J}^{(\mathrm{c})}$$

We now work in the **extreme type II approximation**, $\xi \lll \lambda$. The core is reduced to a point and $\Sigma \setminus \sigma \simeq \Sigma$. Thus

$$\Phi_{\Sigma}(B) \simeq \Lambda \int_{\sigma} d\sigma\, \boldsymbol{\nabla} \times \mathbf{J}^{(\mathrm{c})}$$

We know that magnetic flux is quantized in holes. Since we are dealing with the most basic impurity, which also needs to be the one closest to the transition from the mixed phase to the mean-field solution, it must be

$$\Phi_\Sigma(B) \stackrel{!}{=} \Phi_0$$

which is, we assume **exactly one quantum of flux penetrates the vortex**. This gives a rather simple expression for the core current curl

$$\boldsymbol{\nabla} \times \mathbf{J}^{(\mathrm{c})} = \bar{\mathbf{z}} \frac{\Phi_0}{\Lambda} \delta^{(2)}(\mathbf{r}) \qquad \text{with} \qquad \mathbf{r} = \bar{\mathbf{x}} x + \bar{\mathbf{y}} y$$

Now, since $\mu_0 \mathbf{J} = \boldsymbol{\nabla} \times \mathbf{B}$ (recall: the material is non-magnetic), taking the rotor of the total current $\mathbf{J}$ and using $\boldsymbol{\nabla} \cdot \mathbf{B} = 0$,

$$\begin{aligned} -\frac{1}{\mu_0} \nabla^2 \mathbf{B} &= \boldsymbol{\nabla} \times \mathbf{J} \\ &= \boldsymbol{\nabla} \times \mathbf{J}^{(\mathrm{s})} + \boldsymbol{\nabla} \times \mathbf{J}^{(\mathrm{c})} \\ &\simeq -\frac{1}{\Lambda} \boldsymbol{\nabla} \times \mathbf{A} + \bar{\mathbf{z}} \frac{\Phi_0}{\Lambda} \delta^{(2)}(\mathbf{r}) = \frac{1}{\Lambda} \left[ \bar{\mathbf{z}} \Phi_0 \delta^{(2)}(\mathbf{r}) - \mathbf{B} \right] \end{aligned}$$

then, since $\Lambda / \mu_0 = \lambda^2$

$$\begin{aligned} \lambda^2 \nabla^2 \mathbf{B} - \mathbf{B} &= -\bar{\mathbf{z}} \Phi_0 \delta^{(2)}(\mathbf{r}) \\ \left[ \lambda^2 \frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial}{\partial r} \right) - 1 \right] B(r) &= -\Phi_0 \delta^{(2)}(\mathbf{r}) \end{aligned}$$

where we used the cylindrical symmetry, $\mathbf{B}(\mathbf{r}, z) = \bar{\mathbf{z}} B(r)$. The mystical power of Bessel's functions solves this equation

$$B(r) = \frac{\Phi_0}{2\pi\lambda^2} K_0 \left( \frac{r}{\lambda} \right) \tag{3.3}$$

with $K_0$ a zero-th order modified Bessel function of the second kind, or Hanckel's function with imaginary argument, the form of which the author is essentially obscure of. Qualitatively it goes like

$$\begin{aligned} \xi \ll r \ll \lambda \qquad & K_0 \left( \frac{r}{\lambda} \right) \simeq \log \left( \frac{\lambda}{r} \right) \\ r \gg \lambda \qquad & K_0 \left( \frac{r}{\lambda} \right) \simeq \exp \left( -\frac{r}{\lambda} \right) \end{aligned}$$

In the limit we are considering, reducing the impurity to a zero-dimensional object (a single point) the field would logarithmically diverge. In real world, in the central point of the impurity $B(0) = \mu_0 H$.

We now have an expression for the field strength inside each vortex. Taking the rotor of that field gives us the currents circulating in the vortex. Now we turn to some thermodynamics arguments.

### 3.3.2 Phase diagram of Abrikosov superconductors

As we know from Sec. 3.2.2 and Sec. 3.2.3, the condition $\kappa \gg 1$ defines the region $H_c < H < H_{c2}$ represented in Fig. **??** in which to produce an interface

between a superconducting and a normal phase is thermodynamically convenient. The argument expressed in Sec. 3.2.3 was rather rough and works suitably well in the following scenarios

$$H \ll H_c \qquad \text{and} \qquad H_c \ll H < H_{c2}$$

since in the first case, by physical intuition, the mean-field solution must be preferred, while being the second distant enough from $H_c$ we expect the sample to be essentially in the normal state with some nucleation of local superconductivity.

So, it is natural to expect that the phase diagram in Fig. **??** – with a mean-field, perfectly homogeneous solution evolving into an interfaced solution by crossing the curve $H_c(T)$ – is **not** the real, physical one. To correct this problem, define $F^{(0)}$ as the free energy for the mean-field homogeneous superconductor, and $F^{(1)}$ the free energy for the superconductor **with one Abrikosov vortex**. By Ginzburg-Landau theory (neglect the thermal contribution to free energy)

$$\begin{aligned} F\left[\Psi,\Psi^*,\mathbf{A};\mathbf{H}\right] = &\int d\mathbf{x}\left[\frac{|\boldsymbol{\nabla}\times\mathbf{A}|^2}{2\mu_0} - (\boldsymbol{\nabla}\times\mathbf{A})\cdot\mathbf{H}\right] \\ &+ \int d\mathbf{x}\left[a|\Psi(\mathbf{x})|^2 + \frac{b}{2}|\Psi(\mathbf{x})|^4 + \frac{c}{\hbar^2}|-i\hbar\boldsymbol{\nabla}\Psi(\mathbf{x}) - q\mathbf{A}\Psi(\mathbf{x})|^2\right] \end{aligned}$$

We work in the extreme type II approximation: in both cases, we consider $\Psi(\mathbf{x}) = \Psi_0$ everywhere in the sample (therefore neglecting terms arising from $\boldsymbol{\nabla}\Psi$). Thus

$$\begin{aligned} F\left[\Psi,\Psi^*,\mathbf{A};\mathbf{H}\right] = &\int_{\mathbb{R}^3} d\mathbf{x}\left[\frac{|\boldsymbol{\nabla}\times\mathbf{A}|^2}{2\mu_0} - (\boldsymbol{\nabla}\times\mathbf{A})\cdot\mathbf{H}\right] \\ &+ \int_{\mathbb{R}^3} d\mathbf{x}\left[a\Psi_0^2 + \frac{b}{2}\Psi_0^2 + c\Psi_0^2\frac{q^2}{\hbar^2}|\mathbf{A}|^2\right] \end{aligned}$$

The energy difference is only due to the free magnetic energy penetrating through the vortex and the Peierls contribution in the second line, being for $F^{(1)}$ somewhere $\mathbf{A} \neq \mathbf{0}$:

$$\begin{aligned} F^{(1)} - F^{(0)} &= \int d\mathbf{x}\left[\frac{|\boldsymbol{\nabla}\times\mathbf{A}|^2}{2\mu_0} - (\boldsymbol{\nabla}\times\mathbf{A})\cdot\mathbf{H} + c\frac{q^2}{\hbar^2}|\mathbf{A}|^2\Psi_0^2\right] \\ &= \int d\mathbf{x}\left[\frac{|\mathbf{B}|^2}{2\mu_0} - \mathbf{B}\cdot\mathbf{H} + \frac{\lambda^2}{2\mu_0}|\boldsymbol{\nabla}\times\mathbf{B}|^2\right] \end{aligned}$$

The last term is due to the facts

$$\mathbf{A} = -\Lambda\mathbf{J} = -\frac{\Lambda}{\mu_0}\boldsymbol{\nabla}\times\mathbf{B} = -\lambda^2\boldsymbol{\nabla}\times\mathbf{B}$$

and

$$\lambda^2 = \frac{m}{\mu_0 q^2\Psi_0^2} = \frac{\hbar^2}{q^2}\frac{1}{2\mu_0 c\Psi_0^2} \qquad \Longrightarrow \qquad c\Psi_0^2\frac{q^2}{\hbar^2} = \frac{1}{2\mu_0\lambda^2}$$

Now: the system exhibits $z$ translational invariance. The integration is carried out over the sample volume. We assume the sample to have a $xy$ surface $S$ and a $z$ extension $L$. Then the second term can be reduced to

$$\int d\mathbf{x}\, \mathbf{B}\cdot\mathbf{H} = LH\int_S dS\, \bar{\mathbf{z}}\cdot\mathbf{B} = LH\Phi_S(B)$$

Finally, defining the **vortex line energy per unit length** at some fixed $z$

$$\epsilon^{(1)} \equiv \frac{1}{2\mu_0}\int_S dS\, \left[|\mathbf{B}|^2 + \lambda^2|\boldsymbol{\nabla}\times\mathbf{B}|^2\right]$$

and using once again $z$ invariance, the resulting equation reads

$$F^{(1)} - F^{(0)} = L\left[\epsilon^{(1)} - H\Phi_S(B)\right] \tag{3.4}$$

In order to create one vortex, it must be $F^{(1)} - F^{(0)} < 0$, i.e.

$$H > \frac{\epsilon^{(1)}}{\Phi_S(B)}$$

Since as we know the minimum magnetic flux is one flux quantum, $\Phi_S(B) = \Phi_0$, this defines a **minimum magnetic field to enter the Abrikosov mixed phase $\mathbf{H_{c1}}$**,

$$H_{c1} \equiv \frac{\epsilon^{(1)}}{\Phi_0}$$

In principle, it is not guaranteed $H_{c1} < H_{c2}$. It is now necessary to calculate $\epsilon^{(1)}$. With some algebraic manipulations we can get

$$\epsilon^{(1)} = \frac{\lambda^2}{2\mu_0}\int_{S\setminus\sigma} d\Sigma\, \boldsymbol{\nabla}\cdot[\mathbf{B}\times(\boldsymbol{\nabla}\times\mathbf{B})]$$

using the notation of Fig. 3.8b, with $\sigma$ denoting the surface of the vortex inner core of radius $\xi$. Thanks to the Divergence Theorem and the fact that deep in the bulk the fields are null, the above integral becomes

$$\epsilon^{(1)} = \frac{\lambda^2}{2\mu_0}\oint_{\partial\sigma} d\boldsymbol{\eta}\cdot[\mathbf{B}\times(\boldsymbol{\nabla}\times\mathbf{B})] \qquad \text{where} \qquad d\boldsymbol{\eta} = -\bar{\mathbf{r}}\xi d\theta$$

The integral is performed along the closed circuit $\partial\sigma$, but – accordingly with the Divergence Theorem – the differential element points outside the boundary. Outside, with respect to $S\setminus\sigma$, means towards the center of the inner core, which justifies the proportionality to $-\bar{\mathbf{r}}$. Being $\mathbf{B}\parallel\bar{\mathbf{z}}$ and only dependent on $r$, in cylindrical coordinates

$$\mathbf{B}\times(\boldsymbol{\nabla}\times\mathbf{B}) = \bar{\mathbf{z}}B(r)\times\left(-\bar{\boldsymbol{\theta}}\frac{\partial}{\partial r}B(r)\right) = \bar{\mathbf{r}}B(r)\frac{\partial}{\partial r}B(r)$$

This quantity is obviously constant along $\partial\sigma$, therefore

$$\epsilon^{(1)} = -\frac{\lambda^2}{2\mu_0}2\pi\xi\left[B(r)\frac{\partial}{\partial r}B(r)\right]_{r=\xi}$$

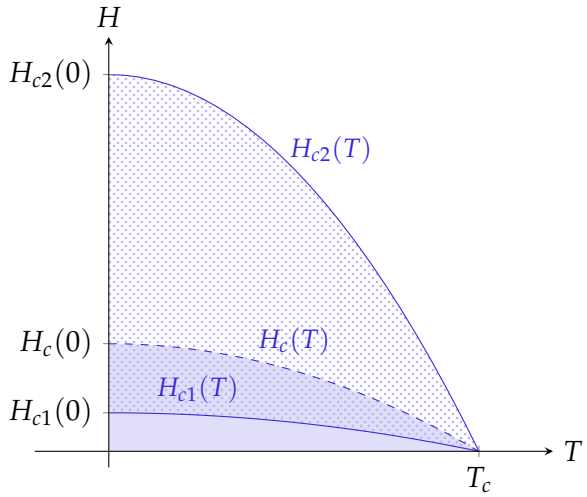



**Figure 3.9:** Phase diagram of an Abrikosov superconductor. The dotted region represents the mixed phase, in which Abrikosov vortices – impurities carrying each one one quantum of flux – can exist. Phenomenologically speaking, the field $H_c(T)$ has no particular behavior.

Using Eq. (3.3), in the case $r \gtrsim \xi$ (neglect the constant term, due to logarithmic dominance),

$$\begin{aligned}\epsilon^{(1)} &\simeq -\frac{\lambda^2}{2\mu_0} 2\pi\xi \frac{\Phi_0}{2\pi\lambda^2} \log\left(\frac{\lambda}{\xi}\right) \left[\frac{\partial}{\partial r} \frac{\Phi_0}{2\pi\lambda^2} \log\left(\frac{\lambda}{r}\right)\right]_{r=\xi} \\ &= \frac{\xi\Phi_0^2}{4\pi\mu_0\lambda^2} \frac{1}{\xi} \log\left(\frac{\lambda}{\xi}\right)\end{aligned}$$

then

$$H_{c1} = \frac{\Phi_0}{4\pi\mu_0\lambda^2} \log\left(\frac{\lambda}{\xi}\right)$$

In doing all these calculations the mathematical rigour is acrobatically ignored. One could argue that the last passage is inconsistent with the extreme type II approximation. One could, but that one is not the author. In Sec. 3.2.2 we used the identity

$$\mu_0 H_c = \frac{1}{\lambda\xi} \frac{\Phi_0}{4\pi}$$

then

$$H_{c1} = \frac{\Phi_0}{4\pi\mu_0\lambda^2} \log\left(\frac{\lambda}{\xi}\right) = \frac{1}{\mu_0} \left(\frac{1}{\lambda\xi} \frac{\Phi_0}{4\pi}\right) \frac{\xi}{\lambda} \log\left(\frac{\lambda}{\xi}\right)$$

which gives

$$H_{c1} = \frac{\log\kappa}{\kappa} H_c$$

Being $\kappa \gg 1$ by hypothesis, this relation confirms that Abrikosov's mixed phase exists in the window

$$H_{c1} < H_c < H_{c2}$$

The corrected phase diagram is depicted in Fig. 3.9, and shows the intermediate region where Abrikosov vortices live.

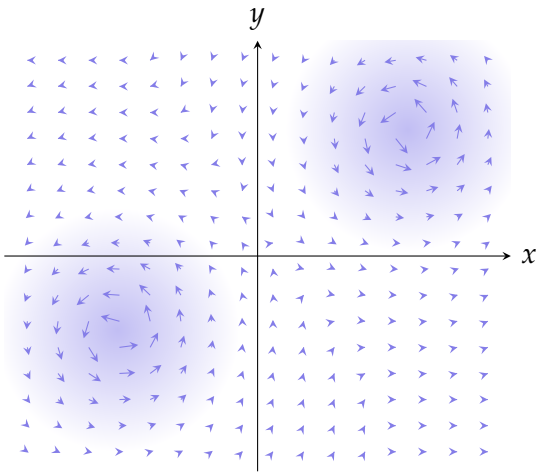



**Figure 3.10:** Pictorial representation of two "distant" fluxons. The arrows on the $xy$ represent current flows, while the shading (of radius $\sim \lambda$) represents field intensity.

## 3.4 THE ABRIKOSOV LATTICE

So, in certain conditions magnetic flux can penetrate the sample in a quantized manner, piercing holes with a specific field-current structure – the Abrikosov vortices. Each of this vortices carries exactly one quantum of flux $\Phi_0$, its "charge". Clearly, inverting the direction of the penetrating field, the sign of the charge inverts. Then two classes of vortices exist – those with charge $+\Phi_0$ and those with charge $-\Phi_0$. The noun "charge" is all but casual: as it turns out, **vortices with same charge repel each other, while vortices with same charge attract each other**. Impressive, isn't it?

This section is devoted to the study of the so-called Abrikosov lattice. The phenomenology is the following: when many vortices with the same charge nucleate in the superconducting sample (maintaining low the vortices density) they create a regular configuration – a triangular lattice. These vortices, also called **fluxons**, behave just like particles with a certain interaction law. As low-energy excited states, they are quasiparticles. First, we take off by studying how two fluxons interact.

### 3.4.1 Interaction of two static fluxons with same charge

Consider two fluxons with same charge $+\Phi_0$ at positions $\mathbf{r}_1 \equiv (x_1, y_1)$ and $\mathbf{r}_2 \equiv (x_2, y_2)$. Each structure has an inner core of radius $\xi$ (which we approximate as a dot) and an outer, superconducting core of radius $\lambda$. A pictorial representation of such situation is given in Fig. 3.10. To keep it simple, we consider the fluxons *distant*. This means

$$|\mathbf{r}_1 - \mathbf{r}_2| \gg \lambda$$

Within this approximation it makes sense to assume that the magnetic flux lines of one vortex do not interact with those of the other. Therefore each vortex has the same field-currents structure a single vortex in the sample

would have. We may refer to the field "in the vortex $j$" as $\mathbf{B}_j(\mathbf{r})$. Then the total magnetic flux density is given by

$$\begin{aligned}\mathbf{B}_{\text{tot}}(\mathbf{r}) &= \mathbf{B}_1(\mathbf{r}) + \mathbf{B}_2(\mathbf{r}) \\ &= \bar{\mathbf{z}}\left[B(|\mathbf{r}-\mathbf{r}_1|) + B(|\mathbf{r}-\mathbf{r}_2|)\right]\end{aligned}$$

where $B(r)$ is the function in Eq. (3.3). Even if the field lines of one fluxon are not modified by the presence of the other fluxon, the two interact. We want to estimate the free energy due to interaction as

$$F^{(\text{int})} \equiv F^{(2)} - 2F^{(1)}$$

where $F^{(2)}$ is the total free energy of the sample with two fluxons and $F^{(1)}$ is, as before, the free energy with one fluxon. The latter we know. As in the case of a single fluxon (see the previous section), $F^{(2)}$ is given by

$$\begin{aligned}F^{(2)} - F^{(0)} &= \int d\mathbf{x}\left[\frac{|\mathbf{B}_{\text{tot}}|^2}{2\mu_0} - \mathbf{B}_{\text{tot}}\cdot\mathbf{H} + \frac{\lambda^2}{2\mu_0}|\boldsymbol{\nabla}\times\mathbf{B}_{\text{tot}}|^2\right] \\ &= L\left[\epsilon^{(2)} - H\Phi_S(\mathbf{B}_{\text{tot}})\right]\end{aligned} \tag{3.5}$$

where, analogously,

$$\begin{aligned}\epsilon^{(2)} &= \frac{\lambda^2}{2\mu_0}\int_{S\setminus(\sigma_1\cup\sigma_2)} d\Sigma\,\boldsymbol{\nabla}\cdot[\mathbf{B}_{\text{tot}}\times(\boldsymbol{\nabla}\times\mathbf{B}_{\text{tot}})] \\ &= \frac{\lambda^2}{2\mu_0}\oint_{\partial\sigma_1} d\boldsymbol{\eta}_1\cdot[\mathbf{B}_{\text{tot}}\times(\boldsymbol{\nabla}\times\mathbf{B}_{\text{tot}})] + \frac{\lambda^2}{2\mu_0}\oint_{\partial\sigma_2} d\boldsymbol{\eta}_2\cdot[\mathbf{B}_{\text{tot}}\times(\boldsymbol{\nabla}\times\mathbf{B}_{\text{tot}})]\end{aligned}$$

In the above passages we have worked exactly as in the above section, integrating over the whole sample $S$ except the two cores $\sigma_1$ (around $\mathbf{r}_1$) and $\sigma_2$ (around $\mathbf{r}_2$) and then used the Divergence Theorem to move the integral over the domain boundary $\partial\sigma_1 \cup \partial\sigma_2$. Remember that the two differentials $d\boldsymbol{\eta}_1$ and $d\boldsymbol{\eta}_2$ point inward. Let us analyze the first term

$$\oint_{\partial\sigma_1} d\boldsymbol{\eta}_1\cdot[\mathbf{B}_{\text{tot}}\times(\boldsymbol{\nabla}\times\mathbf{B}_{\text{tot}})] = \oint_{\partial\sigma_1} d\boldsymbol{\eta}_1\cdot[\mathbf{B}_1\times(\boldsymbol{\nabla}\times\mathbf{B}_1)] \tag{3.6}$$

$$+ \oint_{\partial\sigma_1} d\boldsymbol{\eta}_1\cdot[\mathbf{B}_1\times(\boldsymbol{\nabla}\times\mathbf{B}_2)] \tag{3.7}$$

$$+ \oint_{\partial\sigma_1} d\boldsymbol{\eta}_1\cdot[\mathbf{B}_2\times(\boldsymbol{\nabla}\times\mathbf{B}_1)] \tag{3.8}$$

$$+ \oint_{\partial\sigma_1} d\boldsymbol{\eta}_1\cdot[\mathbf{B}_2\times(\boldsymbol{\nabla}\times\mathbf{B}_2)] \tag{3.9}$$

Here we have labeled each line of the equation to refer precisely to each term in the sum. In the first (3.6) we recognize an expression proportional to $\epsilon^{(1)}$. Consider the second term, (3.7):

$$\oint_{\partial\sigma_1} d\boldsymbol{\eta}_1\cdot[\mathbf{B}_1\times(\boldsymbol{\nabla}\times\mathbf{B}_2)]$$

We are integrating around core 1. Thus $\mathbf{B}_1 = B(\xi)\bar{\mathbf{z}}$ along all the integration path $\partial\sigma_1$. What remains is $\boldsymbol{\nabla}\times\mathbf{B}_2$. We now show

$$\left|\oint_{\partial\sigma_1} d\boldsymbol{\eta}_1\cdot[B(\xi)\bar{\mathbf{z}}\times(\boldsymbol{\nabla}\times\mathbf{B}_2)]\right| = B(\xi)\Phi_{\sigma_1}(\mathbf{B}_2)$$

This can be seen as follows. Set temporarily the coordinates origin at $(\mathbf{r}_1, 0)$ and use cylindrical coordinates. The differential $d\boldsymbol{\eta}_1$ is directed radially, while $B(\xi)\bar{\mathbf{z}} \times (\boldsymbol{\nabla} \times \mathbf{B}_2)$ is in principle a complicated vector in the $xy$ plane, being $\mathbf{B}_2$ off-center. Due to the vector product with $\bar{\mathbf{z}}$, its radial and angular components are interchanged. Then, integrating along $\partial\sigma_1$ with a radial differential is the same as integrating directly $\boldsymbol{\nabla} \times \mathbf{B}_2$ along the same path with an angular differential - that is, computing the circuitation of $\mathbf{B}_2$. Due to Stokes Theorem, we have the result.

An analogous reasoning can be done for the third term, (3.8), now approximating $\mathbf{B}_2 \simeq B(|\mathbf{r}_1 - \mathbf{r}_2|)$ along $\partial\sigma_1$: the fluxons are distant, so one field is approximately constant over the other's core. Therefore this term is proportional to the "on site flux" $\Phi_{\sigma_1}(\mathbf{B}_1) \gg \Phi_{\sigma_1}(\mathbf{B}_2)$. Summarizing, we have

$$\left| \oint_{\partial\sigma_1} d\boldsymbol{\eta}_1 \cdot [\mathbf{B}_1 \times (\boldsymbol{\nabla} \times \mathbf{B}_2)] \right| = B(\xi)\Phi_{\sigma_1}(\mathbf{B}_2)$$
$$\left| \oint_{\partial\sigma_1} d\boldsymbol{\eta}_1 \cdot [\mathbf{B}_2 \times (\boldsymbol{\nabla} \times \mathbf{B}_1)] \right| \simeq B(|\mathbf{r}_1 - \mathbf{r}_2|)\Phi_{\sigma_1}(\mathbf{B}_1)$$

Of course $B(|\mathbf{r}_1 - \mathbf{r}_2|) \ll B(\xi)$. We compare (3.7) and (3.8) as follows: first, we approximate $\Phi_{\sigma_1}(\mathbf{B}_2) \simeq \sigma_1 B(|\mathbf{r}_1 - \mathbf{r}_2|)$. Then

$$\frac{B(\xi)\Phi_{\sigma_1}(\mathbf{B}_2)}{B(|\mathbf{r}_1 - \mathbf{r}_2|)\Phi_{\sigma_1}(\mathbf{B}_1)} \simeq \frac{\sigma_1 B(\xi)}{\Phi_{\sigma_1}(\mathbf{B}_1)} \ll 1$$

because $B(r)$ is a decreasing function of distance and its minimum value in the core is $B(\xi)$, at its border. Moreover, at the core center the field is much stronger than at the border. Thus (3.7) is subleading if confronted with (3.8).

The last term (3.9) can be safely neglected, since the vortices are distant and $\mathbf{B}_2$ is suppressed on the boundary $\partial\sigma_1$, thus such term is subleading if compared to the second (3.7) and the third (3.8). Then

$$\oint_{\partial\sigma_1} d\boldsymbol{\eta}_1 \cdot [\mathbf{B}_{\text{tot}} \times (\boldsymbol{\nabla} \times \mathbf{B}_{\text{tot}})] \simeq \oint_{\partial\sigma_1} d\boldsymbol{\eta}_1 \cdot [\mathbf{B}_1 \times (\boldsymbol{\nabla} \times \mathbf{B}_1)] + \oint_{\partial\sigma_1} d\boldsymbol{\eta}_1 \cdot [\mathbf{B}_2 \times (\boldsymbol{\nabla} \times \mathbf{B}_1)]$$

The same obviously holds for the terms around $\partial\sigma_2$. Putting everything together, we get

$$\epsilon^{(2)} = 2\epsilon^{(1)} + \frac{\lambda^2}{2\mu_0} \oint_{\partial\sigma_1} d\boldsymbol{\eta}_1 \cdot [\mathbf{B}_2 \times (\boldsymbol{\nabla} \times \mathbf{B}_1)] + \frac{\lambda^2}{2\mu_0} \oint_{\partial\sigma_2} d\boldsymbol{\eta}_2 \cdot [\mathbf{B}_1 \times (\boldsymbol{\nabla} \times \mathbf{B}_2)]$$

The two integrals give the same result due to exchange symmetry. Therefore we can write, finally

$$\epsilon^{(2)} = 2\left[\epsilon^{(1)} + \frac{\lambda^2}{2\mu_0} \oint_{\partial\sigma_1} d\boldsymbol{\eta}_1 \cdot [\mathbf{B}_2 \times (\boldsymbol{\nabla} \times \mathbf{B}_1)]\right]$$

We proceed as before to calculate the integral. First, we get rid of the off-center field,

$$\frac{\lambda^2}{2\mu_0} B(|\mathbf{r}_1 - \mathbf{r}_2|) \oint_{\partial\sigma_1} d\boldsymbol{\eta}_1 \cdot [\bar{\mathbf{z}} \cdot (\boldsymbol{\nabla} \times \mathbf{B}_1)]$$

Without loss of generality we can move the origin of the axis in order to get $\mathbf{r}_1 = \mathbf{0}$. This choice is done to make the following notations simpler. Then, by the same logic exposed in Sec. 3.3.2, the field can only have a radial dependence (with respect to the core center, $\mathbf{r}_1$, which is now the coordinate origin $\mathbf{0}$) and thus

$$\bar{\mathbf{z}} \times (\boldsymbol{\nabla} \times \mathbf{B}) = \bar{\mathbf{z}} \times \left( -\bar{\boldsymbol{\theta}} \frac{\partial}{\partial r} B(r) \right) = \bar{\mathbf{r}} \frac{\partial}{\partial r} B(r)$$

This quantity is clearly constant along $\partial\sigma_1$; being $d\boldsymbol{\eta}_1 = -\xi d\theta \bar{\mathbf{r}}$ and making use of Eq. (3.3) it turns out

$$\begin{aligned}
&\frac{\lambda^2}{2\mu_0} B(|\mathbf{r}_2|) \oint_{\partial\sigma_1} d\boldsymbol{\eta}_1 \cdot [\bar{\mathbf{z}} \cdot (\boldsymbol{\nabla} \times \mathbf{B}_1)] \\
&\quad = \frac{\lambda^2}{2\mu_0} \left[ \frac{\Phi_0}{2\pi\lambda^2} K_0 \left( \frac{|\mathbf{r}_2|}{\lambda} \right) \right] \int_0^{2\pi} (-\xi) d\theta \left[ \frac{\partial}{\partial r} \frac{\Phi_0}{2\pi\lambda^2} K_0 \left( \frac{r}{\lambda} \right) \right]_{r=\xi} \\
&\quad \simeq \frac{\Phi_0^2}{8\pi^2\mu_0\lambda^2} K_0 \left( \frac{|\mathbf{r}_2|}{\lambda} \right) \int_0^{2\pi} \xi d\theta \left[ \frac{\partial}{\partial r} \log \left( \frac{r}{\lambda} \right) \right]_{r=\xi} \\
&\quad = \frac{\Phi_0^2}{4\pi^2\mu_0\lambda^2} K_0 \left( \frac{|\mathbf{r}_2|}{\lambda} \right)
\end{aligned}$$

The obvious, general result for $\mathbf{r}_1 \neq \mathbf{0}$ is simply

$$\epsilon^{(2)} = 2 \left[ \epsilon^{(1)} + \frac{\Phi_0^2}{4\pi^2\mu_0\lambda^2} K_0 \left( \frac{|\mathbf{r}_1 - \mathbf{r}_2|}{\lambda} \right) \right]$$

then, going back to Eq. (3.5), and noting that the total flux through $S$ is simply $2\Phi_0$ (each vortex carries one quantum of flux),

$$\begin{aligned}
F^{(2)} - F^{(0)} &= L \left[ \epsilon^{(2)} - 2H\Phi_0 \right] \\
&= 2L \left[ \epsilon^{(1)} + \frac{\Phi_0^2}{4\pi^2\mu_0\lambda^2} K_0 \left( \frac{|\mathbf{r}_1 - \mathbf{r}_2|}{\lambda} \right) - H\Phi_0 \right] \\
&= 2 \left[ F^{(1)} - F^{(0)} \right] + 2L \frac{\Phi_0^2}{4\pi^2\mu_0\lambda^2} K_0 \left( \frac{|\mathbf{r}_1 - \mathbf{r}_2|}{\lambda} \right)
\end{aligned}$$

where we made use of Eq. (3.4). We finally have an expression for the interaction energy $F^{(\mathrm{int})} \equiv F^{(2)} - 2F^{(1)}$,

$$\begin{aligned}
F^{(\mathrm{int})} &= -F^{(0)} + \frac{L\Phi_0^2}{2\pi^2\mu_0\lambda^2} K_0 \left( \frac{|\mathbf{r}_1 - \mathbf{r}_2|}{\lambda} \right) \\
&\simeq -F^{(0)} + \frac{L\Phi_0^2}{2\pi^2\mu_0\lambda^2} \exp \left( -\frac{|\mathbf{r}_1 - \mathbf{r}_2|}{\lambda} \right)
\end{aligned} \tag{3.10}$$

In the last passage we used the approximation for distant fluxons $|\mathbf{r}_1 - \mathbf{r}_2| \gg \lambda$. The term $-F^{(0)}$ is a mere shift and can be neglected. It is evident that two fluxons with the same charge repel each other, being the interaction contribution in Eq. (3.10) inherently positive and steadily decreasing for increasing distance $|\mathbf{r}_1 - \mathbf{r}_2|$.

It is now trivial to compute the force exerted from one vortex onto the other: by defining $\mathbf{r}_{12} \equiv \mathbf{r}_1 - \mathbf{r}_2$, $\bar{\mathbf{r}}_{12} \equiv (\mathbf{r}_1 - \mathbf{r}_2) / |\mathbf{r}_1 - \mathbf{r}_2|$ and $\boldsymbol{\nabla}_{12}$ the gradient taken with respect to $\bar{\mathbf{r}}_{12}$, it is

$$\mathbf{f}_{12} = -\boldsymbol{\nabla}_{12} F^{(\mathrm{int})} = \bar{\mathbf{r}}_{12} \frac{L\Phi_0^2}{2\pi^2\mu_0\lambda^3} \exp \left( -\frac{|\mathbf{r}_1 - \mathbf{r}_2|}{\lambda} \right)$$

which is repulsive, correctly.

### 3.4.2 Interaction of two static fluxons with opposite charge

The computation for two **distant** fluxons with opposite charge is rather trivial at this point. All that changes is the direction of one of the two penetrating fields (say, the second),

$$\begin{aligned}\mathbf{B}_{\text{tot}}(\mathbf{r}) &= \mathbf{B}_1(\mathbf{r}) + \mathbf{B}_2(\mathbf{r}) \\ &= \bar{\mathbf{z}}\left[B(|\mathbf{r}-\mathbf{r}_1|) - B(|\mathbf{r}-\mathbf{r}_2|)\right]\end{aligned}$$

By going through all the above calculation once again, it is easy to see that an **attractive** force emerges between the two fluxons,

$$\mathbf{f}_{12} = -\bar{\mathbf{r}}_{12}\frac{L\Phi_0^2}{2\pi^2\mu_0\lambda^3}\exp\left(-\frac{|\mathbf{r}_1-\mathbf{r}_2|}{\lambda}\right)$$

We need to be careful. While in the above section two fluxons repel each other, letting us use the distant fluxons approximation all the way through the evolution, in this case all our approximations will break at some point when the two fluxons get near enough. To determine the exact behavior of the fluxons is rather difficult, so we shall not go in that direction.

In certain situations, a vortex and an antivortex can form a bound-state: for example, by separating the sample in two superconducting domains with the field applied in opposite directions, we expect that due to magnetic interaction a repulsive force emerges inhibiting the vortex and the antivortex from crossing the wall; this repulsion can balance with the vortex-antivortex attraction. As shown in the article "Bound and stable vortex–antivortex pairs in high-Tc superconductors" [11], this situations leads to a very interesting and special bound state across the domain wall.

### 3.4.3 Interaction of two moving fluxons with same charge

This short section is largely based on the article by Kogan and Prozorov, "Interaction between moving Abrikosov vortices in type-II superconductors" [10], and rapidly treats how two moving fluxons interact. go on.... . .

### 3.4.4 Many fluxons with the same charge: the lattice ground state

Consider now a slab with $N$ identical fluxons, in the **low-density approximation**: any vortex is **distant** (which means, at a distance larger than $\lambda$) from any other vortex. We work in the thermodynamical limit: the volume is large enough to neglect border effect, while the fluxon density remains finite. Experimentally, it is found that fluxons form a stable lattice: our aim is to prove that a square lattice is energetically unfavorable with respect to a triangular lattice. All this section is intended in two dimensions.

Start by taking the $N$ fluxons free energy, $F^{(N)}$. In principle, are to be expected three-fluxons interactions (and possible higher orders). Let's extend what we found for two fluxons. The total magnetic field is

$$\mathbf{B}_{\text{tot}} = \sum_i \mathbf{B}_i(\mathbf{r}) = \sum_i \bar{\mathbf{z}}B\left(\mathbf{r}-\mathbf{r}_i\right)$$

where the sum runs over all the fluxons, at positions $\mathbf{r}_i$. The $N$ fluxons free energy is given by

$$F^{(N)} - F^{(0)} = L\left[\epsilon^{(N)} - H\Phi_S(B)\right]$$

The above equation is the simple extension of Eq. (3.4). Here we have defined

$$\begin{aligned}\epsilon^{(N)} &= \frac{\lambda^2}{2\mu_0}\sum_i \oint_{\partial\sigma_i} d\boldsymbol{\eta}_i \cdot [\mathbf{B}_{\text{tot}} \times (\boldsymbol{\nabla} \times \mathbf{B}_{\text{tot}})] \\ &= \frac{\lambda^2}{2\mu_0}\sum_i \sum_{jl} \oint_{\partial\sigma_i} d\boldsymbol{\eta}_i \cdot \left[\mathbf{B}_j \times (\boldsymbol{\nabla} \times \mathbf{B}_l)\right]\end{aligned}$$

The magnetic corrections to free energy involve line integrals around cores $\partial\sigma_i$ of **two** fields. These fields could be both off-center (which means, fields pinning through vortices $j$ and $l$, with $i \neq j \neq l \neq i$). We neglect such contributions, clearly subleading in the low-density regime, limiting ourselves to two-centers integrals: in the above summation, we ignore all terms where $i$,$j$ and $l$ are all different. By following the same line of reasoning of the above section, only two terms survive

$$\begin{aligned}\epsilon^{(N)} \simeq \frac{\lambda^2}{2\mu_0}\sum_i \oint_{\partial\sigma_i} d\boldsymbol{\eta}_i \cdot [\mathbf{B}_i \times (\boldsymbol{\nabla} \times \mathbf{B}_i)] \\ + \frac{\lambda^2}{2\mu_0}\sum_{i\neq j} \oint_{\partial\sigma_i} d\boldsymbol{\eta}_i \cdot \left[\mathbf{B}_j \times (\boldsymbol{\nabla} \times \mathbf{B}_i)\right]\end{aligned}$$

In the first one we recognize $N\epsilon^{(1)}$. For the second, defining $f(i,j)$ as

$$f(i,j) = \frac{\lambda^2}{2\mu_0}\oint_{\partial\sigma_i} d\boldsymbol{\eta}_i \cdot \left[\mathbf{B}_j \times (\boldsymbol{\nabla} \times \mathbf{B}_i)\right] + \frac{\lambda^2}{2\mu_0}\oint_{\partial\sigma_j} d\boldsymbol{\eta}_j \cdot \left[\mathbf{B}_i \times (\boldsymbol{\nabla} \times \mathbf{B}_j)\right]$$

We get

$$\epsilon^{(N)} = N\epsilon^{(1)} + \sum_{i<j} f(i,j)$$

Check the above section: we already carried out a the calculation for $f$. It is easy to recognize that

$$f(i,j) = \frac{\Phi_0^2}{2\pi^2\mu_0\lambda^2} K_0\left(\frac{|\mathbf{r}_i - \mathbf{r}_j|}{\lambda}\right)$$

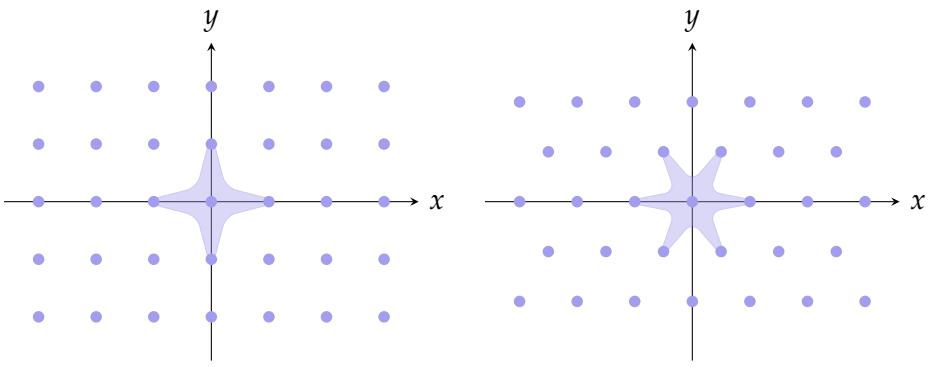
Now: each vortex carries a flux $\Phi_0$. Since $F^{(1)} - F^{(0)} = L[\epsilon^{(1)} - H\Phi_0]$, we have

$$F^{(N)} - F^{(0)} = N\left[F^{(1)} - F^{(0)}\right] + L\sum_{i<j} f(i,j)$$

thus

$$F^{(\text{int})} \equiv F^{(N)} - NF^{(1)} = -(N-1)F^{(0)} + \frac{L}{2}\sum_{i\neq j} f(i,j)$$

In the last passage the sum was corrected by the factor $1/2$ . The geometric contribution of the sample to stability enters in the last term in the sum. With the notation $\langle i,j\rangle$ we indicate the set of sites $j$ of the lattice such that $i$ and $j$ are first nearest neighbors: since the interaction is exponentially suppressed




**(a)** Square lattice. **(b)** Triangular lattice.

**Figure 3.11:** The square lattice in Fig. 3.11a has a coordination number $k = 4$, since 4 are the first nearest neighbors of a given site. The triangular lattice, represented in Fig. 3.11b, has instead $k = 6$. The shading is pictorial and has no physical meaning, except indicating the nearest neighbors for a given site.

with distance, we neglect interaction with neighbors of order higher than first,

$$F^{(\text{int})} \simeq -(N-1)F^{(0)} + \frac{L}{2}\sum_{\langle i,j\rangle} f(i,j)$$

Now: we wish to confront two simple two-dimensional lattices: square and triangular, as represented in Fig. 3.11. Abrikosov's first guess for the fluxon lattice configuration was a square one. Suppose $d$ to be the lattice spacing: the favorite lattice will be the one with the lowest interaction free energy, since the total single-vortex contribution to energy is fixed by the number of fluxons and independent of the lattice geometry. Neglect the shift $-(N-1)F^{(0)}$: due to lattice symmetry

$$F^{(\text{int})} = \frac{L}{2}Nkf^{\star}$$

The factor $N$ comes from translational invariance across the lattice, since the same contribution arises from any of the $N$ sites. Any of those sites will have the same number of nearest neighbors, $k$; finally, for any couple of nearest neighbors $i, j$ the function $f$ has the same value, which we called $f^{\star}$. To calculate $f^{\star}$, we consider two fluxons at a distance $d$. We have

$$f^{\star} = \frac{\Phi_0^2}{2\pi^2\mu_0\lambda^2}K_0\left(\frac{d}{\lambda}\right) \simeq \frac{\Phi_0^2}{2\pi^2\mu_0\lambda^2}\exp\left(-\frac{d}{\lambda}\right)$$

At a first sight, a lattice with a lower coordination number $k$ (i.e. the square one) might seem energetically convenient. However we did not specify the value of $d$, and it needs not to be equal in the two situations. In fact, it is not. For a given lattice, we call $s$ the Wigner-Seitz unit cell – which is, the surface containing exactly one vortex. It is easy to see that its numerical value is

$$|s| = d^2 \qquad \text{for the square lattice}$$

$$|s| = d^2\frac{\sqrt{3}}{2} \qquad \text{for the triangular lattice}$$

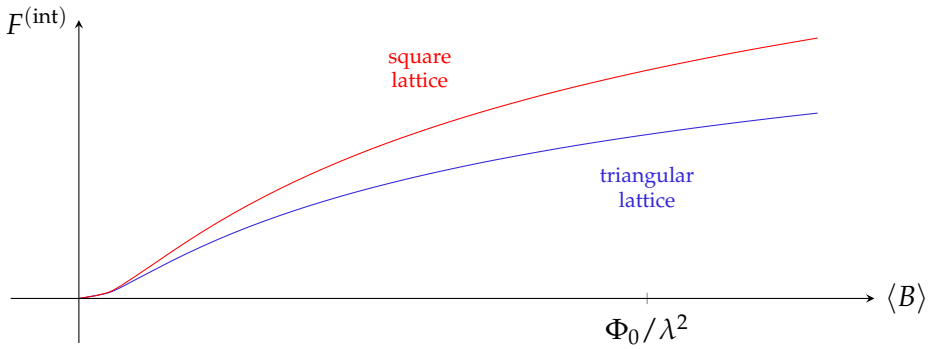



**Figure 3.12:** Plot of the interaction free energy $F^{(\mathrm{int})}$ as a function of the mean magnetic flux $\langle B \rangle$. The general leading order behavior is clear: the triangular lattice has a smaller interaction free energy.

The density, however, is fixed by the thermodynamic limit. Then

$$n = \frac{N}{S} = \frac{1}{|s|}$$

We can also connect $n$ with the average flux density per unit surface across the sample:

$$\langle B \rangle = n\Phi_0$$

a simpler quantity to be measured with respect to the fluxons number. Then, putting everything together,

$$F^{(\mathrm{int})} = NL\frac{\Phi_0^2}{2\pi^2\mu_0\lambda^2}\left[4\exp\left(-\sqrt{\frac{\Phi_0}{\langle B\rangle\,\lambda^2}}\right)\right] \qquad \text{(square lattice)}$$

$$F^{(\mathrm{int})} = NL\frac{\Phi_0^2}{2\pi^2\mu_0\lambda^2}\left[6\exp\left(-\sqrt{\frac{2\Phi_0}{\sqrt{3}\,\langle B\rangle\,\lambda^2}}\right)\right] \qquad \text{(triangular lattice)}$$

In Fig. 3.12 the two functions are plotted. For vanishing average magnetic flux density the two curves collide, as it is to be expected: a "square lattice" with one fluxon is identical to a "triangular lattice" with one fluxon. For low magnetic fields in the region of interest, $\langle B \rangle < \Phi_0/\lambda^2$ (approximately one fluxon per a $\lambda^2$ area, which is, low-density) evidently the triangular configuration (see Fig. 3.11b) is preferred to the square configuration (see Fig. 3.11a).

# 4 COOPER PAIRS: HOW ELECTRONS FORM BOUND STATES



In the last chapters, we treated superconductivity on a purely macroscopical level. Chap. 1 made use of classical electrodynamics, while Chaps. 2 and 3 of Ginzburg-Landau theory; within this last approach, a lot of striking features of superconductivity have been derived, with much focus on phenomenology. It is time to pass to a much shorter length scale, investigating the profound causes of superconductivity in the behavior of the quantum particles involved. The theory describing such regime is the one elaborated by Bardeen, Cooper and Schrieffer (BCS), and published in their historical article "Theory of Superconductivity" [1] in 1957 and will be analyzed in the next chapter.

Briefly, the BCS theory accurately explains superconductivity as the phase of matter in which the dominant interaction acting on a system of electrons inside a material is an effective interaction arising from the crystalline structure of the material itself: couples of electrons experience an **attractive interaction**, mediated by a phonon. The lattice vibrations act as carriers of energy and momentum and make it possible for electron pairs to form bound states. Such bosonic couples, called **Cooper pairs**, enter the superconducting state by forming a Bose-Einsten condensate. We will delve into these concepts in this chapter.

> **G. Grosso and G. Pastori Parravicini**: *Solid State Physics* ([2] @ 18)
>
> The history of superconductivity is full of fascinating surprises and chal-

> lenging developments. The milestone work of Kamerlingh Onnes in 1911 on the electrical resistivity of mercury has opened a new world to the physical investigation, and the discovery of high-$T_c$ superconductivity in barium-doped lanthanum cuprate, by Bednorz and Müller in 1986, has given a novel impetus to the subject.
>
> The microscopic origin of superconductivity is linked to the possible occurrence of a (small) effective attractive interaction between conduction electrons (or valence holes in p-type conductors) and the consequent formation of electron pairs (or hole pairs), at sufficiently low temperatures. The mechanism of electron pairing is at the origin of perfect conductivity, perfect diamagnetism, anomalous specific heat and thermodynamical properties, magnetic flux quantization, coherent tunneling, and several other effects in superconductors. Empirical laws and semi-empirical models have accompanied the accumulation of the wide and rich phenomenology of superconductors. Eventually, the fundamental work of Bardeen, Cooper, and Schrieffer (1957) has transformed an endless list of peculiar effects and conjectures into a logically consistent theoretical framework. [...] without its concepts no serious discussion would be possible at all.

This chapter may appear a little convoluted in its exposition order: we start by analyzing how some kind of interaction should work on top of the Fermi sphere, and then pass to real Physics and find a mechanism producing an interaction of that kind. Although maybe a little inefficient, this order allows to collect some theoretical arguments to support a first-sight bizarre idea, of phonons acting as some strange glue for electrons.

## 4.1 WHAT IF ELECTRONS ATTRACT?

Many experiments exist, showing that the elementary "object" inside a superconductor has charge $q = 2e$. This can be verified, for example, measuring the quantization of the magnetic flux inside a superconducting sample. The flux gets quantized as

$$\Phi = n\frac{h}{|q|} = n\frac{h}{2|e|}$$

This general rule, that seems to be obeyed flawlessly in the superconducting phase, indicates that such object is **a pair of electrons**.

Moreover, the superconducting transition exhibits many similarities with the superfluid transition of liquid Helium, which is well known to be a Bose-Einstein condensation process. As it turns out, a superconductor is a condensate state. To produce a condensate, then, we need bosons. Electron pairs, seen as composite objects, are bosons.

Other arguments point in the same direction: superconductivity is the condensation of a system of electrons pairs. This is the corner stone of the BCS theory. To make a pair, we need an **attractive interaction** between electrons: we know they interact via the (screened) Coulomb interaction and the Pauli principle, so it may seem strange to look for some kind of attraction; we assume they somehow attract, and see if they form bound states – which are, pairs.

### 4.1.1 Bound states

Consider two interacting electrons in $D$ dimensions, with hamiltonian

$$\hat{H} = \frac{\hat{\mathbf{p}}_1^2}{2m} + \frac{\hat{\mathbf{p}}_2^2}{2m} + V(\hat{\mathbf{x}}_1 - \hat{\mathbf{x}}_2)$$

with obvious notation. The same hamiltonian can be decomposed in the sum of the center of mass part and the relative part,

$$\hat{H} = \left[\frac{\hat{\mathbf{P}}^2}{2M}\right] + \left[\frac{\hat{\mathbf{p}}^2}{2\mu} + V(\hat{\mathbf{x}})\right]$$

with

$$\mathbf{P} \equiv \mathbf{p}_1 + \mathbf{p}_2 \qquad \mathbf{X} \equiv \frac{\mathbf{x}_1 + \mathbf{x}_2}{2} \qquad \mathbf{p} \equiv \frac{\mathbf{p}_1 - \mathbf{p}_2}{2} \qquad \mathbf{x} \equiv \mathbf{x}_1 - \mathbf{x}_2$$

and

$$M = 2m \qquad \mu = \frac{m}{2}$$

Assuming overall translational symmetry, the wavefunction can be factorized as

$$\psi(\mathbf{x}_1, \mathbf{x}_2) = \Phi(\mathbf{X})\phi(\mathbf{x})$$

where $\Phi$ is the wavefunction of the center of mass, and $\phi$ is the relative wavefunction. The ground state of the problem, being decoupled, would clearly be the one with zero net momentum,

$$\mathbf{p}_1 + \mathbf{p}_2 = 0$$

for which the relative momentum is just the momentum of one of the two particles.

A brief specification: to say that the ground state is the one with opposite momenta is here a little inessential because it would require to specify a reference frame. In other words, an eventual bound state would not be weaker if equipped with a net momentum, just as a moving hydrogen atom is not less bounded than one at rest. For a two particles system there is not a third actor with respect to which it makes sense to set the zero of the energy; instead, by just performing a Galilean transformation, we can absorb the center of mass contribution to energy.

On the contrary, in Sec. 4.1.3 the ground state will be exactly the one with zero momentum *in the physical sense* that its energy is lower by an amount that can not be eliminated just through a Galilean transformation. This comment will be useful later.

Now, consider a local interaction, on a "small" length scale. We may start by considering the perfectly local contact-attractive interaction,

$$V(\mathbf{x}) \equiv -V_0 \delta^{(D)}(\mathbf{x}) \qquad \text{with} \qquad V_0 > 0$$

Here we are neglecting the Coulomb interaction. It is reasonable to do so if such interaction is screened, as it commonly is in materials. For the Coulomb interaction to be screened we need the whole electron liquid background: for more details on this subject, check the vast book *Quantum Theory of the Electron Liquid* [3] by Giuliani and Vignale. Let us forget for a moment both the

electron liquid and the Coulomb interaction, and proceed with two locally interacting chargeless fermions. This evidently incoherent argument is necessary to highlight, in the following, the essential collective nature of the attractive interaction.

The Schrödinger's Equation for the relative part of the wavefunction is given by

$$\left[\frac{\hat{\mathbf{p}}^2}{2\mu} + V\left(\hat{\mathbf{x}}\right)\right]\phi(\mathbf{x}) = -E^{(\mathrm{b})}\phi(\mathbf{x})$$

where the eigenvalue $-E^{(\mathrm{b})} < 0$ indicates the binding energy. Consider now the complete basis of orthonormal plane waves,

$$w_{\mathbf{k}}\left(\mathbf{x}\right) = L^{-D/2}e^{i\mathbf{k}\cdot\mathbf{x}}$$

with $L^D$ the total volume. The wavefunction can be decomposed as

$$\phi\left(\mathbf{x}\right) = \sum_{\mathbf{k}} \alpha_{\mathbf{k}} w_{\mathbf{k}}\left(\mathbf{x}\right)$$

Thus, inserting the above decomposition in the Schrödinger's Equation and projecting onto the plane wave $w_{\mathbf{k}}\left(\mathbf{x}\right)$, we obtain

$$2\epsilon_{\mathbf{k}}\alpha_{\mathbf{k}} + \sum_{\mathbf{k}'} V_{\mathbf{k}-\mathbf{k}'}\alpha_{\mathbf{k}'} = -E^{(\mathrm{b})}\alpha_{\mathbf{k}} \qquad \text{with} \qquad \epsilon_{\mathbf{k}} = \frac{\hbar^2|\mathbf{k}|^2}{2m}$$

First, note that we defined $\epsilon_{\mathbf{k}}$ as the kinetic energy **of an electron** in state $|\mathbf{k}\rangle$, and not of the particle of mass $\mu = m/2$. This is the reason for the 2 factor in front of the kinetic contribution. The choice was done for notational conformity with the rest of the chapter. Second, the Fourier transform of the interaction potential is intended,

$$V_{\mathbf{k}-\mathbf{k}'} = \frac{1}{L^D}\int_{\mathbb{R}^D} d\mathbf{x}\, V\left(\mathbf{x}\right) e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{x}} = -V_0$$

since the potential is *delta-like*. Then,

$$\left(2\epsilon_{\mathbf{k}} + E^{(\mathrm{b})}\right)\alpha_{\mathbf{k}} = V_0 \sum_{\mathbf{k}'} \alpha_{\mathbf{k}'}$$

Notice that defined as such $V_{\mathbf{k}-\mathbf{k}'}$ is the Fourier component of the potential capable of changing the relative momentum from $\mathbf{k}'$ to $\mathbf{k}$. Remember that, although not necessary to understand the dynamical nature of the bound state, if we use a pair with opposite momenta such Fourier component is the one capable of changing the moment of one particle from $\mathbf{k}'$ to $\mathbf{k}$, and the moment of the other from $-\mathbf{k}'$ to $-\mathbf{k}$. It follows:

$$\alpha_{\mathbf{k}} = \frac{V_0}{2\epsilon_{\mathbf{k}} + E^{(\mathrm{b})}} \sum_{\mathbf{k}'} \alpha_{\mathbf{k}'}$$

then, summing over $\mathbf{k}$, the coefficient $\sum \alpha$ can be simplified both sides, leaving the self-consistency equation

$$\sum_{\mathbf{k}} \frac{V_0}{2\epsilon_{\mathbf{k}} + E^{(\mathrm{b})}} = 1$$

Assuming a large volume we can make approximate the momenta as continuous. A little caution is here needed: we approximate the potential as "perfectly local", which means that the length scale over which it drops to zero is much smaller than any physical length scale involved in the system. Having neglected the Coulomb long-range interaction, we understand the relevant length here cited is of the order of the particle dimension: our fermions are dimensionless points. Briefly, we should integrate on $\mathbb{R}^D \setminus s(2\pi/\Lambda)$ with $s(r)$ the sphere of radius $r$ in $D$ dimensions and $2\pi/\Lambda$ the said length for a properly defined momentum $\Lambda$.

This is equivalent to integrating over $\mathbb{R}^D$ a potential whose Fourier transform is constant for $|\mathbf{k}| < \Lambda$ and (approximately and continuously) drops to zero for bigger momenta. Such potential is strongly localized, and *delta-like* as seen "from distant". Defining $\kappa \equiv (2\pi)^D / L^D V_0$ we have

$$\kappa = \int_{|\mathbf{k}|<\Lambda} d^D\mathbf{k} \, \frac{1}{2\epsilon_{\mathbf{k}} + E^{(\mathrm{b})}}$$

The question is: at varying dimensionality $D$, is there a solution for any given $\kappa$?

1. For $D = 1$, the integral becomes

$$\kappa = \int_{|k|<\Lambda} dk \, \left[ \frac{\hbar^2 k^2}{2\mu} + E^{(\mathrm{b})} \right]^{-1}$$

   The above function is solved by an infinite set of couples $(E^{(\mathrm{b})}, \kappa)$; $\kappa$ is a continuous function of $E^{(\mathrm{b})}$. Moreover, for $E^{(\mathrm{b})} \to 0$ the integral presents an hyperbolic divergence, thus allowing for a $\kappa \to \infty$ solution. Then for any choice of $\kappa \in \mathbb{R}$ a solution exists.

   The bound state is formed regardless of $\kappa$, which is, regardless of the attraction strength $V_0$.

2. For $D = 2$, we get

$$\kappa = \int_{|\mathbf{k}|<\Lambda} d^2k \, \left[ \frac{\hbar^2 k^2}{2\mu} + E^{(\mathrm{b})} \right]^{-1} = \pi \int_{k^2<\Lambda^2} dk^2 \, \left[ \frac{\hbar^2 k^2}{2\mu} + E^{(\mathrm{b})} \right]^{-1}$$

   where we used $d^2k = 2\pi k dk = \pi dk^2$. The same argument of the point above holds: for $E^{(\mathrm{b})} \to 0$ the integral presents a logarithmic divergence, thus allowing for a $\kappa \to \infty$ solution. Then for any choice of $\kappa \in \mathbb{R}$ a solution exists.

   Also for $D = 2$ the bound state is formed regardless of the attraction strength $V_0$.

3. For $D > 2$, we can use

$$d^D\mathbf{k} = \Omega_D k^{D-1} dk$$

   with $\Omega_D$ the $D$-dimensional solid angle. Thus the integral becomes

$$\kappa = \int_{|\mathbf{k}|<\Lambda} d^D\mathbf{k} \, \left[ \frac{\hbar^2 k^2}{2\mu} + E^{(\mathrm{b})} \right]^{-1} = \Omega_D \int_{k<\Lambda} dk \, k^{D-1} \left[ \frac{\hbar^2 k^2}{2\mu} + E^{(\mathrm{b})} \right]^{-1}$$

Being $D-1 \geq 2$, this integral remains finite for any value of $E^{(b)}$, as long as the cutoff $\Lambda$ is finite. Moreover, the maximum value (which is finite and we denote by $\kappa^\star$) of the integral is recovered for $E^{(b)} \to 0$.

For $D = 3$ and in higher dimensions, two electrons form a bound state if $\kappa \leq \kappa^\star$ – or, if the interaction potential $V_0$ exceeds a certain threshold value.

It looks like two chargeless fermions equipped with a local and attractive interaction cannot form a pair in three dimensions. This should limit the phenomenon of superconductivity to two-dimensional materials. Then, why do we have three-dimensional superconductors?

### 4.1.2 Adding particles on top of the Fermi sphere

We are missing something. As anticipated, to neglect the Coulomb interaction between electrons we need the whole electron liquid – which is, we need a great number of electrons plus our two interacting electrons, interacting with all others only by the Pauli principle. This may seem a little modification; it is instead a huge one, because now this kind of attraction allows for bound states also for $D > 2$, and the next section is devoted to showing that. We may say that the electron pair is an object formed by two electrons directly and all others indirectly – a collective configuration.

What now changes a lot is that it makes sense to say that the bound state has zero net momentum. It is not a matter of reference frames anymore, because of the presence of the whole Fermi sphere. To correct for a kinetic contribution of the total momentum would mean to shift the entire Fermi sphere. Then the ground state *must* be searched among those states with opposite momenta (in the reference frame of a still Fermi sphere), in the sense that it *physically has lower energy* independently of how we set the zero.

We now focus on a little formal aspect of Quantum Mechanics that may seem obvious. We know from Pauli exclusion principle that electrons must occupy totally antisymmetric states, which are all eigenstates of any electron couple permutation operator with eigenvalue $-1$. Consider the non-antisymmetric state

$$|\Omega\rangle \equiv \prod_{|\mathbf{k}|<k_F} |\mathbf{k}\rangle$$

which represents the Fermi Sea. Suppose a certain $\mathbf{k}^\star$ exists inside the sphere, $|\mathbf{k}^\star| < k_F$, and consider adding **one** electron in the **normalized** state

$$|\psi\rangle \equiv \alpha^\star |\mathbf{k}^\star\rangle + \sum_{\mathbf{k}\in\mathcal{S}} \alpha_{\mathbf{k}} |\mathbf{k}\rangle$$

with $\alpha^\star$ and $\alpha_{\mathbf{k}}$ the expansion coefficients, and $\mathcal{S}$ a certain set of momenta **outside the sphere**. We will go back to two electrons in a moment. Pauli principle tells us that the physical state is given by

$$\hat{\mathcal{A}}\left[|\Omega\rangle \otimes |\psi\rangle\right]$$

with $\hat{\mathcal{A}}$ the "anti-symmetrizer". Then expanding

$$\alpha^\star \hat{\mathcal{A}}\left[|\Omega\rangle \otimes |\mathbf{k}^\star\rangle\right] + \hat{\mathcal{A}}\left[|\Omega\rangle \otimes \sum_{\mathbf{k}\in\mathcal{S}} \alpha_{\mathbf{k}} |\mathbf{k}\rangle\right]$$

and then $\alpha^\star$ must be zero, since $|\Omega\rangle \otimes |\mathbf{k}^\star\rangle$ exhibits exchange symmetry for the state $|\mathbf{k}^\star\rangle$ and thus is unphysical. The application of $\hat{\mathcal{A}}$ cancels its contribution to . If $|\psi\rangle$ had $\alpha^\star \neq 0$, then

$$\sum_{\mathbf{k}\in\mathcal{S}} |\alpha_{\mathbf{k}}|^2 < 1$$

which means that the global state is not normalized

$$\left[\langle\Omega| \otimes \sum_{\mathbf{k}\in\mathcal{S}} \alpha^*_{\mathbf{k}} \langle\mathbf{k}|\right] \hat{\mathcal{A}}^\dagger\hat{\mathcal{A}} \left[|\Omega\rangle \otimes \sum_{\mathbf{k}\in\mathcal{S}} \alpha_{\mathbf{k}} |\mathbf{k}\rangle\right] = \sum_{\mathbf{k}\in\mathcal{S}} |\alpha_{\mathbf{k}}|^2 < 1$$

as can be easily checked.

This result, rather obvious, helps us understand how the exclusion from the sphere works: when adding one electron, its plane waves expansion is made up exclusively of momenta outside the sphere. The price to pay for having electrons all occupying momentum eigenstates is that anything inside the sphere is inaccessible, even as a superposition.

Take now two electrons with zero net momentum. The statement is the following: for a couple built like this, the center of mass part of the wavefunction is a constant (rather obvious) and the relative part is equivalent to a single electron being added to the Fermi sphere – thus everything said above applies. The reason is: the state of the couple can be expressed as a combination of product states

$$|\psi\rangle = \sum_{\mathbf{k}} \alpha_{\mathbf{k}} |\mathbf{k}\rangle \otimes |-\mathbf{k}\rangle$$

since the set $\{|\mathbf{k}\rangle \otimes |-\mathbf{k}\rangle\}$ is a complete orthonormal basis for the subset of the Hilbert space of wavefunctions with zero total momentum. The wavefunction of the couple can be expanded in plane waves

$$\psi(\mathbf{x}_1, \mathbf{x}_2) = \sum_{\mathbf{k}} \alpha_{\mathbf{k}} w_{\mathbf{k}}(\mathbf{x}_1)\, w_{-\mathbf{k}}(\mathbf{x}_2)$$

By construction

$$\psi(\mathbf{x}_1, \mathbf{x}_2) = L^{-D/2} \sum_{\mathbf{k}} \alpha_{\mathbf{k}} w_{\mathbf{k}}(\mathbf{x}_1 - \mathbf{x}_2) = L^{-D/2} \sum_{\mathbf{k}} \alpha_{\mathbf{k}} w_{\mathbf{k}}(\mathbf{x})$$

Also, it must be $\psi(\mathbf{x}_1, \mathbf{x}_2) = \Phi(\mathbf{X})\phi(\mathbf{x})$. Evidently $L^{-D/2} = \Phi(\mathbf{X})$, since $\Phi$ must be eigenstate of $\hat{\mathbf{P}}$ with null eigenvalue. Then

$$\phi(\mathbf{x}) = \sum_{\mathbf{k}} \alpha_{\mathbf{k}} w_{\mathbf{k}}(\mathbf{x})$$

which means: the relative wavefunction has exactly the same Fourier coefficients as the couple expansion. For a single particle added on top of the sphere the components must lie outside. Then the same holds for the couple wavefunction in, thus by this simple argument for the components $\alpha_{\mathbf{k}}$ of the relative wavefunction.

### 4.1.3 Bound states, considering statistics

Consider a system formed by the filled Fermi sphere plus two electrons, as described in the above paragraph. We want to understand how the ground state is made. All the "Fermi electrons" prevent our two interacting electrons from occupying states inside the Fermi sphere, and the Pauli exclusion principle is **the only way** in which the sphere and the electrons outside interact. We assume to have a **weak** interaction at work. Here "weak" means on energy scales much smaller than $\epsilon_F$, the natural energy scale. This also means that the maximum amount of energy exchange through the interaction is "small".

The problem we present is known as the Cooper's problem: it is a toy problem of two electrons outside the sphere feeling the mutual attractive interaction. It is reasonable to assume the potential to be active only nearby the Fermi surface, outside of it: from one point of view the two added electrons are surely more responsive to such attraction than those inside, compressed by statistics; on the other hand, for states very distant from the surface the potential must be irrelevant if compared with the kinetic contribution.

Then we have an interaction shell around the Fermi sphere, in the sense that two electrons in the shell can interact via this potential. It is reasonable to assume it extends only outside the sphere by an amount $\delta k^\star \ll k_F$ and is negligible anywhere outside, so that the shell is *thin*. A pictorial representation of what we are doing is in Fig. 4.1. We also define a energy scale $\delta\epsilon^\star$

$$\epsilon_F + \delta\epsilon^\star \equiv \frac{\hbar^2}{2m}\left(k_F + \delta k^\star\right)^2 \qquad \Longrightarrow \qquad \delta\epsilon^\star \simeq v_F \hbar \delta k^\star$$

with $v_F \equiv \hbar k_F / m$ the Fermi velocity. Note that this is a toy problem and the real potential we will encounter later on will extend both inside and outside the sphere by the same amount. Don't be confused: both problems have a rigid bulk and an interaction shell, with the difference than now we have 2 electrons in the shell and we can study their pairing, while later we will have a multitude, and we will need to think in terms of many-body Physics.

For now, we stick to two electrons outside the shell. Following the line of reasoning exposed in Sec. 4.1.2, we take them with zero total momentum, and follow an argument similar to the Sec. 4.1.1. First, take the complete Schrödinger's Equation

$$\left[\frac{\hat{\mathbf{p}}_1^2}{2m} + \frac{\hat{\mathbf{p}}_2^2}{2m} + V\left(\hat{\mathbf{x}}_1 - \hat{\mathbf{x}}_2\right)\right]\psi(\mathbf{x}_1, \mathbf{x}_2) = E\psi(\mathbf{x}_1, \mathbf{x}_2)$$

with $\psi$ the total eigenfunction and $E$ the energy. From what we know from Sec. 4.1.2, this gives rise in its relative part to the equation

$$\left[\frac{\hat{\mathbf{p}}^2}{2\mu} + V\left(\hat{\mathbf{x}}\right)\right]\phi(\mathbf{x}) = E\phi(\mathbf{x}) \quad \text{with} \quad \phi\left(\mathbf{x}\right) = \sum_{\mathbf{k}} \alpha_{\mathbf{k}} w_{\mathbf{k}}\left(\mathbf{x}\right)$$

since in this situation the kinetic contribution of the center of mass to energy in zero, so all energy enters the relative equation. We take the relative part and proceed with the plane wave expansion and projection, as in Sec. 4.1.1,

$$2\epsilon_{\mathbf{k}}\alpha_{\mathbf{k}} + \sum_{\mathbf{k}'} V_{\mathbf{k}-\mathbf{k}'}\alpha_{\mathbf{k}'} = E\alpha_{\mathbf{k}} \tag{4.1}$$

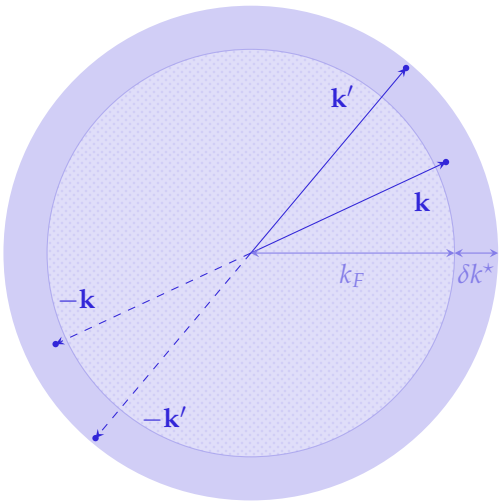



**Figure 4.1:** Representation of the Fermi sphere, of radius $k_F$ in lighter color, and the interaction shell of width $\delta k^\star$ in darker color. The dotted region represents the inaccessible Fermi sphere. The solid dots with momenta $\mathbf{k}$ and $-\mathbf{k}$ are the electrons added to the filled sphere. The couple is scattered to the states $\mathbf{k}'$, $-\mathbf{k}'$.

The potential has nonzero Fourier components only in the shell of width $\delta k^\star$ in Fig. 4.1. The factor 2 in front of $\epsilon_{\mathbf{k}}$, as before, is a consequence of how we defined it. Remember what we discussed in Sec. 4.1.2: this equations makes sense only for components outside the sphere.

Also, given two particles at momenta $\mathbf{k}$ and $-\mathbf{k}$ interact via the potential; the component $V_{\mathbf{q}}$ is a measure of how much the potential is able to change the states of the couple to $\mathbf{k}+\mathbf{q}$ and $-\mathbf{k}-\mathbf{q}$. Notice that the potential of course is active also for couples with nonzero net momentum, but the ground state is to be searched in coupling particles at antipodal sites outside the sphere.

Now we select a specific potential. The simplest potential active only in the shell (which means: felt by particles in the shell and with target states in the shell) is the one constant

$$V_{\mathbf{k}-\mathbf{k}'} = -V_0 A\left(\mathbf{k}\right) A\left(\mathbf{k}'\right)$$

where $A$ is the characteristic function of the shell,

$$A\left(\mathbf{k}\right) \equiv \begin{cases} 0 & \text{if} \quad |\mathbf{k}| > k_F + \delta k^\star \\ 1 & \text{if} \quad 0 < |\mathbf{k}| < k_F + \delta k^\star \\ 0 & \text{if} \quad |\mathbf{k}| < k_F \end{cases}$$

null outside the shell and equal to 1 inside. Notice something very obvious: this definition does **not** mean that the potential has Fourier components only inside the shell; instead it means that it has all the components that can be built by any sum of two vectors, $\mathbf{k}$ and $-\mathbf{k}'$, each of which lies inside the interaction shell. This means that the potential is equipped with all and only the components to scatter two particles inside the shell to two positions still inside the shell. Then:

$$\left(2\epsilon_{\mathbf{k}} - E\right) \alpha_{\mathbf{k}} = V_0 A\left(\mathbf{k}\right) \sum_{\mathbf{k}'} A\left(\mathbf{k}'\right) \alpha_{\mathbf{k}'}$$

It follows:

$$\alpha_{\mathbf{k}} = \frac{V_0 A(\mathbf{k})}{2\epsilon_{\mathbf{k}} - E} \sum_{\mathbf{k}'} A(\mathbf{k}')\, \alpha_{\mathbf{k}'} \tag{4.2}$$

This equation is inconsistent for any $|\mathbf{k}| < k_F$. We multiply both sides by $A(\mathbf{k})$ and sum over $\mathbf{k}$. Of course $A^2 = A$. The coefficient $\sum A\alpha$ can be eliminated by simplification, leaving the self-consistency equation

$$\sum_{|\mathbf{k}|>k_F} \frac{V_0 A(\mathbf{k})}{2\epsilon_{\mathbf{k}} - E} = 1$$

We convert the sum in an integral,

$$1 = \frac{L^D}{(2\pi)^D} \int_{k_F<|\mathbf{k}|<\Lambda} d^D\mathbf{k}\, \frac{V_0 A(\mathbf{k})}{2\epsilon_{\mathbf{k}} - E}$$

Since $\delta k^\star \ll k_F$ by the assumption of thinness of the shell, to this one the integration domain is limited by the function $A$:

$$1 = \frac{L^D}{(2\pi)^D} \int_{|\mathbf{k}|>k_F}^{|\mathbf{k}|<k_F+\delta k^\star} d^D\mathbf{k} \frac{V_0}{2\epsilon_{\mathbf{k}} - E}$$

We make use of the $D$-dimensional density of states $\rho_D(\epsilon)$ to convert this to an energy integral,

$$1 = \int_{\epsilon_F}^{\epsilon_F+\delta\epsilon^\star} d\epsilon\, \frac{\rho_D(\epsilon)}{2} \frac{V_0}{2\epsilon - E}$$

The factor 2 at the denominator is included in order not to count the same states twice. The reason is the following: we are summing over the relative momentum, which is defined as the difference in momenta of the particles. In order to define it, we need to take one as reference; to do so, we need to distinguish them. Even if the states $|\mathbf{k}\rangle \otimes |-\mathbf{k}\rangle$ and $|-\mathbf{k}\rangle \otimes |\mathbf{k}\rangle$ have opposite relative momentum due to our choice of notation, they represent the same state. Thus we must consider half the DoS.

Also, we are using spinless fermions; as it will turn out, including spin the same idea holds: it will turn out that $s$-wave pairing is favorite, thus we will choose as reference the one electron with $\uparrow$ spin and for this reason we will need to count half the states. For $D = 3$, the density of states depends on energy as $\sqrt{\epsilon}$, approximately horizontal around the Fermi energy. Then for any energy in the range of interest we can approximate $\rho_3(\epsilon) \simeq \rho_3(\epsilon_F) \equiv \rho_0$,

$$1 = \frac{\rho_0 V_0}{2} \int_{\epsilon_F}^{\epsilon_F+\delta\epsilon^\star} \frac{d\epsilon}{2\epsilon - E}$$

Since we are near the Fermi surface, it is reasonable to assume that in absence of the interaction both electrons would have energy $\epsilon_F$. Note that here we are assuming to have not only electrons with zero net momentum, but also both **deep in the shell** (which means: far from both the internal and the external shell boundaries). Thus the deviation from the value $2\epsilon_F$ is all binding energy $E^{(\mathrm{b})}$,

$$E \simeq 2\epsilon_F - E^{(\mathrm{b})}$$

It follows

$$1 \simeq \frac{\rho_0 V_0}{2} \int_{\epsilon_F}^{\epsilon_F+\delta\epsilon^\star} \frac{d\epsilon}{2(\epsilon - \epsilon_F) + E^{(\mathrm{b})}}$$

This integral is analogous to the $D = 1$ integral of the Sec. 4.1.1. Then for any given $\kappa$ a binding energy $E^{(b)}$ exists such that the above equation is satisfied. The integral can be solved, giving

$$1 \simeq \frac{\rho_0 V_0}{4} \log\left(1 + \frac{2\delta\epsilon^\star}{E^{(b)}}\right)$$

Then:

$$\frac{2\delta\epsilon^\star}{E^{(b)}} \simeq e^{4/V_0\rho_0} - 1$$

and for small potentials $e^{2/V_0\rho_0}$ dominates, so

$$E^{(b)} \simeq 2\delta\epsilon^\star e^{-4/V_0\rho_0}$$

which is the desired expression for the binding energy. The whole argument holds for $D = 2$, for which the density of states is a constant, and $D = 1$, for which it goes like $\epsilon^{-1/2}$. This section lets us conclude that a "shell interaction" of strength $V_0$ creates electron pairs quite independently of $V_0$, as long as it is not too small. This is an astonishing result: not only it effectively corrects the incoherence of the above section, but it also demonstrates that the pairing of electrons observed in superconductors is a collective phenomenon arising from the presence of an entire electron liquid.

### 4.1.4 Some considerations on the single Cooper pair problem

As seen, the strength of the binding is given by the energy extension of the interaction shell, suppressed exponentially by a factor $V_0^{-1}$. This makes sense: weak interactions produce a strong damping, making it much difficult for the shell thickness to compensate. One can think about the shell width as a measure of *how many* states can couple. On the contrary, strong interactions produce negligible damping, and the strength of the binding is controlled by how much the shell is thick. In such situation

$$\frac{2}{V_0\rho_0} \to 0 \quad \Longrightarrow \quad E^{(b)} \simeq 2\delta\epsilon^\star$$

and this is correct! To break a pair with opposite momenta and just outside the Fermi surface just by pumping energy in the system we must take both electrons outside the shell, which is, give to each of them the same amount of energy $\delta\epsilon^\star$. To imagine to take just one electron outside would require to perturb the system with energy pumping able to distinguish "up from down", thus breaking symmetries.

Identical considerations can be made for $\rho_0$: small densities produce small binding, while big densities strong binding. The reason for that is to searched in Eq. (4.1),

$$2\epsilon_{\mathbf{k}}\alpha_{\mathbf{k}} - V_0 A(\mathbf{k}) \sum_{\mathbf{k}'} A(\mathbf{k}')\alpha_{\mathbf{k}'} = E\alpha_{\mathbf{k}}$$

which gives the equation for energy, using wavefunction normalization,

$$E = \sum_{\mathbf{k}} 2\epsilon_{\mathbf{k}} |\alpha_{\mathbf{k}}|^2 - V_0 \left| \sum_{\mathbf{k}} A(\mathbf{k})\alpha_{\mathbf{k}} \right|^2 \tag{4.3}$$

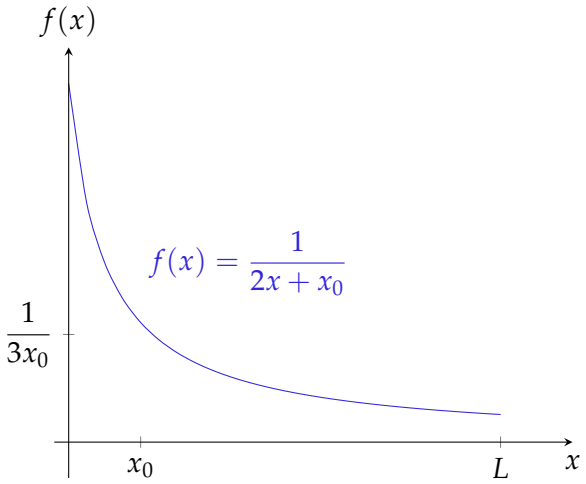



**Figure 4.2:** Plot of the function $f(x)$ as defined in Sec. 4.1.4.

The first term is the normal kinetic one: the wider the wavefunction is distributed in $\mathbf{k}$ space, the bigger its kinetic energy. The wavefunction expansion probabilities $|\alpha_{\mathbf{k}}|^2$ are the distribution over which the mean kinetic energy of the wavepacket is calculated.

For what concerns the second term, consider Eq. (4.2): defining $\xi_{\mathbf{k}} \equiv \epsilon_{\mathbf{k}} - \epsilon_F$, by what we said it reduces to

$$\alpha_{\mathbf{k}} = \frac{V_0 A(\mathbf{k})}{2\xi_{\mathbf{k}} + E^{(\mathrm{b})}} \sum_{\mathbf{k}'} A(\mathbf{k}') \alpha_{\mathbf{k}'}$$

In a narrow shell around the Fermi surface,

$$\xi_{\mathbf{k}} = \frac{\hbar^2}{2m} \left( (k_F + \delta k)^2 - k_F^2 \right) \simeq v_F \hbar \delta k + \mathcal{O}(\delta k)^2$$

as long as $\delta k < k_F$, true for the whole shell of extension $\delta k^\star$. Thus we can pass from $\delta k$ space to $\xi_{\mathbf{k}}$ space, in which the shell has an energy extension of $\delta\epsilon^\star$. Take now the function

$$f(x) \equiv \frac{1}{2x + x_0}$$

represented in Fig. 4.2. The function is limited in space by a cutoff $L \gg x_0$. There $L$ represents $\delta\epsilon^\star$ and $x_0$ the binding energy $E^{(\mathrm{b})}$. Then by what we said

$$\frac{L}{x_0} \simeq \frac{L}{2Le^{-\eta}} = \frac{1}{2} e^{\eta}$$

with $\eta$ some exponent we expect to be bigger than one. The situation in Fig. 4.2 is *a posteriori* reasonable, with $L$ arbitrary chosen as $L = 6x_0$. Now, for such a function it is easy to demonstrate that the majority of the distribution falls in a range $x < 3x_0$. Moreover, for $x < x_0$ the magnitude of the function increases dramatically. This consideration can be mapped on the coefficients $\alpha_{\mathbf{k}}$: the pair wavefunction in our solution has chosen to concentrate the occupation of phase space near the Fermi surface, while states much distant from the surface are lesser populated. Remember that the coefficients have normalization constraints: each individual state is very little populated

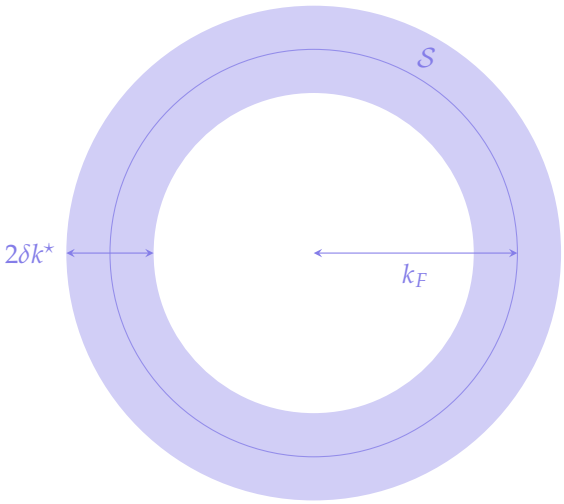



**Figure 4.3:** Sketch of the potential described in Sec. 4.1.5.

– practically empty. But those within a range $E^{(b)}$ from $\epsilon_F$ are more relevant than those more distant.

This makes sense, getting back to the second term on the right side of Eq. (4.3): at first glance, it might seem that the more states are uniformly occupied, the bigger is the term

$$\left|\sum_{\mathbf{k}} A(\mathbf{k})\alpha_{\mathbf{k}}\right|^2$$

and then the more negative is the potential contribution to energy. But to spread the distribution in momentum space too much costs kinetic energy. So, the ground state we have found is the one optimizing the spreading in momentum space of the pair wavefunction.

We can extract two important heuristic lessons from analysis of the pairing toy-problem: first, a single isolated Cooper pair is nothing more esoteric than the best way for a pair wavefunction to distribute in momentum space in order to get the advantage of a contribution $-V_0$ to energy for any state the pair occupies, weighted by the state occupation probability $|\alpha_{\mathbf{k}}|^2$, and not get the disadvantage of an increasing kinetic contribution. The more states are available, the stronger the binding: and that's why $\rho_0$ is there. Second, the binding energy is also a good scale for the distribution of the pair wavefunction in $\mathbf{k}$ space. All these consideration are surely beautiful on a physical intuition level but are very dangerous in the context of a real system, with many electrons feeling the potential: the next section is devoted to this.

### 4.1.5 (In)Stability of the Fermi Sea

So far, we have considered two electrons outside a filled Fermi sphere. The Fermi sphere was just a statistical background preventing the pair from penetrating the surface. Now we do something different and try to work our way out: we have $N$ electrons, provided that the potential acts as

$$V_{\mathbf{k}-\mathbf{k}'} = -V_0 A\left(\mathbf{k}\right) A\left(\mathbf{k}'\right)$$

with

$$A(\mathbf{k}) = \begin{cases} 1 & \text{if} \quad \mathbf{k} \in \mathcal{S} \\ 0 & \text{if} \quad \mathbf{k} \notin \mathcal{S} \end{cases}$$

as in Fig. 4.3. Now the shell $\mathcal{S}$ extends symmetrically inside and outside the Fermi sphere by the same amount $\delta k^\star = \delta\epsilon^\star/\hbar v_F$. So, the number of electrons feeling the interaction is approximately given by

$$N\frac{4\pi k_F^2 \delta k^\star}{\frac{4}{3}\pi k_F^3} = 3N\frac{\delta k^\star}{k_F} = 3N\frac{\delta\epsilon^\star}{\epsilon_F} \ll N$$

The system is not so different from the one before: there is a rigid bulk, this time of radius $k_F - \delta k^\star$, and an interaction shell, this time of width $2\delta k^\star$. If inside the shell the number of electrons was 2, the binding energy should change to

$$E^{(\mathrm{b})} = 2\left[2\delta\epsilon^\star e^{-4/V_0\rho_0}\right]$$

Remember, we are using spinless fermions: if we include spin, for $s$-wave pairing the DoS must be doubled, leading to

$$E^{(\mathrm{b})} = 2\left[2\delta\epsilon^\star e^{-2/V_0\rho_0}\right] \equiv 2\Delta$$

This is in principle the value of the binding energy for coupled electrons in a shell wide $2\delta\epsilon^\star$ excluded from the rigid bulk. The quantity $\Delta$ is the binding energy per electron. But the problem we aim to investigate is different: the number of electrons participating and feeling each other attraction is now very big. Therefore, there is no argument *a priori* that should lead us to conclude that the energy scale $\Delta$ has some relevance. What happens?

The ground state must be equivalent to a perfectly degenerate Fermi gas for $\mathbf{k} \notin \mathcal{S}$: so we expect perfect filling up until the boundary $k_F - \delta k^\star$ and perfect vacuum from $k_F + \delta k^\star$ on. Inside the shell, it is very difficult and misleading to think in terms of individual electrons or isolated couples of electrons, for the phase space availability depends directly on the behavior of all fermions. But some general features of the Cooper pairing mechanism must survive: first, attraction should be enhanced for antipodal electrons. For a single Cooper pair, the best thing to do for the pair was to distribute the wavefunction as much as possible at low wavevector, maintaining a small kinetic contribution to energy and taking the advantage of the many states available in phase space. The resulting wavefunction was made of a distribution spread in momentum space, with major contribution within a range $E^{(\mathrm{b})}$ from the Fermi surface. Said distribution "touched" a lot of states to get potential gain, but did not *fill* many states due to normalization. This means: if other fermions are added, they should not feel any relevant exclusion from the states the couple is spread over.

Before reaching the situation of Fig. 4.3, let us try to understand what should happen if we let the potential penetrates the surface by an infinitesimal amount much smaller than $\delta\epsilon^\star$, ideally letting only the surface fermions interact. Take Fig. 4.4 as reference. Here there are plenty of antipodal electrons. The ground state should find the optimal balance in kinetic distribution of pairs and amount of phase space occupied to extract potential gain.

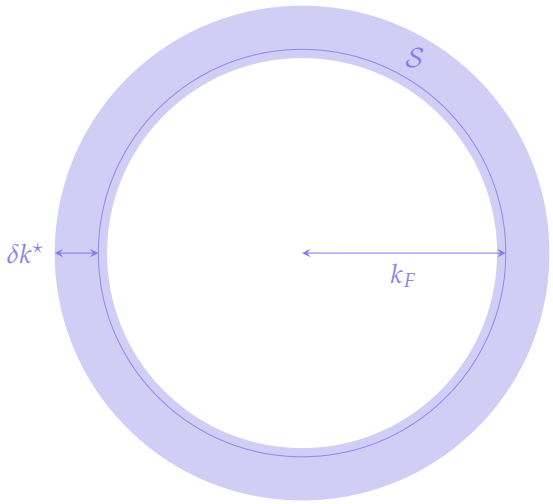



**Figure 4.4:** A shell potential slightly penetrating the sphere. Ideally, only those electrons that would be on the surface in the non-interacting system feel it. Proportions are exaggerated for graphic clarity.

And the number of states available are a lot more than the particle to occupy them. In terms of occupation probabilities, we expect them to be 1 inside the bulk and 0 outside the shell; in the intermediate region, we expect the occupations to be a decreasing function of $\delta k = |\mathbf{k}| - k_F$, due to kinetic optimization. This situation, apart from technicalities, should not be much dissimilar from a system of multiple pairs as we encountered it in Sec. 4.1.3.

Now, we let the interaction shell penetrate a little more the surface. Other electrons are involved now, and the ground state must enhance their pairing. We can still expect a similar situation, with a certain momentum spread of the collective wavefunction and a denser part at lower $|\mathbf{k}|$. Now, the states just outside the rigid bulk are a little more populated (their occupation probabilities are less than 1, surely, but closer than before to unity) and the binding energy is a little bigger because all electrons have access to states at lower kinetic contribution.

If we go on, at some point the states just outside the shell will be filled. This can be understood as if each couple occupies a little portion of the state and is spread in momentum space, but there are a lot of couples. When does this point occur? To answer intuitively this question is now an out-of-reach task for the author. The important thing is that the occupation probabilities can change from 1 to 0 on an energy range much smaller than $\delta\epsilon^\star$, closely related to how many electrons we have and how many states are available. If we let interact the whole sector of $\mathcal{S}$ for

$$k_F - \delta k^\star < |\mathbf{k}| \leq k_F$$

as in Fig. 4.3, it is to be expected that only a small portion of electrons on the surface actually create pairs with a distribution spread in momentum space.

This *naive* and intuitive picture is confirmed by BCS theory: it will turn out in Sec. 5.2 that the occupation probabilities are significant (not too close to 1, neither to 0) over a range

$$-\Delta < \xi_{\mathbf{k}} < \Delta$$

with $\xi_{\mathbf{k}} = \epsilon_{\mathbf{k}} - \epsilon_F$ and $\Delta$ the binding energy per electron as defined as before. Electrons below this shell are compressed by statistics and exclusion, and do not form pairs even if they feel the interaction potential. It is interesting to notice that the width of the distribution has doubled, thus if we think about pairs they have lowered their energy both by acceding to lower kinetic state and by occupying more states to gain potential advantage. Then the conclusion that reveals to be correct (see next Chapter) is that the "binding energy", or its analogous defined as the amount of energy to be given to the system to get two electrons outside of the ground state, is $2\Delta$ – as if we had just two electrons (with spin) in a free interaction shell of double width compared to the one of Fig. 4.1 analyzed in Sec. 4.1.3.

But the fact that the result is the same must not confuse: after all, if the exclusion principle actually compresses many electrons preventing them to form pairs and precluding the access a lot of the intern of the sphere, the accessible part of the interaction shell has not actually doubled. One could imagine that, with respect to the Cooper problem, the interaction shell has penetrated a little the sphere; the electrons, although, are able to form bindings stronger than they would in an equivalently wide and free interaction shell. The key intuition for that, as we will see, is that what we interpret as the binding energy in the many-body problem is actually a collective property, not one of the specific electrons involved.

Even if we cannot see intuitively why the binding energy should be doubled, the fact that it is indeed is a very important clue in the direction of interpreting the role of pairing in this problem. In fact, if the problem could have thought as a system of many independent isolated Cooper pairs (as the one of Sec. 4.1.3) plus a rigid bulk, the binding energy should have remained approximately the same. Instead the way the collective wavefunction is spread in momentum space, and the way the occupation numbers turn out to be, demonstrate that the binding energy is bigger than the one of the single Cooper pair because two mechanisms occur: the Cooper mechanism for allowing pairs populating many states and take advantage of the energy gain, and the self-interacting exclusion principle that shapes the occupation probabilities of the low-momentum states in such a way to let a big number of pairs lower their kinetic energy by an additional amount of order $-\Delta$. It is very difficult and rather futile to search for more intuitive explanations: the key point is that for the collective system is more convenient to have many pairs a little more distributed, but with a lower global kinetic contribution, than a little pairs distributed in the best way possible by themselves.

### 4.1.6 Some final remarks

The next step is to understand how this interaction comes to life at all. We know we are inside a material, a crystal of some kind: it is necessary to screen the Coulomb interaction. It is natural to look for any kind of *attraction* inside the interactions of electrons with the crystal, instead of interactions of electrons within themselves. Moreover, we need some kind of quantum mechanism capable of storing the binding energy of the Cooper pair. As the

electromagnetic field stores the binding energy of an atom with its electrons (pictorially we say they "exchange a photon", although this description is quite misleading), we expect some quantized field of the material to mediate the interaction and store the binding energy of the Cooper pair. In general we may look for any kind of collective excitation of materials – quasiparticles of any kind – but the most general, simple and obvious is the phonon, "the quantum of lattice vibrations".

A very good argument in favor of phonons as mediators of the attraction is the so-called **isotope-effect**: as it is observed in many conventional superconductors, for a given material some constant $\alpha$ exists such that

$$M^{\alpha} T_c = \text{constant}$$

when $M$, the ions mass, is changed. This is achieved by substituting the ions with isotopes. Any phenomenon sensible to the changing of the mass of the constituents of the lattice is known to be related to the lattice dynamics, which is, its phononic properties.

## 4.2 HOW THE PHONON PAIRING WORKS

The phononic field is quantized in crystals: a good source about such quantization is *Solid State Physics* [2] by Grosso and Pastori Parravicini; we will just rapidly sketch the essential concepts. Consider a general crystal with $n$ atoms per unit cell. The index $\lambda = 1, \cdots, n$ varies over the cell atoms. The ion in position $\lambda$ has mass $M_\lambda$. The hamiltonian describing phonons is an harmonic one, with dispersion $\Omega_{\mathbf{k}\nu}$

$$\hat{H}^{(\mathrm{p})} = \sum_{\mathbf{k}} \sum_{\nu} \hbar \Omega_{\mathbf{k}\nu} \left[ \hat{a}^{\dagger}_{\mathbf{k}\nu} \hat{a}_{\mathbf{k}\nu} + \frac{\mathbb{1}}{2} \right] \tag{4.4}$$

with $\hat{a}^{\dagger}_{\mathbf{k}\nu}$ the creation operator for an oscillation of wavevector $\mathbf{k}$ in band $\nu$, and $\hat{a}_{\mathbf{k}\nu}$ the related destruction operator. Such operators obey Bose commutation rules:

$$\left[ \hat{a}_{\mathbf{k}\nu}, \hat{a}^{\dagger}_{\mathbf{h}\mu} \right] = \delta_{\mathbf{k}\mathbf{h}} \delta_{\nu\mu}$$

The next section is devoted to a very rapid description of how this hamiltonian is obtained.

### 4.2.1 Born-Oppenheimer quantization

In order to recover the hamiltonian in Eq. (4.4), **Born-Oppenheimer approximation** is used: the motion of ions is assumed to be much slower than those of electrons, thus their positions $\mathbf{R}_{i\lambda}$ (with $i$ the index for the $i$-th cell and $\lambda$ the index for the atom inside the cell) are assumed as a parameter for the electronic problem and this last is solved, giving out a parametric solution for the energy $E[\mathbf{R}]$. Here $\mathbf{R}$ represents the vector collecting all vectors $\mathbf{R}_{i\lambda}$.

Throughout a variational approach over the parametric solution $E[\mathbf{R}]$ around a energy minimum (a certain ions configuration $\mathbf{R}^{(0)}$) the energy is reduced

to a quadratic expression in terms of the ions displacement from rest positions, $\mathbf{u}_{i\lambda} \equiv \mathbf{R}_{i\lambda} - \mathbf{R}_{i\lambda}^{(0)}$. Including also the kinetic contribution, the overall result is

$$\hat{H}^{(\mathrm{p})} = \sum_i \sum_\lambda \left\{ \frac{|\hat{\mathbf{p}}_{i\lambda}|^2}{2M_\lambda} + \frac{1}{2}\sqrt{M_\lambda}\hat{\mathbf{u}}_{i\lambda} \cdot \sum_j \sum_\eta \left[ \frac{1}{\sqrt{M_\lambda M_\eta}} \frac{\partial^2 E\,[\mathbf{R}]}{\partial \mathbf{R}_{i\lambda} \partial \mathbf{R}_{j\eta}} \right] \cdot \sqrt{M_\eta}\hat{\mathbf{u}}_{j\eta} \right\}$$

The mass factors are included to simplify the following steps. Now: performing a Discrete Fourier Transform over the above equation and recognizing the dynamical matrix $\mathcal{D}$ in the (transformed) matrix in the quadratic term, one gets

$$\hat{H}^{(\mathrm{p})} = \sum_{\mathbf{k}} \sum_\lambda \left\{ \frac{\hat{\mathbf{p}}_{\mathbf{k}\lambda} \cdot \hat{\mathbf{p}}_{-\mathbf{k}\lambda}}{2M_\lambda} + \frac{1}{2}\sqrt{M_\lambda}\hat{\mathbf{u}}_{\mathbf{k}\lambda} \cdot \sum_\eta \mathcal{D}_{\lambda\eta}\,(\mathbf{k}) \cdot \sqrt{M_\eta}\hat{\mathbf{u}}_{-\mathbf{k}\eta} \right\}$$

Diagonalization of $\mathcal{D}$ is possible. Such diagonalization is intended over the sublattice index $\lambda$, and provides a set of bands parameterized by the index $\nu = 1, \cdots, n$,

$$\mathcal{D}\,(\mathbf{k})\,\mathbf{w}_{\mathbf{k}}^{(\nu)} = \Omega_{\mathbf{k}\nu}^2 \mathbf{w}_{\mathbf{k}}^{(\nu)}$$

The eigenvectors

$$\left[\mathbf{w}_{\mathbf{k}}^{(\nu)}\right]_\lambda$$

are $n$ in number, and each one has $n$ components and is a specific mixed oscillations of the sublattices. The eigenvalues at fixed wavevector $\Omega_{\mathbf{k}\nu}^2$ are also $n$ in general. Now: define $\mathrm{U}_{\mathbf{k}\nu}$ such that

$$\mathrm{U}_{\mathbf{k}\nu} \equiv \sum_\lambda \mathbf{u}_{\mathbf{k}\lambda} \cdot \mathbf{w}_{\mathbf{k}\lambda}^{(\nu)}$$

Remember: two scalar products are involved here. One is the spatial one, indicated by "$\cdot$". The other is the $\lambda$ scalar product, indicated by the sum and the index contraction. Then $\mathrm{U}_{\mathbf{k}\nu}$ is defined as the projection of the vector $\mathbf{u}_{\mathbf{k}}$ (the vector which has as components the displacement of the $\lambda$-th sublattice) onto the eigenvector $\mathbf{w}_{\mathbf{k}}^{(\nu)}$. It follows obviously

$$\mathbf{u}_{\mathbf{k}} = \sum_\nu \mathrm{U}_{\mathbf{k}\nu} \mathbf{w}_{\mathbf{k}}^{(\nu)} \tag{4.5}$$

Analogously

$$\mathrm{P}_{\mathbf{k}\nu} \equiv \sum_\lambda \mathbf{p}_{\mathbf{k}\lambda} \cdot \mathbf{w}_{\mathbf{k}\lambda}^{(\nu)} \qquad \Longrightarrow \qquad \mathbf{p}_{\mathbf{k}} = \sum_\nu \mathrm{P}_{\mathbf{k}\nu} \mathbf{w}_{\mathbf{k}}^{(\nu)}$$

Using orthonormality of the eigenvectors one gets by substitution

$$\hat{H}^{(\mathrm{p})} = \sum_{\mathbf{k}} \sum_\lambda \left\{ \frac{\hat{\mathrm{P}}_{\mathbf{k}\lambda}\hat{\mathrm{P}}_{-\mathbf{k}\lambda}}{2M_\lambda} + \frac{1}{2} M_\lambda \Omega_{\mathbf{k}\lambda}^2 \hat{\mathrm{U}}_{\mathbf{k}\lambda}\hat{\mathrm{U}}_{-\mathbf{k}\lambda} \right\}$$

This one is the hamiltonian of $n$ independent harmonic oscillators. Each oscillator has a specific frequency $\Omega_{\mathbf{k}\lambda}$ at fixed wavevector. Then an expression for the displacements and momenta in terms of the Bose operators is possible,

$$\hat{\mathrm{U}}_{\mathbf{k}\lambda} = \sqrt{\frac{\hbar}{2M_\lambda \Omega_{\mathbf{k}\lambda}}} \left( \hat{a}_{\mathbf{k}\lambda}^\dagger + \hat{a}_{\mathbf{k}\lambda} \right) \qquad \hat{\mathrm{P}}_{\mathbf{k}\lambda} = \sqrt{\frac{\hbar}{2M_\lambda \Omega_{\mathbf{k}\lambda}}} \left( \hat{a}_{\mathbf{k}\lambda}^\dagger - \hat{a}_{\mathbf{k}\lambda} \right)$$

It is of our particular interest the expression for the three-dimensional displacement, thus we reconstruct $\mathbf{u_k}$ through Eq. (4.5),

$$\hat{\mathbf{u}}_{\mathbf{k}} = \sum_{\lambda} \hat{\mathrm{U}}_{\mathbf{k}\lambda} \mathbf{w}_{\mathbf{k}}^{(\lambda)} = \sum_{\lambda} \sqrt{\frac{\hbar}{2M_\lambda \Omega_{\mathbf{k}\lambda}}} \left( \hat{a}_{\mathbf{k}\lambda}^{\dagger} + \hat{a}_{\mathbf{k}\lambda} \right) \mathbf{w}_{\mathbf{k}}^{(\lambda)}$$

Notice that for simple crystals ($n = 1$) the $\lambda$ sum vanishes. This will be applied later on.

### 4.2.2 Electron-phonon interactions

The key idea to model electron-phonon interactions is that ions do not get "too far" from the equilibrium positions $\mathbf{R}_{i\lambda}^{(0)}$. Then, coherently with Born-Oppenheimer approximation, the total hamiltonian

$$\hat{H} = \hat{H}^{(\mathrm{e})} + \hat{H}^{(\mathrm{ep})} + \hat{H}^{(\mathrm{p})}$$

contains an interaction term $H^{(\mathrm{ep})}$ which depends on the electron positions $\mathbf{r}_\alpha$ and the ions positions $\mathbf{R}_{i\lambda}$ and can be expanded as

$$\hat{H}^{(\mathrm{ep})}\left(\mathbf{r}, \mathbf{R}\right) \simeq \hat{H}^{(\mathrm{ep})}\left(\mathbf{r}, \mathbf{R}^{(0)}\right) + \sum_{i} \boldsymbol{\nabla}_{\mathbf{R}_{i\lambda}} \hat{H}^{(\mathrm{ep})}\big|_{\mathbf{R}_{i\lambda}=\mathbf{R}_{i\lambda}^{(0)}} \cdot \hat{\mathbf{u}}_{i\lambda} + \cdots$$

The above expression includes phonon operators and electron operators, both indicated by the "hat". We already have $\hat{\mathbf{u}}_{i\lambda}$ from last section. The first term is identically null, because the contribution arising from rest positions of the ions is already included in the bare phonon hamiltonian. The interaction is Coulomb-like, thus

$$\boldsymbol{\nabla}_{\mathbf{R}_{i\lambda}} \hat{H}^{(\mathrm{ep})}\big|_{\mathbf{R}_{i\lambda}=\mathbf{R}_{i\lambda}^{(0)}} = \boldsymbol{\nabla}_{\mathbf{R}_{i\lambda}} \sum_{\alpha} v\left(\left|\hat{\mathbf{r}}_\alpha - \hat{\mathbf{R}}_{i\lambda}^{(0)}\right|\right) \qquad \text{with} \qquad v(r) = -\frac{Ze^2}{r}$$

Moving to a field formalism, this single-particle operator becomes

$$\sum_{\alpha} v\left(\left|\hat{\mathbf{r}}_\alpha - \mathbf{R}_{i\lambda}^{(0)}\right|\right) \to \int_{\mathbb{R}^D} d^D\mathbf{r}\, \hat{\psi}^{\dagger}(\mathbf{r}) v\left(\left|\mathbf{r} - \mathbf{R}_{i\lambda}^{(0)}\right|\right) \hat{\psi}(\mathbf{r})$$

where $\hat{\psi}$ is the electronic field operator, acting on the electrons Hilbert space. The algebraic notation "†" now is intended upon such space. Such field can be expanded through a complete set of single particle Bloch spin-wavefunctions,

$$\hat{\psi}(\mathbf{r}) = \sum_{\mathbf{k}\sigma} \sum_{\lambda} \varphi_{\mathbf{k}\lambda}(\mathbf{r}) \hat{c}_{\mathbf{k}\sigma\lambda} \qquad \text{with} \qquad \varphi_{\mathbf{k}\lambda}(\mathbf{r}) = \frac{1}{L^{D/2}} u_{\mathbf{k}\lambda}(\mathbf{r}) e^{i\mathbf{k}\cdot\mathbf{r}}$$

and $u$ a lattice-periodic function. For simplicity we consider simple crystals, for which $n = 1$. For such crystals the dispersion of phonons has an energy extension of approximately $\hbar\omega_D$, with $\omega_D$ the Debye frequency; in the language of the Sec. 4.1.3, $\delta\epsilon^\star \simeq \hbar\omega_D$. For composite crystals, due to the presence of optical bands and polarization effects, the argument must be corrected (sometimes, fatally).

A long and tedious calculation no one can convince the author to perform leads the rather simple result

$$\hat{H}^{(\mathrm{ep})} = \sum_{\mathbf{k}\sigma} \sum_{\mathbf{q}} g_{\mathbf{k},\mathbf{q}} \hat{c}_{\mathbf{k}+\mathbf{q}\sigma}^{\dagger} \hat{c}_{\mathbf{k}\sigma} \left( \hat{a}_{\mathbf{q}} + \hat{a}_{-\mathbf{q}}^{\dagger} \right) \tag{4.6}$$

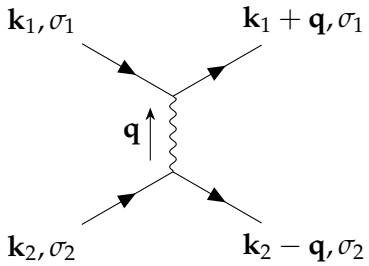



**Figure 4.5:** Second order process: scattering of two electrons of momenta $\mathbf{k}_1$ and $\mathbf{k}_2$ with exchange of momentum $\mathbf{q}$. No spin interaction is considered in our theory.

with

$$g_{\mathbf{k},\mathbf{q}} \equiv \left(-i\mathbf{q}\cdot\mathbf{w}_{\mathbf{q}} v_{\mathbf{q}}\right)\sqrt{\frac{\hbar}{2NM\Omega_{\mathbf{q}}}}\left\langle u^*_{\mathbf{k}+\mathbf{q}}(\mathbf{r})u_{\mathbf{k}}(\mathbf{r})\right\rangle_c \tag{4.7}$$

where $v_{\mathbf{q}}$ is the $\mathbf{q}$ component of the Coulomb potential, $N$ is the number of cells and $\langle\cdots\rangle_c$ is the spatial mean over a single cell. Aside from how $g$ turns out to be, we are interested in the general form of Eq. (4.6). It is made of two contributions: the first contribution describes the scattering of one electron with annihilation of a phonon with wavevector $\mathbf{q}$; the second is the same process, but with creation of one phonon of inverse wavevector. Both have amplitude $g_{\mathbf{k},\mathbf{q}}$. To get Eq. (4.7), the necessary assumptions are:

1. phonons mediated interaction cannot couple different bands. This is reasonable because interband spacing is often bigger than any phononic energy scale;
2. the Fermi level is crossed by one single band;
3. we ignore Umklapp scattering and consider scattering only near the Fermi surface;
4. we assume a monoatomic simple crystal.

In common situations **at very low temperature** the phonons density is close to zero: thus the rate of these first order processes is negligible. The relevant processes are those at second order, in which the phonon is a mediator, as the one represented in Fig. 4.5.

### 4.2.3 The effective interaction hamiltonian: Schrieffer-Wolff transformation

We now have something of the form

$$\hat{H} = \left(\hat{H}^{(\mathrm{e})} + \hat{H}^{(\mathrm{p})}\right) + \left(\hat{H}^{(\mathrm{ep})}\right)$$

where the first term is the non-interacting, leading contribution, while the second is of order $g$ and is a perturbative correction. We temporarily introduce a fictitious perturbative parameter $\lambda$ in order to distinguish different orders, and rewrite the hamiltonian as

$$\hat{H} = \hat{H}_0 + \lambda\hat{H}_1$$

where the (e) and (p) contributions are in $\hat{H}_0$, while the (ep) contribution is in $\lambda\hat{H}_1$. The key idea is the following: performing a unitary transformation upon the hamiltonian does not change the spectrum of the problem, thus we can search from some operator $\hat{S}$ and use it to transform $\hat{H}$,

$$\hat{\boldsymbol{H}} = e^{-\lambda\hat{S}}\hat{H}e^{\lambda\hat{S}}$$

We call the bold hamiltonian $\hat{\boldsymbol{H}}$ "effective hamiltonian". For the transformation to be unitary $\hat{S}$ needs to be anti-hermitian, $\hat{S}^\dagger = -\hat{S}$. Apart from this, we still have maximum freedom on $\hat{S}$. Notice that we used the perturbative parameter as the scale parameter of the unitary transformation: in our problem, $\lambda$ takes the role of $g$. Inserting the perturbative expansion and making use of the Baker-Campbell-Hausdorff (BCH) formula we get

$$\begin{aligned} e^{-\lambda\hat{S}}\left(\hat{H}_0+\lambda\hat{H}_1\right)e^{\lambda\hat{S}} = &\left(\hat{H}_0+\lambda\hat{H}_1\right) \\ &+\left[\left(\hat{H}_0+\lambda\hat{H}_1\right),\lambda\hat{S}\right] \\ &+\frac{1}{2}\Big[\left[\left(\hat{H}_0+\lambda\hat{H}_1\right),\lambda\hat{S}\right],\lambda\hat{S}\Big]+\cdots \end{aligned} \tag{4.8}$$

We collect the terms of order $\lambda$,

$$\lambda\left(\hat{H}_1+\left[\hat{H}_0,\hat{S}\right]\right)$$

and those of order $\lambda^2$

$$\lambda^2\left(\left[\hat{H}_1,\hat{S}\right]+\frac{1}{2}\Big[\left[\hat{H}_0,\hat{S}\right],\hat{S}\Big]\right)$$

These last terms are those of our interest: at second order we have phonon mediation. Now: if we impose $\mathcal{O}(\lambda)$ terms to be null,

$$\hat{H}_1+\left[\hat{H}_0,\hat{S}\right]\overset{!}{=}0 \tag{4.9}$$

we also get for those at $\mathcal{O}(\lambda^2)$

$$\lambda^2\left(\left[\hat{H}_1,\hat{S}\right]+\frac{1}{2}\Big[\left[\hat{H}_0,\hat{S}\right],\hat{S}\Big]\right)=\frac{\lambda^2}{2}\left[\hat{H}_1,\hat{S}\right]$$

then in Eq. (4.8) we have, up to second order,

$$e^{-\lambda\hat{S}}\left(\hat{H}_0+\lambda\hat{H}_1\right)e^{\lambda\hat{S}}=\hat{H}_0+\frac{\lambda^2}{2}\left[\hat{H}_1,\hat{S}\right]+\cdots \tag{4.10}$$

Then we succeeded in reducing a $\mathcal{O}(\lambda)$ interaction to a $\mathcal{O}(\lambda^2)$ one via a unitary transformation, subject to the constraints of anti-hermiticity and of Eq. (4.9). All this procedure is often called **Schrieffer-Wolff transformation**. It can be shown using standard properties of bosonic and fermionic operators that a suitable $\hat{S}$ exists in a form mimicking Eq. (4.6)

$$\hat{S}=\sum_{\mathbf{k}\sigma}\sum_{\mathbf{q}}g_{\mathbf{k},\mathbf{q}}\hat{c}^\dagger_{\mathbf{k}+\mathbf{q}\sigma}\hat{c}_{\mathbf{k}\sigma}\left(\frac{\hat{a}_{\mathbf{q}}}{(\epsilon_{\mathbf{k}}-\epsilon_{\mathbf{k}+\mathbf{q}})+\hbar\Omega_{\mathbf{q}}}+\frac{\hat{a}^\dagger_{-\mathbf{q}}}{(\epsilon_{\mathbf{k}}-\epsilon_{\mathbf{k}+\mathbf{q}})-\hbar\Omega_{\mathbf{q}}}\right) \tag{4.11}$$

Such operator is anti-hermitian:

$$\left[\sum_{\mathbf{q}} g_{\mathbf{k},\mathbf{q}} \hat{c}^\dagger_{\mathbf{k}+\mathbf{q}\sigma} \hat{c}_{\mathbf{k}\sigma} \left( \frac{\hat{a}_{\mathbf{q}}}{(\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}+\mathbf{q}}) + \hbar\Omega_{\mathbf{q}}} + \frac{\hat{a}^\dagger_{-\mathbf{q}}}{(\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}+\mathbf{q}}) - \hbar\Omega_{\mathbf{q}}} \right)\right]^\dagger$$
$$= \sum_{\mathbf{q}} g^*_{\mathbf{k},\mathbf{q}} \hat{c}^\dagger_{\mathbf{k}\sigma} \hat{c}_{\mathbf{k}+\mathbf{q}\sigma} \left( \frac{\hat{a}^\dagger_{\mathbf{q}}}{(\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}+\mathbf{q}}) + \hbar\Omega_{\mathbf{q}}} + \frac{\hat{a}_{-\mathbf{q}}}{(\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}+\mathbf{q}}) - \hbar\Omega_{\mathbf{q}}} \right)$$

From direct inspection of Eq. (4.7) we see $g^*_{\mathbf{k},\mathbf{q}} = g_{\mathbf{k}-\mathbf{q},-\mathbf{q}}$. Thanks to the presence of the symmetric sum over $\mathbf{q}$, we may change sign of the variable inside the sum

$$\sum_{\mathbf{q}} g_{\mathbf{k}-\mathbf{q},\mathbf{q}} \hat{c}^\dagger_{\mathbf{k}\sigma} \hat{c}_{\mathbf{k}-\mathbf{q}\sigma} \left( \frac{\hat{a}^\dagger_{-\mathbf{q}}}{(\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}-\mathbf{q}}) + \hbar\Omega_{-\mathbf{q}}} + \frac{\hat{a}_{-\mathbf{q}}}{(\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}-\mathbf{q}}) - \hbar\Omega_{-\mathbf{q}}} \right)$$

The dispersion exhibits crystal symmetry, $\Omega_{\mathbf{q}} = \Omega_{-\mathbf{q}}$. Finally, we are summing over the whole Brillouin Zone, which means we can shift $\mathbf{k}$ by an amount $+\mathbf{q}$,

$$\sum_{\mathbf{q}} g_{\mathbf{k},\mathbf{q}} \hat{c}^\dagger_{\mathbf{k}+\mathbf{q}\sigma} \hat{c}_{\mathbf{k}\sigma} \left( \frac{\hat{a}^\dagger_{-\mathbf{q}}}{(\epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}}) + \hbar\Omega_{\mathbf{q}}} + \frac{\hat{a}_{-\mathbf{q}}}{(\epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}}) - \hbar\Omega_{\mathbf{q}}} \right) = -\hat{S}$$

as requested. Moreover, defined as in Eq. (4.11) such $\hat{S}$ satisfies Eq. (4.9), with the following definitions

$$\hat{H}_0 \equiv \sum_{\mathbf{k}\sigma} \epsilon_{\mathbf{k}} \hat{c}^\dagger_{\mathbf{k}\sigma} \hat{c}_{\mathbf{k}\sigma} + \sum_{\mathbf{q}} \hbar\Omega_{\mathbf{q}} \hat{a}^\dagger_{\mathbf{q}} \hat{a}_{\mathbf{q}}$$
$$\hat{H}_1 \equiv \sum_{\mathbf{k}\sigma} \sum_{\mathbf{q}} g_{\mathbf{k},\mathbf{q}} \hat{c}^\dagger_{\mathbf{k}+\mathbf{q}\sigma} \hat{c}_{\mathbf{k}\sigma} \left( \hat{a}_{\mathbf{q}} + \hat{a}^\dagger_{-\mathbf{q}} \right)$$

where for the phononic part in $\hat{H}_0$ we neglected the constant shift.

We are ready to calculate the second order expression for $\hat{\boldsymbol{H}}$ via Eq. (4.10). As always, the calculation is long enough for the lazy author to acrobatically jump to the result

$$\begin{aligned} \hat{\boldsymbol{H}} = & \sum_{\mathbf{k}\sigma} \epsilon_{\mathbf{k}} \hat{c}^\dagger_{\mathbf{k}\sigma} \hat{c}_{\mathbf{k}\sigma} \\ & + \sum_{\mathbf{k}_1\sigma_1} \sum_{\mathbf{k}_2\sigma_2} \sum_{\mathbf{q}} \left[ \frac{\hbar\Omega_{\mathbf{q}}}{(\epsilon_{\mathbf{k}_1+\mathbf{q}} - \epsilon_{\mathbf{k}_1})^2 - (\hbar\Omega_{\mathbf{q}})^2} \right. \\ & \left. \cdot \left( g_{\mathbf{k}_1,\mathbf{q}} \hat{c}^\dagger_{\mathbf{k}_1+\mathbf{q}\sigma_1} \hat{c}_{\mathbf{k}_1\sigma_1} \right) \left( g_{\mathbf{k}_2,-\mathbf{q}} \hat{c}^\dagger_{\mathbf{k}_2-\mathbf{q}\sigma_2} \hat{c}_{\mathbf{k}_2\sigma_2} \right) \right] \end{aligned} \tag{4.12}$$

We found an expression for $\hat{\boldsymbol{H}}$. The next section is devoted to analyzing its effects.

### 4.2.4 Effective form of the phonon-assisted interaction

Let's comment the second term of Eq. (4.12). First, it involves four fermionic operators, and describes exactly the process in Fig. 4.5. The process amplitude has two $g$ contributions, one for the electron gaining momentum $\mathbf{q}$,

the other for the electron losing it. The amplitude is modulated by a quasi-lorentzian factor (notice the minus sign) of width $\hbar\Omega_{\mathbf{q}}$. For a fixed $\mathbf{q}$, this modulation tends to favor those processes "inside the peak".

By energy conservation $\epsilon_{\mathbf{k}_1+\mathbf{q}} - \epsilon_{\mathbf{k}_1} = -\epsilon_{\mathbf{k}_2-\mathbf{q}} + \epsilon_{\mathbf{k}_2}$, so the said processes are those such that

$$\left|\epsilon_{\mathbf{k}_1+\mathbf{q}} - \epsilon_{\mathbf{k}_1}\right| = \left|\epsilon_{\mathbf{k}_2-\mathbf{q}} - \epsilon_{\mathbf{k}_2}\right| < \hbar\Omega_{\mathbf{q}} \tag{4.13}$$

Now: in the language of Quantum Field Theory the $g$ factors are the vertices contributions, while the lorentzian factor is the propagator contribution – the potential. We can make an important simplification: suppose the Bloch eigenstates to be essentially plane waves, a reasonable description for many simple crystals. Then in Eq. (4.7) we have

$$\left\langle u^*_{\mathbf{k}+\mathbf{q}}(\mathbf{r}) u_{\mathbf{k}}(\mathbf{r}) \right\rangle_c \simeq 1 \quad \Longrightarrow \quad g_{\mathbf{k},\mathbf{q}} \simeq g_{\mathbf{q}}$$

The vertex strength is practically independent of the electron momentum $\mathbf{k}$. Then

$$\hat{H} = \sum_{\mathbf{k}\sigma} \epsilon_{\mathbf{k}} \hat{c}^\dagger_{\mathbf{k}\sigma} \hat{c}_{\mathbf{k}\sigma} + \sum_{\mathbf{q}} |g_{\mathbf{q}}|^2 \sum_{\mathbf{k}_1\sigma_1} \sum_{\mathbf{k}_2\sigma_2} \mathcal{M}_{\mathbf{k}_1,\mathbf{q}} \left(\hat{c}^\dagger_{\mathbf{k}_1+\mathbf{q}\sigma_1} \hat{c}_{\mathbf{k}_1\sigma_1}\right) \left(\hat{c}^\dagger_{\mathbf{k}_2-\mathbf{q}\sigma_2} \hat{c}_{\mathbf{k}_2\sigma_2}\right)$$

where the modulation is given by

$$\mathcal{M}_{\mathbf{k}_1,\mathbf{q}} = \frac{\hbar\Omega_{\mathbf{q}}}{\left(\epsilon_{\mathbf{k}_1+\mathbf{q}} - \epsilon_{\mathbf{k}_1}\right)^2 - \left(\hbar\Omega_{\mathbf{q}}\right)^2}$$

We could also have defined $\mathcal{M}$ via $\mathbf{k}_2$. It is irrelevant, because of what we are going to assume. Under this approximation, the energies in Eq. (4.13) also assume a very special meaning: they define the energy region for which the interaction potential becomes negative – which means, **attractive**. And the potential is practically zero everywhere outside due to damping, just as in in Sec. 4.1.3. This suggests a bound state is possible: we are on the verge of finding quite formally Cooper pairs mediated by phonons.

Now it is the right time to ask: what does this all mean? How phonons actually let electrons form bound states? What is the answer to life, the universe, and everything? 42, of course. Apart from that, a pictorial, semi-classical image one can use to understand this behavior is the following:

> **D. L. Goodstein**: ***States of Matter*** ([4] @ 5.3.c)
>
> The direct interaction between electrons is, of course, the Coulomb electrostatic force, which is repulsive. That force is largely screened by the positively charged medium in which each electron finds itself, but how does an attractive force between them arise?
>
> [...] Imagine two people on an old, sagging, nonlinear mattress. They tend to roll toward the middle, even if they don't like each other. That is, there is an attractive interaction. The cause of this interaction (remember, we are ignoring all features that are not important to superconductivity) is that the people create distortions in the mattress, and the distortions are attracted to each other and try to merge. The attractive interaction between electrons occurs in somewhat the same way. The negatively charged electrons cause distortions of the lattice of positive ions, and the real attraction occurs between these distortions.

> Pursuing the idea further, the electrons in the metal do not stand still but rather zip through the lattice at something like the Fermi velocity. The ions are attracted to the electrons but, owing to their large mass, move very slowly compared to the much lighter electrons. By the time the ions respond the electron is long gone, but it has, in effect, left behind a trail of positive charge, which is the lattice distortion we mentioned above. Another electron, traversing the same path, would find that its way had been prepared with the positive charge that it finds so attractive. We can imagine, if we wish, that the first electron created a phonon, which the second happily absorbs. This is the nature of the interaction between the two electrons. Notice that the interaction is strongest if the two electrons traverse exactly the same path – that is, if they have, say, equal and opposite momenta.

To imagine what a single, isolated Cooper pair looks like, we should not think about a strongly localized particle-like object. A Cooper pair is a large object containing two quasi-free electrons with opposite momenta, moving apparently as unbounded, but correlated on a statistical level – which means, the two change their momentum due to lattice interaction maintaining a zero net momentum. This description works for the single pair of Sec. 4.1.3.

Getting back to the Schrieffer-Wolff transformation used, the most important point to understand about this whole procedure is that we assumed the phononic populations to be negligible, which means, the lattice to vibrate in a few controlled ways, those supporting the passage of Cooper pairs. If the temperature rises we expect the thermal contribution to increase, thus the number of "free" phonons to get bigger. Then the first order contributions (energy loss by phonon emission, energy gain by phonon absorption) and spurious second order contributions (electron-phonon scattering with electronic propagator) are no longer negligible and the superconducting contributions lose importance.

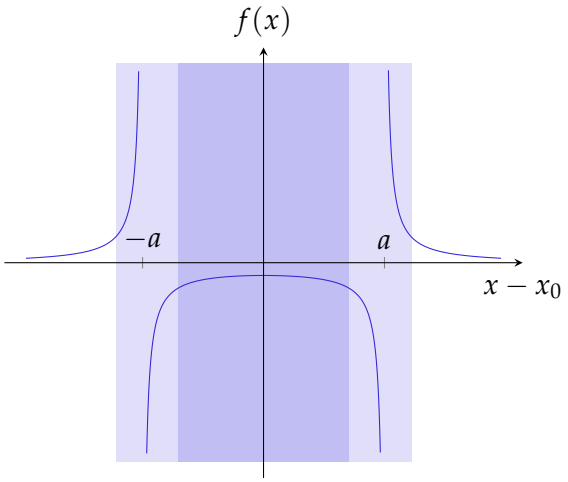
### 4.2.5 A few assumptions, and Cooper pairs

We now simplify further more the hamiltonian in our search of the ground state. Here we analyze on intuitive grounds how a Cooper pair looks like; however, what we find in this chapter is **not** the system ground-state wavefunction, which is strongly correlated between different Cooper pairs. We will analyze it later. Because of arguments similar to those exposed in Sec. 4.1.3, we can limit to pairs with opposite momenta,

$$\mathbf{k}_1 \stackrel{!}{=} -\mathbf{k}_2$$

For a process moving a pair from states $|\mathbf{k}\rangle \otimes |-\mathbf{k}\rangle$ to $|\mathbf{k}+\mathbf{q}\rangle \otimes |-\mathbf{k}-\mathbf{q}\rangle$ the amplitude has the form

$$\begin{aligned} |g_{\mathbf{q}}|^2 \mathcal{M}_{\mathbf{k},\mathbf{q}} &= |g_{\mathbf{q}}|^2 \frac{\hbar\Omega_{\mathbf{q}}}{\left(\epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}}\right)^2 - \left(\hbar\Omega_{\mathbf{q}}\right)^2} \\ &= \hbar\Omega_{\mathbf{q}} \frac{\kappa_{\mathbf{q}}^2 \left(\hbar\Omega_{\mathbf{q}}\right)^2}{\left(\epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}}\right)^2 - \left(\hbar\Omega_{\mathbf{q}}\right)^2} \equiv V_{\mathbf{k},\mathbf{q}} \end{aligned}$$




$$f(x) = \frac{\kappa^2 a^2}{(x - x_0)^2 - a^2}$$
$$f(x_0) = \kappa^2$$
$$\kappa < 1$$

**Figure 4.6:** Plot of the quasi-lorentzian function $f(x)$, which resembles the modulation $\mathcal{M}$ for some fixed $\mathbf{q}$. Two parameter are included in the definition of $f$: $x_0$, which represents the Fermi energy, and $a$, which takes the part of $\hbar\Omega_{\mathbf{q}}$. For graphic clarity the function is plotted shifting the axis to $x_0$. As described in Sec. 4.2.5, the lighter region indicates the regions of mathematical divergence of the function, while the darker region represents a sort of *plateau* region. In the remaining region the function is close to zero.

where we defined

$$\kappa_{\mathbf{q}} \equiv \frac{|g_{\mathbf{q}}|^2}{\hbar\Omega_{\mathbf{q}}}$$

We can assume $\kappa_{\mathbf{q}} < 1$ for any exchanged vector of interest: in building the Schrieffer-Wolff transformation we used that the (ep) hamiltonian, of energy scale $g$, is perturbative with respect to the (p) hamiltonian, of energy scale $\hbar\Omega$.

Look at Fig. 4.6: it is the plot of the above function, for a fixed exchanged vector $\mathbf{q}$. This function clearly can be approximated as flat in the dark region and null in the white regions, while the lighter regions are "small" and can be neglected. Then the mathematical structure reminds of the potential defined in Sec. 4.1.3.

Be careful, now: $\hbar\Omega_{\mathbf{q}}$ ($a$ in the figure) controls the width of the darker region $\hbar\Omega_{\mathbf{q}}$; $\kappa_{\mathbf{q}}^2$ is the approximate value of the plateau. The function $f(x)$ is dimensionless and modulates the value $\hbar\omega_D \ll \epsilon_F$. Remember that we are in a simple monoatomic crystal. For such crystal the acoustic phononic band vanishes linearly at the center of the Brillouin Zone and is approximately flat at the boundaries. If, say, the relevant part of the dispersion is at low wavevector, then $\hbar\Omega_{\mathbf{q}\to\mathbf{0}} \to 0$ and the width of the interaction zone vanishes. You can imagine that this is not the case, if we have gone through this whole chapter. So, we need to understand why the mapping onto the situation in Sec. 4.1.3 makes sense. The argument is rather rough, but reasonable.

For typical crystals the Fermi energy is of the order of $1 \div 10$eV; the lattice spacing is $1 \div 10$ Å. Performing the calculation, it is easy to see that the Fermi wavevector is typically comparable in size to the extension of the Brillouin Zone. Consider now Fig. 4.1: apart from the states near the starting states, processes scattering the particles around the shell exchange momen-

tum comparable with the Fermi wavevector, which means, at the boundary of the Brillouin Zone. And these processes are *a lot*. To get nearly everywhere around the shell a momentum this large is needed. Then it makes sense to approximate everything contains $\mathbf{q}$ in the last equation to its value at the boundary of the phononic band, in particular

$$\hbar\Omega_{\mathbf{q}} \simeq \hbar\omega_D$$

With such approximation the amplitude becomes

$$\begin{aligned} V_{\mathbf{k}-\mathbf{k}'} &\simeq \kappa^2\hbar\omega_D \frac{(\hbar\omega_D)^2}{(\epsilon_{\mathbf{k}'}-\epsilon_{\mathbf{k}})^2 - (\hbar\omega_D)^2} \\ &= -V_0 \frac{(\hbar\omega_D)^2}{(\hbar\omega_D)^2 - (\xi_{\mathbf{k}'}-\xi_{\mathbf{k}})^2} \end{aligned}$$

where $V_0 \equiv \kappa^2\hbar\omega_D \ll \epsilon_F$ and $\xi_{\mathbf{k}} \equiv \epsilon_{\mathbf{k}} - \epsilon_F$. This quantity measures the deviation from the Fermi energy of the state's energy. If we take a closer look, this expression $V_{\mathbf{k}-\mathbf{k}'}$ could be approximated as constant for couple of states whose distance in energy is smaller than $\hbar\omega_D$.

Now we make an assumption: the ground state of the system is made up of Cooper pairs (not isolated), and the collective state is mostly made of states just outside the Fermi surface. By this, we mean: we only consider for pairing those states with $\xi_{\mathbf{k}} \ll \hbar\omega_D$. In other words, an underlying structure exists in the pairing mechanism that privileges the states near the surface for creating a pair. To justify it we shall wait until Sec. 5.3.1; it is coherent with the pictorial description of Sec. 4.1.5, which predicts pairing only in a range

$$-\hbar\omega_D \ll -\Delta < \xi_{\mathbf{k}} < \Delta \ll \hbar\omega_D$$

as we will see. Then, if the relevant states are all around the Fermi energy and must be less distant than $\hbar\omega_D$, feeling a constants potential, the potential may as well be approximated by

$$V_{\mathbf{k}-\mathbf{k}'} = -V_0 A(\mathbf{k}) A(\mathbf{k}')$$

with

$$A(\mathbf{k}) = \begin{cases} 1 & \text{if} \quad \mathbf{k} \in \mathcal{S} \\ 0 & \text{if} \quad \mathbf{k} \notin \mathcal{S} \end{cases}$$

as in Sec. 4.1.5. Then, by what we know from Sec. 4.1.3, the Cooper pair has a binding energy of

$$E^{(\mathrm{b})} \simeq 2\Delta = 4\hbar\omega_D e^{-2/\rho_0\left(\kappa^2\hbar\omega_D\right)}$$

having set $\delta\epsilon^\star = \hbar\omega_D$. What remains to couple is the spin of the particles. We shall do it in the next section.

### 4.2.6 What does a Cooper pair look like?

For two spin 1/2 particles, the possible states are the singlet and the triplet

$$|s\rangle = \frac{|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle}{\sqrt{2}} \qquad \begin{aligned} |t_1\rangle &= |\uparrow\uparrow\rangle \\ |t_0\rangle &= \frac{|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle}{\sqrt{2}} \\ |t_{-1}\rangle &= |\downarrow\downarrow\rangle \end{aligned}$$

The singlet state exhibits anti-symmetry, while the triplet states are symmetric. Due to natural rotational invariance, it is reasonable to expect the pair to be formed preferentially in states $|s\rangle$ and $|t_0\rangle$. In the book *Solid State Physics* ([2] @ 18.2) a by Grosso and Pastori Parravicini a calculation is performed, showing that the singlet state is energetically favorite in our situation. The reason is essentially the spherical symmetry of the potential:

$$V_{\mathbf{k}-\mathbf{k}'} = V_{\mathcal{R}\mathbf{k}-\mathcal{R}\mathbf{k}'}$$

with $\mathcal{R}$ a rotation matrix. The potential can scatter with equal amplitude $\mathbf{k}$ to $\mathbf{k}'$ and $\mathcal{R}\mathbf{k}$ to $\mathcal{R}\mathbf{k}'$. Then the natural state must exhibit the same spherical symmetry. We are in the context of the so-called *s*-wave superconductivity. This is not valid in general, many materials exhibiting different Cooper coupling exist: continuing on the atomic physics derived terminology, *p*-wave pairing or *d*-wave pairing exist in those materials.

With this spin pairing we can finally write the effective hamiltonian:

$$\hat{H} = \sum_{\mathbf{k}} \epsilon_{\mathbf{k}} \left( \hat{c}^{\dagger}_{\mathbf{k}\uparrow} \hat{c}_{\mathbf{k}\uparrow} + \hat{c}^{\dagger}_{\mathbf{k}\downarrow} \hat{c}_{\mathbf{k}\downarrow} \right) + \sum_{\mathbf{k}\mathbf{k}'} V_{\mathbf{k}-\mathbf{k}'} \left( \hat{c}^{\dagger}_{\mathbf{k}\uparrow} \hat{c}_{\mathbf{k}'\uparrow} \right) \left( \hat{c}^{\dagger}_{-\mathbf{k}\downarrow} \hat{c}_{-\mathbf{k}'\downarrow} \right)$$

with $V_{\mathbf{k}-\mathbf{k}'}$ defined as before. In this sum we considered every Cooper pairing and assigned the positive momentum to the spin-up particle. This hamiltonian is the quintessential element of BCS theory.

If the spin part in in singlet state, the orbital part must exhibit symmetry to obtain global anti-symmetry. Remember what we said in Sec. 4.1.2: the wavefunction of the pair can be written as

$$\psi(\mathbf{x}_1, \mathbf{x}_2) = \sum_{\mathbf{k}} \alpha_{\mathbf{k}} w_{\mathbf{k}}(\mathbf{x}_1) w_{-\mathbf{k}}(\mathbf{x}_2) = L^{-D/2} \sum_{\mathbf{k}} \alpha_{\mathbf{k}} w_{\mathbf{k}}(\mathbf{x})$$

and in order to implement symmetry it must be $\alpha_{\mathbf{k}} = \alpha_{-\mathbf{k}}$. This also implies $\alpha_{\mathbf{k}} = \alpha^*_{\mathbf{k}}$, thus $\alpha_{\mathbf{k}} \in \mathbb{R}$. Then, limiting the sum to a half of the Fermi sphere (say, in $D = 3$, $k_z > 0$) and indicating such constraint on the sum as $(\star)$ we get

$$\psi(\mathbf{x}_1, \mathbf{x}_2) = L^{-D} \sum_{\mathbf{k}}^{(\star)} 2\alpha_{\mathbf{k}} \cos(\mathbf{k} \cdot \mathbf{x})$$

The wavefunction is really a one-variable function, and is a real superposition of plane waves distributed in momentum all over the interaction shell. We may estimate the momentum (magnitude) dispersion of the wavepacket, not so sure!

$$\Delta|\mathbf{k}| \simeq \delta k^{\star} \simeq \frac{\delta \epsilon^{\star}}{\hbar v_F}$$

then the spatial width of the wavepacket can be estimated via the Heisenberg principle,

$$\Delta|\mathbf{x}| \sim \frac{1}{|\mathbf{k}|} \simeq \frac{\hbar v_F}{\delta \epsilon^{\star}} = \frac{v_F}{\omega_D}$$

A very, *very* rough estimation of the size of a Cooper pair can be done using that for an acoustic phononic branch

$$\Omega_{\mathbf{q}} \sim v_S |\mathbf{q}|$$

having assumed isotropy, and calling $v_S$ the sound velocity on the phononic branch. We said before that the acoustic branch becomes somewhat horizontal at the boundary of the Brillouin Zone but generally the linear relation holds for a good portion of the Zone and so is not very distant from the real value of the dispersion at the boundary. Then $\omega_D \sim v_S q^\star$, with $q^\star$ the size of the Zone, and

$$\Delta|\mathbf{x}| \sim \frac{v_F}{v_S}\frac{1}{q^\star}$$

But $1/q^\star$ is roughly the lattice spacing, and $v_F \gg v_S$ in materials. Then we understand: a Cooper pair is *big*, compared to the ions of the crystal. This makes sense intuitively: the attraction is quite weak, the electrons move fast, and being separated enough in space the couple can explore a larger portion of the phononic band (in the sense that they become sensible to lattice vibration on larger scales than a single cell).

So, over all this chapter we have derived an effective mechanism for the pairing of electrons. We never talked about superconductors, however; the next chapter treats the central part of BCS theory, and explains how all this infernal machinery allows for charges to flow with zero resistance and expel magnetic fields.

# 5 THE BCS THEORY



It is time to develop formally and completely the theory by Bardeen, Cooper and Schrieffer. We saw in the last chapter how electrons form Cooper pairs, because apparently in superconductors charges flow coupled. The question now is: why is the coupling of electrons (a rather weak one, also) necessary for the exotic phenomena of superconductivity, like resistanceless flow of charge and Meissner effect?

## 5.1 BCS THEORY: SETUP

This section is devoted to the formal, quantum-mechanical treatment of BCS theory. The analysis is brought with two complementary methods.

### 5.1.1 The BCS hamiltonian and ground state

From last chapter we know that the phonon-mediated effective hamiltonian is given by

$$\hat{H} = \sum_{\mathbf{k}} \epsilon_{\mathbf{k}} \left[ \hat{c}^{\dagger}_{\mathbf{k}\uparrow} \hat{c}_{\mathbf{k}\uparrow} + \hat{c}^{\dagger}_{\mathbf{k}\downarrow} \hat{c}_{\mathbf{k}\downarrow} \right] + \sum_{\mathbf{k}\mathbf{k}'} V_{\mathbf{k}-\mathbf{k}'} \left[ \hat{c}^{\dagger}_{\mathbf{k}\uparrow} \hat{c}_{\mathbf{k}'\uparrow} \right] \left[ \hat{c}^{\dagger}_{-\mathbf{k}\downarrow} \hat{c}_{-\mathbf{k}'\downarrow} \right]$$

We also know that the Fermi Sea,

$$|\mathrm{F}\rangle \equiv \bigotimes_{|\mathbf{k}|<k_F} \hat{c}^\dagger_{\mathbf{k}\uparrow}\hat{c}^\dagger_{\mathbf{k}\downarrow}\,|0\rangle = \bigotimes_{|\mathbf{k}|<k_F} \hat{c}^\dagger_{\mathbf{k}\uparrow}\hat{c}^\dagger_{-\mathbf{k}\downarrow}\,|\Omega\rangle$$

with $|\Omega\rangle$ the state with 0 electrons, is an unstable state. This means, as it is obvious now, that the Fermi Sea configuration is not the ground state of the system. We know that the ground state will be in some measure populated by a mixture of Cooper pairs; note that, even if we have a general idea of how a *single* Cooper pair looks like, we have no way of predicting the real correlated state with multiple pairs.

The key idea is: the interaction involves a small portion of the Fermi sphere, mainly a thin shell around the surface. The total charge is a physically conserved quantity; however by defining the BCS ground state $|\Psi\rangle$ as

$$|\Psi\rangle \equiv \bigotimes_{\mathbf{k}} \left[u_{\mathbf{k}} + v_{\mathbf{k}}\hat{c}^\dagger_{\mathbf{k}\uparrow}\hat{c}^\dagger_{-\mathbf{k}\downarrow}\right]|\Omega\rangle \quad \text{with} \quad u_{\mathbf{k}}, v_{\mathbf{k}} \in \mathbb{C} \tag{5.1}$$

we shall not make a big mistake if $u_{\mathbf{k}}$ vanishes rapidly enough inside the Fermi sphere and $v_{\mathbf{k}}$ outside. Such condition is similar to the Fermi sphere configuration,

$$u_{\mathbf{k}} = \theta\left(k_F - |\mathbf{k}|\right) \qquad \text{and} \qquad v_{\mathbf{k}} = \theta\left(|\mathbf{k}| - k_F\right)$$

that can be taken as its limiting case. Defined as it is, $|\Psi\rangle$ is a superposition of many states with different number of particles,

$$\begin{aligned}|\Psi\rangle = &\prod_{\mathbf{k}} u_{\mathbf{k}}\,|\Omega\rangle \\ &+ \sum_{\mathbf{k}_1}\prod_{\mathbf{k}\neq\mathbf{k}_1} u_{\mathbf{k}}v_{\mathbf{k}_1}\left[\hat{c}^\dagger_{\mathbf{k}_1\uparrow}\hat{c}^\dagger_{-\mathbf{k}_1\downarrow}\right]|\Omega\rangle \\ &+ \sum_{\mathbf{k}_1\mathbf{k}_2}\prod_{\mathbf{k}\neq\mathbf{k}_1,\mathbf{k}_2} u_{\mathbf{k}}v_{\mathbf{k}_1}\left[\hat{c}^\dagger_{\mathbf{k}_1\uparrow}\hat{c}^\dagger_{-\mathbf{k}_1\downarrow}\right]v_{\mathbf{k}_2}\left[\hat{c}^\dagger_{\mathbf{k}_2\uparrow}\hat{c}^\dagger_{-\mathbf{k}_2\downarrow}\right]|\Omega\rangle\end{aligned}$$

The first term has zero particles; the second has one pair; the third has two pairs, and so on. For our description to be coherent, we expect in thermodynamic limit the number of particles to localize, which is, to negligibly fluctuate around its mean value. Next section deals with this problem.

### 5.1.2 How many particles?

As said, we want the mean number of particles $\langle\hat{N}\rangle$ in the ground state to be a well-defined quantity, at least in the thermodynamic limit. So, consider the number operators,

$$\hat{N}_\uparrow \equiv \sum_{\mathbf{k}}\hat{c}^\dagger_{\mathbf{k}\uparrow}\hat{c}_{\mathbf{k}\uparrow} \qquad \hat{N}_\downarrow \equiv \sum_{\mathbf{k}}\hat{c}^\dagger_{\mathbf{k}\downarrow}\hat{c}_{\mathbf{k}\downarrow} \qquad \hat{N} \equiv \hat{N}_\uparrow + \hat{N}_\downarrow$$

which count the number of particles for a given state. Simple calculations lead us to

$$\langle\Psi|\hat{N}_\uparrow|\Psi\rangle = \sum_{\mathbf{k}}|v_{\mathbf{k}}|^2 \qquad \langle\Psi|\hat{N}_\downarrow|\Psi\rangle = \sum_{\mathbf{k}}|v_{-\mathbf{k}}|^2$$

This result is rather obvious, once seen the form of Eq. (5.1). The parameter $v_{\mathbf{k}}$ is the probability amplitude for the pair occupation of the states $|\mathbf{k}\uparrow\rangle$

and $|-\mathbf{k}\downarrow\rangle$. We impose spin balance, so that $|v_{-\mathbf{k}}|^2 = |v_{\mathbf{k}}|^2$. Notice that requiring $v_{-\mathbf{k}} = v_{\mathbf{k}}$ implies the parameter to be real and the above condition to be satisfied. We will make that assumption. Thus, we have

$$\langle \hat{N} \rangle = 2 \sum_{\mathbf{k}} |v_{\mathbf{k}}|^2$$

To get the number fluctuations, we need to extract

$$\langle \hat{N}^2 \rangle = \langle \hat{N}_\uparrow^2 \rangle + \langle \hat{N}_\uparrow \hat{N}_\downarrow \rangle + \langle \hat{N}_\downarrow \hat{N}_\uparrow \rangle + \langle \hat{N}_\downarrow^2 \rangle$$

that turns out to be

$$\langle \hat{N}^2 \rangle = 4 \sum_{\mathbf{k} \neq \mathbf{k}'} |v_{\mathbf{k}}|^2 |v_{\mathbf{k}'}|^2 + 2 \langle \hat{N} \rangle$$

and this implies

$$\begin{aligned} \frac{\sqrt{\langle \hat{N}^2 \rangle - \langle \hat{N} \rangle^2}}{\langle \hat{N} \rangle} &= \frac{\sqrt{2 \langle \hat{N} \rangle + 4 \sum_{\mathbf{k} \neq \mathbf{k}'} |v_{\mathbf{k}}|^2 |v_{\mathbf{k}'}|^2 - 4 \sum_{\mathbf{k}} |v_{\mathbf{k}}|^2 \sum_{\mathbf{k}'} |v_{\mathbf{k}'}|^2}}{\langle \hat{N} \rangle} \\ &= \frac{\sqrt{2 \langle \hat{N} \rangle - 4 \sum_{\mathbf{k}} |v_{\mathbf{k}}|^4}}{\langle \hat{N} \rangle} < \sqrt{\frac{2}{\langle \hat{N} \rangle}} \end{aligned}$$

thus in thermodynamic limit the number of particles is a well-defined quantity. Now: to extract the BCS value of $u_{\mathbf{k}}$, $v_{\mathbf{k}}$ and all other features of the BCS hamiltonian two ways are the most commonly used. Both are interesting, so we shall explore them separately.

## 5.2 THE VARIATIONAL METHOD

The key idea is: the energy is a functional of $u_{\mathbf{k}}$, $v_{\mathbf{k}}$. To extract the energy we need to minimize the functional with respect to parameters variations. To simplify, we expect one parameter to be "fictitious" – in the sense that due to normalization of $|\Psi\rangle$ some relation between $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ must exist, limiting our variational approach to a single parameter. So we compute $\langle \Psi | \Psi \rangle$,

$$\begin{aligned} \langle \Psi | \Psi \rangle &= \bigotimes_{\mathbf{k}} \langle \Omega | \left[ u_{\mathbf{k}} + v_{\mathbf{k}}^* \hat{c}_{-\mathbf{k}\downarrow} \hat{c}_{\mathbf{k}\uparrow} \right] \bigotimes_{\mathbf{k}'} \left[ u_{\mathbf{k}'} + v_{\mathbf{k}'} \hat{c}_{\mathbf{k}'\uparrow}^\dagger \hat{c}_{-\mathbf{k}'\downarrow}^\dagger \right] |\Omega\rangle \\ &= \bigotimes_{\mathbf{k}} \langle \Omega | \Big[ |u_{\mathbf{k}}|^2 + u_{\mathbf{k}}^* v_{\mathbf{k}} \hat{c}_{\mathbf{k}\uparrow}^\dagger \hat{c}_{-\mathbf{k}\downarrow}^\dagger + u_{\mathbf{k}} v_{\mathbf{k}}^* \hat{c}_{-\mathbf{k}\downarrow} \hat{c}_{\mathbf{k}\uparrow} \\ &\qquad + |v_{\mathbf{k}}|^2 \hat{c}_{-\mathbf{k}\downarrow} \hat{c}_{\mathbf{k}\uparrow} \hat{c}_{\mathbf{k}\uparrow}^\dagger \hat{c}_{-\mathbf{k}\downarrow}^\dagger \Big] |\Omega\rangle \\ &= \bigotimes_{\mathbf{k}} \left[ |u_{\mathbf{k}}|^2 + |v_{\mathbf{k}}|^2 \right] \stackrel{!}{=} 1 \end{aligned}$$

since mixed terms vanish and having used fermionic commutation rules. The above condition is solved by

$$\forall \mathbf{k} \quad : \quad |u_{\mathbf{k}}|^2 + |v_{\mathbf{k}}|^2 = 1$$

This relations allows us to define two parameters $\theta_{\mathbf{k}}$, $\varphi_{\mathbf{k}}$,

$$u_{\mathbf{k}} \equiv \cos\theta_{\mathbf{k}} \qquad \text{and} \qquad v_{\mathbf{k}} \equiv \sin\theta_{\mathbf{k}} e^{i\varphi_{\mathbf{k}}}$$

This definition is equivalent to taking both complex and collecting a global phase outside the state, and a local relative phase inside the $\mathbf{k}$ term. With further elaboration, one can show $\varphi_{\mathbf{k}} = \varphi$, all phases equal one constant phase. It can be proven to be basically the local phase of the condensate, the one associated to local spontaneous $\mathrm{U}(1)$ symmetry breaking. For now it is not so important, we can implement it later, and take $\varphi = 0$,

$$u_{\mathbf{k}} \equiv \cos\theta_{\mathbf{k}} \qquad \text{and} \qquad v_{\mathbf{k}} \equiv \sin\theta_{\mathbf{k}} \tag{5.2}$$

### 5.2.1 The variational solutions for the BCS parameters

Now we want to find some expression for the functional $E[\theta_{\mathbf{k}}]$. We consider shifting the energies up to $\epsilon_F$, thus considering the operator

$$\hat{\boldsymbol{H}} - \epsilon_F \hat{N} \qquad \text{with} \qquad \hat{N} \equiv \sum_{\mathbf{k}} \left[ \hat{c}^\dagger_{\mathbf{k}\uparrow} \hat{c}_{\mathbf{k}\uparrow} + \hat{c}^\dagger_{\mathbf{k}\downarrow} \hat{c}_{\mathbf{k}\downarrow} \right]$$

and define the functional as

$$\begin{aligned} E\left[\{\theta_{\mathbf{q}}\}\right] &\equiv \langle\Psi| \left[\hat{\boldsymbol{H}} - \epsilon_F \hat{N}\right] |\Psi\rangle \\ &= \sum_{\mathbf{k}} \xi_{\mathbf{k}} \langle\Psi| \left[ \hat{c}^\dagger_{\mathbf{k}\uparrow} \hat{c}_{\mathbf{k}\uparrow} + \hat{c}^\dagger_{\mathbf{k}\downarrow} \hat{c}_{\mathbf{k}\downarrow} \right] |\Psi\rangle \\ &\qquad + \sum_{\mathbf{k}\mathbf{k}'} V_{\mathbf{k}-\mathbf{k}'} \langle\Psi| \left[ \hat{c}^\dagger_{\mathbf{k}\uparrow} \hat{c}^\dagger_{-\mathbf{k}\downarrow} \right] \left[ \hat{c}_{-\mathbf{k}'\downarrow} \hat{c}_{\mathbf{k}'\uparrow} \right] |\Psi\rangle \end{aligned}$$

with $\xi_{\mathbf{k}} \equiv \epsilon_{\mathbf{k}} - \epsilon_F$, and where some fermionic rules have been used. Some straightforward calculations lead us to

$$\begin{aligned} E\left[\{\theta_{\mathbf{q}}\}\right] &= 2\sum_{\mathbf{k}} \xi_{\mathbf{k}} |v_{\mathbf{k}}|^2 + \sum_{\mathbf{k}\mathbf{k}'} V_{\mathbf{k}-\mathbf{k}'} \left[v^*_{\mathbf{k}} u_{\mathbf{k}}\right] \left[u^*_{\mathbf{k}'} v_{\mathbf{k}'}\right] \\ &= 2\sum_{\mathbf{k}} \xi_{\mathbf{k}} \sin^2\theta_{\mathbf{k}} + \sum_{\mathbf{k}\mathbf{k}'} V_{\mathbf{k}-\mathbf{k}'} \cos\theta_{\mathbf{k}} \sin\theta_{\mathbf{k}} \sin\theta_{\mathbf{k}'} \cos\theta_{\mathbf{k}'} \\ &= 2\sum_{\mathbf{k}} \xi_{\mathbf{k}} \sin^2\theta_{\mathbf{k}} + \frac{1}{4}\sum_{\mathbf{k}\mathbf{k}'} V_{\mathbf{k}-\mathbf{k}'} \sin 2\theta_{\mathbf{k}} \sin 2\theta_{\mathbf{k}'} \end{aligned}$$

where we used $2\sin\theta\cos\theta = \sin 2\theta$. To avoid confusion with the notation, we highlight that $E\left[\{\theta_{\mathbf{q}}\}\right]$ indicates a functional of all angles, $E\left[\theta_{\mathbf{q}_1}, \theta_{\mathbf{q}_2} \cdots\right]$ while the indices $\mathbf{k}$ and $\mathbf{k}'$ are mute and have nothing to do with the argument fo $E$. This means that to derive the functional means to derive both $\sin 2\theta_{\mathbf{k}}$ and $\sin 2\theta_{\mathbf{k}'}$ in the last term. Then, deriving with respect to one precise angle $\theta_{\mathbf{q}}$,

$$\begin{aligned} 0 &\overset{!}{=} \frac{\partial}{\partial\theta_{\mathbf{q}}} E\left[\{\theta_{\mathbf{q}'}\}\right] \\ &= 2\xi_{\mathbf{q}} \sin 2\theta_{\mathbf{q}} + \frac{1}{2}\cos 2\theta_{\mathbf{q}} \sum_{\mathbf{k}'} V_{\mathbf{q}-\mathbf{k}'} \sin 2\theta_{\mathbf{k}'} + \frac{1}{2}\cos 2\theta_{\mathbf{q}} \sum_{\mathbf{k}} V_{\mathbf{k}-\mathbf{q}} \sin 2\theta_{\mathbf{k}} \\ &= 2\xi_{\mathbf{q}} \sin 2\theta_{\mathbf{q}} + \frac{1}{2}\cos 2\theta_{\mathbf{q}} \sum_{\mathbf{k}'} V_{\mathbf{q}-\mathbf{k}'} \sin 2\theta_{\mathbf{k}'} + \frac{1}{2}\cos 2\theta_{\mathbf{q}} \sum_{\mathbf{k}} V^*_{\mathbf{q}-\mathbf{k}} \sin 2\theta_{\mathbf{k}} \end{aligned}$$

since $\partial_\theta \sin^2\theta = \sin 2\theta$ and $\partial_\theta \sin 2\theta = 2\cos 2\theta$. In the last passage we used the relation $V_{\mathbf{k}-\mathbf{q}} = V^*_{\mathbf{q}-\mathbf{k}}$. Time-reversal symmetry, as well as what we said in the previous chapter about the interaction potential, allows us to conclude $V_{\mathbf{q}} = V^*_{\mathbf{q}}$. Then the last two sums in the above equation are equal and sum up. We change label to conform to standard notation, $\mathbf{q} \to \mathbf{k}$, and get

$$2\xi_{\mathbf{k}} \sin 2\theta_{\mathbf{k}} + \cos 2\theta_{\mathbf{k}} \sum_{\mathbf{k}'} V_{\mathbf{k}-\mathbf{k}'} \sin 2\theta_{\mathbf{k}'} = 0$$

We define

$$\Delta_{\mathbf{k}} \equiv -\frac{1}{2} \sum_{\mathbf{k}'} V_{\mathbf{k}-\mathbf{k}'} \sin 2\theta_{\mathbf{k}'} \qquad \Longrightarrow \qquad \xi_{\mathbf{k}} \sin 2\theta_{\mathbf{k}} = \Delta_{\mathbf{k}} \cos 2\theta_{\mathbf{k}} \tag{5.3}$$

Expanding the above result and taking its square it is easy to get to the equation

$$4\left(\xi_{\mathbf{k}} + \Delta_{\mathbf{k}}\right)^2 \cos^4\theta_{\mathbf{k}} - 4\left(\xi_{\mathbf{k}} + \Delta_{\mathbf{k}}\right)^2 \cos^2\theta_{\mathbf{k}} + \Delta_{\mathbf{k}}^2 = 0$$

or, substituting

$$|v_{\mathbf{k}}|^4 - |v_{\mathbf{k}}|^2 + \frac{\Delta_{\mathbf{k}}^2}{4\left(\xi_{\mathbf{k}} + \Delta_{\mathbf{k}}\right)^2} = 0$$

Implementing the condition

$$\lim_{\xi_{\mathbf{k}} \gg \Delta_{\mathbf{k}}} v_{\mathbf{k}} \stackrel{!}{=} 0$$

which guarantees that the ground state has no population for states very distant from the Fermi surface, we get the solution

$$|v_{\mathbf{k}}|^2 = \frac{1}{2}\left(1 - \frac{\xi_{\mathbf{k}}}{\sqrt{\xi_{\mathbf{k}}^2 + \Delta_{\mathbf{k}}^2}}\right) \tag{5.4}$$

and due to normalization

$$|u_{\mathbf{k}}|^2 = \frac{1}{2}\left(1 + \frac{\xi_{\mathbf{k}}}{\sqrt{\xi_{\mathbf{k}}^2 + \Delta_{\mathbf{k}}^2}}\right) \tag{5.5}$$

In Fig. 5.1 the two solutions are plotted as functions of the energy $\xi_{\mathbf{k}}$ in the special case $\Delta_{\mathbf{k}} = \Delta$. The actual solution will be very similar. As evident, deep inside the sphere (for $\xi_{\mathbf{k}} \leq 0$) we have $v_{\mathbf{k}} \simeq 1$, and $u_{\mathbf{k}} \simeq 0$, leading to a state similar to the non-interacting perfect Fermi sphere. Far outside the situation is the opposite, $v_{\mathbf{k}} \simeq 0$, and $u_{\mathbf{k}} \simeq 1$, which correctly means that far states are not populated. Overall the state is pretty similar to the Fermi sphere, with significant variations only in the energy range

$$-\Delta < \xi_{\mathbf{k}} < \Delta$$

This observation is quite important in understating the nature of the collective state. We anticipated in Chap. 4 that the relevant states for pairing are those within a range $\pm\Delta$ from the Fermi energy. This result, although obtained for $\Delta_{\mathbf{k}} = \Delta$ (constant) holds very well for the weak coupling in general. We will come back on this in a moment, in Sec. 5.3.1.

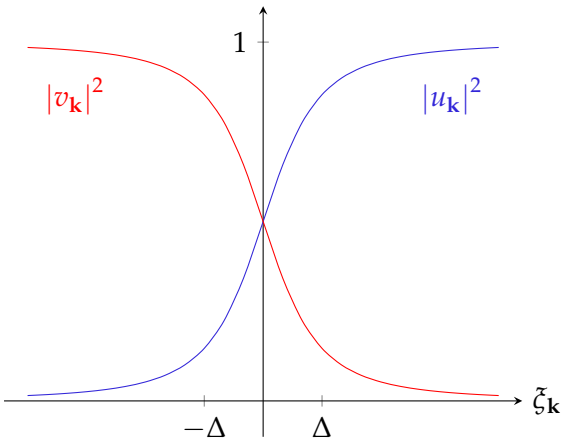



**Figure 5.1:** Plot of the variational solutions for the problem parameters $u_{\mathbf{k}}$, in Eq. (5.5), and $v_{\mathbf{k}}$, in Eq. (5.4). This plot was realized using a constant Ansatz, $\Delta_{\mathbf{k}} = \Delta$. This plot is in general reasonable but not completely coherent with the theory; instead, the plot in Fig. 5.3 is.

It is useful to understand the relative sign of $u_{\mathbf{k}}$, $v_{\mathbf{k}}$. From Eq. (5.3),

$$\frac{\Delta_{\mathbf{k}}}{\xi_{\mathbf{k}}} = \tan 2\theta_{\mathbf{k}} = \frac{2\tan\theta_{\mathbf{k}}}{1-\tan^2\theta_{\mathbf{k}}}$$

We anticipate $\Delta_{\mathbf{k}} > 0$, as we are going to see in Sec. 5.2.2 Take $\xi_{\mathbf{k}} > 0$, the right part of the graph in Fig. 5.1. Here

$$|v_{\mathbf{k}}|^2 > |u_{\mathbf{k}}|^2 \quad \Longrightarrow \quad \tan^2\theta_{\mathbf{k}} < 1$$

It follows

$$\tan\theta_{\mathbf{k}} = \left(1-\tan^2\theta_{\mathbf{k}}\right)\frac{\Delta_{\mathbf{k}}}{2\xi_{\mathbf{k}}} > 0$$

which means that $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ have equal sign. The same holds for $\xi_{\mathbf{k}} < 0$. Then we know

$$\begin{aligned} u_{\mathbf{k}}v_{\mathbf{k}} &= \sqrt{|u_{\mathbf{k}}|^2|v_{\mathbf{k}}|^2} \\ &= \left[\frac{1}{4}\left(1+\frac{\xi_{\mathbf{k}}}{\sqrt{\xi_{\mathbf{k}}^2+\Delta_{\mathbf{k}}^2}}\right)\left(1-\frac{\xi_{\mathbf{k}}}{\sqrt{\xi_{\mathbf{k}}^2+\Delta_{\mathbf{k}}^2}}\right)\right]^{1/2} \\ &= \frac{1}{2}\frac{\Delta_{\mathbf{k}}}{\sqrt{\xi_{\mathbf{k}}^2+\Delta_{\mathbf{k}}^2}} \end{aligned} \tag{5.6}$$

This quantity is often referred to as the **pairing amplitude**. Although now unclear, soon the reason for such a name will be explained. A plot is reported in Fig. 5.2. We use this result immediately.

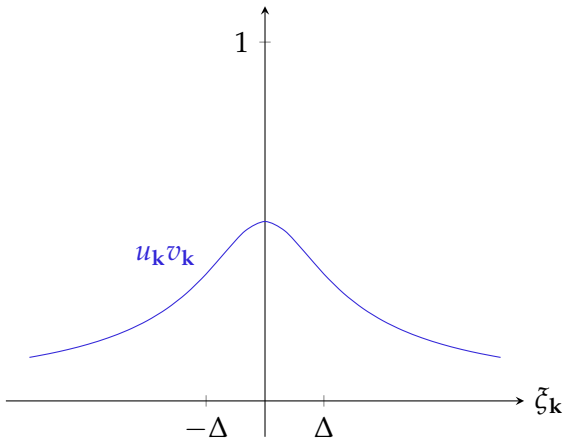



**Figure 5.2:** Plot of the pairing amplitude, $u_{\mathbf{k}}v_{\mathbf{k}}$, as described at the end of Sec. 5.2.1. As evident this quantity has a peak in the region $-\Delta < \xi_{\mathbf{k}} < \Delta$.

### 5.2.2 The self-consistency equation

We aim to find an expression for $\Delta_{\mathbf{k}}$, using the found solutions. Substituting the obtained solutions in the definition of $\Delta_{\mathbf{k}}$, we get

$$\begin{aligned}\Delta_{\mathbf{k}} &\equiv -\frac{1}{2}\sum_{\mathbf{k}'} V_{\mathbf{k}-\mathbf{k}'}\sin 2\theta_{\mathbf{k}'} \\ &= -\sum_{\mathbf{k}'} V_{\mathbf{k}-\mathbf{k}'}v_{\mathbf{k}'}u_{\mathbf{k}'} \\ &= -\frac{1}{2}\sum_{\mathbf{k}'} V_{\mathbf{k}-\mathbf{k}'}\frac{\Delta_{\mathbf{k}'}}{\sqrt{\xi_{\mathbf{k}'}^2+\Delta_{\mathbf{k}'}^2}}\end{aligned} \tag{5.7}$$

having we used Eq. (5.6). This is the so-called **self-consistency equation**. Consider now the potential analyzed in Sec. 4.1.3. The phonon effective potential of Sec. 4.2.5 is well approximated by

$$V_{\mathbf{k}-\mathbf{k}'} = -V_0 A\left(\mathbf{k}\right) A\left(\mathbf{k}'\right)$$

with $A$, as always, the characteristic function of the shell. Then

$$\Delta_{\mathbf{k}} = \frac{V_0 A\left(\mathbf{k}\right)}{2}\sum_{\mathbf{k}'} A\left(\mathbf{k}'\right)\frac{\Delta_{\mathbf{k}'}}{\sqrt{\xi_{\mathbf{k}'}^2+\Delta_{\mathbf{k}'}^2}}$$

The right part of the equation is zero for $\mathbf{k}$ outside the interaction shell, and constant inside. Then

$$\Delta_{\mathbf{k}} = \Delta A\left(\mathbf{k}\right)$$

with $\Delta$ a constant term. It follows, inside the shell,

$$\Delta = \frac{V_0}{2}\sum_{|\xi_{\mathbf{k}}|<\delta\epsilon^\star}\frac{\Delta}{\sqrt{\xi_{\mathbf{k}}^2+\Delta^2}}$$

with $\delta\epsilon^\star = \hbar\omega_D$, as explained in Sec. 4.2.5. We now convert the sum into an energy integral and approximate the density of states constant as in Sec. 4.1.3. Be *very* careful now: the operation we are trying to perform is

$$\sum_{\xi_{\mathbf{k}}}(\cdots) \to \int d\xi\tilde{\rho}(\epsilon_F+\xi)\,(\cdots)$$

with $\tilde{\rho}$ the density of states and $(\cdots)$ is *something*. The most important detail to notice here, that caused a lot of trouble to the lazy and distracted author when first writing this notes, is that $\tilde{\rho}(\epsilon_F + \xi)$ counts how many states available for the pair there are in the energy range $[\xi, \xi + \delta\xi]$. Take the standard single-electron density of states $\rho$. The point is that $\xi_{\mathbf{k}} = \xi_{-\mathbf{k}}$, so the states $|\mathbf{k}\rangle$ and $|-\mathbf{k}\rangle$ contribute separately to $\rho$, since the electron finds independently the two states; counting spin, the electron finds independently the four states. It is not the same here. Here we are dealing with pairs, for which the couple $|\mathbf{k}\uparrow\rangle \otimes |-\mathbf{k}\downarrow\rangle$ represents **one** state and so does $|\mathbf{k}\downarrow\rangle \otimes |-\mathbf{k}\uparrow\rangle$. This means that here we need to use

$$\tilde{\rho}(\epsilon_F + \xi) = \frac{\rho(\epsilon_F + \xi)}{2}$$

Now we can approximate $\rho(\epsilon_F + \xi) \simeq \rho(\epsilon_F) \equiv \rho_0$, the density of states at the Fermi level. We get

$$\Delta \simeq \frac{\rho_0 V_0}{4} \int_{-\hbar\omega_D}^{\hbar\omega_D} d\xi \frac{\Delta}{\sqrt{\xi^2 + \Delta^2}} = \frac{\rho_0 V_0}{2} \int_0^{\hbar\omega_D} d\xi \frac{\Delta}{\sqrt{\xi^2 + \Delta^2}}$$

where in the first passage the factor $1/4$ is the product of the already present factor $1/2$ and the correct density of states. Then, changing variable $s = \xi/\Delta$, we recognize the derivative of $\sinh^{-1} s$,

$$1 = \frac{\rho_0 V_0}{2} \int_0^{\hbar\omega_D/\Delta} ds \frac{\Delta}{\sqrt{1 + s^2}} = \frac{\rho_0 V_0}{2} \sinh^{-1}\left(\frac{\hbar\omega_D}{\Delta}\right)$$

which implies:

$$\Delta = \frac{\hbar\omega_D}{\sinh\left(\dfrac{2}{\rho_0 V_0}\right)}$$

Since:

$$\lim_{x\to+\infty} \sinh x = \lim_{x\to+\infty} \frac{e^x - e^{-x}}{2} = \lim_{x\to+\infty} \frac{e^x}{2}$$

and we have seen that the potential $V_0$ expressed by phonon mediation is weak, we may approximate

$$\sinh\left(\frac{2}{\rho_0 V_0}\right) \simeq \frac{e^{2/\rho_0 V_0}}{2}$$

which finally gives

$$\Delta = 2\hbar\omega_D e^{-2/\rho_0 V_0}$$

Familiar? That is precisely half the binding energy of the Cooper pair of Sec. 4.1.5 and Sec. 4.2.5. We interpret as the binding energy of the pair, in the many body problem, $2\Delta$. The reason may seem obscure: we omit for a second more comments about the meaning of $\Delta$, which becomes clear enough in Sec. 4.2.5.

### 5.2.3 Cooper pairs in the collective state

Take the BCS ground state:

$$|\Psi\rangle \equiv \bigotimes_{\mathbf{k}} \left[u_{\mathbf{k}} + v_{\mathbf{k}} \hat{c}^\dagger_{\mathbf{k}\uparrow} \hat{c}^\dagger_{-\mathbf{k}\downarrow}\right] |\Omega\rangle$$

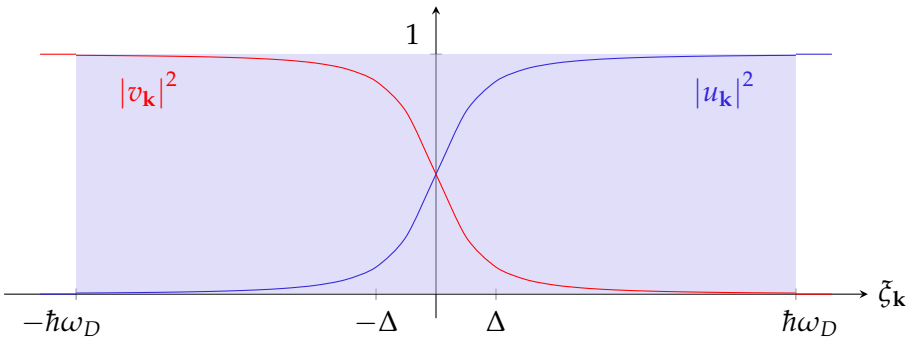



**Figure 5.3:** Plot of the amplitudes $|u_{\mathbf{k}}|^2$ and $|v_{\mathbf{k}}|^2$ as reported in Eqns. (5.8) and (5.9). The shaded region $|\xi_{\mathbf{k}}| < \hbar\omega_D$ is the interaction region, out of which the solution is correctly given by the Fermi state. In this plot we arbitrarily set $\Delta = \hbar\omega_D/6$, inspired by the idea of weakness of the attracting potential. In general, for $\Delta \ll \hbar\omega_D$, as $\Delta$ decreases the plot becomes rather continuous and resembles more closely the one in Fig. 5.1.

Evidently the quantity $|u_{\mathbf{k}}|^2$ measures the probability for a given couple of states, $|\mathbf{k}\uparrow\rangle$ and $|-\mathbf{k}\downarrow\rangle$, of being unoccupied. Complementarily $|v_{\mathbf{k}}|^2$ measures the probability of finding them both occupied. Both them being occupied clearly is not enough to declare both electrons to be participating in a pairing via the Cooper mechanism: after all, the states deep in the sphere have perfect couple occupation. Then, were are Cooper pairs in the BCS state?

For the BCS ground state, inserting what we know from last section about $\Delta_{\mathbf{k}} = \Delta A(\mathbf{k})$, we have

$$|u_{\mathbf{k}}|^2 = \begin{cases} 0 & \xi_{\mathbf{k}} < -\hbar\omega_D \\ \frac{1}{2}\left(1 + \frac{\xi_{\mathbf{k}}}{\sqrt{\xi_{\mathbf{k}}^2 + \Delta^2}}\right) & -\hbar\omega < \xi_{\mathbf{k}} < \hbar\omega_D \\ 1 & \xi_{\mathbf{k}} > \hbar\omega_D \end{cases} \tag{5.8}$$

and

$$|v_{\mathbf{k}}|^2 = \begin{cases} 1 & \xi_{\mathbf{k}} < -\hbar\omega_D \\ \frac{1}{2}\left(1 - \frac{\xi_{\mathbf{k}}}{\sqrt{\xi_{\mathbf{k}}^2 + \Delta^2}}\right) & -\hbar\omega < \xi_{\mathbf{k}} < \hbar\omega_D \\ 0 & \xi_{\mathbf{k}} > \hbar\omega_D \end{cases} \tag{5.9}$$

Notice that $\Delta < 2\hbar\omega_D$, but in principle it is not guaranteed $\Delta < \hbar\omega_D$. However weak phononic coupling, necessary for the whole procedure to work, implies small interactions and thus $\Delta \ll 2\hbar\omega_D$. Check Fig. 5.3: the above functions are there plotted, with the arbitrary choice $\Delta = \hbar\omega_D/6$. Such choice was made for reasons of graphic clarity. Don't take too seriously the discontinuities at the boundaries of the shaded region: those are generated by the approximation $\Delta_{\mathbf{k}} = \Delta A(\mathbf{k})$.

We comment now the nature of these amplitudes. First, recall what we already said at the end of Sec. 5.2.1: all the relevant differences from the free Fermi amplitudes occur in the shell

$$-\Delta < \xi_{\mathbf{k}} < \Delta$$

Pursuing the very pictorial ideas of Sec. 4.1.5, we see that the BCS solution is indeed the optimal distribution of occupation probabilities in the situation of many paired electrons. Being

$$E = 2\sum_{\mathbf{k}} \xi_{\mathbf{k}} |v_{\mathbf{k}}|^2 + \sum_{\mathbf{k}\mathbf{k}'} V_{\mathbf{k}-\mathbf{k}'} \left[v_{\mathbf{k}}^* u_{\mathbf{k}}\right] \left[u_{\mathbf{k}'}^* v_{\mathbf{k}'}\right]$$
$$= 2\sum_{\mathbf{k}} \xi_{\mathbf{k}} |v_{\mathbf{k}}|^2 - V_0 \left[\sum_{\mathbf{k}} A\left(\mathbf{k}\right) \frac{\Delta_{\mathbf{k}}^2}{\xi_{\mathbf{k}}^2 + \Delta_{\mathbf{k}}^2}\right]^2$$

having we used Eq. (5.6), we see that the variational approach to BCS theory is ultimately the task of finding a balance between two behaviors: decreasing the kinetic contribution by enhancing occupation of states with smaller $\xi_{\mathbf{k}}$ (which would tend to accumulate all the occupation for $\xi_{\mathbf{k}} < 0$) and increasing the width of the distribution, $\Delta$, in order to fill many states and gaining from each one the potential gain $-V_0$. To balance these trends is precisely the task of the variational approach. In the following extract, the notation is a little different than here: replace $v_{\mathbf{k}} \leftrightarrow b_{\mathbf{q}}$, $-V_0 \leftrightarrow -V$ and $\Delta \leftrightarrow \Delta_0$.

> **D. L. Goodstein**: ***States of Matter*** ([4] @ 5.3.d)
>
> $$2b_{\mathbf{q}}^2 = 1 - \frac{\xi_{\mathbf{q}}}{\sqrt{\xi_{\mathbf{q}}^2 + \Delta_0^2}}$$
>
> The $b_{\mathbf{q}}^2$'s are the probabilities that the original single-particle states are occupied. Obviously, for $\xi_{\mathbf{q}} \ll 0$, $b_{\mathbf{q}}^2 = 1$ and for $\xi_{\mathbf{q}} \gg 0$, $b_{\mathbf{q}}^2 = 0$. Between these limits, over a region whose width is of order $\Delta_0$, $b_{\mathbf{q}}^2$ changes smoothly from 1 to 0 The distribution [...] is reminiscent of that [...] for the thermal probability of the occupation of these same states for normal electrons at finite temperature. The $b_{\mathbf{q}}^2$'s, however, are quantum probabilities in the ground state, at $T = 0$. In retrospect, it is not hard to see why this form has occurred. The influence of the interaction, $-V$, depends intimately on the number of unoccupied states nearby into which a pair might be scattered. [...] a weak potential caused pairs to be bound simply because many states were available. The effect of the interaction is to cause some states that are unoccupied in the normal ground state (those above the Fermi level) to have some amplitude for being occupied and to leave states below the Fermi level with some amplitude for being unoccupied. This result, in turn, enhances the effect of $-V$ on other pairs, wince there are now more electrons within $\hbar\omega_D$ of states with at least some amplitude for being unoccupied and therefore available for scattering into. The more the distribution spreads out to lower its energy by taking advantage of $-V$ operating between pairs of electrons, the more it raises its kinetic energy. The calculation we have done has found the optimum compromise, the delicate balance.

So, lastly: where are Cooper pairs? We begun by saying that to have a couple of antipodal states filled is not enough to declare them actually part of a Cooper pair. And this is sensate: otherwise, we should think about the filled bulk part of the sphere, unaltered by interaction, as filled by pairs. The key point is that in this many-body context the pairing is more a property of

the states rather than specific couples of electrons. This means: for electrons, to be in some states enhances to probability to be scattered to some other states. And that is what defines a pair as we know it: the optimal distribution of the couple state in momentum space in order to minimize energy. The trivial distribution, a couple stable in its state, is not a pair. It's not interacting. The reason is that states deep below the surface, $-\hbar\omega_D < \xi_{\mathbf{k}} < -\Delta$, are very compressed by statistics and should increase their kinetic energy a lot to scatter to far states; states far from the surface, $\Delta < \xi_{\mathbf{k}} < \hbar\omega_D$ may easily couple to many states, but are difficult to reach in the first place due to the great kinetic cost. The following example quantifies this trend.

We now indicate as $|\mathbf{k}\rangle$ the state of an antipodal couple

$$|\mathbf{k}\rangle \equiv |\mathbf{k}\uparrow\rangle \otimes |-\mathbf{k}\downarrow\rangle$$

The amplitude for the process $|\mathbf{k}'\rangle \to |\mathbf{k}\rangle$ is given by

$$\langle \mathbf{k}|\hat{\boldsymbol{H}} - \epsilon_F \hat{N}|\mathbf{k}'\rangle = V_{\mathbf{k}-\mathbf{k}'} \left[u_{\mathbf{k}} v_{\mathbf{k}}\right] \left[u_{\mathbf{k}'} v_{\mathbf{k}'}\right]$$

being the parameters real. Two pairing amplitudes appear: the one of the starting state, $[u_{\mathbf{k}'} v_{\mathbf{k}'}]$, and the one for the target state, $[u_{\mathbf{k}} v_{\mathbf{k}}]$. Consider now the plot of Fig. 5.2: the pairing amplitude is significantly different from zero only within a range $\pm\Delta$ from the Fermi energy. So, if the starting couple $[u_{\mathbf{k}'} v_{\mathbf{k}'}]$ is deep below the surface, $-\hbar\omega_D < \xi_{\mathbf{k}'} < -\Delta$, the state it occupies is poorly connected to other states via the potential. The same holds if the target state is below the surface, $-\hbar\omega_D < \xi_{\mathbf{k}} < -\Delta$: it is almost filled, so by exclusion principle it is difficult to reach via the potential. And the same holds for starting and target states at high energy.

So: we may say that, if we start with a non-interacting filled sphere, turn on the interaction and wait for the system to collapse on its ground state, we will find a smeared distribution of occupation probabilities, confronted with the starting $T = 0$ Fermi-Dirac distribution, and that really only the fraction of the electrons initially in the range

$$-\Delta < \xi_{\mathbf{k}} < 0$$

is now *de facto* in a complicated state with a significant distribution in momentum space via the Cooper potential, while for the others the situation is pretty similar to before. And this can be pictorially interpreted, finally, as if only those electrons are actually using the Cooper pairing mechanism (in a collective way).

Finally, the point here is that the global state presents couple symmetry in a certain region of single-particle momentum space. This means, it is more likely to scatter a pair from a couple of antipodal states to another couple, than single electrons. For those electrons deep below the Fermi energy, in the ground state, that likelihood is zero via the kinetic part of the hamiltonian and very low as well for the potential part. When saying that Cooper pairs form in the range

$$-\Delta < \xi_{\mathbf{k}} < \Delta$$

this is what we mean. This is as far as Cooper pair can be imagined in the many-body context. Now we turn to a very beautiful approach to BCS theory, a little abstract but with the capability of giving some important intuition on how the BCS state is made.

## 5.3 THE MEAN-FIELD METHOD

We have seen in Sec. 5.2 that how to build self-consistently the BCS ground state via a variational approach – which is, minimizing the energy functional. To do so we assumed a certain parametric form for the BCS ground state, in terms of $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$, which we were able to identify. Now we turn to a somewhat more sophisticated method, which relies on mean-field theory and Bogoliubov transformations for quadratic fermionic hamiltonian. This method allows for a crystalline interpretation of $\Delta$, and shows the emergence of a gap in the energy spectrum. The existence itself of said gap is the quintessence of superconductivity. We will work our way through.

First, define the **un-pairing operator**

$$\hat{\phi}_{\mathbf{k}} \equiv \hat{c}_{-\mathbf{k}\downarrow}\hat{c}_{\mathbf{k}\uparrow}$$

that un-pairs a pair with opposite momenta and spins. The order of the operators in the definition is important. Its conjugate is the **pairing operator** often encountered in literature

$$\hat{\phi}^{\dagger}_{\mathbf{k}} \equiv \hat{c}^{\dagger}_{\mathbf{k}\uparrow}\hat{c}^{\dagger}_{-\mathbf{k}\downarrow}$$

It lets us rewrite the hamiltonian as

$$\hat{\boldsymbol{H}} - \mu\hat{N} = \sum_{\mathbf{k}} \xi_{\mathbf{k}}\hat{n}_{\mathbf{k}} + \sum_{\mathbf{k}\mathbf{k}'} V_{\mathbf{k}-\mathbf{k}'}\hat{\phi}^{\dagger}_{\mathbf{k}}\hat{\phi}_{\mathbf{k}'}$$

with $\hat{n}_{\mathbf{k}} \equiv \hat{c}^{\dagger}_{\mathbf{k}\uparrow}\hat{c}_{\mathbf{k}\uparrow} + \hat{c}^{\dagger}_{\mathbf{k}\downarrow}\hat{c}_{\mathbf{k}\downarrow}$. In order to get this expression fermionic commutation rules have been used. Analogously, the BCS ground state is given by

$$|\Psi\rangle \equiv \bigotimes_{\mathbf{k}} \left[u_{\mathbf{k}} + v_{\mathbf{k}}\hat{\phi}^{\dagger}_{\mathbf{k}}\right]|\Omega\rangle \tag{5.10}$$

### 5.3.1 The pairing amplitude

Let us spend some words about the un-pairing operator,

$$\hat{\phi}_{\mathbf{k}} \equiv \hat{c}_{-\mathbf{k}\downarrow}\hat{c}_{\mathbf{k}\uparrow}$$

We see that, applying it on the BCS state,

$$\hat{\phi}_{\mathbf{k}}|\Psi\rangle = \hat{\phi}_{\mathbf{k}}\bigotimes_{\mathbf{k}'}\left[u_{\mathbf{k}'} + v_{\mathbf{k}'}\hat{\phi}^{\dagger}_{\mathbf{k}'}\right]|\Omega\rangle = v_{\mathbf{k}}\hat{\phi}_{\mathbf{k}}\hat{\phi}^{\dagger}_{\mathbf{k}}\bigotimes_{\mathbf{k}'\neq\mathbf{k}}\left[u_{\mathbf{k}'} + v_{\mathbf{k}'}\hat{\phi}^{\dagger}_{\mathbf{k}'}\right]|\Omega\rangle$$

Using anti-commutation relations for fermionic operators, it is easy to see

$$\hat{\phi}_{\mathbf{k}}\hat{\phi}^{\dagger}_{\mathbf{k}} = \left(1 - \hat{n}_{\mathbf{k}\uparrow}\right)\left(1 - \hat{n}_{-\mathbf{k}\downarrow}\right) \qquad \text{with} \qquad \hat{n}_{\mathbf{k}\sigma} = \hat{c}^{\dagger}_{\mathbf{k}\sigma}\hat{c}_{\mathbf{k}\sigma}$$

This means

$$\hat{\phi}_{\mathbf{k}}|\Psi\rangle = v_{\mathbf{k}}|\Omega\rangle$$

So, at first glance, this operator extracts information on how much a given couple of antipodal state is occupied. It's expectation value, however, tells us something more:

$$\begin{aligned}\langle\Psi|\hat{\phi}_{\mathbf{k}}|\Psi\rangle &= \langle\Omega|\bigotimes_{\mathbf{q}}\left[u^*_{\mathbf{q}}+v^*_{\mathbf{q}}\hat{\phi}_{\mathbf{q}}\right]\hat{\phi}^\dagger_{\mathbf{k}}\bigotimes_{\mathbf{q}'}\left[u_{\mathbf{q}'}+v_{\mathbf{q}'}\hat{\phi}^\dagger_{\mathbf{q}'}\right]|\Omega\rangle \\ &= \langle\Omega|\left[u^*_{\mathbf{k}}+v^*_{\mathbf{k}}\hat{\phi}_{\mathbf{k}}\right]\hat{\phi}_{\mathbf{k}}\left[u_{\mathbf{k}}+v_{\mathbf{k}}\hat{\phi}^\dagger_{\mathbf{k}}\right]|\Omega\rangle \\ &= \langle\Omega|\left[u^*_{\mathbf{k}}+v^*_{\mathbf{k}}\hat{\phi}_{\mathbf{k}}\right]v_{\mathbf{k}}|\Omega\rangle = u^*_{\mathbf{k}}v_{\mathbf{k}}\end{aligned}$$

We are using real parameters. So, it turns out

$$\langle\hat{\phi}_{\mathbf{k}}\rangle = \langle\hat{\phi}^\dagger_{\mathbf{k}}\rangle = u_{\mathbf{k}}v_{\mathbf{k}}$$

and we have already encountered this product back in Sec. 5.2.1! We called it the **pairing amplitude**, and argued back in Sec. 5.3.1 that it is a sort of measure of how much a given couple of antipodal states is connected to other pairs around the interaction shell via the potential. Here we understand completely that comment: to compute the expectation value

$$\langle\hat{c}_{-\mathbf{k}\downarrow}\hat{c}_{\mathbf{k}\uparrow}\rangle$$

is a measure of how effective we expect the couple of operators $\hat{c}_{-\mathbf{k}\downarrow}\hat{c}_{\mathbf{k}\uparrow}$ to be on the antipodal states. This has to be explained better: for a general system if we consider

$$\langle\hat{c}_{\mathbf{k}\sigma}\rangle$$

if this expectation value is zero, the state is meanly empty; if not, the state has in mean some occupation.

### 5.3.2 Mean-field approach

We substitute the pairing operator by its fluctuation around the mean value,

$$\hat{\phi} = \langle\hat{\phi}\rangle + \delta\hat{\phi}$$

Then, substituting in the potential term and neglecting quadratic contributions,

$$V_{\mathbf{k}-\mathbf{k}'}\hat{\phi}^\dagger_{\mathbf{k}}\hat{\phi}_{\mathbf{k}'} = V_{\mathbf{k}-\mathbf{k}'}\langle\hat{\phi}^\dagger_{\mathbf{k}}\rangle\langle\hat{\phi}_{\mathbf{k}'}\rangle + V_{\mathbf{k}-\mathbf{k}'}\delta\hat{\phi}^\dagger_{\mathbf{k}}\langle\hat{\phi}_{\mathbf{k}'}\rangle + V_{\mathbf{k}-\mathbf{k}'}\langle\hat{\phi}^\dagger_{\mathbf{k}}\rangle\delta\hat{\phi}_{\mathbf{k}'} + \cdots$$

The next step is quite of a turnaround: substituting only in the linear terms the same expression, $\delta\hat{\phi} = \hat{\phi} - \langle\hat{\phi}\rangle$, we get

$$V_{\mathbf{k}-\mathbf{k}'}\hat{\phi}^\dagger_{\mathbf{k}}\hat{\phi}_{\mathbf{k}'} = -V_{\mathbf{k}-\mathbf{k}'}\langle\hat{\phi}^\dagger_{\mathbf{k}}\rangle\langle\hat{\phi}_{\mathbf{k}'}\rangle + V_{\mathbf{k}-\mathbf{k}'}\hat{\phi}^\dagger_{\mathbf{k}}\langle\hat{\phi}_{\mathbf{k}'}\rangle + V_{\mathbf{k}-\mathbf{k}'}\langle\hat{\phi}^\dagger_{\mathbf{k}}\rangle\hat{\phi}_{\mathbf{k}'} + \cdots$$

This kind of argument may seem circular, as it is, and it only holds if higher-than-linear terms are in effect negligible. This kind of approach falls under **mean-field theory**. We get

$$\hat{\boldsymbol{H}} - \mu\hat{N} \simeq \sum_{\mathbf{k}}\xi_{\mathbf{k}}\hat{n}_{\mathbf{k}} - \sum_{\mathbf{k}\mathbf{k}'}V_{\mathbf{k}-\mathbf{k}'}\langle\hat{\phi}^\dagger_{\mathbf{k}}\rangle\langle\hat{\phi}_{\mathbf{k}'}\rangle + \sum_{\mathbf{k}\mathbf{k}'}V_{\mathbf{k}-\mathbf{k}'}\hat{\phi}^\dagger_{\mathbf{k}}\langle\hat{\phi}_{\mathbf{k}'}\rangle + \sum_{\mathbf{k}\mathbf{k}'}V_{\mathbf{k}-\mathbf{k}'}\langle\hat{\phi}^\dagger_{\mathbf{k}}\rangle\hat{\phi}_{\mathbf{k}'}$$

Now we define

$$\Delta_{\mathbf{k}} \equiv -\sum_{\mathbf{k}'} V_{\mathbf{k}-\mathbf{k}'} \langle \hat{\phi}_{\mathbf{k}'} \rangle \quad \Longrightarrow \quad \Delta^*_{\mathbf{k}} = -\sum_{\mathbf{k}'} V^*_{\mathbf{k}-\mathbf{k}'} \langle \hat{\phi}^\dagger_{\mathbf{k}'} \rangle \tag{5.11}$$

being $V^*_{\mathbf{k}-\mathbf{k}'} = V_{\mathbf{k}'-\mathbf{k}}$. Defined this way, based on how we interpret the pairing operator from Sec. 5.3.1, we can already understand the meaning of $\Delta_{\mathbf{k}}$: Going on, we have

$$\hat{\boldsymbol{H}} - \mu \hat{N} \simeq \sum_{\mathbf{k}} \xi_{\mathbf{k}} \hat{n}_{\mathbf{k}} - \sum_{\mathbf{k}} \Delta_{\mathbf{k}} \hat{\phi}^\dagger_{\mathbf{k}} - \sum_{\mathbf{k}} \Delta^*_{\mathbf{k}} \hat{\phi}_{\mathbf{k}} + \sum_{\mathbf{k}} \Delta_{\mathbf{k}} \langle \hat{\phi}^\dagger_{\mathbf{k}} \rangle$$

where in the last term relabeling $\mathbf{k}' \to \mathbf{k}$ has been used. We now define the shifted hamiltonian $\hat{\mathcal{H}} \equiv \hat{\boldsymbol{H}} - \mu \hat{N} - \sum \Delta \langle \hat{\phi}^\dagger \rangle$. Getting the equation compact,

$$\hat{\mathcal{H}} = \sum_{\mathbf{k}} \left[ \xi_{\mathbf{k}} \hat{c}^\dagger_{\mathbf{k}\uparrow} \hat{c}_{\mathbf{k}\uparrow} + \xi_{\mathbf{k}} \hat{c}^\dagger_{\mathbf{k}\downarrow} \hat{c}_{\mathbf{k}\downarrow} - \Delta_{\mathbf{k}} \hat{c}^\dagger_{\mathbf{k}\uparrow} \hat{c}^\dagger_{-\mathbf{k}\downarrow} - \Delta^*_{\mathbf{k}} \hat{c}_{-\mathbf{k}\downarrow} \hat{c}_{\mathbf{k}\uparrow} \right]$$

We now use some fermionic commutation relations and the symmetry of the dispersion relation $\xi_{\mathbf{k}} = \xi_{-\mathbf{k}}$,

$$\begin{aligned}
\hat{\mathcal{H}} &= \sum_{\mathbf{k}} \left[ \xi_{\mathbf{k}} \hat{c}^\dagger_{\mathbf{k}\uparrow} \hat{c}_{\mathbf{k}\uparrow} + \xi_{\mathbf{k}} \left( 1 - \hat{c}_{\mathbf{k}\downarrow} \hat{c}^\dagger_{\mathbf{k}\downarrow} \right) - \Delta_{\mathbf{k}} \hat{c}^\dagger_{\mathbf{k}\uparrow} \hat{c}^\dagger_{-\mathbf{k}\downarrow} - \Delta^*_{\mathbf{k}} \hat{c}_{-\mathbf{k}\downarrow} \hat{c}_{\mathbf{k}\uparrow} \right] \\
&= \sum_{\mathbf{k}} \left[ \xi_{\mathbf{k}} \hat{c}^\dagger_{\mathbf{k}\uparrow} \hat{c}_{\mathbf{k}\uparrow} - \xi_{\mathbf{k}} \hat{c}_{-\mathbf{k}\downarrow} \hat{c}^\dagger_{-\mathbf{k}\downarrow} - \Delta_{\mathbf{k}} \hat{c}^\dagger_{\mathbf{k}\uparrow} \hat{c}^\dagger_{-\mathbf{k}\downarrow} - \Delta^*_{\mathbf{k}} \hat{c}_{\mathbf{k}\uparrow} \hat{c}_{-\mathbf{k}\downarrow} \right] + \sum_{\mathbf{k}} \xi_{\mathbf{k}} \\
&= \sum_{\mathbf{k}} \begin{bmatrix} \hat{c}^\dagger_{\mathbf{k}\uparrow} & \hat{c}_{-\mathbf{k}\downarrow} \end{bmatrix} \begin{bmatrix} \xi_{\mathbf{k}} & -\Delta_{\mathbf{k}} \\ -\Delta^*_{\mathbf{k}} & -\xi_{\mathbf{k}} \end{bmatrix} \begin{bmatrix} \hat{c}_{\mathbf{k}\uparrow} \\ \hat{c}^\dagger_{-\mathbf{k}\downarrow} \end{bmatrix} + \sum_{\mathbf{k}} \xi_{\mathbf{k}} \\
&= \sum_{\mathbf{k}} \hat{\Phi}^\dagger_{\mathbf{k}} D_{\mathbf{k}} \hat{\Phi}_{\mathbf{k}} + \sum_{\mathbf{k}} \xi_{\mathbf{k}}
\end{aligned}$$

with $\hat{\Phi}_{\mathbf{k}}$ the spinorial operator in vector form and $D_{\mathbf{k}}$ the central matrix. We define

$$\begin{aligned}
\hat{\mathcal{K}} &\equiv \hat{\mathcal{H}} - \sum_{\mathbf{k}} \xi_{\mathbf{k}} \\
&= \hat{\boldsymbol{H}} - \mu \hat{N} - \left[ \sum_{\mathbf{k}} \Delta_{\mathbf{k}} \langle \hat{\phi}^\dagger_{\mathbf{k}} \rangle + \sum_{\mathbf{k}} \xi_{\mathbf{k}} \right]
\end{aligned} \tag{5.12}$$

We work with the new hamiltonian $\hat{\mathcal{K}}$. The (infinite) energy shift $\sum \xi$ will be re-absorbed later. The next step is to perform a very common procedure in mean-field quadratic hamiltonian.

### 5.3.3 Bogoliubov-Vitalin transformation

Rewrite $D_{\mathbf{k}}$ in terms of Pauli matrices,

$$D_{\mathbf{k}} = -\operatorname{Re}\{\Delta_{\mathbf{k}}\} \sigma^1 - \operatorname{Im}\{\Delta_{\mathbf{k}}\} \sigma^2 + \xi_{\mathbf{k}} \sigma^3$$

with $\sigma^i$ the $i$-th Pauli matrix. Thus this hamiltonian is the one for a spin in a (pseudo)magnetic field $\mathbf{b}_{\mathbf{k}}$ given by

$$\mathbf{b}_{\mathbf{k}} = \begin{bmatrix} -\operatorname{Re}\{\Delta_{\mathbf{k}}\} \\ -\operatorname{Im}\{\Delta_{\mathbf{k}}\} \\ \xi_{\mathbf{k}} \end{bmatrix}$$

Analogously we define the (pseudo)spin components as

$$[\hat{\sigma}_{\mathbf{k}}]^i \equiv \hat{\Phi}^\dagger_{\mathbf{k}} \sigma^i \hat{\Phi}_{\mathbf{k}}$$

and the hamiltonian is reduced to the simple form

$$\hat{\mathcal{K}} \equiv \sum \mathbf{b}_{\mathbf{k}} \cdot \hat{\sigma}_{\mathbf{k}}$$

This is the problem of a spin of magnitude 1 in a magnetic field (with a general $-$ sign, responsible of exchanging the eigenstates if compared to a real spin in a real field). The eigenvalues are well-known to be plus or minus the intensity of the field,

$$\pm\lambda_{\mathbf{k}} = \pm\sqrt{\xi^2_{\mathbf{k}} + |\Delta_{\mathbf{k}}|^2}$$

as can be seen easily starting from $D_{\mathbf{k}}$. Now $\pm\lambda_{\mathbf{k}}$ represent the spectrum of the system. We will dedicate the next section to comment the result. First, indicating by $U_{\mathbf{k}}$ the matrix that diagonalizes $D_{\mathbf{k}}$,

$$\Lambda_{\mathbf{k}} \equiv U_{\mathbf{k}} D_{\mathbf{k}} U^\dagger_{\mathbf{k}} = \begin{bmatrix} \lambda_{\mathbf{k}} & \\ & -\lambda_{\mathbf{k}} \end{bmatrix}$$

it is clear that

$$\hat{\Phi}^\dagger_{\mathbf{k}} D_{\mathbf{k}} \hat{\Phi}_{\mathbf{k}} = \hat{\Phi}^\dagger_{\mathbf{k}} U^\dagger_{\mathbf{k}} U_{\mathbf{k}} D_{\mathbf{k}} U^\dagger_{\mathbf{k}} U_{\mathbf{k}} \hat{\Phi}_{\mathbf{k}} = \hat{\Gamma}^\dagger_{\mathbf{k}} \Lambda_{\mathbf{k}} \hat{\Gamma}_{\mathbf{k}}$$

where we defined the spinor in the eigenvectors basis,

$$\hat{\Gamma}_{\mathbf{k}} \equiv U_{\mathbf{k}} \hat{\Phi}_{\mathbf{k}}$$

This kind of approach is called **Bogoliubov-Valatin transformation**. Now: the matrix $U_{\mathbf{k}}$ will surely mix up the operators $\hat{c}_{\mathbf{k}\uparrow}$ and $\hat{c}^\dagger_{-\mathbf{k}\downarrow}$ inside $\hat{\Phi}_{\mathbf{k}}$. The essence of the Bogoliubov approach to quadratic hamiltonian is, in fact, to find an optimal linear combination of second quantization operators (and fields) that reduces the whole problem to a system of new free fermions, born by some kind of combination of "physical" particles. This will become clear in a moment.

So, we need $U_{\mathbf{k}}$. Since this problem is physically equivalent to finding the eigenstates of a spin in tilted magnetic field, we already know that the diagonal form of the matrix is obtained by applying a rotation that aligns the $z$ axis with the field. To write the rotation, we use the result of Sec. 5.2, which embeds a real $\Delta_{\mathbf{k}}$. Then the (pseudo)magnetic field is rotated on the $zx$ plane by an angle $\zeta_{\mathbf{k}}$ with respect to the $z$ axis, defined such that

$$\frac{\Delta_{\mathbf{k}}}{\xi_{\mathbf{k}}} \equiv \tan 2\theta_{\mathbf{k}} \tag{5.13}$$

The notation choice is not casual. We have already seen a similar relation, in Eq. (5.3)! In Fig. 5.4 a sketch of the field and its eigenvectors is reported. To align the $z$ axis with the field we need to rotate the $zx$ plane by an angle $2\theta_{\mathbf{k}}$ clockwise; this is equivalent to rotating the vectors by the same angle anti-clockwise, which is, by a positive amount $2\theta_{\mathbf{k}}$. The SO(3) representation of this rotation is

$$U^{\mathrm{SO}(3)}_{\mathbf{k}} = \begin{bmatrix} \cos 2\theta_{\mathbf{k}} & 0 & -\sin 2\theta_{\mathbf{k}} \\ 0 & 1 & 0 \\ \sin 2\theta_{\mathbf{k}} & 0 & \cos 2\theta_{\mathbf{k}} \end{bmatrix}$$

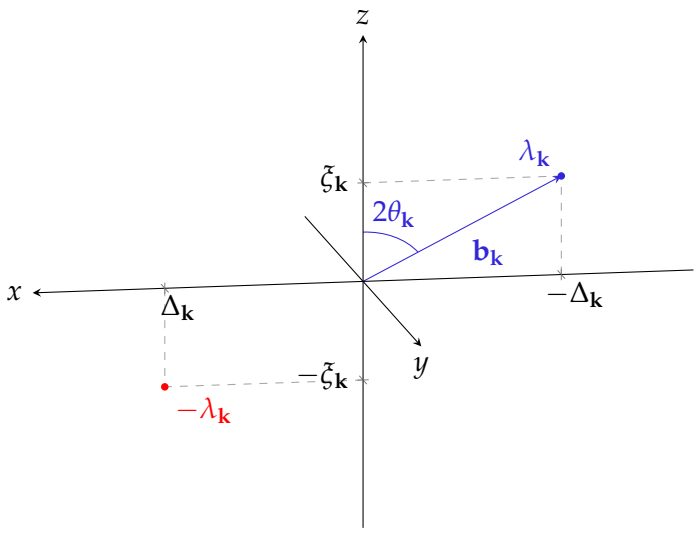



**Figure 5.4:** Representation of the (pseudo)magnetic field $\mathbf{b}_{\mathbf{k}}$. Obviously the notation $xyz$ is unphysical and in any way related to real space. The dots represent the eigenvectors of the problem, and the relative eigenvalue is indicated. As in any problem of this kind, the highest eigenvalue $\lambda_{\mathbf{k}}$ is reached for a combination of spinors perfectly aligned with the field, while the lowest eigenvalue is its antipodal point.

We need its mapping onto its SU(2) version. In general, a rotation of angle $\alpha$ around the versor $\bar{\mathbf{n}}$ is represented in the group by

$$\exp\left\{-i\frac{\alpha}{2}\bar{\mathbf{n}}\cdot\boldsymbol{\sigma}\right\}$$

and for us $\bar{\mathbf{n}} = \bar{\mathbf{y}}$, $\alpha = 2\theta_{\mathbf{k}}$; expanding:

$$\begin{aligned} U_{\mathbf{k}}^{\mathrm{SU}(2)} &= \exp\left\{-i\theta_{\mathbf{k}}\sigma^2\right\} \\ &= \mathbb{1}\cos\theta_{\mathbf{k}} - i\sigma^2\sin\theta_{\mathbf{k}} = \begin{bmatrix}\cos\theta_{\mathbf{k}} & -\sin\theta_{\mathbf{k}} \\ \sin\theta_{\mathbf{k}} & \cos\theta_{\mathbf{k}}\end{bmatrix} \end{aligned}$$

We omit now the representation superscript. It can be checked easily that

$$\begin{bmatrix}\cos\theta_{\mathbf{k}} & -\sin\theta_{\mathbf{k}} \\ \sin\theta_{\mathbf{k}} & \cos\theta_{\mathbf{k}}\end{bmatrix}\begin{bmatrix}\xi_{\mathbf{k}} & -\Delta_{\mathbf{k}} \\ -\Delta_{\mathbf{k}} & -\xi_{\mathbf{k}}\end{bmatrix}\begin{bmatrix}\cos\theta_{\mathbf{k}} & \sin\theta_{\mathbf{k}} \\ -\sin\theta_{\mathbf{k}} & \cos\theta_{\mathbf{k}}\end{bmatrix} = \begin{bmatrix}\lambda_{\mathbf{k}} & \\ & -\lambda_{\mathbf{k}}\end{bmatrix}$$

Then we can read the spinor in the eigenvectors basis just by applying the rotation,

$$\begin{aligned} \hat{\Gamma}_{\mathbf{k}} &= U_{\mathbf{k}}\hat{\Phi}_{\mathbf{k}} \\ &= \begin{bmatrix}\cos\theta_{\mathbf{k}} & -\sin\theta_{\mathbf{k}} \\ \sin\theta_{\mathbf{k}} & \cos\theta_{\mathbf{k}}\end{bmatrix}\begin{bmatrix}\hat{c}_{\mathbf{k}\uparrow} \\ \hat{c}^{\dagger}_{-\mathbf{k}\downarrow}\end{bmatrix} \\ &= \begin{bmatrix}\cos\theta_{\mathbf{k}}\hat{c}_{\mathbf{k}\uparrow} - \sin\theta_{\mathbf{k}}\hat{c}^{\dagger}_{-\mathbf{k}\downarrow} \\ \sin\theta_{\mathbf{k}}\hat{c}_{\mathbf{k}\uparrow} + \cos\theta_{\mathbf{k}}\hat{c}^{\dagger}_{-\mathbf{k}\downarrow}\end{bmatrix} \equiv \begin{bmatrix}\hat{\gamma}_{\mathbf{k}\uparrow} \\ \hat{\gamma}^{\dagger}_{-\mathbf{k}\downarrow}\end{bmatrix} \end{aligned}$$

where in the last step we defined a pair of new fermionic operators, $\hat{\gamma}_{\mathbf{k}\uparrow}$ and $\hat{\gamma}^{\dagger}_{-\mathbf{k}\downarrow}$. It makes sense to define them this way, because in

$$\hat{\gamma}_{\mathbf{k}\uparrow} \equiv \cos\theta_{\mathbf{k}}\hat{c}_{\mathbf{k}\uparrow} - \sin\theta_{\mathbf{k}}\hat{c}^{\dagger}_{-\mathbf{k}\downarrow}$$

to create an excitation with momentum $-\mathbf{k}$ and spin $\downarrow$ is kind of equivalent to annihilating an excitation with momentum $\mathbf{k}$ and spin $\uparrow$. The two operations are not equivalent with respect to the number of electrons in the system, however the change in total momentum and spin is the same. An analogous consideration holds for

$$\hat{\gamma}^\dagger_{-\mathbf{k}\downarrow} \equiv \sin\theta_\mathbf{k}\hat{c}_{\mathbf{k}\uparrow} + \cos\theta_\mathbf{k}\hat{c}^\dagger_{-\mathbf{k}\downarrow}$$

For what concerns the $\hat{\gamma}$ operators, the subscripts must be intended in this way and do not have the physical meaning they have for the $\hat{c}$ operators. It can be checked that the $\hat{\gamma}$ operators obey the common anti-commutation rules. Thanks to these transformations the hamiltonian reads

$$\hat{\mathcal{K}} = \sum_\mathbf{k} \hat{\Gamma}^\dagger_\mathbf{k} \begin{bmatrix} \lambda_\mathbf{k} & \\ & -\lambda_\mathbf{k} \end{bmatrix} \hat{\Gamma}_\mathbf{k} = \sum_\mathbf{k} \left[ \lambda_\mathbf{k}\hat{\gamma}^\dagger_{\mathbf{k}\uparrow}\hat{\gamma}_{\mathbf{k}\uparrow} - \lambda_\mathbf{k}\hat{\gamma}_{-\mathbf{k}\downarrow}\hat{\gamma}^\dagger_{-\mathbf{k}\downarrow} \right] \tag{5.14}$$

Next section is devoted to commenting and further developing this hamiltonian.

### 5.3.4 The Bogoliubov fermions

We adopt the following notation to indicate sums over the shell and outside of it

$$\sum_\mathcal{S} \equiv \sum_{|\xi_\mathbf{k}|<\hbar\omega_D} \qquad \text{and} \qquad \sum_{\mathbb{R}^D\setminus\mathcal{S}} \equiv \sum_{|\xi_\mathbf{k}|<\hbar\omega_D}$$

Take Eq. (5.14). Using fermionic rules ($\hat{\gamma}_{-\mathbf{k}\downarrow}\hat{\gamma}^\dagger_{-\mathbf{k}\downarrow} = 1 - \hat{\gamma}^\dagger_{-\mathbf{k}\downarrow}\hat{\gamma}_{-\mathbf{k}\downarrow}$), the dispersion symmetry ($\lambda_\mathbf{k} = \lambda_{-\mathbf{k}}$), and recalling the definition of $\hat{\mathcal{K}}$ in Eq. (5.12) we have

$$\hat{\mathcal{H}} - \sum_\mathbf{k} \xi_\mathbf{k} = -\sum_\mathbf{k} \lambda_\mathbf{k} + \sum_\mathbf{k} \lambda_\mathbf{k} \left[ \hat{\gamma}^\dagger_{\mathbf{k}\uparrow}\hat{\gamma}_{\mathbf{k}\uparrow} + \hat{\gamma}^\dagger_{\mathbf{k}\downarrow}\hat{\gamma}_{\mathbf{k}\downarrow} \right]$$

Since:

$$\sum_\mathbf{k} \xi_\mathbf{k} = \sum_{\mathbb{R}^D\setminus\mathcal{S}} \xi_\mathbf{k} + \sum_\mathcal{S} \xi_\mathbf{k}$$

$$\sum_\mathbf{k} \lambda_\mathbf{k} = \sum_{\mathbb{R}^D\setminus\mathcal{S}} \xi_\mathbf{k} + \sum_\mathcal{S} \sqrt{\xi^2_\mathbf{k} + \Delta^2_\mathbf{k}}$$

the final form of the hamiltonian is

$$\hat{\mathcal{H}} = \sum_\mathcal{S} \left[ \xi_\mathbf{k} - \sqrt{\xi^2_\mathbf{k} + \Delta^2_\mathbf{k}} \right] + \sum_\mathbf{k} \lambda_\mathbf{k} \left[ \hat{\gamma}^\dagger_{\mathbf{k}\uparrow}\hat{\gamma}_{\mathbf{k}\uparrow} + \hat{\gamma}^\dagger_{\mathbf{k}\downarrow}\hat{\gamma}_{\mathbf{k}\downarrow} \right]$$

Now everything is clear. To use the solemn and inspiring words of one of the major art pieces of the author's country, "Nessuno è più basito, chiaro? Nessuno è più basito, nessuno è sorpreso, ognuno di voi ha capito tutto. Nei primi piani fate sì con la testa, che avete capito e state sereni". As long as the un-pairing operator fluctuates negligibly, the BCS hamiltonian can be mapped on a system of free fermions described by the $\hat{\gamma}$ operators. Those fermions are divided in two classes, $\uparrow$ and $\downarrow$, distinguished by the change in total spin the system obtains when one of these fermions is added or removed. Both classes have dispersion

$$\lambda_\mathbf{k} = \sqrt{\xi^2_\mathbf{k} + \Delta^2_\mathbf{k}}$$

These fermions are collective excitations, **quasiparticles**. Note that the orientation in abstract space of the (pseudo)field obviously depends on $\mathbf{k}$. Take Fig. 5.4: if we use the known value

$$\Delta_{\mathbf{k}} = \Delta A(\mathbf{k}) \qquad \text{with} \qquad \Delta = 2\hbar\omega_D e^{-2/\rho_0 V_0}$$

(be careful! We need to justify the identification of this $\Delta_{\mathbf{k}}$ with the self-consistent parameter derived in the above section via the variational approach; we will do this in Sec. 5.3.5) inside the interaction shell $\Delta_{\mathbf{k}} = \Delta$, while the $\xi_{\mathbf{k}}$ component will increase in magnitude as move $\mathbf{k}$ away from the Fermi surface, in both directions. Eventually we exit the shell: for $|\xi_{\mathbf{k}}| > \hbar\omega_D$, we have $\Delta_{\mathbf{k}} = 0$ and therefore the magnetic field has only the $z$ component and $\theta_{\mathbf{k}} = 0$. This means, as it is evident from the beginning, that the Bogoliubov $\hat{\gamma}$ operators coincide with the electron $\hat{c}$ operators, and the hamiltonian can be written as

$$\hat{\mathcal{H}} = \sum_{\mathbb{R}^D \setminus \mathcal{S}} \xi_{\mathbf{k}} \left[\hat{c}^\dagger_{\mathbf{k}\uparrow}\hat{c}_{\mathbf{k}\uparrow} + \hat{c}^\dagger_{\mathbf{k}\downarrow}\hat{c}_{\mathbf{k}\downarrow}\right] + \sum_{\mathcal{S}} \left[\xi_{\mathbf{k}} - \sqrt{\xi^2_{\mathbf{k}} + \Delta^2}\right] + \sum_{\mathcal{S}} \sqrt{\xi^2_{\mathbf{k}} + \Delta^2} \left[\hat{\gamma}^\dagger_{\mathbf{k}\uparrow}\hat{\gamma}_{\mathbf{k}\uparrow} + \hat{\gamma}^\dagger_{\mathbf{k}\downarrow}\hat{\gamma}_{\mathbf{k}\downarrow}\right]$$

It is not necessary and sometime confusing to distinguish $c$ particles from $\gamma$ particles, but the above equation allows us to note that the mean-field approach really has some effect only inside the interaction shell, as it makes sense to. Outside the interaction shell the Sommerfeld free electron model works well.

### 5.3.5 Gap equation

Recall how we defined $\Delta_{\mathbf{k}}$, back in Eq. (5.11):

$$\Delta_{\mathbf{k}} \equiv -\sum_{\mathbf{k}'} V_{\mathbf{k}-\mathbf{k}'} \langle \hat{\phi}_{\mathbf{k}'} \rangle$$

We already know $\langle \hat{\phi}_{\mathbf{k}'} \rangle$ from Sec. 5.3.1, but we go now through an equivalent and instructive derivation. We defined the (pseudo)spin components as

$$[\hat{\sigma}_{\mathbf{k}}]^i \equiv \hat{\Phi}^\dagger_{\mathbf{k}} \sigma^i \hat{\Phi}_{\mathbf{k}}$$

and it is easy to check

$$\hat{\sigma}^1_{\mathbf{k}} = \hat{\phi}^\dagger_{\mathbf{k}} + \hat{\phi}_{\mathbf{k}} \qquad \hat{\sigma}^2_{\mathbf{k}} = -i\hat{\phi}^\dagger_{\mathbf{k}} + i\hat{\phi}_{\mathbf{k}}$$

This implies:

$$\langle \hat{\phi}_{\mathbf{k}} \rangle = \frac{1}{2}\left[\langle \hat{\sigma}^1_{\mathbf{k}} \rangle - i \langle \hat{\sigma}^2_{\mathbf{k}} \rangle\right] \tag{5.15}$$

$$\langle \hat{\phi}^\dagger_{\mathbf{k}} \rangle = \frac{1}{2}\left[\langle \hat{\sigma}^1_{\mathbf{k}} \rangle + i \langle \hat{\sigma}^2_{\mathbf{k}} \rangle\right] \tag{5.16}$$

We also know the transformed version of the spinors,

$$\hat{\Gamma}_{\mathbf{k}} = U_{\mathbf{k}} \hat{\Phi}_{\mathbf{k}} \quad \Longrightarrow \quad [\hat{\sigma}_{\mathbf{k}}]^i \equiv \hat{\Gamma}^\dagger_{\mathbf{k}} U_{\mathbf{k}} \sigma^i U^\dagger_{\mathbf{k}} \hat{\Gamma}_{\mathbf{k}}$$

Being $U_{\mathbf{k}} = e^{-i\theta_{\mathbf{k}}\sigma^2}$, we have

$$\begin{aligned} U_{\mathbf{k}}\sigma^1 U_{\mathbf{k}}^\dagger &= e^{-i\theta_{\mathbf{k}}\sigma^2}\sigma^1 e^{i\theta_{\mathbf{k}}\sigma^2} \\ &= \left(\cos\theta_{\mathbf{k}} - i\sin\theta_{\mathbf{k}}\sigma^2\right)\sigma^1\left(\cos\theta_{\mathbf{k}} + i\sin\theta_{\mathbf{k}}\sigma^2\right) \\ &= \cos^2\theta_{\mathbf{k}}\sigma^1 + \sin^2\theta_{\mathbf{k}}\sigma^2\sigma^1\sigma^2 - i\sin\theta_{\mathbf{k}}\cos\theta_{\mathbf{k}}\left[\sigma^2,\sigma^1\right] \\ &= \left(\cos^2\theta_{\mathbf{k}} - \sin^2\theta_{\mathbf{k}}\right)\sigma^1 - 2\sin\theta_{\mathbf{k}}\cos\theta_{\mathbf{k}}\sigma^3 \\ &= \cos 2\theta_{\mathbf{k}}\sigma^1 - \sin 2\theta_{\mathbf{k}}\sigma^3 \end{aligned}$$

having used commutation and anti-commutations properties of the Pauli matrices. It follows

$$\langle\hat{\sigma}^1_{\mathbf{k}}\rangle = \cos 2\theta_{\mathbf{k}}\langle\hat{\Gamma}^\dagger_{\mathbf{k}}\sigma^1\hat{\Gamma}_{\mathbf{k}}\rangle - \sin 2\theta_{\mathbf{k}}\langle\hat{\Gamma}^\dagger_{\mathbf{k}}\sigma^3\hat{\Gamma}_{\mathbf{k}}\rangle$$

As $\langle\hat{\Phi}^\dagger_{\mathbf{k}}\sigma^i\hat{\Phi}_{\mathbf{k}}\rangle$ represents the $i$-th component of the (pseudo)spin in the original frame (the one in Fig. 5.4), now $\langle\hat{\Gamma}^\dagger_{\mathbf{k}}\sigma^i\hat{\Gamma}_{\mathbf{k}}\rangle$ does the same in the tilted frame, with the $z$ axis aligned to the (pseudo)magnetic field $\mathbf{b}_{\mathbf{k}}$. Then

$$\begin{aligned} \langle\hat{\Gamma}^\dagger_{\mathbf{k}}\sigma^1\hat{\Gamma}_{\mathbf{k}}\rangle &= \langle\hat{\gamma}^\dagger_{\mathbf{k}\uparrow}\hat{\gamma}^\dagger_{-\mathbf{k}\downarrow}\rangle + \langle\hat{\gamma}_{-\mathbf{k}\downarrow}\hat{\gamma}_{\mathbf{k}\uparrow}\rangle \\ \langle\hat{\Gamma}^\dagger_{\mathbf{k}}\sigma^2\hat{\Gamma}_{\mathbf{k}}\rangle &= -i\langle\hat{\gamma}^\dagger_{\mathbf{k}\uparrow}\hat{\gamma}^\dagger_{-\mathbf{k}\downarrow}\rangle + i\langle\hat{\gamma}_{-\mathbf{k}\downarrow}\hat{\gamma}_{\mathbf{k}\uparrow}\rangle \\ \langle\hat{\Gamma}^\dagger_{\mathbf{k}}\sigma^3\hat{\Gamma}_{\mathbf{k}}\rangle &= \langle\hat{\gamma}^\dagger_{\mathbf{k}\uparrow}\hat{\gamma}_{\mathbf{k}\uparrow}\rangle - \langle\hat{\gamma}_{-\mathbf{k}\downarrow}\hat{\gamma}^\dagger_{-\mathbf{k}\downarrow}\rangle \end{aligned}$$

We are working now at $T = 0$. Calculations at finite temperature will be drawn later. Making use of Fermionic rules and performing all calculations on the BCS ground state, it is easy to see

$$\begin{aligned} \langle\hat{\Gamma}^\dagger_{\mathbf{k}}\sigma^1\hat{\Gamma}_{\mathbf{k}}\rangle &= 0 \\ \langle\hat{\Gamma}^\dagger_{\mathbf{k}}\sigma^2\hat{\Gamma}_{\mathbf{k}}\rangle &= 0 \\ \langle\hat{\Gamma}^\dagger_{\mathbf{k}}\sigma^3\hat{\Gamma}_{\mathbf{k}}\rangle &= -1 \end{aligned}$$

In Sec. 5.4.1 we will see in detail how the $\hat{\gamma}$ quasiparticles operators act on $|\Psi\rangle$; for now, it suffices to see that the quasiparticle populations are zero in the ground state, and the pairing and un-pairing amplitude of quasiparticles is therefore zero. Then,

$$\langle\hat{\sigma}^1_{\mathbf{k}}\rangle = \sin 2\theta_{\mathbf{k}} = \frac{\Delta_{\mathbf{k}}}{\lambda_{\mathbf{k}}}$$

Heuristically, this makes sense. Recalling Fig. 5.4, the lowest energy state for a real spin in a real magnetic field with such coupling (remember the $+$ sign in the hamiltonian, implementing anti-ferromagnetic coupling) is the one with this expectation value for the $x$ component. By the same identical arguments one shows

$$\langle\hat{\sigma}^2_{\mathbf{k}}\rangle = 0$$

By inserting these result in Eq. (5.15) and then in Eq. (5.11), we have

$$\Delta_{\mathbf{k}} = -\frac{1}{2}\sum_{\mathbf{k}'} V_{\mathbf{k}-\mathbf{k}'}\frac{\Delta_{\mathbf{k}'}}{\lambda_{\mathbf{k}'}}$$

and this is precisely Eq. (5.7). From now on, the calculation is the same and leads to the identical result as the variational solution. This argument allows us to perfectly identify the two procedures and use them interchangeably. It represents also a important self-consistency check of the mean-field approach.

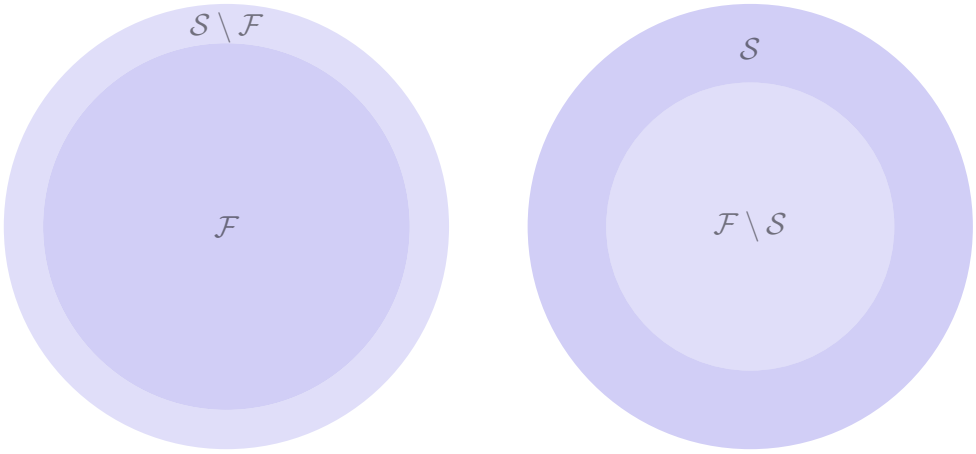



**Figure 5.5:** Sketch of the two sum domains cited in Sec. 5.3.6: $\mathcal{F}$ is the Fermi sphere, and $\mathcal{S}$ is the interaction shell of width $2\hbar\omega_D$.

### 5.3.6 The condensation energy

Take $\hat{\mathcal{H}}$: we reintroduce the mean-field constant contribution to energy,

$$\hat{\mathcal{H}} + \sum_{\mathbf{k}} \Delta_{\mathbf{k}} \langle \hat{\phi}^\dagger_{\mathbf{k}} \rangle = \hat{\boldsymbol{H}} - \mu \hat{N}$$

Then, based on our knowledge of $\langle \hat{\phi}^\dagger_{\mathbf{k}} \rangle$ from last section and Sec. 5.3.1, we expand

$$\sum_{\mathbf{k}} \Delta_{\mathbf{k}} \langle \hat{\phi}^\dagger_{\mathbf{k}} \rangle = \sum_{\mathbf{k}} \frac{\Delta^2_{\mathbf{k}}}{2\lambda_{\mathbf{k}}} = \sum_{\mathcal{S}} \frac{\Delta^2}{2\sqrt{\xi^2_{\mathbf{k}} + \Delta^2}}$$

It follows:

$$\begin{aligned} \hat{\boldsymbol{H}} - \mu \hat{N} = \sum_{\mathcal{S}} \left[ \xi_{\mathbf{k}} - \sqrt{\xi^2_{\mathbf{k}} + \Delta^2} + \frac{\Delta^2}{2\sqrt{\xi^2_{\mathbf{k}} + \Delta^2}} \right] \\ + \sum_{\mathbb{R}^D \setminus \mathcal{S}} \xi_{\mathbf{k}} \left[ \hat{c}^\dagger_{\mathbf{k}\uparrow} \hat{c}_{\mathbf{k}\uparrow} + \hat{c}^\dagger_{\mathbf{k}\downarrow} \hat{c}_{\mathbf{k}\downarrow} \right] + \sum_{\mathcal{S}} \sqrt{\xi^2_{\mathbf{k}} + \Delta^2} \left[ \hat{\gamma}^\dagger_{\mathbf{k}\uparrow} \hat{\gamma}_{\mathbf{k}\uparrow} + \hat{\gamma}^\dagger_{\mathbf{k}\downarrow} \hat{\gamma}_{\mathbf{k}\downarrow} \right] \end{aligned}$$

If we evaluate over the BCS ground-state, we get the superconductor ground-state energy $E^{(\mathrm{s})}$

$$\begin{aligned} E^{(\mathrm{s})} &\equiv \langle \hat{\boldsymbol{H}} - \mu \hat{N} \rangle \\ &= \sum_{\mathcal{S}} \left[ \xi_{\mathbf{k}} - \sqrt{\xi^2_{\mathbf{k}} + \Delta^2} + \frac{\Delta^2}{2\sqrt{\xi^2_{\mathbf{k}} + \Delta^2}} \right] + 2 \sum_{\mathcal{F} \setminus \mathcal{S}} \xi_{\mathbf{k}} \end{aligned}$$

where $\mathcal{F} \setminus \mathcal{S}$ indicates the internal portion of the Fermi sphere, obtained by excluding from the full sphere $\mathcal{F}$ the internal part of the shell. These domains are represented in Fig. 5.5. For a normal metal the ground state energy $E^{(\mathrm{n})}$ is of course

$$E^{(\mathrm{n})} = 2 \sum_{\mathcal{F}} \xi_{\mathbf{k}}$$

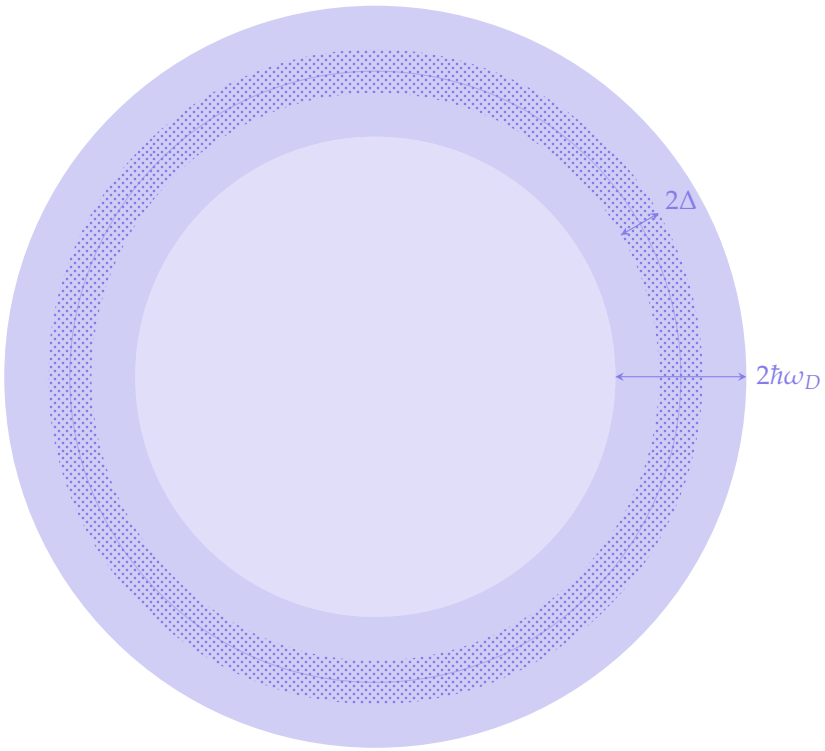



**Figure 5.6:** The shell and sub-shell discussed in Sec. 5.3.6; the onion-like shell of width $2\Delta$ includes those single-particle states highly connected via the Cooper pairing mechanism. Proportions are exaggerated.

Their difference is the **condensation energy**: we use again the continuous approximation,

$$E^{(s)} - E^{(n)} = \sum_{\mathcal{S}} \left[ \xi_{\mathbf{k}} - \sqrt{\xi_{\mathbf{k}}^2 + \Delta^2} + \frac{\Delta^2}{2\sqrt{\xi_{\mathbf{k}}^2 + \Delta^2}} \right] + 2 \sum_{\mathcal{S}} \xi_{\mathbf{k}}$$
$$\simeq \rho_0 \int_{-\hbar\omega_D}^{\hbar\omega_D} d\xi \left[ \xi - \sqrt{\xi^2 + \Delta^2} + \frac{\Delta^2}{2\sqrt{\xi^2 + \Delta^2}} \right] + 2\rho_0 \int_{-\hbar\omega_D}^{0} d\xi\, \xi$$

The first term of the first integral vanishes. We rewrite

$$-\sqrt{\xi^2 + \Delta^2} + \frac{\Delta^2}{2\sqrt{\xi^2 + \Delta^2}} = -\frac{\xi^2 + \Delta^2}{\sqrt{\xi^2 + \Delta^2}} + \frac{\Delta^2}{2\sqrt{\xi^2 + \Delta^2}}$$
$$= -\frac{2\xi^2 + \Delta^2}{2\sqrt{\xi^2 + \Delta^2}}$$

and since

$$\int \frac{2\xi^2 + \Delta^2}{2\sqrt{\xi^2 + \Delta^2}} = \frac{\xi}{2}\sqrt{\xi^2 + \Delta^2}$$

we solve the integral, implementing a Taylor expansion for the weak binding limit $\Delta \ll \hbar\omega_D$

$$E^{(s)} - E^{(n)} = \rho_0 \left[ -\hbar\omega_D \sqrt{(\hbar\omega_D)^2 + \Delta^2} + (\hbar\omega_D)^2 \right] \simeq -\frac{1}{2}\rho_0 \Delta^2$$

This result is of great importance (and can be obtained equivalently by the means of the variational approach). First, it highlights a quadratic dependence of the condensation energy on the gap $\Delta$, and this will become relevant later. Second, it gives us another proof for interpreting the energy range

$$-\Delta < \xi_{\mathbf{k}} < \Delta$$

as the one relevant for pairing. In fact, consider Fig. 5.6: momentum space is represented. The darker shell of width $2\hbar\omega_D$ is the interaction shell $\mathcal{S}$, while the dotted onion-like sub-shell of width $2\Delta$ is the region of momentum space where the pairing amplitude is significantly different from zero, as discussed in Sec. 5.2.1; from Sec. 5.2.3 we know that it is in this region – which means: for the single-particle states inside this region – that the Cooper pairing mechanism is mainly at work.

Now, apart from the factor $1/2$, of order 1, for a filled shell in a range $\Delta$ just below the Fermi energy there are approximately $\rho_0\Delta$ electrons; if those are the electrons scattered in Cooper bound states, each one gains an energy $-\Delta$, leading to a global energy gain

$$-\rho\Delta^2$$

while if the whole big shell down to $-\hbar\omega_D$ participated, the functional form of the condensation energy should have been something like

$$-\rho\hbar\omega_D\Delta$$

which is much smaller, being $\hbar\omega_D \ll \Delta$. The functional form would have been linear in $\Delta$. So, to represent the BCS ground state as something pretty similar to a filled Fermi sphere up to $\xi_{\mathbf{k}} = -\Delta$ and a thin shell $-\Delta < \xi_{\mathbf{k}} < \Delta$ of complicated collective Cooper pairing with statistical couple correlation between states is pictorial, but reasonable. But, while reasonable, this picture can also be misleading, if we confuse condensation with conduction properties. After all we are talking about a superconductor – so, are the electrons in the shell of width $2\Delta$ those carrying superconducting current?

As Goodstein correctly points out (his notation is: $\Delta \leftrightarrow \Delta_0$):

> **D. L. Goodstein**: ***States of Matter*** ([4] @ 5.3.d)
>
> If $\Delta_0$ were the binding energy per electron in the pairing process, we would expect the energy of the superconducting ground state to be lower than the normal ground state by $N\Delta_0$. Instead the difference in energy [...] go on...

## 5.4 THE IMPORTANCE OF BEING GAPPED

### (A TRIVIAL COMEDY FOR SUPERCONDUCTING PEOPLE)

The first thing we need to understand is what the Bogoliubov $\gamma$ fermions are and why their appearance makes sense in the context of superconductivity. So: how to imagine them?

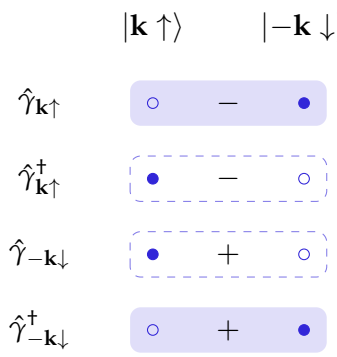



**Figure 5.7:** Pictorial representation of the action of the $\hat{\gamma}$ operators.

### 5.4.1 A closer look to the quasiparticle operators

Take the $\hat{\gamma}$ operators:

$$\begin{cases} \hat{\gamma}_{\mathbf{k}\uparrow} = \cos\theta_{\mathbf{k}}\hat{c}_{\mathbf{k}\uparrow} - \sin\theta_{\mathbf{k}}\hat{c}^\dagger_{-\mathbf{k}\downarrow} \\ \hat{\gamma}^\dagger_{\mathbf{k}\uparrow} = \cos\theta_{\mathbf{k}}\hat{c}^\dagger_{\mathbf{k}\uparrow} - \sin\theta_{\mathbf{k}}\hat{c}_{-\mathbf{k}\downarrow} \end{cases} \qquad \begin{cases} \hat{\gamma}_{-\mathbf{k}\downarrow} = \cos\theta_{\mathbf{k}}\hat{c}_{-\mathbf{k}\downarrow} + \sin\theta_{\mathbf{k}}\hat{c}^\dagger_{\mathbf{k}\uparrow} \\ \hat{\gamma}^\dagger_{-\mathbf{k}\downarrow} = \cos\theta_{\mathbf{k}}\hat{c}^\dagger_{-\mathbf{k}\downarrow} + \sin\theta_{\mathbf{k}}\hat{c}_{\mathbf{k}\uparrow} \end{cases}$$

We want to deepen our understanding of them. We said, at the end of Sec. 5.3.4, that outside the shell $\mathcal{S}$ they can be identified with the ordinary $\hat{c}$ operators. But how are they made in general?

One good starting point is their action on the BCS state $|\Psi\rangle$. The system ground state will be evidently one with zero $\gamma$ particles (we have already used this result), so this property $|\Psi\rangle$ should present. Take, say, $\hat{\gamma}_{\mathbf{k}\uparrow}$,

$$\begin{aligned} \hat{\gamma}_{\mathbf{k}\uparrow} \bigotimes_{\mathbf{q}} \left[u_{\mathbf{q}} + v_{\mathbf{q}}\hat{\phi}^\dagger_{\mathbf{q}}\right]|\Omega\rangle \\ = \bigotimes_{\mathbf{q}\neq\mathbf{k}} \left[u_{\mathbf{q}} + v_{\mathbf{q}}\hat{\phi}^\dagger_{\mathbf{q}}\right] \left[\cos\theta_{\mathbf{k}}\hat{c}_{\mathbf{k}\uparrow} - \sin\theta_{\mathbf{k}}\hat{c}^\dagger_{-\mathbf{k}\downarrow}\right] \left[u_{\mathbf{k}} + v_{\mathbf{k}}\hat{\phi}^\dagger_{\mathbf{k}}\right]|\Omega\rangle \end{aligned}$$

with $\hat{\phi}^\dagger_{\mathbf{k}} = \hat{c}^\dagger_{\mathbf{k}\uparrow}\hat{c}^\dagger_{-\mathbf{k}\downarrow}$ the paring operator. The cosine term can only couple with the $v$ term, otherwise it would annihilate the vacuum; the sine term cannot couple with the $v$ term, because it fills the $-\mathbf{k}\downarrow$ state. Then we are left with

$$\bigotimes_{\mathbf{q}\neq\mathbf{k}} \left[u_{\mathbf{q}} + v_{\mathbf{q}}\hat{\phi}^\dagger_{\mathbf{q}}\right] \left[-\sin\theta_{\mathbf{k}}u_{\mathbf{k}}\hat{c}^\dagger_{-\mathbf{k}\downarrow} + \cos_{\mathbf{k}} v_{\mathbf{k}}\hat{c}_{\mathbf{k}\uparrow}\hat{c}^\dagger_{\mathbf{k}\uparrow}\hat{c}^\dagger_{-\mathbf{k}\downarrow}\right]|\Omega\rangle$$

Since $\hat{c}_{\mathbf{k}\uparrow}\hat{c}^\dagger_{\mathbf{k}\uparrow}\hat{c}^\dagger_{-\mathbf{k}\downarrow} = (1 - \hat{c}^\dagger_{\mathbf{k}\uparrow}\hat{c}_{\mathbf{k}\uparrow})\hat{c}^\dagger_{-\mathbf{k}\downarrow}$ and $|\Omega\rangle$ is empty in the state $\mathbf{k}\uparrow$,

$$\bigotimes_{\mathbf{q}\neq\mathbf{k}} \left[u_{\mathbf{q}} + v_{\mathbf{q}}\hat{\phi}^\dagger_{\mathbf{q}}\right] \left[-u_{\mathbf{k}}\sin\theta_{\mathbf{k}} + v_{\mathbf{k}}\cos\theta_{\mathbf{k}}\right]\hat{c}^\dagger_{-\mathbf{k}\downarrow}|\Omega\rangle$$

Then $|\Psi\rangle$ is the ground state if

$$u_{\mathbf{k}}\sin\theta_{\mathbf{k}} = v_{\mathbf{k}}\cos\theta_{\mathbf{k}}$$

which is evidently satisfied by what we know from the variational solution in Sec. 5.2. Thus $\hat{\gamma}_{\mathbf{k}\uparrow}|\Psi\rangle = 0$, as expected. Similarly one shows $\hat{\gamma}_{-\mathbf{k}\downarrow}|\Psi\rangle = 0$.

What about $\hat{\gamma}^\dagger_{\mathbf{k}\uparrow}$ and $\hat{\gamma}^\dagger_{-\mathbf{k}\downarrow}$? By inspection

$$\hat{\gamma}^\dagger_{\mathbf{k}\uparrow} \bigotimes_{\mathbf{q}} \left[u_{\mathbf{q}} + v_{\mathbf{q}}\hat{\phi}^\dagger_{\mathbf{q}}\right] |\Omega\rangle$$
$$= \bigotimes_{\mathbf{q}\neq\mathbf{k}} \left[u_{\mathbf{q}} + v_{\mathbf{q}}\hat{\phi}^\dagger_{\mathbf{q}}\right] \left[\cos\theta_{\mathbf{k}}\hat{c}^\dagger_{\mathbf{k}\uparrow} - \sin\theta_{\mathbf{k}}\hat{c}_{-\mathbf{k}\downarrow}\right] \left[u_{\mathbf{k}} + v_{\mathbf{k}}\hat{\phi}^\dagger_{\mathbf{k}}\right] |\Omega\rangle$$

and with analogous arguments one finds

$$\hat{\gamma}^\dagger_{\mathbf{k}\uparrow} |\Psi\rangle = \bigotimes_{\mathbf{q}\neq\mathbf{k}} \left[u_{\mathbf{q}} + v_{\mathbf{q}}\hat{\phi}^\dagger_{\mathbf{q}}\right] \hat{c}^\dagger_{\mathbf{k}\uparrow} |\Omega\rangle \qquad \hat{\gamma}^\dagger_{-\mathbf{k}\downarrow} |\Psi\rangle = \bigotimes_{\mathbf{q}\neq\mathbf{k}} \left[u_{\mathbf{q}} + v_{\mathbf{q}}\hat{\phi}^\dagger_{\mathbf{q}}\right] \hat{c}^\dagger_{-\mathbf{k}\downarrow} |\Omega\rangle$$

Then to add one $\gamma$ excitation to the ground-state is precisely to substitute one Cooper pair occupying two antipodal single-particle states, with just one electron precisely in one of the two states. The other state is empty (there is one hole): if it was not, a Cooper pair would form.

In Fig. 5.7 the action of the Bogoliubov operators is represented, omitting the amplitude. The filled dot represents a particle creation, the hollow dot represents a hole creation. The operators are grouped by the following classification: $\{\hat{\gamma}^\dagger_{\mathbf{k}\uparrow}, \hat{\gamma}_{-\mathbf{k}\downarrow}\}$ increase the total momentum by $\mathbf{k}$ and the total spin by $\hbar/2$, then create a particle-like excitation. On the contrary $\{\hat{\gamma}_{\mathbf{k}\uparrow}, \hat{\gamma}^\dagger_{-\mathbf{k}\downarrow}\}$ decrease the total momentum by $\mathbf{k}$ and the total spin by $\hbar/2$, then create a hole-like excitation.

Now it is time to check if the Bogoliubov approach is self-consistent. In order for the mean-field approach to be coherent, we want the pairing operator to fluctuate negligibly

### 5.4.2 Elementary excitations

We start by taking a normal metal at zero temperature, described by a filled Fermi sphere. In this context we have a band,

$$\epsilon_{\mathbf{k}} = \frac{\hbar^2|\mathbf{k}|^2}{2m} = \frac{\hbar^2 k_F^2}{2m} + v_F\hbar\delta k + \mathcal{O}\left(\delta k^2\right)$$

where $|\mathbf{k}| = k_F + \delta k$. The system ground state is $|\mathrm{F}\rangle$, defined at the beginning of this chapter. Consider now adding an electron outside the Fermi sphere, at some state with $\delta k > 0$. We work in the thermodynamic limit, so the Fermi radius for the system of $N$ and $N+1$ electrons is practically the same. The state we obtained is an excited state, and we will find one regardless of where we put the additional electron: the band $\epsilon_{\mathbf{k}}$ covers all energies up to infinity, with no discontinuities.

The same thing can be said if, starting from $|\mathrm{F}\rangle$, we annihilate one electron inside the sphere. The resulting state is not the ground state of the system of $N-1$ electrons. The same operation is interpreted as adding one hole to the Fermi sphere. Taking the zero of energy at $\epsilon_F$, the holes can only be added below the surface (where there are electrons to remove) and have energy $-\xi_{\mathbf{k}} = |\xi_{\mathbf{k}}|$.

The elementary excitations of the free Fermi gas in the low spectrum are these two. To move an electron from inside the Fermi sphere to outside,

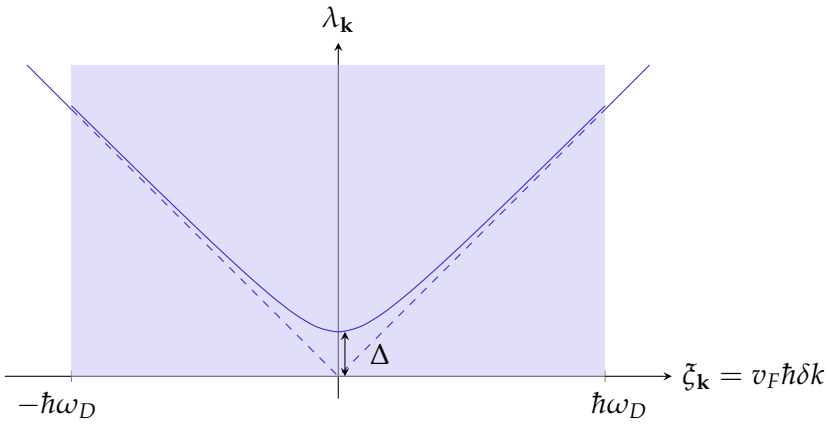



**Figure 5.8:** Sketch of the elementary excitation spectrum for the superconductor (solid line) compared to the spectrum of a normal metal (dashed line). This plot, as in Fig. 5.3, was realized with th arbitrary choice $\Delta = \hbar\omega_D/6$.

even if it seems more "elementary" of modifying the number of particles, is in grand-canonical formalism equivalent to adding one electron outside the sphere and one hole inside. The energy costs for doing both the operations add up to the energy difference of the starting and the target state for the moving electron. Notice that, even if the formalism allows for fluctuations of the particle number, a very different thing are fluctuations of total charge: in absence of external interactions, the electric charge is conserved. So the excitation of a filled Fermi sphere occurs in real world, as one correctly expects, by moving one electron from inside to outside. The real-world basic excitation can be thought formally as the insurgence of a particle-hole couple. So, by "elementary excitation" we intend some configuration that has the lowest amount of additional energy above the ground state, that can happen under certain conditions, but that is not necessarily possible under the symmetries of the problem. In other words: we can describe coherently an isolated Fermi sea as something where the number of particles fluctuates, as well as the number of holes, but their difference (the charge) is preserved.

Consider Fig. 5.8: in the immediate nearby of the Fermi surface,

$$\xi_{\mathbf{k}} = v_F \hbar \delta k$$

so to plot functions of $\xi_{\mathbf{k}}$ is equivalent to plot functions of $\delta k$, at least for $|\xi_{\mathbf{k}}| < \hbar\omega_D$. The dashed line $f(\xi_{\mathbf{k}}) = |\xi_{\mathbf{k}}|$ is the excitation spectrum of the normal metal: the region for $\delta k < 0$ is inside the sphere in momentum space, so its elementary excitation is a hole with energy $-\xi_{\mathbf{k}}$; the region for $\delta k > 0$ is outside the sphere, so its excitation is an electron with energy $+\xi_{\mathbf{k}}$. The *very* important thing to notice is that, for any amount of (small) energy we pump into the system, for the normal metal many excited configurations capable of absorbing such energy exist. Apart from the electrons jumping outside the sphere, which we said are to be interpreted as the excitation of two modes (hole inside, electron outside), the system can absorb the energy by creating one electron outside the sphere or one hole inside (or a superposition of the two). In this context a state with both is not considered an elementary excitation.

Now, in Fig. 5.8 the solid line represents the excitation spectrum obtained via the BCS theory around the Fermi surface. Apart from the discontinuities at the region boundaries, a mere consequence of the approximations we did on the interaction potential, deep in the shell ($|\xi_{\mathbf{k}}| \ll \hbar\omega_D$, near the surface) the excitation of spectrum is *gapped*. This means that the system does **not** excite if the energy we pump in is smaller than $\Delta$. And we can anticipate why, even if we need some more calculations to demonstrate this: if the charge carriers are Cooper pairs, to excite the system means to populate an electron-hole state of the excitation spectrum (to add one $\gamma$ fermion). To do so we need to break a Cooper pair.

But the binding energy of a Cooper pair is $2\Delta$, so, why does the pair even breaks if the available energy is $\Delta$? The reason is our mean-field approach. $\Delta$ is the binding energy per electron, which means, each electron is bound to something by an energy $\Delta$. An external action capable of changing the charge of the system will excite it by converting one pair to a free electron, now not bound to anything. This is an *elementary excitation*, it is born by the vanishing of one Cooper pair, and has the correct energy gap $\Delta$. If such external action is not at work, as it is for an isolated superconductor, charge must be conserved, and the physical excitation we observe by breaking a Cooper pair is made of two particle-hole excitations floating somewhere. That is a broken pair as we intend it naively, and to reach such state we need at least an energy $2\Delta$, and that makes sense with physical intuition. The final state will be populated by two elementary excitations: with this perspective Fig. 5.8 must be interpreted. The fact that the first excited state of the system is not a single elementary excitation is not at all incoherent, on the contrary it is the manifestation of symmetry.

### 5.4.3 Why do we even need a gap?

To be fair, during the very long explanation of the basics of BCS theory throughout Chap. 4 and this chapter the only reference to superconductivity and all the concepts explained in the first part, more phenomenological, was the observation that the charge-carrier in superconductors seems to be a couple of electrons. We then developed a theory describing electrons coupling and the microscopic structure of the collective state, which lead us to the existence of the discussed gap $\Delta$. A very well posed question now would be: if a gap in the excitation spectrum is the typical feature of insulators, those systems that do not conduct electrical current, why the existence of a gap would ever be a symptom of superconductivity?

go on...

# 6 FROM BCS TO CONVENTIONAL SUPERCONDUCTORS



The last chapters were devoted to a quite formal derivation of BCS theory. We used mainly basic Quantum Mechanics, and showed that – starting from the idea of a phononic attractive interaction – the favored energy configuration for a system of electrons under some conditions was the BCS state, $|\Psi\rangle$. To be fair, never we actually specified those conditions, leaving everything quite general: weak binding, negligible phononic population under some temperature, and so on. Unfortunately, this is not the chapter where we specify such conditions, neither there is one in these notes. But here we try to connect the BCS theory we developed at zero temperature with the essential features of superconductivity: the existence of a critical temperature (Sec. 6.1), and perfect diamagnetic screening (Sec. **??**).

## 6.1 BCS THEORY AT FINITE TEMPERATURE

When analyzing a fermionic system at finite temperature $\beta \equiv 1/k_B T$, if the system can be described via the occupation of single-particle states, the essential feature is the distribution of the occupation probabilities. If the fermions are free the probability distribution for a given energy $\xi \equiv \epsilon - \mu$ is given by the Fermi-Dirac function

$$f(\epsilon;\beta) \equiv \frac{1}{e^{\beta\xi}+1}$$

where $\mu$ is the chemical potential. In last chapters we used the notation $\epsilon_F$ assuming to be in a metal, but it is more correct in general to stick to the notation $\mu$. The distribution reduces to the inverted Heaviside distribution at zero temperature,

$$\lim_{\beta\to\infty} f(\xi;\beta) = \theta(-\xi)$$

However we saw that the occupation probability distribution for the electrons in BCS theory at zero temperature is not this one. This of course is a

consequence of the fact that the electrons in BCS theory are interacting via the potential $V$. But in Sec. 5.3, via the Bogoliubov-Valatin transformation, we were able to map the system to one of free fermions, the discussed $\gamma$ particles. For those, what we said here holds. We now develop this idea.

### 6.1.1 Another look to Bogoliubov-Vitalin solution

Recall Eq. (5.14):

$$\hat{\mathcal{K}} = \sum_{\mathbf{k}} \left[ \lambda_{\mathbf{k}} \hat{\gamma}^{\dagger}_{\mathbf{k}\uparrow} \hat{\gamma}_{\mathbf{k}\uparrow} - \lambda_{\mathbf{k}} \hat{\gamma}_{-\mathbf{k}\downarrow} \hat{\gamma}^{\dagger}_{-\mathbf{k}\downarrow} \right]$$

In Sec. 5.3.4 we used anti-commutation rules, collected constant energy shift to recognize condensation energy, and analyzed the common $\lambda_{\mathbf{k}}$ where both classes of $\gamma$ quasiparticles live. Those classes are essentially distinguished by their spin contribution. It is however equivalent to define the following operators

$$\begin{cases} \hat{a}_{\mathbf{k}} \equiv \hat{\gamma}_{\mathbf{k}\uparrow} \\ \hat{a}^{\dagger}_{\mathbf{k}} \equiv \hat{\gamma}^{\dagger}_{\mathbf{k}\uparrow} \end{cases} \qquad \begin{cases} \hat{b}_{\mathbf{k}} \equiv \hat{\gamma}^{\dagger}_{-\mathbf{k}\downarrow} \\ \hat{b}^{\dagger}_{\mathbf{k}} \equiv \hat{\gamma}_{-\mathbf{k}\downarrow} \end{cases}$$

which obviously still satisfy fermion rules, thus represent fermions, and the bands

$$A_{\mathbf{k}} = \sqrt{\xi^2_{\mathbf{k}} + \Delta^2_{\mathbf{k}}} \qquad B_{\mathbf{k}} = -\sqrt{\xi^2_{\mathbf{k}} + \Delta^2_{\mathbf{k}}}$$

for which

$$\hat{\mathcal{K}} = \sum_{\mathbf{k}} \left[ A_{\mathbf{k}} a^{\dagger}_{\mathbf{k}} a_{\mathbf{k}} + B_{\mathbf{k}} b^{\dagger}_{\mathbf{k}} b_{\mathbf{k}} \right]$$

which represents a system of spinless, non-interacting fermions. The $a$ class fermions have positive dispersion, the $b$ class negative dispersion. This is nothing new: also an hamiltonian of free electrons can be mapped to a mixed electron-holes hamiltonian with similar bands. The fact that the $b$ class has negative dispersion is nothing esoteric: we already saw that using $\gamma$ fermions, everything related to energy adds up nice. Here goes a little comment about the electron-hole characters of these fermions.

The important thing is that these fermions are *free*. Thus normal Fermi-Dirac distribution can be used for both. And the meaning of the Fermi-Dirac distribution is precisely to give the mean occupation number for a given state at some energy at thermal equilibrium and nonzero temperature, which means

$$\langle a^{\dagger}_{\mathbf{k}} a_{\mathbf{k}} \rangle = f(A_{\mathbf{k}}; \beta) \qquad \langle b^{\dagger}_{\mathbf{k}} b_{\mathbf{k}} \rangle = f(B_{\mathbf{k}}; \beta)$$

Of course the $b$ quasiparticles populate a Fermi-Dirac with negative dispersion, $f(B_{\mathbf{k}}; \beta) = 1 - f(-B_{\mathbf{k}}; \beta)$, with $f(-B_{\mathbf{k}}; \beta)$ the commonly shaped Fermi-Dirac distribution for $-B_{\mathbf{k}} > 0$. We use immediately this result.

### 6.1.2 Gap equation, again

In Sec. 5.3.5 we derived the gap equation at zero temperature. At finite temperature, the same derivation holds as well, so we can use immediately Eq. (5.11) and Eq. (5.15) combined

$$\Delta_{\mathbf{k}} \equiv -\frac{1}{2}\sum_{\mathbf{k}'} V_{\mathbf{k}-\mathbf{k}'}\left[\langle\hat{\sigma}^1_{\mathbf{k}'}\rangle - i\langle\hat{\sigma}^2_{\mathbf{k}'}\rangle\right]$$

We also derived

$$\langle\hat{\sigma}^1_{\mathbf{k}}\rangle = -\sin 2\theta_{\mathbf{k}}\left[\langle\hat{\gamma}^\dagger_{\mathbf{k}\uparrow}\hat{\gamma}_{\mathbf{k}\uparrow}\rangle - \langle\hat{\gamma}_{-\mathbf{k}\downarrow}\hat{\gamma}^\dagger_{-\mathbf{k}\downarrow}\rangle\right] \qquad \langle\hat{\sigma}^2_{\mathbf{k}}\rangle = 0$$

Here we recognize fermionic occupations at finite temperature,

$$\langle\hat{\sigma}^1_{\mathbf{k}}\rangle = -\sin 2\theta_{\mathbf{k}}\left[\langle a^\dagger_{\mathbf{k}} a_{\mathbf{k}}\rangle - \langle b^\dagger_{\mathbf{k}} b_{\mathbf{k}}\rangle\right]$$

thus, substituting

$$\begin{aligned}
\Delta_{\mathbf{k}} &= \frac{1}{2}\sum_{\mathbf{k}'} V_{\mathbf{k}-\mathbf{k}'}\sin 2\theta_{\mathbf{k}'}\left[f(A_{\mathbf{k}'};\beta) - f(B_{\mathbf{k}'};\beta)\right] \\
&= \frac{1}{2}\sum_{\mathbf{k}'} V_{\mathbf{k}-\mathbf{k}'}\sin 2\theta_{\mathbf{k}'}\left[\frac{1}{e^{\beta A_{\mathbf{k}'}}+1} - \frac{1}{e^{\beta B_{\mathbf{k}'}}+1}\right] \\
&= \frac{1}{2}\sum_{\mathbf{k}'} V_{\mathbf{k}-\mathbf{k}'}\sin 2\theta_{\mathbf{k}'}\left[\frac{1}{e^{\beta\lambda_{\mathbf{k}'}}+1} - \frac{1}{e^{-\beta\lambda_{\mathbf{k}'}}+1}\right] \\
&= -\frac{1}{2}\sum_{\mathbf{k}'} V_{\mathbf{k}-\mathbf{k}'}\sin 2\theta_{\mathbf{k}'}\tanh\left(\frac{\beta\lambda_{\mathbf{k}'}}{2}\right)
\end{aligned}$$

thus we get the **gap equation at finite temperature**

$$\Delta_{\mathbf{k}}(\beta) = -\frac{1}{2}\sum_{\mathbf{k}'} V_{\mathbf{k}-\mathbf{k}'}\sin 2\theta_{\mathbf{k}'}\tanh\left(\frac{\beta}{2}\sqrt{\xi^2_{\mathbf{k}'} + \Delta^2_{\mathbf{k}'}(\beta)}\right) \tag{6.1}$$

For $\beta \to \infty$,

$$\lim_{\beta\to\infty}\tanh\left(\frac{\beta\lambda_{\mathbf{k}}}{2}\right) = 1$$

and the equation reduces to the correct one, at zero temperature. To solve the equation is a very difficult task, especially for the author. But we can use it to do something really important.

### 6.1.3 The critical temperature

Superconductors are fermionic system with broken symmetry under a critical temperature, as we know from Chap. 1 and Chap. 2. From Chap. 5 we know BCS systems are characterized by a gap in excitation spectrum which, together with the quasi-bosonic nature of charge carriers (as we know from Chap. 4) allows for resistance-less transport. To finally show that BCS theory implies superconductivity, we need to show that:

1. a special temperature, namely the critical temperature, above which the gap closes and the excitation spectrum becomes continuous;

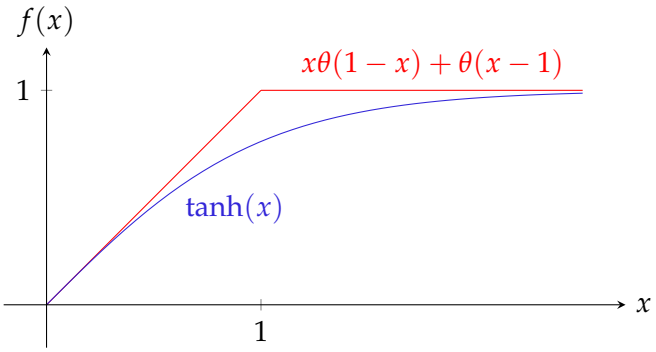



**Figure 6.1:** The approximation for $\tanh(x)$ described in text.

2. the system exhibits perfect diamagnetic response,

but this last is the matter of Sec. 6.2. First, take Eq. (6.1), with the explicit form of $\sin 2\theta_{\mathbf{k}'}$ included

$$\Delta_{\mathbf{k}}(\beta) = -\frac{1}{2}\sum_{\mathbf{k}'} V_{\mathbf{k}-\mathbf{k}'} \frac{\Delta_{\mathbf{k}'}(\beta)}{\sqrt{\xi^2_{\mathbf{k}'} + \Delta^2_{\mathbf{k}'}(\beta)}} \tanh\left(\frac{\beta}{2}\sqrt{\xi^2_{\mathbf{k}'} + \Delta^2_{\mathbf{k}'}(\beta)}\right)$$

As usual, insert our phononic potential

$$V_{\mathbf{k}-\mathbf{k}'} = -V_0 A\left(\mathbf{k}\right) A\left(\mathbf{k}'\right)$$

Proceeding precisely as in Sec. 5.2.2, we get the equation

$$\frac{4}{\rho_0 V_0} = \int_{-\hbar\omega_D}^{\hbar\omega_D} \frac{d\xi}{\sqrt{\xi^2 + \Delta^2(\beta)}} \tanh\left(\frac{\beta}{2}\sqrt{\xi^2 + \Delta^2(\beta)}\right)$$

with $\rho_0$, as always, the spin-wise single-particle DoS at the Fermi energy. We can halve the factor 4 on the left side and limit the integration to $[0, \hbar\omega_D]$, being even in energy the integrand. For any energy and temperature, the hyperbolic tangent is a modulation factor smaller than 1. It follows that

$$\frac{2}{\rho_0 V_0} = \int_0^{\hbar\omega_D} \frac{d\xi}{\sqrt{\xi^2 + \Delta^2(\beta)}} \tanh\left(\frac{\beta}{2}\sqrt{\xi^2 + \Delta^2(\beta)}\right) < \int_0^{\hbar\omega_D} \frac{d\xi}{\sqrt{\xi^2 + \Delta^2(\beta)}}$$

We now use the following notation

$$\Delta_0 \equiv \lim_{\beta\to\infty} \Delta(\beta)$$

for the gap at zero temperature. It must hold

$$\frac{2}{\rho_0 V_0} = \int_0^{\hbar\omega_D} \frac{d\xi}{\sqrt{\xi^2 + \Delta_0^2}}$$

which implies

$$\int_0^{\hbar\omega_D} \frac{d\xi}{\sqrt{\xi^2 + \Delta_0^2}} < \int_0^{\hbar\omega_D} \frac{d\xi}{\sqrt{\xi^2 + \Delta^2(\beta)}} \quad \Longrightarrow \quad \Delta_0 > \Delta(\beta)$$

and this makes sense: increasing temperature, the gap reduces. Suppose it closes at some critical temperature $\beta_c$,

$$\lim_{\beta\to\beta_c^+}\Delta(\beta)=0$$

then

$$\frac{2}{\rho_0 V_0}=\int_0^{\hbar\omega_D}\frac{d\xi}{\xi}\tanh\left(\frac{\beta_c\xi}{2}\right)=\int_0^{x_0}dx\frac{\tanh(x)}{x}$$

with $x_0\equiv\beta_c\hbar\omega_D/2$. We may estimate this integral as follows: for $x>0$ the hyperbolic tangent is roughly approximated by

$$\tanh(x)\simeq x\theta(1-x)+\theta(x-1)$$

as sketched in Fig. 6.1. Then, solving the integrals,

$$\begin{aligned}\int_0^{x_0}dx\frac{\tanh(x)}{x}&\simeq\int_0^1 dx+\int_1^{x_0}\frac{dx}{x}\\&=1+\log x_0\\&=\log(ex_0)=\log\left(\kappa\beta_c\hbar\omega_D\right)\quad\text{with}\quad\kappa\equiv\frac{e}{2}\end{aligned}$$

Better approximation lead to $\kappa\approx 1.13$. We can invert the relation to find an expression for $\beta_c$,

$$\frac{2}{\rho_0 V_0}\simeq\log\left(\kappa\beta_c\hbar\omega_D\right)\quad\Longrightarrow\quad k_B T_c\simeq\kappa\hbar\omega_D e^{-2/\rho_0 V_0}$$

Comparison with the gap value at zero temperature, $2\Delta_0=4\hbar\omega_D e^{-2\rho_0/V_0}$, lead to the ratio

$$\frac{2\Delta_0}{k_B T_c}\approx\frac{4}{1.13}\approx 3.54$$

which appears to be respected experimentally by most conventional superconductors. This prediction is of great beauty and scientific importance, because allows for a rather simple measure of a quantity directly connected with the overall validity of the theory. But the important part of this chapter is yet to come: next section is devoted to deriving the arising of perfect diamagnetism in the BCS system.

## 6.2 IT QUACKS LIKE A SUPERCONDUCTOR, BUT IS IT?

We saw from the very beginning in Sec. 1.1.2 that perfect diamagnetism embeds both resistanceless conduction and Meissner effect. Such phenomenon was synthetically summarized in Eq. 1.6,

$$\mathbf{A}=-\Lambda\mathbf{J}$$

which showed that superconducting current responds to a pure gauge field. The apparent contradiction with gauge invariance was resolved via gauge fixing, Eq. (1.5). We aim to show that BCS theory implies this macroscopic relation.

The starting point is the BCS hamiltonian,

$$\hat{H} = \sum_{\mathbf{k}} \epsilon_{\mathbf{k}} \left[ \hat{c}^{\dagger}_{\mathbf{k}\uparrow} \hat{c}_{\mathbf{k}\uparrow} + \hat{c}^{\dagger}_{\mathbf{k}\downarrow} \hat{c}_{\mathbf{k}\downarrow} \right] + \sum_{\mathbf{k}\mathbf{k}'} V_{\mathbf{k}-\mathbf{k}'} \left[ \hat{c}^{\dagger}_{\mathbf{k}\uparrow} \hat{c}^{\dagger}_{-\mathbf{k}\downarrow} \right] \left[ \hat{c}_{-\mathbf{k}'\downarrow} \hat{c}_{\mathbf{k}'\uparrow} \right]$$

with

$$\epsilon_{\mathbf{k}} = \frac{|\hbar \mathbf{k}|^2}{2m}$$

and $m$ the band effective mass. To include magnetic effects we must pass to a field description, defining the fermionic field operators

$$\hat{\psi}_{\sigma}(\mathbf{x}) \equiv \frac{1}{L^{3/2}} \sum_{\mathbf{k}} \hat{c}_{\mathbf{k}\sigma} e^{i\mathbf{k}\cdot\mathbf{x}} \qquad \hat{\psi}^{\dagger}_{\sigma}(\mathbf{x}) \equiv \frac{1}{L^{3/2}} \sum_{\mathbf{k}} \hat{c}^{\dagger}_{\mathbf{k}\sigma} e^{-i\mathbf{k}\cdot\mathbf{x}}$$

destroying and creating one electron at position $\mathbf{x}$. The kinetic part of the hamiltonian is a one-body operator, the second part a two-body operator, thus

$$\begin{aligned} \hat{H} = & \sum_{\sigma} \int d^3\mathbf{x}\, \hat{\psi}^{\dagger}_{\sigma}(\mathbf{x}) \frac{[-i\hbar\boldsymbol{\nabla}]^2}{2m} \hat{\psi}_{\sigma}(\mathbf{x}) \\ & + \sum_{\sigma\sigma'} \int d^3\mathbf{x}\, d^3\mathbf{x}'\, \hat{\psi}^{\dagger}_{\sigma}(\mathbf{x}) \hat{\psi}^{\dagger}_{\sigma'}(\mathbf{x}') V_{\sigma\sigma'}\left(\mathbf{x}-\mathbf{x}'\right) \hat{\psi}_{\sigma'}(\mathbf{x}') \hat{\psi}_{\sigma}(\mathbf{x}) \end{aligned} \tag{6.2}$$

The specific form of the quantity $V_{\sigma\sigma'}\left(\mathbf{x}-\mathbf{x}'\right)$ can be calculated by the means of Fourier transforming the constant shell potential of BCS Theory. We now wish to include turn on the gauge field $\mathbf{A}(\mathbf{x})$, and include its effects in the above hamiltonian.

### 6.2.1 Peierls approach

The standard approach, already used in Chap. 2, is by Peierls substitution: minimal coupling of the gauge field $\mathbf{A}(\mathbf{x})$ with physical momentum is given by

$$-i\hbar\boldsymbol{\nabla} \rightarrow -i\hbar\boldsymbol{\nabla} - q\mathbf{A}(\mathbf{x}) = -i\hbar\boldsymbol{\nabla} + e\mathbf{A}(\mathbf{x})$$

Evidently this correction leaves unchanged the potential part of the hamiltonian, and does not act on the spin degree of freedom. Indicating by $\hat{H}[\mathbf{A}(\mathbf{x})]$ the parametric hamiltonian, we can substitute in Eq. (6.2)

$$\begin{aligned} \hat{H}[\mathbf{A}(\mathbf{x})] = \hat{H} & - \frac{i\hbar e}{m} \sum_{\sigma} \int d^3\mathbf{x}\, \hat{\psi}^{\dagger}_{\sigma}(\mathbf{x}) \left[\mathbf{A}(\mathbf{x}) \cdot \boldsymbol{\nabla}\right] \hat{\psi}_{\sigma}(\mathbf{x}) \\ & + \frac{e^2}{m} \sum_{\sigma} \int d^3\mathbf{x}\, \hat{\psi}^{\dagger}_{\sigma}(\mathbf{x}) |\mathbf{A}(\mathbf{x})|^2 \hat{\psi}_{\sigma}(\mathbf{x}) \end{aligned} \tag{6.3}$$

We have already seen a very similar procedure in Sec. 2.4.1. Further elaboration allows to define a current $\hat{\mathbf{J}}(\mathbf{x})$, providing an expression for the current operator equivalent to the second Ginzburg-Landau Equation, (2.11): such current is defined as the quantity minimally coupled to the field in the above equation. In order to derive it, we transform the kinetic part

$$\int d^3\mathbf{x}\, \hat{\psi}^{\dagger}_{\sigma}(\mathbf{x}) \left[\mathbf{A}(\mathbf{x}) \cdot \boldsymbol{\nabla} \hat{\psi}_{\sigma}(\mathbf{x})\right] = \mathcal{B} - \int d^3\mathbf{x} \left[\mathbf{A}(\mathbf{x}) \cdot \boldsymbol{\nabla} \hat{\psi}^{\dagger}_{\sigma}(\mathbf{x})\right] \hat{\psi}_{\sigma}(\mathbf{x})$$

with $\mathcal{B}$ a boundary term, that we set to zero having null fields at infinite distance. Then the kinetic term of Eq. (6.3) can be written as

$$-\frac{i\hbar e}{2m}\sum_\sigma \int d^3\mathbf{x}\,\hat{\psi}^\dagger_\sigma(\mathbf{x})\left[\mathbf{A}(\mathbf{x})\cdot\boldsymbol{\nabla}\hat{\psi}_\sigma(\mathbf{x})\right] + \frac{i\hbar e}{2m}\sum_\sigma \int d^3\mathbf{x}\left[\mathbf{A}(\mathbf{x})\cdot\boldsymbol{\nabla}\hat{\psi}^\dagger_\sigma(\mathbf{x})\right]\hat{\psi}_\sigma(\mathbf{x})$$

and we recognize the second part to be the hermitian conjugate of the first one. If we define

$$\hat{\mathbf{j}}^{(\mathrm{p})}(\mathbf{x}) \equiv -\frac{e}{2m}\sum_\sigma\left\{\hat{\psi}^\dagger_\sigma(\mathbf{x})\left[-i\hbar\boldsymbol{\nabla}\hat{\psi}_\sigma(\mathbf{x})\right]-\left[-i\hbar\boldsymbol{\nabla}\hat{\psi}^\dagger_\sigma(\mathbf{x})\right]\hat{\psi}_\sigma(\mathbf{x})\right\} \tag{6.4}$$

as the **paramagnetic current operator**, and

$$\hat{\mathbf{j}}^{(\mathrm{d})}(\mathbf{x}) \equiv -\frac{e^2}{m}\sum_\sigma \hat{\psi}^\dagger_\sigma(\mathbf{x})\mathbf{A}(\mathbf{x})\,\hat{\psi}_\sigma(\mathbf{x}) \tag{6.5}$$

as the **diamagnetic current operator**, the hamiltonian becomes minimally coupled

$$\hat{H}[\mathbf{A}(\mathbf{x})] = \hat{H} - \int d^3\mathbf{x}\,\mathbf{A}(\mathbf{x})\,\hat{\mathbf{j}}(\mathbf{x}) \tag{6.6}$$

where

$$\hat{\mathbf{j}}(\mathbf{x}) \equiv \hat{\mathbf{j}}^{(\mathrm{p})}(\mathbf{x}) + \hat{\mathbf{j}}^{(\mathrm{d})}(\mathbf{x})$$

### 6.2.2 First order expansion of the hamiltonian

We now limit ourselves to the first order paramagnetic contribution. The hamiltonian up to first order is given by

$$\begin{aligned}\hat{H}[\mathbf{A}(\mathbf{x})] &\simeq \hat{H} - \frac{i\hbar e}{m}\sum_\sigma\int d^3\mathbf{x}\,\hat{\psi}^\dagger_\sigma(\mathbf{x})\left[\mathbf{A}(\mathbf{x})\cdot\boldsymbol{\nabla}\right]\hat{\psi}_\sigma(\mathbf{x})\\ &= \hat{H} - \frac{i\hbar e}{m}\sum_\sigma\int d^3\mathbf{x}\,\frac{1}{L^{3/2}}\sum_{\mathbf{k}}\hat{c}^\dagger_{\mathbf{k}\sigma}e^{-i\mathbf{k}\cdot\mathbf{x}}\left[\mathbf{A}(\mathbf{x})\cdot\boldsymbol{\nabla}\right]\frac{1}{L^{3/2}}\sum_{\mathbf{k}'}\hat{c}_{\mathbf{k}'\sigma}e^{i\mathbf{k}'\cdot\mathbf{x}}\\ &= \hat{H} + \frac{\hbar e}{m}\sum_\sigma\sum_{\mathbf{k}\mathbf{k}'}\left[\mathbf{k}'\cdot\frac{1}{L^3}\int d^3\mathbf{x}\,\mathbf{A}(\mathbf{x})\,e^{-i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{x}}\right]\hat{c}^\dagger_{\mathbf{k}\sigma}\hat{c}_{\mathbf{k}'\sigma}\end{aligned}$$

We recognize the field Fourier transform $\mathbf{A}_{\mathbf{k}-\mathbf{k}'}$,

$$\hat{H}[\mathbf{A}(\mathbf{x})] \simeq \hat{H} + \frac{\hbar e}{m}\sum_{\mathbf{k}\mathbf{k}'}\left[\mathbf{k}'\cdot\mathbf{A}_{\mathbf{k}-\mathbf{k}'}\right]\left[\hat{c}^\dagger_{\mathbf{k}\uparrow}\hat{c}_{\mathbf{k}'\uparrow}+\hat{c}^\dagger_{\mathbf{k}\downarrow}\hat{c}_{\mathbf{k}'\downarrow}\right]$$

This expression is easily comprehensible: the contribution to the hamiltonian for scattering an electron from state $|\mathbf{k}'\sigma\rangle$ to $|\mathbf{k}\sigma\rangle$ is given by the necessary momentum component of the gauge field, $\mathbf{k}-\mathbf{k}'$. The same shift occurs in $\hat{\mathcal{K}}$, so we may use this last. Using the general relation $\mathbf{A}^*_{\mathbf{q}} = \mathbf{A}_{-\mathbf{q}}$ ensured by the reality condition of the gauge potential,

$$\begin{aligned}\hat{\mathcal{K}}[\mathbf{A}(\mathbf{x})] \simeq \hat{\mathcal{K}} &+ \frac{\hbar e}{m}\sum_{\mathbf{k}\mathbf{k}'}\left[\mathbf{k}'\cdot\mathbf{A}_{\mathbf{k}-\mathbf{k}'}\right]\hat{c}^\dagger_{\mathbf{k}\uparrow}\hat{c}_{\mathbf{k}'\uparrow}\\ &- \frac{\hbar e}{m}\sum_{\mathbf{k}\mathbf{k}'}\left[\mathbf{k}'\cdot\mathbf{A}^*_{\mathbf{k}-\mathbf{k}'}\right]\hat{c}^\dagger_{-\mathbf{k}\downarrow}\hat{c}_{-\mathbf{k}'\downarrow}\end{aligned}$$

The second term can be written also as

$$\begin{aligned}\sum_{\mathbf{k}\mathbf{k}'}\left[\mathbf{k}'\cdot\mathbf{A}^*_{\mathbf{k}-\mathbf{k}'}\right]\hat{c}^\dagger_{-\mathbf{k}\downarrow}\hat{c}_{-\mathbf{k}'\downarrow} &= \sum_{\mathbf{k}\mathbf{k}'}\left[\mathbf{k}'\cdot\mathbf{A}^*_{\mathbf{k}-\mathbf{k}'}\right]\delta_{\mathbf{k}\mathbf{k}'} - \sum_{\mathbf{k}\mathbf{k}'}\left[\mathbf{k}'\cdot\mathbf{A}^*_{\mathbf{k}-\mathbf{k}'}\right]\hat{c}_{-\mathbf{k}'\downarrow}\hat{c}^\dagger_{-\mathbf{k}\downarrow}\\ &= \sum_{\mathbf{k}}\mathbf{k}\cdot\mathbf{A}^*_{\mathbf{0}} - \sum_{\mathbf{k}\mathbf{k}'}\left[\mathbf{k}'\cdot\mathbf{A}^*_{\mathbf{k}-\mathbf{k}'}\right]\hat{c}_{-\mathbf{k}'\downarrow}\hat{c}^\dagger_{-\mathbf{k}\downarrow}\\ &= -\sum_{\mathbf{k}\mathbf{k}'}\left[\mathbf{k}'\cdot\mathbf{A}^*_{\mathbf{k}-\mathbf{k}'}\right]\hat{c}_{-\mathbf{k}'\downarrow}\hat{c}^\dagger_{-\mathbf{k}\downarrow}\end{aligned}$$

being obviously $\sum\mathbf{k}=\mathbf{0}$. Then

$$\begin{aligned}\hat{\mathcal{K}}\left[\mathbf{A}\left(\mathbf{x}\right)\right] \simeq \hat{\mathcal{K}} &+ \frac{\hbar e}{m}\sum_{\mathbf{k}\mathbf{k}'}\left[\mathbf{k}'\cdot\mathbf{A}_{\mathbf{k}-\mathbf{k}'}\right]\hat{c}^\dagger_{\mathbf{k}\uparrow}\hat{c}_{\mathbf{k}'\uparrow}\\ &+ \frac{\hbar e}{m}\sum_{\mathbf{k}\mathbf{k}'}\left[\mathbf{k}'\cdot\mathbf{A}^*_{\mathbf{k}-\mathbf{k}'}\right]\hat{c}_{-\mathbf{k}'\downarrow}\hat{c}^\dagger_{-\mathbf{k}\downarrow}\end{aligned}$$

From Sec. 5.3.4 we know that the BCS state is really described via the Bogoliubov $\gamma$ quasiparticles, or the equivalent $a$ and $b$ quasiparticles. Their expressions in terms of the $c$ operators can be easily found by applying the inverse Bogoliubov-Valatin transformation as described in Sec. 5.3,

$$\hat{\Phi}_{\mathbf{k}} = e^{i\theta_{\mathbf{k}}\sigma^2}\hat{\Gamma}_{\mathbf{k}} = \begin{bmatrix}u_{\mathbf{k}} & v_{\mathbf{k}}\\ -v_{\mathbf{k}} & u_{\mathbf{k}}\end{bmatrix}\begin{bmatrix}\hat{a}_{\mathbf{k}}\\ \hat{b}_{\mathbf{k}}\end{bmatrix} \qquad \text{with} \qquad \hat{\Phi}_{\mathbf{k}} = \begin{bmatrix}\hat{c}_{\mathbf{k}\uparrow}\\ \hat{c}^\dagger_{-\mathbf{k}\downarrow}\end{bmatrix}$$

Substituting in the precedent equation it is easy to get

$$\begin{aligned}\hat{\mathcal{K}}\left[\mathbf{A}\left(\mathbf{x}\right)\right] \simeq \hat{\mathcal{K}} &+ \frac{\hbar e}{m}\sum_{\mathbf{k}\mathbf{k}'}\left[\mathbf{k}'\cdot\mathbf{A}_{\mathbf{k}-\mathbf{k}'}\right]\left[u^*_{\mathbf{k}}u_{\mathbf{k}'}\hat{a}^\dagger_{\mathbf{k}}\hat{a}_{\mathbf{k}'} + v^*_{\mathbf{k}}v_{\mathbf{k}'}\hat{b}^\dagger_{\mathbf{k}}\hat{b}_{\mathbf{k}'}\right]\\ &+ \frac{\hbar e}{m}\sum_{\mathbf{k}\mathbf{k}'}\left[\mathbf{k}'\cdot\mathbf{A}_{\mathbf{k}-\mathbf{k}'}\right]\left[u^*_{\mathbf{k}}v_{\mathbf{k}'}\hat{a}^\dagger_{\mathbf{k}}\hat{b}_{\mathbf{k}'} + v^*_{\mathbf{k}}u_{\mathbf{k}'}\hat{b}^\dagger_{\mathbf{k}}\hat{a}_{\mathbf{k}'}\right]\\ &+ \frac{\hbar e}{m}\sum_{\mathbf{k}\mathbf{k}'}\left[\mathbf{k}'\cdot\mathbf{A}^*_{\mathbf{k}-\mathbf{k}'}\right]\left[u^*_{\mathbf{k}}u_{\mathbf{k}'}\hat{b}^\dagger_{\mathbf{k}}\hat{b}_{\mathbf{k}'} + v^*_{\mathbf{k}}v_{\mathbf{k}'}\hat{a}^\dagger_{\mathbf{k}}\hat{a}_{\mathbf{k}'}\right]\\ &- \frac{\hbar e}{m}\sum_{\mathbf{k}\mathbf{k}'}\left[\mathbf{k}'\cdot\mathbf{A}^*_{\mathbf{k}-\mathbf{k}'}\right]\left[v^*_{\mathbf{k}}u_{\mathbf{k}'}\hat{a}^\dagger_{\mathbf{k}}\hat{b}_{\mathbf{k}'} + u^*_{\mathbf{k}}v_{\mathbf{k}'}\hat{b}^\dagger_{\mathbf{k}}\hat{a}_{\mathbf{k}'}\right]\end{aligned}$$

We now implement real parameters $u, v \in \mathbb{R}$. Also, we make the physical assumption $\mathbf{A}_{\mathbf{q}} = \mathbf{A}_{-\mathbf{q}} = \mathbf{A}^*_{\mathbf{q}}$. Collecting the various terms we get finally

$$\begin{aligned}\hat{\mathcal{K}}\left[\mathbf{A}\left(\mathbf{x}\right)\right] \simeq &\sum_{\mathbf{k}}\left[A_{\mathbf{k}}a^\dagger_{\mathbf{k}}a_{\mathbf{k}} + B_{\mathbf{k}}b^\dagger_{\mathbf{k}}b_{\mathbf{k}}\right]\\ &+ \frac{\hbar e}{m}\sum_{\mathbf{k}\mathbf{k}'}\left[\mathbf{k}'\cdot\mathbf{A}_{\mathbf{k}-\mathbf{k}'}\right]\left[u_{\mathbf{k}}u_{\mathbf{k}'} + v_{\mathbf{k}}v_{\mathbf{k}'}\right]\left[\hat{a}^\dagger_{\mathbf{k}}\hat{a}_{\mathbf{k}'} + \hat{b}^\dagger_{\mathbf{k}}\hat{b}_{\mathbf{k}'}\right]\\ &+ \frac{\hbar e}{m}\sum_{\mathbf{k}\mathbf{k}'}\left[\mathbf{k}'\cdot\mathbf{A}_{\mathbf{k}-\mathbf{k}'}\right]\left[v_{\mathbf{k}}u_{\mathbf{k}'} - u_{\mathbf{k}}v_{\mathbf{k}'}\right]\left[\hat{a}^\dagger_{\mathbf{k}}\hat{b}_{\mathbf{k}'} + \hat{b}^\dagger_{\mathbf{k}}\hat{a}_{\mathbf{k}'}\right]\end{aligned}$$

The second and third term we collect in $\delta\hat{\mathcal{K}}$. The second term is responsible for intraband scattering (from $A$ to $A$ and from $B$ to $B$), the third for interband scattering (from $A$ to $B$ and from $B$ to $A$).

Now, we aim to find perfect diamagnetism in presence of a perfectly uniform field; thus we are interested to the linear response to a uniform static gauge field; such field only has the long-wavelength component,

$$\lim_{\mathbf{q}\to\mathbf{0}}\mathbf{A}_{\mathbf{q}} \equiv \mathbf{A}_0 \qquad\qquad \mathbf{A}\left(\mathbf{x}\right) = \sum_{\mathbf{q}}\mathbf{A}_{\mathbf{q}}e^{i\mathbf{q}\cdot\mathbf{x}} = \mathbf{A}_0$$

thus in $\delta\hat{\mathcal{K}}$ the only relevant terms are those with $\mathbf{k} = \mathbf{k}'$. This gives the greatly simplified hamiltonian

$$\begin{aligned}\hat{\mathcal{K}}\left[\mathbf{A}_0\right] &\simeq \sum_{\mathbf{k}}\left[A_{\mathbf{k}}a^{\dagger}_{\mathbf{k}}a_{\mathbf{k}} + B_{\mathbf{k}}b^{\dagger}_{\mathbf{k}}b_{\mathbf{k}}\right] + \frac{\hbar e}{m}\mathbf{A}_0\cdot\sum_{\mathbf{k}}\mathbf{k}\left[\hat{a}^{\dagger}_{\mathbf{k}}\hat{a}_{\mathbf{k}} + \hat{b}^{\dagger}_{\mathbf{k}}\hat{b}_{\mathbf{k}}\right] \\ &= \sum_{\mathbf{k}}\underbrace{\left[\lambda_{\mathbf{k}} + \frac{\hbar e}{m}\mathbf{A}_0\cdot\mathbf{k}\right]}_{A_{\mathbf{k}}[\mathbf{A}_0]}a^{\dagger}_{\mathbf{k}}a_{\mathbf{k}} + \sum_{\mathbf{k}}\underbrace{\left[-\lambda_{\mathbf{k}} + \frac{\hbar e}{m}\mathbf{A}_0\cdot\mathbf{k}\right]}_{B_{\mathbf{k}}[\mathbf{A}_0]}b^{\dagger}_{\mathbf{k}}b_{\mathbf{k}}\end{aligned} \tag{6.7}$$

where the new perturbed bands were defined. At fixed wavevector $\mathbf{k}$, both bands get shifted by the same amount. Since by using Bogoliubov fermions we are dealing with a free fermionic system, at finite temperature it is over the new shifted bands that the Fermi-Dirac function must be computed.

### 6.2.3 First order contribution: paramagnetic current

First order is responsible for paramagnetic effects. Inserting field expansions into the paramagnetic current definition, Eq. (6.4), it is easy to get

$$\begin{aligned}\hat{\mathbf{j}}^{(\mathrm{p})}(\mathbf{x}) &\equiv -\frac{e}{2m}\frac{1}{L^3}\sum_{\sigma}\sum_{\mathbf{k}\mathbf{k}'}\left[\hbar\mathbf{k} + \hbar\mathbf{k}'\right]\hat{c}^{\dagger}_{\mathbf{k}\sigma}\hat{c}_{\mathbf{k}'\sigma}e^{-i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{x}} \\ &= -\frac{e}{2m}\frac{1}{L^3}\sum_{\sigma}\sum_{\mathbf{k}\mathbf{k}'}\hbar\mathbf{k}\left[\hat{c}^{\dagger}_{\mathbf{k}\sigma}\hat{c}_{\mathbf{k}'\sigma}e^{-i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{x}} + \hat{c}^{\dagger}_{\mathbf{k}'\sigma}\hat{c}_{\mathbf{k}\sigma}e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{x}}\right] \\ &= \frac{1}{L^3}\sum_{\mathbf{q}}\hat{\mathbf{j}}^{(\mathrm{p})}_{\mathbf{q}}e^{-i\mathbf{q}\cdot\mathbf{x}} \qquad \text{with} \qquad \mathbf{q} \equiv \mathbf{k} - \mathbf{k}'\end{aligned}$$

Here we have recognized the paramagnetic current Fourier transform,

$$\hat{\mathbf{j}}^{(\mathrm{p})}_{\mathbf{q}}(\mathbf{x}) \equiv -\frac{e}{2m}\sum_{\sigma}\sum_{\mathbf{k}\sigma}\hbar\mathbf{k}\left[\hat{c}^{\dagger}_{\mathbf{k}\sigma}\hat{c}_{\mathbf{k}-\mathbf{q}\sigma}e^{-i\mathbf{q}\cdot\mathbf{x}} + \hat{c}^{\dagger}_{\mathbf{k}-\mathbf{q}\sigma}\hat{c}_{\mathbf{k}\sigma}e^{i\mathbf{q}\cdot\mathbf{x}}\right]$$

For the response to a uniform field we take the long wavelength limit; this last reduces to

$$\lim_{\mathbf{q}\to\mathbf{0}}\hat{\mathbf{j}}^{(\mathrm{p})}_{\mathbf{q}}(\mathbf{x}) \equiv -\frac{e}{m}\sum_{\mathbf{k}\sigma}\hbar\mathbf{k}\left[\hat{c}^{\dagger}_{\mathbf{k}\sigma}\hat{c}_{\mathbf{k}\sigma}\right] = -\frac{e}{m}\sum_{\mathbf{k}\sigma}\hbar\mathbf{k}\left[\hat{c}^{\dagger}_{\mathbf{k}\uparrow}\hat{c}_{\mathbf{k}\uparrow} + \hat{c}_{-\mathbf{k}\downarrow}\hat{c}^{\dagger}_{-\mathbf{k}\downarrow}\right]$$

To get the second term in the last part of the equation we used anti-commutation rules for the $\downarrow$ part, eliminated the term $\sum\mathbf{k}$ due to symmetry, and changed the sum variable from $\mathbf{k}$ to $-\mathbf{k}$. We have already calculated exactly the same quantity. This leads us to

$$\sum_{\mathbf{k}\sigma}\hbar\mathbf{k}\left[\hat{c}^{\dagger}_{\mathbf{k}\uparrow}\hat{c}_{\mathbf{k}\uparrow} + \hat{c}_{-\mathbf{k}\downarrow}\hat{c}^{\dagger}_{-\mathbf{k}\downarrow}\right] = \sum_{\mathbf{k}\sigma}\hbar\mathbf{k}\left[\hat{a}^{\dagger}_{\mathbf{k}}\hat{a}_{\mathbf{k}} + \hat{b}^{\dagger}_{\mathbf{k}}\hat{b}_{\mathbf{k}}\right]$$

which gives

$$\begin{aligned}\left\langle\lim_{\mathbf{q}\to\mathbf{0}}\hat{\mathbf{j}}^{(\mathrm{p})}_{\mathbf{q}}(\mathbf{x})\right\rangle &= -\frac{e}{m}\frac{1}{L^3}\sum_{\mathbf{k}\sigma}\hbar\mathbf{k}\left[\left\langle\hat{a}^{\dagger}_{\mathbf{k}}\hat{a}_{\mathbf{k}}\right\rangle + \left\langle\hat{b}^{\dagger}_{\mathbf{k}}\hat{b}_{\mathbf{k}}\right\rangle\right] \\ &= -\frac{e}{m}\frac{1}{L^3}\sum_{\mathbf{k}\sigma}\hbar\mathbf{k}\left[f(A_{\mathbf{k}};\beta) + f(B_{\mathbf{k}};\beta)\right]\end{aligned}$$

The paramagnetic current, as expressed in Eq. (6.4), is independent of the field; so, for low uniform fields its expectation value is simply

$$\left\langle \lim_{\mathbf{q}\to\mathbf{0}} \hat{\mathbf{j}}_{\mathbf{q}}^{(\mathrm{p})}(\mathbf{x}) \right\rangle = -\frac{e}{m}\sum_{\mathbf{k}\sigma}\hbar\mathbf{k}\left[f\left(A_{\mathbf{k}}\left[\mathbf{A}_0\right];\beta\right)+f\left(B_{\mathbf{k}}\left[\mathbf{A}_0\right];\beta\right)\right] \tag{6.8}$$

The first order approximation is coherent for small perturbation fields, so we expect the energy shift on the bands to be small; this means,

$$A_{\mathbf{k}}\left[\mathbf{A}_0\right] < 0 \qquad\qquad B_{\mathbf{k}}\left[\mathbf{A}_0\right] > 0$$

We have set the chemical potential $\mu$ as the energy zero; therefore for small perturbing fields the $B$ band shall be filled, the $A$ band empty. To quantify such fields, consider that

$$B_{\mathbf{k}}\left[\mathbf{0}\right] < B_{|\mathbf{k}|=k_F}\left[\mathbf{0}\right] \simeq -\Delta(\beta) \qquad\qquad A_{\mathbf{k}}\left[\mathbf{0}\right] > A_{|\mathbf{k}|=k_F}\left[\mathbf{0}\right] \simeq \Delta(\beta)$$

(both bands have a gap of order $\Delta$) so the perturbation shall be smaller,

$$\left|\frac{\hbar e}{m}\mathbf{A}_0\cdot\mathbf{k}\right| < \frac{\hbar e}{m}|\mathbf{A}_0|k_F < \Delta(\beta)$$

Now: in such condition, as said, the $B$ band is well below the chemical potential an thus filled, the $A$ band empty. It follows

$$\underbrace{f\left(A_{\mathbf{k}}\left[\mathbf{A}_0\right];\beta\right)}_{=0}+\underbrace{f\left(B_{\mathbf{k}}\left[\mathbf{A}_0\right];\beta\right)}_{=1}=1$$

so, from Eq. (6.8),

$$\left\langle \lim_{\mathbf{q}\to\mathbf{0}} \hat{\mathbf{j}}_{\mathbf{q}}^{(\mathrm{p})}(\mathbf{x}) \right\rangle = -\frac{e}{m}\sum_{\mathbf{k}\sigma}\hbar\mathbf{k} = \mathbf{0}$$

being $\sum\mathbf{k} = \mathbf{0}$. So, the presence of a gap is enough to make the paramagnetic contribution to current vanish. The gap, however, manifests itself in a system of Bogoliubov fermions, not electrons – in other words, the paramagnetic current vanishes also for an insulator. We now check the second order diamagnetic contribution to current.

### 6.2.4 Second order: diamagnetic current

The diamagnetic current is given by Eq. (6.5). Inserting the fields expansions

$$\begin{aligned}\hat{\mathbf{j}}^{(\mathrm{d})}(\mathbf{x}) &\equiv -\frac{e^2}{m}\frac{1}{L^3}\sum_{\sigma}\sum_{\mathbf{k}\mathbf{k}'}\mathbf{A}(\mathbf{x})\,\hat{c}^{\dagger}_{\mathbf{k}\sigma}\hat{c}_{\mathbf{k}'\sigma}e^{-i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{x}}\\ &= \frac{1}{L^3}\sum_{\mathbf{q}}\hat{\mathbf{j}}^{(\mathrm{d})}_{\mathbf{q}}e^{-i\mathbf{q}\cdot\mathbf{x}} \qquad \text{with} \qquad \mathbf{q}\equiv\mathbf{k}-\mathbf{k}'\end{aligned}$$

Here we have recognized the diamagnetic current Fourier transform,

$$\hat{\mathbf{j}}^{(\mathrm{d})}_{\mathbf{q}}(\mathbf{x}) \equiv -\frac{e^2}{m}\sum_{\mathbf{k}\sigma}\mathbf{A}(\mathbf{x})\,\hat{c}^{\dagger}_{\mathbf{k}\sigma}\hat{c}_{\mathbf{k}-\mathbf{q}\sigma}$$

We take the long wavelength limit for a uniform gauge field, $\mathbf{A}(\mathbf{x}) = \mathbf{A}_0$, and evaluate on the BCS ground state,

$$\left\langle \lim_{\mathbf{q}\to\mathbf{0}} \hat{\mathbf{j}}_{\mathbf{q}}^{(\mathrm{d})}(\mathbf{x}) \right\rangle = -\frac{e^2}{m}\mathbf{A}_0\sum_{\mathbf{k}\sigma}\left\langle \hat{c}^{\dagger}_{\mathbf{k}}\hat{c}_{\mathbf{k}}\right\rangle$$

which implies, in the same limit,

$$\left\langle \hat{\mathbf{j}}^{(\mathrm{d})}(\mathbf{x}) \right\rangle = -\frac{e^2}{m}\mathbf{A}_0 \sum_{\mathbf{k}\sigma} \frac{\langle \hat{c}^\dagger_{\mathbf{k}} \hat{c}_{\mathbf{k}} \rangle}{L^3} = -\frac{ne^2}{m}\mathbf{A}_0$$

with $n$ the electronic density. Being null the paramagnetic contribution, we have an expression for the physical current $\mathbf{J} \equiv \langle \hat{\mathbf{j}} \rangle$,

$$\mathbf{J} = -\frac{ne^2}{m}\mathbf{A}_0 \quad \Longrightarrow \quad \mathbf{A} = -\Lambda \mathbf{J}$$

with $\Lambda = m/ne^2$. This is precisely Eq. (1.6), the London Equation for the perfect diamagnet. Thus, we conclude that BCS theory implies perfect diamagnetism, and then:


**Figure 6.2:** (a) Wait, is it a superconductor?
(b) Always has been.

# BIBLIOGRAPHY


[1] J. Bardeen, L. N. Cooper, and J. R. Schrieffer. "Theory of Superconductivity". In: *Phys. Rev.* 108 (5 Dec. 1957), pp. 1175–1204. DOI: 10.1103/PhysRev.108.1175. URL: https://link.aps.org/doi/10.1103/PhysRev.108.1175.

[2] G. Grosso and G. Pastori Parravicini. *Solid State Physics*. Elsevier Science, 2000. ISBN: 9780080481029. URL: https://books.google.it/books?id=L5RrQbbvWn8C.

[3] G. Giuliani and G. Vignale. *Quantum Theory of the Electron Liquid*. Cambridge University Press, 2008. ISBN: 9781139471589. URL: https://books.google.it/books?id=FFydOyv_r78C.

[4] D. L. Goodstein. *States of Matter*. Dover Books on Physics. Dover Publications, 2014. ISBN: 9780486795515. URL: https://books.google.it/books?id=OfEIBAAAQBAJ.

[5] M. Tinkham. *Introduction to Superconductivity*. Dover Books on Physics Series. Dover Publications, 2004. ISBN: 9780486134727. URL: https://books.google.it/books?id=VpUk3NfwDIkC.

[6] K. Huang. *Statistical Mechanics, 2nd Ed*. Wiley India Pvt. Limited, 2008. ISBN: 9788126518494. URL: https://books.google.it/books?id=ZHl8HLk-K3AC.

[7] A. Altland and B. D. Simons. *Condensed Matter Field Theory*. Cambridge books online. Cambridge University Press, 2010. ISBN: 9780521769754. URL: https://books.google.it/books?id=GpF0Pgo8CqAC.

[8] F. London, H. London, and Frederick Alexander Lindemann. "The electromagnetic equations of the supraconductor". In: *Proceedings of the Royal Society of London. Series A - Mathematical and Physical Sciences* 149.866 (1935), pp. 71–88. DOI: 10.1098/rspa.1935.0048. URL: https://royalsocietypublishing.org/doi/abs/10.1098/rspa.1935.0048.

[9] A. A. Abrikosov. "On the Magnetic properties of superconductors of the second group". In: *Sov. Phys. JETP* 5 (1957), pp. 1174–1182.

[10] V. G. Kogan and R. Prozorov. "Interaction between moving Abrikosov vortices in type-II superconductors". In: *Phys. Rev. B* 102 (2 July 2020), p. 024506. DOI: 10.1103/PhysRevB.102.024506. URL: https://link.aps.org/doi/10.1103/PhysRevB.102.024506.

[11] J Simmendinger et al. "Bound and stable vortex–antivortex pairs in high-Tc superconductors". In: *New Journal of Physics* 22.12 (Dec. 2020), p. 123035. DOI: 10.1088/1367-2630/abd123. URL: https://dx.doi.org/10.1088/1367-2630/abd123.