Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Natural Sciences
Department of Physics

Shahzeb Talib Syed

Superconductivity in the attractive Haldane-Hubbard model

Master's thesis in Applied Physics and Mathematics

Supervisor: Asle Sudbø

June 2020

NTNU
Norwegian University of
Science and Technology

Shahzeb Talib Syed

Superconductivity in the attractive Haldane-Hubbard model

Master's thesis in Applied Physics and Mathematics
Supervisor: Asle Sudbø
June 2020

Norwegian University of Science and Technology
Faculty of Natural Sciences
Department of Physics

NTNU
Norwegian University of
Science and Technology

Abstract

Motivated by recent interest in interacting topological insulators, we study the superconductive properties of an attractive Haldane-Hubbard model. The Haldane-Hubbard model displays a phase transition between a Chern insulator and a superconductor governed by the second-nearest neighbor hopping parameter t' and the chemical potential μ. In this thesis, we study the superconductive phase in both the weak and strong-coupling regimes using BCS and Eliashberg theory of superconductivity, respectively. The main difference between these regimes is the fact that the electron self-energy S is accounted for in the strong-coupling regime through Eliashberg theory. The inclusion of S in the gap equation means we get a shift in the quasi-particle spectrum which can significantly alter the solutions to the gap equation. We show through Eliashberg theory that as the coupling strength λ is increased, the increase of the energy shift S can become detrimental to superconductivity in the Haldane-Hubbard model. In this regime, the critical temperature T_c and superconductive gap $W(T = 0)$ are lower than what is predicted by BCS theory. Although T_c shows an upward trend as λ is increased, $W(T = 0)$ seems to fall to zero when both S and λ are large.

Sammendrag

Med nylig økt interesse for vekselvirkende topologiske isolatorer som bakteppe ser vi på superledende egenskaper til en attraktiv Haldane-Hubbard-modell. Haldane-Hubbard-modellen har en faseovergang mellom en Chern-isolator og en superleder styrt av nest næremeste nabo hoppeparameteren t' og det kjemiske potensialet μ. I denne avhandlingen ser vi på den superledende fasen i både det svake og sterke vekselvirkningregimet ved bruk av henholdsvis BCS- og Eliashberg-teori. Den største forskjellen mellom disse regimene er at selvenergien S til elektronet er tatt hensyn til i det sterke vekselvirkningregimet gjennom Eliashberg-teori. Å inkludere S i gapligningen vil gi et skift i kvasipartikkel-spekteret som kan lede til at løsningene til gapligningen endrer seg betydelig. Ved bruk av Eliashberg-teori viser vi at når koblingsstyrken λ økes, vil en økning av energiskiftet S vise seg å være ufordelaktig for superledning i Haldane-Hubbard-modellen. Både den kritiske temperaturen T_c og det superledende gapet $W(T = 0)$ er lavere enn antatt av BCS-teori i det sterke vekselvirkningregimet, og selv om T_c viser en oppadgående trend når λ øker, viser det seg at $W(T = 0)$ faller til null når både S og λ er store.

Abbreviations

TI	Topological insulator
CI	Chern insulator
SC	Superconductor
NN	Nearest neighbor
2NN	Second-nearest neighbor
LHS	Left hand side
RHS	Right hand side
DOS	Density of states
1PI	One-particle irreducible (diagrams)

Nomenclature

N	Number of lattice sites (unit cells)
t	Nearest neighbor hopping parameter
t'	Second-nearest neighbor hopping parameter magnitude
σ	Spin quantum number
U	Hubbard interaction strength
μ	Chemical potential (Fermi level)
λ	Effective electron-electron coupling constant
k_B	Boltzmann's constant
T_c	Critical temperature at which superconductivity occurs
Δ	Superconducting gap (in BCS theory)
W	Superconducting gap (in Eliashberg theory)
S	Electron exchange self-energy
$D(\epsilon_F)$	Density of states at the Fermi level
ω_D	Debye frequency

Preface

This Master's thesis is the result of research conducted primarily between January and June 2020, in addition to work done between August and December 2019 as part of a so-called Project thesis.

Leading up to and during my five year stay at the Norwegian University of Science and Technology (NTNU), I have been lucky enough to have my mother by my side for the entire ride. Deserving of all my gratitude, I would like to thank her for her support and affection, and always taking pride in my path.

During my time at NTNU, I have had the privilege to learn the way of the universe alongside people I hope to call lifelong friends. I have also had the privilege to learn from inspiring educators such as my supervisor Prof. Asle Sudbø, whom I am proud to have worked with. I would like to thank Prof. Asle Sudbø for the support and guidance I have received the past year, in addition to the many hours of well-prepared and inspiring lectures in both thermal and statistical physics.

Lastly, I would like to thank PhD candidates Eirik Erlandsen and Even Thingstad for sacrificing their time, without hesitation, to answer my every question.

Shahzeb Syed
Trondheim, Norway
June 2020

Contents

Chapter 1

Introduction

1.1 Background and motivation

The story of superconductivity dates back to 1908 — three years before its discovery. Scientists at the time thought that electrons in conductors came to a standstill near absolute zero temperatures. Thus causing up to infinite metal resistivity. Some scientists, including Dutch physicist H. K. Onnes, believed otherwise. Onnes was especially interested in liquefaction of helium and conductive properties of cooled metals. In 1908, Onnes was able to liquefy helium by cooling it to its boiling point of 4.2 K. Along with this experimental feat came the discovery of superconductivity. Onnes discovered in 1911 that cooling a solid wire made of mercury to near 4 K by submerging it in liquid helium, caused the resistance in the wire to plummet [1]. The drop in resistance was so large that he soon recognized he had discovered a new state of matter. Onnes received recognition for his ground breaking work in 1913 when he was awarded the Nobel prize in physics.

Since its discovery, superconductivity has been and continues to be of immense interest. Continued work has revealed new properties of superconductors. A defining magnetic property was discovered in 1933 by W. Meissner and R. Ochsenfeld now known as the Meissner effect [2]. Meissner and Ochsenfeld showed that superconductors have the ability to fully screen external magnetic fields from the inside of the metal. This phenomenon laid the foundation for the thermodynamic treatment of superconductors.

In the following decades, the number of remarkable properties of superconductors were ever increasing. However, the underlying mechanism of superconductivity remained a mystery. In 1950, Soviet physicists V. L. Ginzburg and L. D. Landau published a phenomenological theory for superconductivity, now known as Ginzburg-Landau theory [3]. Though greatly beneficial to research, it was only a macroscopic theory on its initial form, and failed to explain the details of superconductivity. In 1957, 46 years after the discovery of superconductivity, the first quantum theory able to explain both zero electrical resistance and the Meissner effect was published. For which a Nobel prize was awarded 15 years later in 1972. The theory was developed by the American physicists J. Bardeen, L. N. Cooper and J. R. Schrieffer, and is now known as the BCS theory of superconductivity [4].

The BCS theory accelerated research on superconductivity. Already in 1960, the Norwegian physicist I. Giæver published work on experiments regarding quantum tunneling in superconductors giving direct evidence of the existence of the so-called energy gap Δ in superconductors, which was predicted by the BCS theory [5–7]. Inspired by the experimental work of Giæver, the English physicist B. D. Josephson analyzed the theoretical description in 1962. Josephson's work lead to theoretical predictions of new phenomena in superconductors. One influential effect that bears his name is the so-called Josephson's effect [8]. Giæver and Josephson's work lead to a Nobel prize in physics in 1973.

Soon after the BCS paper appeared, there was concern among physicists about the universality of BCS theory, i.e. its ability to describe all superconductors by the same mechanism. The universality was, at first, the strength of the theory. However, due to this universality, BCS theory did not seem able to distinguish between superconductors. And as deviations from this universality became more apparent, the blame was put on our incomplete understanding of the electron-phonon mechanism. Almost in parallel, L. Gor'kov [9] developed a Green's function method from which both the BCS results and the Ginzburg-Landau phenomenology could be derived. Although there were many other formalisms and methods in use, Gor'kov's formalism proved to be the most useful. Especially for the purposes of generalizing BCS theory to the case where the electron-phonon interaction was properly taken into account. This generalization was done by G. Eliashberg in 1961 [10, 11] and was a successful attempt at fixing the universality of BCS theory by including retardation effects in the electron-phonon interaction. Retardation effects, along with the inclusion of the electron exchange self-energy, is what makes Eliashberg theory a more complete theory.

A little more than a couple of decades later a new class of superconductors were discovered. In 1986, the German and Swiss physicists J. G. Bednorz and K. A. Müller opened the door to high(er) temperature superconductors [12]. They were able to produce superconductivity in oxide materials involving copper (cuprates) at temperatures 12 K higher than the previously known highest temperature superconductor. This new class of superconductors are today called *unconventional* superconductors. Conventional superconductors were originally explained by the BCS theory as an exchange of phonons between electrons. Unconventional superconductors, however, do not necessarily exchange phonons but rather e.g. magnons [13]. Another observation was that this new class of superconductors had stronger effective coupling between the electrons. This discovery was so profound that it took superconductors out of the confinements of physics laboratories and into the world, making it an almost household word. Bednorz and Müller's advances towards high-temperature superconductors lead to a Nobel prize in physics the very next year in 1987 [14].

Room temperature superconductivity is one of the most sought after discoveries in condensed matter physics. As almost everything in a modern circuit loses energy to heat, substituting regular conductors with superconductors will eliminate build up of heat in such systems. This implies better computing power and energy consumption. Even though room temperature superconductivity is somewhere in the future, low-temperature superconductivity is still widely used in modern technology. The impact of superconductivity on modern medicine and physics research can not be downplayed. Two great examples of superconductors used as powerful electromagnets are magnetic resonance imaging (MRI) and the beam-steering and focusing magnets used in particle accelerators, e.g. LHC at CERN.

The phase transition between a metal and a superconductor is an example of a second order phase transition. A paradigm of a second order phase transition was given by the so-called Landau theory [15]. In Landau theory, one can define an order parameter which is finite and non-zero in the ordered state (superconductor) and zero in the disordered state (metal). Moreover, Landau theory laid the foundation for the development of Ginzburg-Landau theory, which in turn played a major role in the development of BCS theory. Thus, according to Ginzburg-Landau theory, it is customary to regard the superconducting gap Δ as the order parameter in BCS theory[1]. For a long time it was believed that Landau theory could describe all such phase transitions. However, in 1980 the German physicist K. v. Klitzing discovered a new state that lead to the realization of a new type of phase transition, namely topological phase transitions. Klitzing found

[1] Δ is not really the order parameter of superconductivity. Phase transitions given by Landau theory are described by spontaneous symmetry breaking. Because superconductivity is a gauge theory, Elitzur's theorem [16] forbids the local gauge symmetry of superconductivity to be spontaneously broken. This is why Δ, a local parameter, cannot be the order parameter of superconductivity. However, Elitzur's theorem does allow for spontaneous breaking of a global symmetry within a theory that has a local gauge symmetry. In superconductivity this manifests itself through the Higgs mechanism which assigns the gauge field (photons) a mass in the superconducting state. In other words, the true order parameter of superconductivity is the mass of the photon. It is zero in metals and non-zero in superconductors. This is also the origin of the Meissner effect.

that confining electrons to a two-dimensional plane and turning on a strong magnetic field resulted in new, outstanding behavior called the quantum Hall effect (QHE) [17], for which he received a Nobel prize in 1985. Classically a setup like this gives rise to the classical Hall effect, discovered by E. Hall in 1879 [18]. The QHE differs from the classical case in that the temperature is low and the magnetic field strong. Klitzing found that the Hall conductance was quantized, i.e. the conductance σ_{xy} grew in integer steps of e^2/h, where e is the elementary charge and h Planck's constant. The thought of things being quantized at the microscopic level is not that surprising today. However, the conductance is not generally thought of as a microscopic quantity. It is a macroscopic quantity emerging from a large and messy system of many electrons. The fact that a quantity like that can be quantized is remarkable. The explanation of the QHE required something new, and it turned out to be topology in quantum many-body systems.

Another property of the QHE is the existence of chiral conduction states along the edges. In this context, chiral means that the electron transport is one-way. The topology of a quantum Hall system is the reason these edge states exist. This was shown by D. J. Thouless et al. in 1982 where they showed that the quantum Hall system could be characterized by a topological invariant integer called the TKNN invariant [19], which is today known as the Chern number.

The QHE requires a strong magnetic field, which is partly why it is a technologically impractical system. From this came the idea of a similar system exhibiting QHE in the absence of an external magnetic field. This model, which would be the birth of topological insulators (TIs), was presented by F. D. M. Haldane in 1988 [20]. Thouless and Haldane's work on topological phases lead to a Nobel prize in physics in 2016 [21].

The work on topological insulators is of high interest today, see e.g. review by M. Z. Hasan and C. L. Kane [22]. And due to the study of topological systems being relatively new, there is still room for exploration. This is especially true for the case of interacting topological insulators. Haldane presented a system without interactions, and in recent years there has been an increased interest in interacting topological insulators [23, 24]. Motivated by this, we aim to learn more about attractively interacting TIs. In the following, we will include attractive Hubbard interactions in the spinful Haldane model. From BCS theory we know that attraction between electrons can lead to superconductivity. Keeping this in mind, we will study the superconductive properties of the attractive Haldane-Hubbard model[2]. And because strongly coupled electrons is regarded as one of the ingredients for high-temperature superconductors, we will study both the weak and strong-coupling regimes of the Haldane-Hubbard model using BCS and Eliashberg theory, respectively.

[2]Strictly focused on the bulk and not the edges.

1.2 Thesis structure

We begin by giving a brief introduction to some important concepts in Chapter 2. We introduce graphene, tight-binding models, the BCS and Eliashberg theory of superconductivity, and topological insulators.

Chapter 3 is the main part of this thesis where we study both BCS and Eliashberg theory. In Section 3.2, we derive the gap equation from BCS theory for the Haldane-Hubbard model. The gap equation lays the basis for studying the superconductive properties of the Haldane-Hubbard model in the following sections.

Section 3.3 is where we derive the Eliashberg equations which determine the gap W, T_c and the electron self-energy S. We proceed to present the solutions to the Eliashberg equations in Section 3.4 and compare them to the BCS case.

We conclude our findings in Chapter 4 followed by a few appendices. Appendix A includes a calculation of the Chern number of the Haldane model. In Appendix B, we derive the diagonalized Hamiltonian followed by a derivation of the density of states at the Fermi level in Appendix C. Lastly, Appendix D includes a detailed derivation of the Eliashberg equations.

Chapter 2

Preliminaries

2.1 Graphene

Graphene is a one-atom layer thick sheet of carbon atoms arranged as a two dimensional honeycomb lattice with a single atom at each vertex of the lattice, shown in Fig. 2.1. The three sets of nearest neighbors of each carbon atom is given by the following vectors

$$\begin{array}{lll} \boldsymbol{e}_1 = (0,1)a & \boldsymbol{\delta}_1 = (\sqrt{3},0)a & \boldsymbol{b}_1 = (\sqrt{3},1)a \\ \boldsymbol{e}_2 = (-\frac{\sqrt{3}}{2},-\frac{1}{2})a & \boldsymbol{\delta}_2 = (-\frac{\sqrt{3}}{2},\frac{3}{2})a & \boldsymbol{b}_2 = (-\sqrt{3},1)a \\ \boldsymbol{e}_3 = (\frac{\sqrt{3}}{2},-\frac{1}{2})a & \boldsymbol{\delta}_3 = (-\frac{\sqrt{3}}{2},-\frac{3}{2})a & \boldsymbol{b}_3 = (0,-2)a \end{array} \tag{2.1}$$

where $\boldsymbol{e}_j$, $\boldsymbol{\delta}_j$ and $\boldsymbol{b}_j$ are the first, second, and third nearest neighbor vectors, respectively. These vectors are shown in Fig. 2.2. The distance between each nearest neighbor carbon atom is about $a = 1.42\,\text{Å}$. However, we will set $a = 1$ for brevity of notation. Graphene has been a starting point for many condense matter models. Some of the most popular models will be presented next.

2.2 Tight-binding model

The goal of the tight-binding model is to simplify the quantum mechanics of crystals. Many models are built on the tight-binding approximation as it is a simple model of how electrons are arranged on periodic lattices.

Consider a periodic lattice comprised of atoms, where each atom's equilibrium position is at the lattice sites, e.g. the vertices of graphene. The tight-binding model approximates that electrons are *tightly* bound to atoms in so-called Wannier orbitals[1]. These orbitals are localized around particular atoms with a small amplitude of tunneling (or hopping) to a neighboring atom. We often restrict ourselves to nearest and second-nearest neighbors (NN and 2NN, respectively). The process of hopping between nearby atoms can be described using creation ($c^\dagger$) and annihilation (c) operators for electrons. A Hamiltonian for the kinetic energy associated with hopping in the tight-binding approximation has the form

$$H = -t \sum_{\substack{\langle i,j \rangle \\ \sigma}} \left(c^\dagger_{i\sigma} c_{j\sigma} + c^\dagger_{j\sigma} c_{i\sigma} \right). \tag{2.2}$$

Above, the sum over $\langle i,j \rangle$ is taken over pairs of NN atoms, namely atom i and j. In a one-dimensional lattice, e.g. evenly spaced beads on a necklace, the NN of atom i would be $j = i + 1$. The sum over σ is a spin sum, where the two possible spin values are $\sigma = \uparrow, \downarrow$. The summand

[1] See chapter 10 of [25] for more on Wannier orbitals.

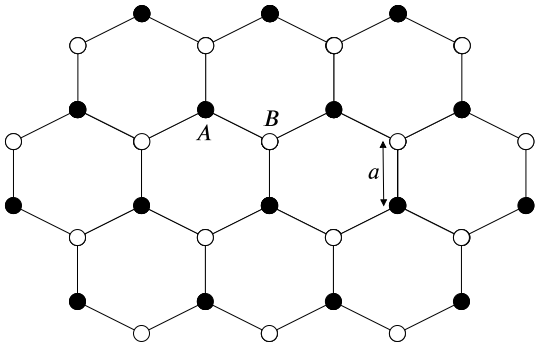


Figure 2.1: Graphene's honeycomb structure. The lattice is comprised of two sub-lattices A and B, with lattice constant a.

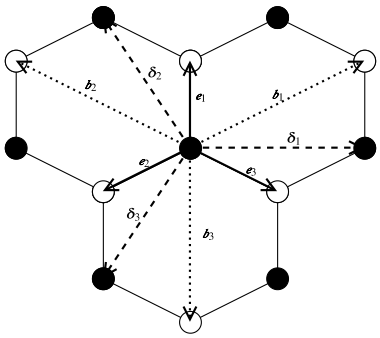


Figure 2.2: Honeycomb lattice with the three nearest neighbor vectors drawn in, namely $\boldsymbol{e}_j$, $\boldsymbol{\delta}_j$, and $\boldsymbol{b}_j$.

describes a two-step process of electron hopping. The annihilation operator $c_{j\sigma}$ removes an electron from the orbital around atom j with spin σ. Meanwhile, the creation operator $c^\dagger_{i\sigma}$ adds an electron to the orbital around atom i with spin σ. In other words, the electron that was removed from one atom, gets added to the neighboring atom, which is an effective description of electron hopping. The second term of the summand describes the reverse process: hopping from atom i to j. t is the small tunneling amplitude that describes the likelihood of hopping. If we add a second sum to H but with $\langle\langle i, j\rangle\rangle$ instead, we would have a Hamiltonian that describes both NN and 2NN hopping.

In order to diagonalize the Hamiltonian, we introduce the Fourier transformed operators

$$c_{i\sigma} = \frac{1}{\sqrt{N}} \sum_{\boldsymbol{k}} \mathrm{e}^{i\boldsymbol{k}\cdot\boldsymbol{r}_i} c_{\boldsymbol{k}\sigma}, \tag{2.3}$$

where N is the number of unit cells and $\boldsymbol{r}_i$ is the real-space vector to lattice site i, who's component in one dimension would be $r_i = ia$, where a is the lattice constant. The annihilation operator $c_{\boldsymbol{k}\sigma}$ removes an electron in the momentum-state $\boldsymbol{k}$ in reciprocal space with spin σ. Transforming H to the $\boldsymbol{k}$-basis gives (see Appendix B for a thorough derivation)

$$H = \sum_{\boldsymbol{k}\sigma} \epsilon_{\boldsymbol{k}} c^\dagger_{\boldsymbol{k}\sigma} c_{\boldsymbol{k}\sigma}, \tag{2.4}$$

where $\epsilon_{\boldsymbol{k}}$ is the energy of a given $\boldsymbol{k}$-state, often referred to as the dispersion relation. For a one-dimensional lattice, the dispersion relation becomes $\epsilon_k = -2t\cos(ka)$. Going forward, we will mostly work in the $\boldsymbol{k}$-basis.

2.3 Hubbard model

The Hubbard model is an extension of the tight-binding model discussed earlier. It was first proposed by J. Hubbard in 1963 [26] as an oversimplified but useful interacting model of electrons on periodic lattices. On its on-site form, the Hubbard model reads

$$H = -t \sum_{\substack{\langle i,j\rangle \\ \sigma}} \left(c^\dagger_{i\sigma} c_{j\sigma} + c^\dagger_{j\sigma} c_{i\sigma} \right) + U \sum_i c^\dagger_{i\uparrow} c_{i\uparrow} c^\dagger_{i\downarrow} c_{i\downarrow}, \tag{2.5}$$

where U is the strength of the interaction. We call the second sum the Hubbard term. The combination of $c^\dagger_{i\sigma} c_{i\sigma}$ in the Hubbard term is called a number operator. If an electron occupies the state (i, σ), the eigenvalue of the number operator would be 1. If the electron in question is

in another state, however, the number operator's eigenvalue is 0. In other words, $c^\dagger_{i\sigma}c_{i\sigma}$ counts whether there is an electron in state (i, σ) or not. The Hubbard term describes electron-electron interaction between two electrons on the same site with opposite spin. Note that the Hubbard term only contributes if both spin up and spin down states are filled, in effect counting the cumulative energy for each filled lattice site.

The sign of U determines whether it is favorable to fill the orbital with both spin up and spin down. If $U > 0$, as originally proposed by Hubbard, the energy of the system will increase as more orbitals are filled with electrons, hence making filling unfavorable. However, if $U < 0$, filling becomes favorable because the energy decreases for every filled orbital. Due to the Pauli principle and the repulsive Coulomb interaction, it makes sense to consider $U > 0$: the electrons want to stay far away from each other. However, we will consider the case of $U < 0$, called the attractive Hubbard model. The attractive Hubbard model is a simplified model that describes an effective attraction between electrons. An example where attraction between electrons is important is the BCS theory of superconductivity, where electrons form so-called Cooper pairs.

2.4 Superconductivity

Two main features define a superconductor: zero electrical DC resistance and the Meissner effect. We will explain the former after a primer on the BCS theory. While we make due with a heuristic explanation of the latter.

The Meissner effect is the ability of a superconductor to fully expel an external magnetic field from the inside of the metal. A regular conductor placed in an external magnetic field has practically the same magnetic flux both inside and outside the metal. A superconductor, on the other hand, expels it completely due to induced currents on the surface of the metal. The surface current induces a magnetic field exactly opposite of the external field canceling it completely, i.e. a perfect diamagnetic response. Though, there exists two types of superconductors, namely type I and type II. Type I is characterized by the fact that the internal magnetic flux of the superconductor can jump to a finite value if the external field strength surpasses a critical value. While type II superconductors see a continuous increase in internal magnetic flux if the external field is strong enough. Although the Meissner effect is a remarkable effect on its own, we will not spend more time discussing it. See chapter 1.3 of [27] for more on the Meissner effect.

These properties, zero electrical resistance and the Meissner effect, were for a long time only understood empirically. A quantum theory able to explain both effects was not developed until 1957 by Bardeen, Cooper and Schrieffer, namely the BCS theory of superconductivity.

2.4.1 Introduction to BCS theory

Electrons in a crystal are subject to two main interactions, namely electron-electron interactions and electron-ion interactions. The Coulomb interaction between electrons, along with the Pauli principle, make electrons repel each other. However, the story of superconductivity is of attraction between electrons. We will discuss how effective attraction between electrons arises to form Cooper pairs. But first, we will look at a toy model Cooper thought of in 1956 [28] (referred to as the Cooper problem) leading up to the full BCS theory in 1957.

The Cooper problem

The Cooper problem uses a simplified model to show how electrons can form bound states just by introducing a weak attractive interaction in the Hamiltonian, without considering the source of this attraction. As told in chapter 3 of [27], the story goes as follows: Imagine a system with a sea of electrons where all states up to the Fermi level are occupied, i.e. a Fermi sea. Consider also that the electrons in the Fermi sea are non-interacting. If we now add *two* electrons near the

surface of the Fermi sea, we claim these two electrons only interact with each other and not with the Fermi sea. However, their interaction is not repulsive, like the Coulomb interaction. Instead they interact attractively, given that the electrons are within a small energy ω_0 from the Fermi surface and on opposite sides of the Fermi surface. Otherwise they do not interact at all. Being on the opposite sides of the Fermi surface means that if one electron is in the free state $|\boldsymbol{k}\rangle$, the other must be in a free state $|-\boldsymbol{k}\rangle$. In the absence of this attractive interaction, the two-particle state is denoted $|\boldsymbol{k}, -\boldsymbol{k}\rangle$. Meanwhile, in the presence of the interaction, we denote it $|1, 2\rangle$. We write the Hamiltonian of this system as

$$H = H_0 + V_{\text{eff}}, \tag{2.6}$$

where H_0 is the free Hamiltonian and V_{eff} describes the attractive interaction. The energy eigenvalue of $|\boldsymbol{k}, -\boldsymbol{k}\rangle$ can be found from the Schrödinger equation[2]

$$H_0|\boldsymbol{k}, -\boldsymbol{k}\rangle = 2\epsilon_{\boldsymbol{k}}|\boldsymbol{k}, -\boldsymbol{k}\rangle, \tag{2.7}$$

where $\epsilon_{\boldsymbol{k}}$ is the energy of a single non-interacting electron. Note that the energy eigenvalue of the two non-interacting electrons in state $|\boldsymbol{k}, -\boldsymbol{k}\rangle$ is $2\epsilon_{\boldsymbol{k}}$, and that for free electrons $2\epsilon_{\boldsymbol{k}} > 2\epsilon_{\text{F}}$, where ϵ_{F} is the Fermi energy. If we include the attractive interaction, the Schrödinger equation becomes

$$H|1, 2\rangle = (H_0 + V_{\text{eff}})|1, 2\rangle = E|1, 2\rangle, \tag{2.8}$$

where E is the two-particle energy of the attractively interacting electrons above the Fermi sea. We want to find out if $E < 2\epsilon_{\boldsymbol{k}}$. If it is, then it means that the electrons have lower energy when interacting than when free. Which means that the two electrons have formed a bound state. The problem at hand is to find E. The first step towards goal is to assume the states $|\boldsymbol{k}, -\boldsymbol{k}\rangle$ form a complete basis such that we can expand $|1, 2\rangle$ in this basis,

$$|1, 2\rangle = \sum_{\boldsymbol{k}} a_{\boldsymbol{k}}|\boldsymbol{k}, -\boldsymbol{k}\rangle, \tag{2.9}$$

where $a_{\boldsymbol{k}}$ are the expansion coefficients. Determining $a_{\boldsymbol{k}}$ involves using that the interaction between the electrons is negative and constant $-V$, where $V > 0$, only when they are within an energy ω_0 from the Fermi surface, and zero otherwise. Then, one introduces the density of states $D(\epsilon) = \sum_{\boldsymbol{k}} \delta(\epsilon - \epsilon_{\boldsymbol{k}})$ to find the form of $a_{\boldsymbol{k}}$. The actual calculation is not that informative for our purposes but can be found in chapter 3 of [27]. Hence, we jump to the equation which determines E,

$$\frac{1}{\lambda} = \ln\left[1 + \frac{2\omega_0}{\Delta}\right]. \tag{2.10}$$

Above, $\lambda \equiv VD(\epsilon_{\text{F}})$ and $\Delta \equiv 2\epsilon_{\text{F}} - E$. Because both $V > 0$ and $D(\epsilon_{\text{F}}) > 0$, it means that $\lambda > 0$. The only way Eq. (2.10) can have a solution is if $\Delta > 0$, which means $E < 2\epsilon_{\text{F}}$. Recall, we wanted to know if $E < 2\epsilon_{\boldsymbol{k}}$ because then the interacting electrons would have formed a bound state. We just found that E not only is smaller than $2\epsilon_{\boldsymbol{k}}$ but also $2\epsilon_{\text{F}}$. The electrons have formed a Cooper pair. This result was surprising at the time. Mostly because there were no known mechanisms that caused attraction between electrons, but Cooper did the calculations anyway. What makes the result even more odd, is that the collective two-particle energy is less than $2\epsilon_{\text{F}}$. This result can be thought of as a slight violation of the Pauli principle. The short explanation is that by forming a Cooper pair, the emerging "particle" has a slightly different statistic than their individual fermion statistics. Each electron in a Cooper pair has opposite spin relative to each other, which means their total spin is 0. Thus, Cooper pairs can be thought of as boson-like particles. And as we know, only fermions are subject to the Pauli principle.

[2] We set $V_{\text{eff}} = 0$ when considering the non-interacting case.

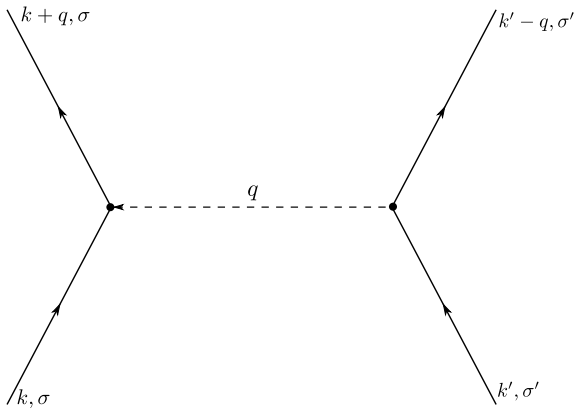


Figure 2.3: Feynman diagram describing electrons in states $(\boldsymbol{k},\sigma)$ and $(\boldsymbol{k}',\sigma')$ being scattered into states $(\boldsymbol{k}+\boldsymbol{q},\sigma)$ and $(\boldsymbol{k}'-\boldsymbol{q},\sigma')$ with a momentum transfer $\boldsymbol{q}$. Figure is inspired from [27].

The BCS mechanism

We will now introduce a mechanism that realizes Cooper pairs by including two known interactions: electron-electron and electron-ion interactions. We know already that electrons interact repulsively through the Coulomb interaction. If we include electron-ion interactions, the emerging effective interaction between electrons turns out to be attractive. Electrons interact with ions by exchange of phonons. Phonons are quasi-particles and defined as the quantization of energy modes of ions in a lattice. The Hamiltonian for this system reads [27]

$$\begin{aligned} H = &\sum_{\boldsymbol{k}\sigma} \epsilon_{\boldsymbol{k}} c^\dagger_{\boldsymbol{k}\sigma} c_{\boldsymbol{k}\sigma} + \sum_{\boldsymbol{k}\boldsymbol{k}'\boldsymbol{q}\sigma\sigma'} \frac{1}{4\pi\epsilon_0}\frac{2\pi e^2}{q^2} c^\dagger_{\boldsymbol{k}+\boldsymbol{q},\sigma} c^\dagger_{\boldsymbol{k}'-\boldsymbol{q},\sigma'} c_{\boldsymbol{k}\sigma} c_{\boldsymbol{k}'\sigma'} \\ &+ \sum_{\boldsymbol{k}\boldsymbol{q}\sigma} M_{\boldsymbol{q}} (a^\dagger_{-\boldsymbol{q}} + a_{\boldsymbol{q}}) c^\dagger_{\boldsymbol{k}+\boldsymbol{q},\sigma} c_{\boldsymbol{k}\sigma}, \end{aligned} \tag{2.11}$$

where the second term (Coulomb term) describes Coulomb repulsion between electrons and the last term (phonon term) is the electron-phonon interaction with interaction coupling $M_{\boldsymbol{q}}$. The $\boldsymbol{q}$-momentum in both sums describes a momentum transfer in a scattering process between two electrons and between an electron and a phonon, respectively. Moreover, $a^\dagger$ and a are the bosonic creation and annihilation operators describing the phonons. We can understand the Coulomb and phonon sums in terms of a simple Feynman diagram. Figure 2.3 shows a scattering process of two electrons mediated by a virtual particle with momentum $\boldsymbol{q}$. The Coulomb term describes two electrons in state $(\boldsymbol{k},\sigma)$ and $(\boldsymbol{k}',\sigma')$ approaching their vertex (black dot). The electron in state $(\boldsymbol{k}',\sigma')$ loses a momentum $\boldsymbol{q}$ and enters a state $(\boldsymbol{k}'-\boldsymbol{q},\sigma')$. Meanwhile, the electron in state $(\boldsymbol{k},\sigma)$ acquires a momentum $\boldsymbol{q}$ and enters state $(\boldsymbol{k}+\boldsymbol{q},\sigma)$. In the Coulomb case, the carrier of $\boldsymbol{q}$ is a photon. This diagram can also be used to describe the phonon term. If we cut the dashed line into two pieces, we are left with two vertices. Each half corresponds to their term in the phonon summand, i.e. the $a^\dagger_{-\boldsymbol{q}}$- and $a_{\boldsymbol{q}}$-term. This time, the carrier of $\boldsymbol{q}$ is a phonon instead of a photon, and the two vertices describe electron-phonon scattering. Next, we will see Fig. 2.3 describing the entire Hamiltonian instead of only describing pieces of it. Chapter 3 in [27] shows that we can rewrite the phonon term such that we obtain a single effective interaction term describing electron-electron scattering. The Hamiltonian becomes

$$H = \sum_{\boldsymbol{k}\sigma} \epsilon_{\boldsymbol{k}} c^\dagger_{\boldsymbol{k}\sigma} c_{\boldsymbol{k}\sigma} + \sum_{\boldsymbol{k}\boldsymbol{k}'\boldsymbol{q}\sigma\sigma'} \tilde{V}_{\text{eff}}(\boldsymbol{q},\omega) c^\dagger_{\boldsymbol{k}+\boldsymbol{q},\sigma} c^\dagger_{\boldsymbol{k}'-\boldsymbol{q},\sigma'} c_{\boldsymbol{k}\sigma} c_{\boldsymbol{k}'\sigma'}, \tag{2.12}$$

where

$$\tilde{V}_{\text{eff}}(\boldsymbol{q},\omega) = \frac{2|M_{\boldsymbol{q}}|^2\omega_{\boldsymbol{q}}}{\omega^2-\omega_{\boldsymbol{q}}^2} + \frac{1}{4\pi\epsilon_0}\frac{2\pi e^2}{q^2}. \tag{2.13}$$

In $\tilde{V}_{\text{eff}}$, $\omega_{\boldsymbol{q}}$ is the eigenvalue of a phonon in a plane-wave state $|\boldsymbol{q}\rangle$ and ω is the energy transfer in the electron-electron scattering. Recall that Eq. (2.11) includes both Coulomb and phonon interactions as two separate effects, while in Eq. (2.12) we have combined the two interactions into one effective interaction. This effective interaction is on a similar form as the original Coulomb interaction. Which means it describes an electron-electron scattering exactly like the one shown in Fig. 2.3. As we suggested earlier, this effective interaction can be attractive, i.e. $\tilde{V}_{\text{eff}} < 0$. We see that $\tilde{V}_{\text{eff}} < 0$ if ω^2 approaches $\omega_{\boldsymbol{q}}^2$ from below, regardless of the strength of the repulsive Coulomb interaction. In other words, it does not matter how small the electron-phonon coupling is, or how strong the Coulomb repulsion is, there will be a range of ω that will almost guarantee effective attraction between electrons. And considering what we learned from the Cooper problem, this means electrons can form bound states such as Cooper pairs. Recall that we did not specify the source of attraction between electrons in the Cooper problem. And in all honesty, it did not seem like any known interactions could give us the attraction. However, we have just seen that combining two quite reasonable interactions leads us to an effective attractive interaction between electrons. This is the motivation behind the BCS theory.

The results we have achieved so far are mathematically appealing. However, we are yet to gain a physical understanding of attraction between electrons. This is what we will try to do next. Imagine an electron whizzing past a heavier ion in a lattice. The massive ion, like the electron, will feel an attraction and get displaced from its equilibrium position. Because the ion is so massive, it will relax back to its equilibrium relatively slowly. Simultaneously giving the electron time to get far away. Because the ion is still displaced from equilibrium, and in the vicinity of a second approaching electron, the second electron interacts with the ion just like the first one. Effectively, this describes an attractive interaction between electrons mediated by a phonon, because the second electron interacted with the ion only because the ion was displaced from its equilibrium. Note, however, for the electrons to be able to interact like this, the second electron has to interact with the ion before the ion has had time to relax. Moreover, the second electron must also wait an appropriate amount of time before approaching to give the first electron the opportunity to get as far away as possible to minimize Coulomb repulsion. We see from this that an effective interaction is only realized under certain conditions. These conditions are analogous to being on opposite sides and an energy ω_0 from the Fermi surface in the Cooper problem, and ω^2 approaching $\omega_{\boldsymbol{q}}^2$ from below in Eq. (2.13).

Although this is the physical picture of the interaction, BCS theory only implicitly includes these details. The interaction, as described above, is local in space and retarded in time. And as described in the introduction, BCS theory does not account for retardation effects explicitly. These effects are only treated in detail by theories such as Eliashberg theory.

How does the BCS theory explain zero electrical resistance?

So far, we have developed an understanding of what the BCS theory is built on. In short, electrons go together to form Cooper pairs with a bound energy less than the Fermi energy. The fact that their bound energy is less than the Fermi energy tells us that their fermion statistics has been altered. To be clear, Cooper pairs are not bosons. However, they do have some boson-like properties. The main property we will focus on is called Cooper pair condensation, comparable to Bose-Einstein condensation [29, 30]. A Bose-Einstein condensate (BEC) is a state of matter at near zero temperatures where bosons condense into the ground state of the system. Bosons, unlike fermions, can occupy the same quantum state, and in a BEC they all get trapped in the ground state. Similarly, Cooper pairs get trapped in a condensate below the Fermi surface. In this context, *trapped* means that many Cooper pairs are in energy states 2Δ below the Fermi surface.

Δ is called the superconducting order parameter, or the superconducting gap. The explanation of zero electrical resistance relies on this condensate and goes as follows [27, 31]: The electrons of the condensate have highly correlated momenta, i.e. the entire condensate moves like a unit. If we apply an electric field, generating a current in a direction, in practice all electrons in the condensate will move opposite of that field. Which means there is an increase of momenta in the direction of the flow. In regular conductors, resistance comes from electrons being scattered in the opposite direction of the current due to e.g. imperfections in the metal. In superconductors, however, this nearly never happens. For electrons to be scattered opposite of the flow, the scattering energy must be enough to break a Cooper pair and excite the electrons to the opposite side of the Fermi surface. There are no scattering mechanisms in the superconducting state that have enough energy Δ to do this. Hence, superconductors have zero resistance.

Bardeen used a football field invaded by a crowd as an analogy for the electron pairs in a superconductor. We can use a slightly different analogy communicating the same point. The electron pairs can be thought of as leafcutter ants carrying leaves back to their nest. Such ants move in large numbers, all in the same direction. The colony as a whole manages to supply their nest with leaves at a steady rate without interruption in spite of their path being imperfect. The flow of leaves can be compared to current in the superconductor unimpaired by relatively small obstacles, such as twigs and rocks.

We will derive an equation which determines the superconducting gap Δ in Chapter 3, called the gap equation. As we will see, the gap equation can be used to determine both Δ and the critical temperature T_c.

2.4.2 Eliashberg theory

The BCS theory works well for many elemental superconductors. One example is aluminium, which has a critical temperature of $T_c \approx 1.2\,\mathrm{K}$ [32]. Common for these elemental superconductors is the fact that the interaction strength, or coupling, between the electrons and the phonons is relatively weak. A common number describing coupling strength is the positive coupling constant λ, which is defined as $\lambda \equiv VD(\epsilon_\mathrm{F})$ in BCS theory. Here V is the effective interaction strength and $D(\epsilon_\mathrm{F})$ is the density of states at the Fermi level. BCS theory is based on the simple approximation that V can be taken to be small and constant close to the Fermi level, and zero otherwise. Generally, however, this approximation is inadequate. This was made clear in the 1960s when the first discrepancies between experimental results and theoretical predictions surfaced. The reason for the discrepancy was that many superconductors have quite strong electron-phonon coupling. Such superconductors are not well suited for BCS theory. In fact, it has been shown that one has to limit $\lambda \ll 1$ if the BCS theory is to have decent predictive power. Aluminium, for instance, has a coupling constant of $\lambda \approx 0.4$ [32], and is thus well described by BCS theory. Mercury, on the other hand, being the first superconductor ever discovered is ironically not suited to be described by the BCS theory due to its large coupling constant of $\lambda \approx 1.6$ [32]. Any superconductor with a larger coupling constant would be inaccurately described by BCS theory. Hence the need for a more general theory.

A theory better equipped to describe superconductors like mercury and lead is called Eliashberg theory named after G. Eliashberg [10, 11]. The goal of Eliashberg theory is to more accurately describe the electron-phonon interaction. The BCS theory assumes weak, constant coupling which is instantaneous, with little to no regard for the inertial difference between electrons and ions. Therefore, to get the whole picture of the interaction, one must consider the different time scales of the electron-phonon interaction and the Coulomb interaction. The photon mediated Coulomb interaction is practically instant. Meanwhile, the phonon mediated attractive electron-electron interaction is very much frequency dependent. Which means that the electron-phonon interaction term must be fully retarded. The intuition for this is the fact that ions are much more massive than electrons and thus have more inertia. And by inspecting the periodic table of elements, one can easily see that mercury and lead ions are much heavier than an aluminium ion. Another difference

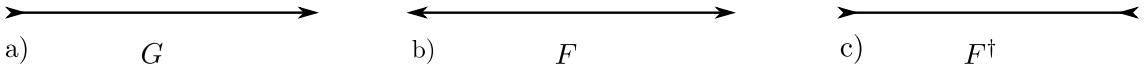


Figure 2.4: Diagram of fermionic propagators in Eliashberg theory defined in Eq. (2.14).

between BCS and Eliashberg theory is the self-energy functions of the electron. BCS neglects these with the assumption that their contribution is negligible, which in fairness is a good approximation if the coupling is weak. However, now that the coupling between electrons and phonons is strong, the electron self-energy should not be ignored.

The difference in approach between the theories is that in Eliashberg theory, it is much more expedient to work in the imaginary-time (Matsubara) Green's function formalism [33]. In this formalism, one defines a few Green's functions with imaginary-time τ

$$\begin{aligned} G(\boldsymbol{k},\tau) &\equiv -\langle T_\tau c_{\boldsymbol{k}\sigma}(\tau) c^\dagger_{\boldsymbol{k}\sigma}(0)\rangle \\ F(\boldsymbol{k},\tau) &\equiv -\langle T_\tau c_{\boldsymbol{k}\uparrow}(\tau) c_{-\boldsymbol{k}\downarrow}(0)\rangle \\ F^\dagger(\boldsymbol{k},\tau) &\equiv -\langle T_\tau c^\dagger_{-\boldsymbol{k}\downarrow}(\tau) c^\dagger_{\boldsymbol{k}\uparrow}(0)\rangle \end{aligned} \tag{2.14}$$

where T_τ is a time-ordering operator defined as

$$T_\tau A(\tau)B(\tau') = \Theta(\tau-\tau')A(\tau)B(\tau') - \Theta(\tau'-\tau)B(\tau')A(\tau), \tag{2.15}$$

and $\Theta(\tau)$ is the Heaviside step function and A, B are fermion operators. The time-ordering operator keeps track of the order of the operators with regards to present and past. The three Green's functions are all single-particle Green's functions. However, the first function, G, is a normal Green's function which diagrammatically describes an electron line pointing in one direction, i.e. an electron with momentum and spin $(\boldsymbol{k},\sigma)$ created at one point in time and destroyed at another, see Fig. 2.4a. The last two functions, F and $F^\dagger$, are called anomalous Green's functions. Diagrammatically, they describe electron lines pointing in towards center ($F^\dagger$) or out from center (F). In the case of $F^\dagger$, an electron is created at each end of the line resulting in two electron lines pointing inward, see Fig. 2.4c. While for F, an electron is destroyed at each end resulting in two electron lines pointing outward, see Fig. 2.4b. We call these functions anomalous because, unlike G, they do not conserve particle number.

Eliashberg theory is built on these Green's functions. And from these Green's functions emerges two vital quantities called self-energy functions. The self-energy functions of this theory are S and W, called the exchange self-energy and the superconducting gap, respectively. In BCS theory, Δ is equivalent to W while S is unique to Eliashberg theory. We can define S and W in Eliashberg theory as [34]

$$S(\boldsymbol{p}, i\omega_n) = -\frac{1}{\beta}\sum_{\boldsymbol{q},m} V_{\text{eff}}(\boldsymbol{q}, i\omega_m) G(\boldsymbol{p}+\boldsymbol{q}, i\omega_n + i\omega_m) \tag{2.16}$$

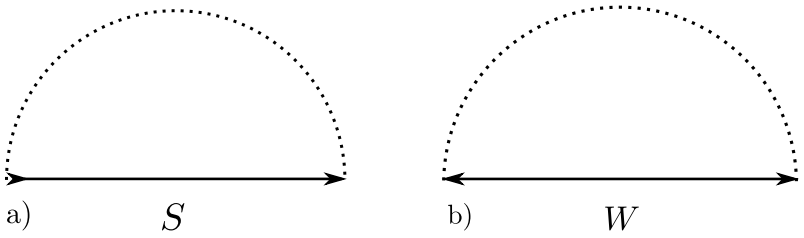


Figure 2.5: Diagram of self-energies in Eliashberg theory. The solid lines are fermionic propagators while the dashed lines are bosonic propagators represented as V_{eff} in Eqs. (2.16) and (2.17).

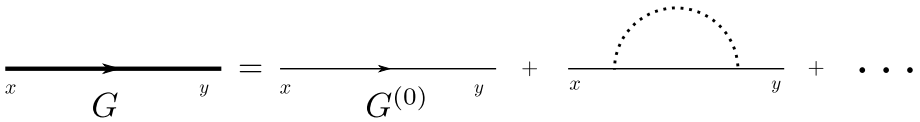


Figure 2.6: Diagram of the full propagator taking an electron from state x to y. The thick electron line on the LHS represents the full propagator G. The RHS is a sum of the constituents of the full propagator: the bare propagator $G^{(0)}$ and an infinite sum of $G^{(0)}$ with self-interactions.

$$W(\boldsymbol{p}, i\omega_n) = -\frac{1}{\beta}\sum_{\boldsymbol{q},m} V_{\text{eff}}(\boldsymbol{q}, i\omega_m) F(\boldsymbol{p}+\boldsymbol{q}, i\omega_n + i\omega_m), \tag{2.17}$$

where V_{eff} is the effective attractive interaction strength. A diagrammatic interpretation of S and W is shown in Fig. 2.5.

A note on self-energy

The concept of self-energy can be tricky to understand, which is why we will spend a little time on it here. We will use Richard Feynman's diagrammatic approach as our main tool. In fact, we have already used Feynman diagrams in Figs. 2.3 to 2.5 to explain interactions, propagators and, lastly, self-energies. In short, self-energies are defined as corrections to bare propagators. Because bare propagators are only valid in an interaction-free theory, we need to correct these propagators when moving to a theory with interactions. The quantum theory of superconductivity is an example of an exclusively interacting theory, which is why it is important to understand self-energy.

Imagine an electron completely alone in the universe where nothing is able to affect its state. Without interactions, this electron is described by a so-called bare electron propagator. We denote bare propagators as $G^{(0)}$. For an electron, the bare propagator is often written as [35]

$$G^{(0)}(p) = \frac{i(\not{p} + m_0)}{p^2 - m_0^2 + i\epsilon}. \tag{2.18}$$

In reality, however, electrons are never alone in the universe, and thus they are always subject to interactions. Thus, the bare propagator is only an approximation of reality. To improve our approximation we have to correct the bare propagator by including the possibility of self-interactions. Remember that in any interaction, we can only observe the initial state and the final state. The intermediate states are virtual and cannot be observed. This is captured by Feynman diagrams by drawing internal lines, like the boson line in Fig. 2.3. One can then imagine a range of possible ways an electron can go from state x to y by interacting with itself. Fig. 2.6 shows some of the possibilities. We can be more succinct when defining the possible self-interactions by considering *irreducible* diagrams. We define a one-particle irreducible (1PI) diagram to be any diagram that

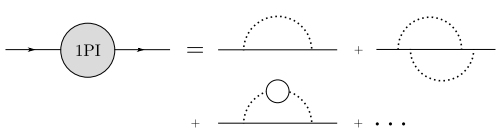


Figure 2.7: Diagram of some of the one-particle irreducible (1PI) diagrams. Note that the definition of 1PI diagrams is without external lines, i.e. only the gray circle on the LHS represents all 1PI diagrams.

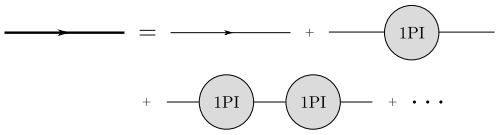


Figure 2.8: Diagram of the full propagator and its constituents. The thick line represents the full propagator. A sum of infinite combinations of all 1PI diagrams with the bare propagator defines the full propagator. This is a diagrammatic representation of Eq. (2.19).

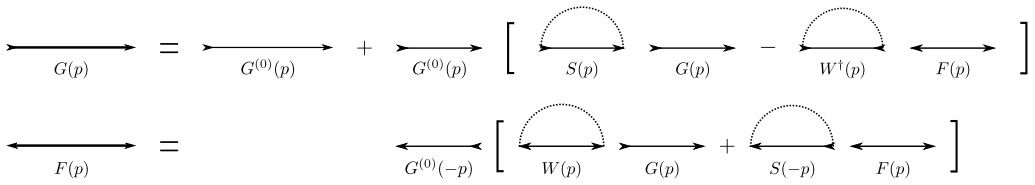


Figure 2.9: Diagrammatic representation of Eq. (2.21). An important step when constructing these diagrams is to always conserve particle number at each vertex. Note that only one self-energy diagram is used for both S and W.

cannot be split in two by cutting a single internal line [35], see Fig. 2.7 for examples of such diagrams. The important part is that the sum of all infinite 1PI diagrams defines the self-energy Σ, i.e. the sum of all diagrams on the right hand side in Fig. 2.7. To fix the problem of the bare propagator we need to account for all infinite 1PI diagrams and their combinations. The full propagator G is a sum of the bare propagator and all infinite combinations of 1PI diagrams as shown in Fig. 2.8. We can write this as

$$\begin{aligned} G(p) &= G^{(0)}(p) + G^{(0)}(p)\Sigma(p)G^{(0)}(p) + G^{(0)}(p)\Sigma(p)G^{(0)}(p)\Sigma(p)G^{(0)}(p) + \dots \\ &= G^{(0)}(p) + G^{(0)}(p)\Sigma(p)\left[G^{(0)}(p) + G^{(0)}(p)\Sigma(p)G^{(0)}(p) + \dots\right] \\ &= G^{(0)}(p) + G^{(0)}(p)\Sigma(p)G(p). \end{aligned} \tag{2.19}$$

The expression inside the square brackets is just the full propagator G itself. The emerging equation determines the full propagator and is famously known as the Dyson equation

$$G(p) = G^{(0)}(p)[1 + \Sigma(p)G(p)]. \tag{2.20}$$

In Eliashberg theory, S corresponds to Σ while W does not have any direct equivalent in our discussion so far. This is due to the anomalous Green's functions F and $F^\dagger$ and the fact that $F^{(0)} = F^{\dagger(0)} = 0$ [34], i.e. superconductivity cannot occur in a non-interacting theory. However, if we were to do a similar analysis of S and W, we would end up with something analogous to the Dyson equation [34]

$$\begin{aligned} G(p) &= G^{(0)}(p)[1 + S(p)G(p) - W^\dagger(p)F(p)] \\ F(p) &= G^{(0)}(-p)[W(p)G(p) + S(-p)F(p)]. \end{aligned} \tag{2.21}$$

These equations are depicted diagrammatically in Fig. 2.9. The construction of Eq. (2.21) from Eq. (2.20) can, to begin with, be done only by inspection. Equation (2.20) does not have any contribution from the anomalous functions F and $F^\dagger$. When including W (or $W^\dagger$) in Eq. (2.20), the trick is to remember the conservation of particle number at each vertex. This rule will determine the allowed vertices. Constructing a similar equation for F (or $F^\dagger$) is done by first realizing $F^{(0)} = 0$ and then determining which propagators need to be paired with S and W (while conserving particle number at each vertex). This method of constructing Eq. (2.21) is only meant as a sketch. A more rigorous method is to derive the corresponding Dyson equations by e.g. the equations of motion approach. We will derive Dyson equations similar to Eq. (2.21) that are specific to our model in Section 3.3.

In principle one is required to sum all infinite self-energy diagrams to get the proper full propagator. Because this is not possible, one has to make due with only a few diagrams and argue why higher order diagrams do not contribute significantly. From Fig. 2.9 we see that the only 1PI diagram included in S and W is the first diagram on the right hand side in Fig. 2.7. The second diagram in Fig. 2.7 is a vertex correction and is ignored in Eliashberg theory (along with higher order diagrams). The reason for ignoring vertex corrections is due to a "theorem" by Migdal [36]

stating that all vertex corrections are $\mathcal{O}(\sqrt{m^*/M})$, where m^* is the effective electron mass and M the ion mass. As long as $\sqrt{m^*/M} \simeq \lambda\omega_D/\epsilon_\mathrm{F} \ll 1$, Migdal argues that all vertex diagrams can be neglected. However, Mahan [34] and Ummarino [32] discuss that ignoring higher order diagrams does put a limit on Eliashberg theory's validity. Mahan poses the following difficulty with Migdal's argument: since superconductivity itself occurs due to a vertex diagram (scattering between electrons by exchange of e.g. phonons), relying on a "theorem" which neglects all vertex corrections is unreliable. Additionally, due to the limitation $\lambda\omega_D/\epsilon_\mathrm{F} \ll 1$ introduced by Migdal, one can even argue that Eliashberg theory is a weak-coupling theory in the sense that the Fermi energy ϵ_F is still the dominant energy scale. Despite the controversy of Migdal's "theorem", neglecting all vertex corrections is customary in Eliashberg theory.

2.5 Topological insulators

Topological insulators (TIs) share some of their properties with regular insulators. The obvious property being that such materials can not conduct electricity due to the energy gap between valence and conduction bands being too large. However, the main practical difference between a TI and a topologically trivial one is that TIs have exotic metallic states on their surfaces [37], much like in the QHE. In practice this means that, unlike a regular insulator, a TI can conduct electricity along the surface or edges of the material. An incomplete but practical description of such edge states is that a TI behaves like a regular insulator coated with a metal – insulating in the bulk and conducting on the surface. The reason the above description is incomplete is because these edge states are protected by abstract symmetries of the TI, unlike conducting states in regular metals which are prone to perturbations. This is where the topology of TIs come into play.

Edge states are robust because of topological invariants. A topological invariant is a quantity, usually an integer, that remains the same (invariant) under a continuous change of the system. We can understand such invariants by comparing a doughnut to a football. In topology, one can define a quantity called the genus g. It describes the number of holes in a closed shape. A doughnut, for instance, has exactly one hole $g = 1$. Meanwhile, a football, taken to be sphere, has no holes $g = 0$. Topologically speaking, these shapes cannot by unified, because we cannot deform a doughnut into a sphere without having to cut the doughnut into a cylinder or punch a hole through the sphere. In other words, there is no continuous transformation that can transform a doughnut to a sphere, or vice versa. Another popular example of topology is the comparison of a coffee cup to a doughnut. Since both shapes have exactly one hole $g = 1$, they are topologically equivalent. We can transform a coffee cup into a doughnut without needing to cut the object, i.e. a continuous transformation.

We have seen how topological invariants can be defined for physical objects. One such topological invariant in the context of topologically non-trivial insulators is called the Chern number. The genus was invariant under continuous physical deformations of the shape. Meanwhile the Chern number is invariant under smooth variations of the Hamiltonian. As explained by Hasan and Kane [22], the Chern number can be understood in terms of the Berry phase [38]. The Berry phase is a phase the Bloch functions $|u(\boldsymbol{k})\rangle$ of the Hamiltonian acquire around a closed loop[3]. This phase is given through the line integral of the Berry connection $\boldsymbol{A} = -i\langle u(\boldsymbol{k})|\partial_{\boldsymbol{k}}|u(\boldsymbol{k})\rangle$. And the Chern number is proportional to the closed Brillouin zone integral of the Berry curvature $\boldsymbol{\Omega} \equiv \nabla \times \boldsymbol{A}$

$$C = \frac{1}{2\pi} \oiint_{\mathrm{BZ}} \mathrm{d}\boldsymbol{k} \cdot \boldsymbol{\Omega}. \tag{2.22}$$

The Chern number can be a positive or negative integer. If $C = 0$, we say the system is topologically trivial, e.g. a regular insulator. If $C \neq 0$, however, we have a topologically non-trivial system. We call insulators characterized by Chern numbers Chern insulators (CI).

[3] See [39] for more on the Berry phase.

2.5.1 Haldane model

A CI of considerable importance was presented by F. D. M. Haldane in 1988 [20]. The Haldane model is an interaction-free tight-binding model on a honeycomb lattice where we allow electrons to hop to NN and 2NN sites. Figure 2.2 shows the arrangement of NN and 2NN lattice sites. The spinful version of the model is described by the following Hamiltonian

$$\begin{aligned} H = &- t \sum_{\langle i,j\rangle\sigma} \left(c^\dagger_{i\sigma} d_{j\sigma} + d^\dagger_{j\sigma} c_{i\sigma} \right) \\ &- t' \sum_{\langle\langle i,j\rangle\rangle\sigma} \left(c^\dagger_{i\sigma} c_{j\sigma} \mathrm{e}^{i\phi} + c^\dagger_{j\sigma} c_{i\sigma} \mathrm{e}^{-i\phi} + d^\dagger_{i\sigma} d_{j\sigma} \mathrm{e}^{-i\phi} + d^\dagger_{j\sigma} d_{i\sigma} \mathrm{e}^{i\phi} \right), \end{aligned} \tag{2.23}$$

where the first term describes hopping to NN sites with t the NN hopping parameter, and the second term describes 2NN hopping with the complex hopping parameter $t'\mathrm{e}^{\pm i\phi}$. ϕ is a phase the electron wavefunctions acquire when hopping between 2NNs and $(c^\dagger, c)$ and $(d^\dagger, d)$ are creation and annihilation operators for sub-lattices A and B respectively, see Fig. 2.1. The sign of the phase is based on the conventions of Haldane. The sign is positive (negative) when hopping $j \to i$ in sublattice A (B), and changes sign when hopping $i \to j$. The Chern number of this Hamiltonian is $C = \pm 1$ and has been calculated in Appendix A. Because the Chern number is non-zero, we know that the Haldane model is indeed topologically non-trivial. And a study of its edges reveals topologically protected conduction states.

Chapter 3

Superconductivity in the Haldane-Hubbard model

One of the most interesting fields in modern condensed matter physics is the study of topological states of matter. Topological insulators are well-described by band theory and typically regarded as a system of non-interacting fermions, e.g. the Haldane model in Eq. (2.23). However, many fascinating phenomena in condensed matter physics exist simply due to electron-electron interactions, e.g. superconductivity. This fact has caused a natural increase in focus on interacting topological insulators in recent years [23], with hopes of finding novel topological states.

The Haldane model describes a non-interacting Chern insulator (CI) with nearest and second-nearest neighbor hopping. The goal of this chapter is to study topological phase transitions of this CI induced by attractive electron-electron interactions. More specifically, the transition between a CI and a superconductor in the weak and strong coupling regimes. One of the simplest examples of electron interactions on a lattice is the on-site Hubbard interaction. By adding a Hubbard term for each sub-lattice A and B of graphene to the spinful Haldane model in Eq. (2.23), we get the following Hamiltonian

$$\begin{aligned} H = &- t \sum_{\substack{\langle i,j\rangle \\ \sigma}} \left(c_{i\sigma}^\dagger d_{j\sigma} + d_{j\sigma}^\dagger c_{i\sigma} \right) - t' \sum_{\substack{\langle\langle i,j\rangle\rangle \\ \sigma}} \left(c_{i\sigma}^\dagger c_{j\sigma} \mathrm{e}^{i\phi} + c_{j\sigma}^\dagger c_{i\sigma} \mathrm{e}^{-i\phi} + d_{i\sigma}^\dagger d_{j\sigma} \mathrm{e}^{-i\phi} + d_{j\sigma}^\dagger d_{i\sigma} \mathrm{e}^{i\phi} \right) \\ &+ U \sum_{i\in A} c_{i\uparrow}^\dagger c_{i\uparrow} c_{i\downarrow}^\dagger c_{i\downarrow} + U \sum_{j\in B} d_{j\uparrow}^\dagger d_{j\uparrow} d_{j\downarrow}^\dagger d_{j\downarrow}, \end{aligned} \tag{3.1}$$

called the attractive Haldane-Hubbard model. The diagonalization of H, outlined in Appendix B, results in

$$H = \sum_{\boldsymbol{k}\sigma} (\epsilon_{\boldsymbol{k}} - \mu) c_{\boldsymbol{k}\sigma}^\dagger c_{\boldsymbol{k}\sigma} + \frac{U}{4N} \sum_{\boldsymbol{k}\boldsymbol{k}'\boldsymbol{q},\sigma} c_{\boldsymbol{k}+\boldsymbol{q}\sigma}^\dagger c_{\boldsymbol{k}'-\boldsymbol{q},-\sigma}^\dagger c_{\boldsymbol{k}',-\sigma} c_{\boldsymbol{k}\sigma}, \tag{3.2}$$

where μ is defined as the chemical potential (Fermi level), N is the number of unit cells, and $\epsilon_{\boldsymbol{k}}$ is the lower band of the dispersion relation defined as

$$\epsilon_{\boldsymbol{k}} = -2t' \cos\phi \sum_j \cos(\boldsymbol{k}\cdot\boldsymbol{\delta}_j) - \left[t^2 |S(\boldsymbol{k})|^2 + 4t'^2 \sin^2\phi \Big(\sum_j \sin(\boldsymbol{k}\cdot\boldsymbol{\delta}_j) \Big)^2 \right]^{1/2}, \tag{3.3}$$

where $S(\boldsymbol{k}) \equiv \sum_j \exp(i\boldsymbol{k}\cdot\boldsymbol{e}_j)$. Both NN and 2NN vectors $\boldsymbol{e}_j$ and $\boldsymbol{\delta}_j$, respectively, are defined in Eq. (2.1). Going forward we will measure energy in units of t, thus setting $t = 1$. Moreover, the

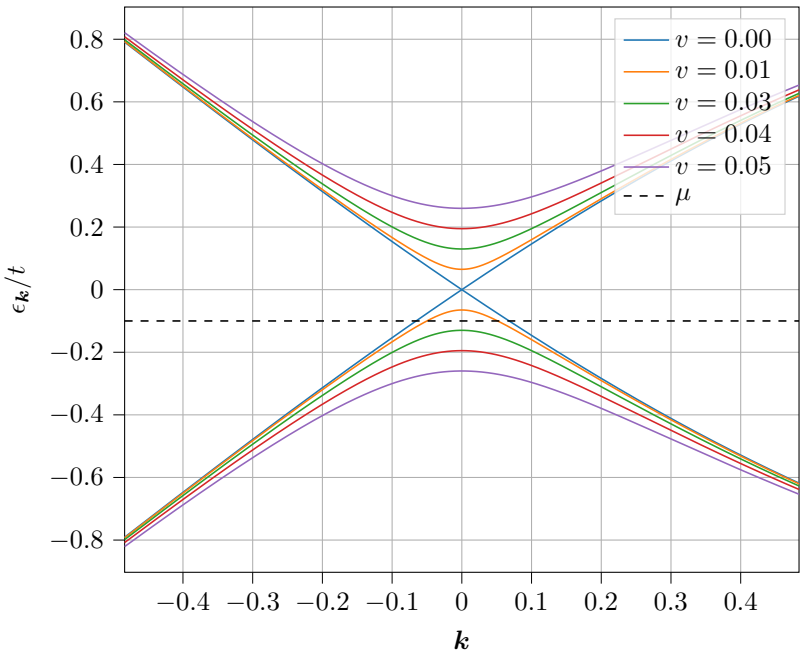


Figure 3.1: Plot of the upper and lower band of the dispersion relation $\epsilon_{\boldsymbol{k}}$ for varying values of $v \equiv t'/t$. The lower band is given in Eq. (3.3). The Fermi level μ is drawn in as a dashed line to show how the system can transition between insulating and superconducting states depending on t'.

phase ϕ will be set to $\phi = \pi/2$. This is motivated by the common choice of 2NN hopping parameter $\pm it'$. Note that the creation and annihilation operators in Eq. (3.2) are pseudo-particle fermion operators unlike the ones in Eq. (3.1).

3.1 Superconducting in the bulk

The CI described by the Haldane model is insulating in the bulk. However, by introducing attractive electron coupling, correlated electron states like Cooper pairs emerge. Which in turn makes the CI superconductive in the bulk. This phase transition has a critical point determined by t' and μ. For a given Fermi level μ, we can tune the 2NN hopping parameter t' such that the shape and size of the gap in the dispersion relation $\epsilon_{\boldsymbol{k}}$ changes, shown in Fig. 3.1. As long as μ resides inside the energy gap, the bulk remains insulating. At a critical value t'_c, however, μ will intersect with an energy band resulting in a transition to the superconductive state. The critical t'_c can be calculated in terms of μ by setting the dispersion relation $\epsilon_{\boldsymbol{k}}$ equal to μ at the Dirac point $\boldsymbol{k} = \boldsymbol{K} \equiv (4\pi/3\sqrt{3}, 0)$

$$t'_c = \frac{\mu}{3(\cos\phi - \sqrt{3}|\sin\phi|)} = \frac{|\mu|}{3\sqrt{3}}, \tag{3.4}$$

where we have chosen the lower band $\mu < 0$ and set $\phi = \pi/2$, like in Fig. 3.1.

3.2 BCS theory

The Bardeen-Cooper-Schrieffer (BCS) theory is the first and a highly successful quantum theory for describing superconductors with weak electron correlation. Which is why BCS is a natural

starting point for studying superconductivity in our interacting Haldane model. We quantify weak coupling by $\lambda \ll 1$, where $\lambda \equiv VD(\epsilon_{\mathrm{F}})$ is called the coupling constant. λ is an appropriate measure of coupling strength with V defined as an effective interaction strength and $D(\epsilon_{\mathrm{F}})$ the density of states at the Fermi level. The next step is to derive the BCS gap equation with Eq. (3.2) as our Hamiltonian.

The interaction term in Eq. (3.2) describes a scattering process between two electrons with momenta $(\boldsymbol{k}, \boldsymbol{k}')$ scattered to momenta $(\boldsymbol{k}+\boldsymbol{q}, \boldsymbol{k}'-\boldsymbol{q})$. Generally we allow all $\boldsymbol{k}'$ in the sum. However, one of the important assumptions of BCS theory is that only electrons close to the Fermi surface with opposite momenta and spin can interact attractively. If we say that $\boldsymbol{k}$ is within a thin shell around the Fermi surface, then only a handful of $\boldsymbol{k}'$ and $\boldsymbol{q}$ momenta are relevant to the attraction. Thus, we need to limit $\boldsymbol{k}'$ such that if $\boldsymbol{k}+\boldsymbol{q}$ is close to the Fermi surface, $\boldsymbol{k}'-\boldsymbol{q}$ also stays close to the Fermi surface. The choice of $\boldsymbol{k}'$ that satisfies this constraint for all $\boldsymbol{q}$ is $\boldsymbol{k}' = -\boldsymbol{k}$ [27]. Only including $\boldsymbol{k}' = -\boldsymbol{k}$ and rewriting $\boldsymbol{k}+\boldsymbol{q} \to \boldsymbol{k}$ and $\boldsymbol{k} \to \boldsymbol{k}'$ in the interaction term, we get

$$H = \sum_{\boldsymbol{k}\sigma} (\epsilon_{\boldsymbol{k}} - \mu) c^\dagger_{\boldsymbol{k}\sigma} c_{\boldsymbol{k}\sigma} + \frac{U}{2N} \sum_{\boldsymbol{k}\boldsymbol{k}'} c^\dagger_{\boldsymbol{k}'\uparrow} c^\dagger_{-\boldsymbol{k}\downarrow} c_{-\boldsymbol{k}'\downarrow} c_{\boldsymbol{k}'\uparrow}. \tag{3.5}$$

3.2.1 Mean-field theory

The BCS theory is a mean-field theory, which is why we will simplify the interaction term in Eq. (3.5), denoted H_U, by rewriting it using

$$c_{-\boldsymbol{k}\downarrow} c_{\boldsymbol{k}\uparrow} = b_{\boldsymbol{k}} + \underbrace{c_{-\boldsymbol{k}\downarrow} c_{\boldsymbol{k}\uparrow} - b_{\boldsymbol{k}}}_{=\delta_{b_{\boldsymbol{k}}}}, \tag{3.6}$$

where $b_{\boldsymbol{k}} = \langle c_{-\boldsymbol{k}\downarrow} c_{\boldsymbol{k}\uparrow} \rangle$ is a statistical mean [27]. The mean-field approximation is that the difference $\delta_{b_{\boldsymbol{k}}}$ is small. Thus, we insert Eq. (3.6) back into H_U and neglect all terms $\mathcal{O}(\delta^2_{b_{\boldsymbol{k}}})$

$$H_U = \frac{U}{2N} \sum_{\boldsymbol{k}\boldsymbol{k}'} \left(b_{\boldsymbol{k}'} c^\dagger_{\boldsymbol{k}\uparrow} c^\dagger_{-\boldsymbol{k}\downarrow} + b^\dagger_{\boldsymbol{k}} c_{-\boldsymbol{k}'\downarrow} c_{\boldsymbol{k}'\uparrow} - b^\dagger_{\boldsymbol{k}} b_{\boldsymbol{k}'} \right). \tag{3.7}$$

H_U can be simplified further by introducing the superconducting order parameter

$$\Delta \equiv -\frac{U}{2N} \sum_{\boldsymbol{k}} b_{\boldsymbol{k}}. \tag{3.8}$$

This is the superconducting gap and the solution to the gap equation. Inserting Δ and $\Delta^\dagger$ in H_U gives

$$H_U = -\sum_{\boldsymbol{k}} \left(\Delta c^\dagger_{\boldsymbol{k}\uparrow} c^\dagger_{-\boldsymbol{k}\downarrow} + \Delta^\dagger c_{-\boldsymbol{k}\downarrow} c_{\boldsymbol{k}\uparrow} \right) + \Delta \sum_{\boldsymbol{k}} b_{\boldsymbol{k}}. \tag{3.9}$$

This is the final form of the mean-field approximated H_U. All that remains now is to diagonalize the full Hamiltonian and find an equation for Δ self-consistently.

3.2.2 Diagonalizing the Hamiltonian

The full Hamiltonian H up to this point is

$$H = \sum_{\boldsymbol{k}\sigma} (\epsilon_{\boldsymbol{k}} - \mu) c^\dagger_{\boldsymbol{k}\sigma} c_{\boldsymbol{k}\sigma} - \sum_{\boldsymbol{k}} \left(\Delta c^\dagger_{\boldsymbol{k}\uparrow} c^\dagger_{-\boldsymbol{k}\downarrow} + \Delta^\dagger c_{-\boldsymbol{k}\downarrow} c_{\boldsymbol{k}\uparrow} \right) + \Delta \sum_{\boldsymbol{k}} b^\dagger_{\boldsymbol{k}}. \tag{3.10}$$

Before proceeding with diagonalizing H, we should carry out the σ-sum in the first term

$$\sum_{\boldsymbol{k}\sigma}(\epsilon_{\boldsymbol{k}}-\mu)c^{\dagger}_{\boldsymbol{k}\sigma}c_{\boldsymbol{k}\sigma}=\sum_{\boldsymbol{k}}(\epsilon_{\boldsymbol{k}}-\mu)\Big(c^{\dagger}_{\boldsymbol{k}\uparrow}c_{\boldsymbol{k}\uparrow}-c_{-\boldsymbol{k}\downarrow}c^{\dagger}_{-\boldsymbol{k}\downarrow}+1\Big), \tag{3.11}$$

here we have used the anti-commutation relation and the fact that $\epsilon_{\boldsymbol{k}}=\epsilon_{-\boldsymbol{k}}$.

A transformation that diagonalizes the Hamiltonian for given $u_{\boldsymbol{k}}$ and $v_{\boldsymbol{k}}$ is[1]

$$\begin{aligned}\eta_{\boldsymbol{k}}&=u_{\boldsymbol{k}}c_{\boldsymbol{k}\uparrow}+v_{\boldsymbol{k}}c^{\dagger}_{-\boldsymbol{k}\downarrow}\\ \gamma_{\boldsymbol{k}}&=u_{\boldsymbol{k}}c^{\dagger}_{-\boldsymbol{k}\downarrow}-v_{\boldsymbol{k}}c_{\boldsymbol{k}\uparrow}.\end{aligned} \tag{3.12}$$

We wish to diagonalize H such that it can be written on a form similar to a free electron gas, i.e. using only $\eta^{\dagger}_{\boldsymbol{k}}\eta_{\boldsymbol{k}}$ and $\gamma^{\dagger}_{\boldsymbol{k}}\gamma_{\boldsymbol{k}}$. The anti-commutation relation for $\eta_{\boldsymbol{k}}$ (and similarly for $\gamma_{\boldsymbol{k}}$) is

$$\left\{\eta_{\boldsymbol{k}},\eta^{\dagger}_{\boldsymbol{k}'}\right\}=u_{\boldsymbol{k}}u_{\boldsymbol{k}'}\left\{c_{\boldsymbol{k}\uparrow},c^{\dagger}_{\boldsymbol{k}'\uparrow}\right\}+v_{\boldsymbol{k}}v_{\boldsymbol{k}'}\left\{c^{\dagger}_{-\boldsymbol{k}\downarrow},c_{-\boldsymbol{k}'\downarrow}\right\}=\delta_{\boldsymbol{k}\boldsymbol{k}'}. \tag{3.13}$$

For the above anti-commutation relation to be satisfied, the pre-factors $u_{\boldsymbol{k}}$ and $v_{\boldsymbol{k}}$ have to satisfy

$$u^2_{\boldsymbol{k}}+v^2_{\boldsymbol{k}}=1. \tag{3.14}$$

We can write the transformation on matrix form

$$\begin{pmatrix}\eta_{\boldsymbol{k}}\\ \gamma_{\boldsymbol{k}}\end{pmatrix}=\underbrace{\begin{pmatrix}u_{\boldsymbol{k}} & v_{\boldsymbol{k}}\\ -v_{\boldsymbol{k}} & u_{\boldsymbol{k}}\end{pmatrix}}_{\equiv M}\begin{pmatrix}c_{\boldsymbol{k}\uparrow}\\ c^{\dagger}_{-\boldsymbol{k}\downarrow}\end{pmatrix}. \tag{3.15}$$

The determinant of the above matrix is $\det(M)=u^2_{\boldsymbol{k}}+v^2_{\boldsymbol{k}}=1$. Which means the transformation is unitary. This is why we need the minus sign in front of $v_{\boldsymbol{k}}$ in $\gamma_{\boldsymbol{k}}$. Because M is unitary, its inverse is equal to its transpose $M^{-1}=M^T$. Using this we can easily invert Eq. (3.15)

$$\begin{pmatrix}c_{\boldsymbol{k}\uparrow}\\ c^{\dagger}_{-\boldsymbol{k}\downarrow}\end{pmatrix}=\begin{pmatrix}u_{\boldsymbol{k}} & -v_{\boldsymbol{k}}\\ v_{\boldsymbol{k}} & u_{\boldsymbol{k}}\end{pmatrix}\begin{pmatrix}\eta_{\boldsymbol{k}}\\ \gamma_{\boldsymbol{k}}\end{pmatrix} \tag{3.16}$$

and similarly for the conjugate

$$\begin{pmatrix}c^{\dagger}_{\boldsymbol{k}\uparrow}\\ c_{-\boldsymbol{k}\downarrow}\end{pmatrix}=\begin{pmatrix}u_{\boldsymbol{k}} & -v_{\boldsymbol{k}}\\ v_{\boldsymbol{k}} & u_{\boldsymbol{k}}\end{pmatrix}\begin{pmatrix}\eta^{\dagger}_{\boldsymbol{k}}\\ \gamma^{\dagger}_{\boldsymbol{k}}\end{pmatrix}. \tag{3.17}$$

Using these inverse expressions we can express the Hamiltonian with the new fermion operators $\eta_{\boldsymbol{k}}$ and $\gamma_{\boldsymbol{k}}$

$$\begin{aligned}H=\sum_{\boldsymbol{k}}\Big\{&(\xi_{\boldsymbol{k}}+\Delta b^{\dagger}_{\boldsymbol{k}})+\left[\xi_{\boldsymbol{k}}(u^2_{\boldsymbol{k}}-v^2_{\boldsymbol{k}})-u_{\boldsymbol{k}}v_{\boldsymbol{k}}(\Delta+\Delta^{\dagger})\right](\eta^{\dagger}_{\boldsymbol{k}}\eta_{\boldsymbol{k}}-\gamma^{\dagger}_{\boldsymbol{k}}\gamma_{\boldsymbol{k}})\\ &+\left[-2\xi_{\boldsymbol{k}}u_{\boldsymbol{k}}v_{\boldsymbol{k}}+\Delta v^2_{\boldsymbol{k}}-\Delta^{\dagger}u^2_{\boldsymbol{k}}\right]\gamma^{\dagger}_{\boldsymbol{k}}\eta_{\boldsymbol{k}}+\left[-2\xi_{\boldsymbol{k}}u_{\boldsymbol{k}}v_{\boldsymbol{k}}+\Delta^{\dagger}v^2_{\boldsymbol{k}}-\Delta u^2_{\boldsymbol{k}}\right]\eta^{\dagger}_{\boldsymbol{k}}\gamma_{\boldsymbol{k}}\Big\}.\end{aligned} \tag{3.18}$$

Above, we have set $\epsilon_{\boldsymbol{k}}-\mu\equiv\xi_{\boldsymbol{k}}$ and collected all combinations of $\eta_{\boldsymbol{k}}$ and $\gamma_{\boldsymbol{k}}$. For the transformation to be diagonal, i.e. only $\eta^{\dagger}_{\boldsymbol{k}}\eta_{\boldsymbol{k}}$ and $\gamma^{\dagger}_{\boldsymbol{k}}\gamma_{\boldsymbol{k}}$ terms in H, the pre-factors in front of $\gamma^{\dagger}_{\boldsymbol{k}}\eta_{\boldsymbol{k}}$ and $\eta^{\dagger}_{\boldsymbol{k}}\gamma_{\boldsymbol{k}}$ have to be zero. This means we need to find expressions for $u_{\boldsymbol{k}}$ and $v_{\boldsymbol{k}}$ that satisfies

$$-4\xi_{\boldsymbol{k}}u_{\boldsymbol{k}}v_{\boldsymbol{k}}=(u^2_{\boldsymbol{k}}-v^2_{\boldsymbol{k}})(\Delta^{\dagger}+\Delta). \tag{3.19}$$

The above equation is simply the sum of the two pre-factors set equal to zero. Considering the relation from Eq. (3.14), we are free to set $u_{\boldsymbol{k}}=\sin\theta$ and $v_{\boldsymbol{k}}=\cos\theta$. Additionally, we can assume Δ to be real without any loss of generality. These changes give us

[1] The minus sign in front of $v_{\boldsymbol{k}}$ in $\gamma_{\boldsymbol{k}}$ makes the transformation unitary.

$$-2\xi_{\boldsymbol{k}}\sin\theta\cos\theta = (\sin^2\theta - \cos^2\theta)\Delta. \tag{3.20}$$

Using the trigonometric identities $\sin 2\theta = 2\sin\theta\cos\theta$ and $\cos 2\theta = \sin^2\theta - \cos^2\theta$, we arrive at the equation which determines θ, and thus also $u_{\boldsymbol{k}}$ and $v_{\boldsymbol{k}}$,

$$\tan 2\theta = -\frac{\Delta}{\xi_{\boldsymbol{k}}}. \tag{3.21}$$

This equation sets the pre-factors of $\eta^\dagger_{\boldsymbol{k}}\gamma_{\boldsymbol{k}}$ and $\gamma^\dagger_{\boldsymbol{k}}\eta_{\boldsymbol{k}}$ to zero. If we assume $\Delta \geq 0$, we know $\tan 2\theta > 0$ if $\xi_{\boldsymbol{k}} < 0$ and $\tan 2\theta < 0$ if $\xi_{\boldsymbol{k}} > 0$. Moreover, if $\sin 2\theta > 0$, i.e. $2\theta \in [0, \pi)$, we can use

$$\tan^2 2\theta = \frac{\sin^2 2\theta}{\cos^2 2\theta} = b^2 \tag{3.22}$$

to write

$$\cos 2\theta = -\frac{\mathrm{sgn}(\xi_{\boldsymbol{k}})}{\sqrt{1+b^2}}, \tag{3.23}$$

where $b \equiv \Delta/\xi_{\boldsymbol{k}}$. Our next step is to use this result to determine the expression inside the square brackets in front of $\eta^\dagger_{\boldsymbol{k}}\eta_{\boldsymbol{k}}$ in Eq. (3.18). The expression inside the square brackets reads

$$\begin{aligned}
\xi_{\boldsymbol{k}}(u^2_{\boldsymbol{k}} - v^2_{\boldsymbol{k}}) - u_{\boldsymbol{k}}v_{\boldsymbol{k}}(\Delta + \Delta^\dagger) &= \xi_{\boldsymbol{k}}(\sin^2\theta - \cos^2\theta) - 2\Delta\sin\theta\cos\theta \\
&= \xi_{\boldsymbol{k}}\cos 2\theta - \Delta\sin 2\theta \\
&= \xi_{\boldsymbol{k}}\cos 2\theta - \Delta|b||\cos 2\theta| \\
&= -\frac{1}{\sqrt{1+b^2}}\left(\xi_{\boldsymbol{k}}\mathrm{sgn}(\xi_{\boldsymbol{k}}) + \frac{\Delta^2}{|\xi_{\boldsymbol{k}}|}\right) \\
&= -\frac{1}{\sqrt{\xi^2_{\boldsymbol{k}} + \Delta^2}}\left(\xi^2_{\boldsymbol{k}} + \Delta^2\right) \\
&= -\sqrt{\xi^2_{\boldsymbol{k}} + \Delta^2}
\end{aligned} \tag{3.24}$$

Inserting this result back into H, we get the following diagonal Hamiltonian

$$H = H_0 + \sum_{\boldsymbol{k}} E_{\boldsymbol{k}}\left(-\eta^\dagger_{\boldsymbol{k}}\eta_{\boldsymbol{k}} + \gamma^\dagger_{\boldsymbol{k}}\gamma_{\boldsymbol{k}}\right) \tag{3.25}$$

where $H_0 \equiv \sum_{\boldsymbol{k}}(\xi_{\boldsymbol{k}} + \Delta b^\dagger_{\boldsymbol{k}})$ and $E_{\boldsymbol{k}} \equiv \sqrt{\xi^2_{\boldsymbol{k}} + \Delta^2}$. This Hamiltonian has the same form as a free electron gas with energies $\pm E_{\boldsymbol{k}}$.

3.2.3 The BCS gap equation

Our goal now is to derive the BCS gap equation using H from Eq. (3.25) and $b_{\boldsymbol{k}} \equiv \langle c_{-\boldsymbol{k}\downarrow}c_{\boldsymbol{k}\uparrow}\rangle$. We re-write $b_{\boldsymbol{k}}$ using $\eta_{\boldsymbol{k}}$ and $\gamma_{\boldsymbol{k}}$

$$b_{\boldsymbol{k}} = \langle(v_{\boldsymbol{k}}\eta^\dagger_{\boldsymbol{k}} + u_{\boldsymbol{k}}\gamma^\dagger_{\boldsymbol{k}})(u_{\boldsymbol{k}}\eta_{\boldsymbol{k}} - v_{\boldsymbol{k}}\gamma_{\boldsymbol{k}})\rangle. \tag{3.26}$$

The only terms which contribute are $\langle\eta^\dagger_{k}\eta_{k}\rangle$ and $\langle\gamma^\dagger_{k}\gamma_{k}\rangle$

$$b_{\boldsymbol{k}} = u_{\boldsymbol{k}}v_{\boldsymbol{k}}\left(\langle\eta^\dagger_{k}\eta_{k}\rangle - \langle\gamma^\dagger_{k}\gamma_{k}\rangle\right). \tag{3.27}$$

The mean values in the above equation are simply Fermi-Dirac distributions with energy $-E_{\boldsymbol{k}}$ and $+E_{\boldsymbol{k}}$, respectively, given by H

$$\begin{aligned} b_{\boldsymbol{k}} &= u_{\boldsymbol{k}} v_{\boldsymbol{k}} \left(\frac{1}{\mathrm{e}^{-E_{\boldsymbol{k}}\beta} + 1} - \frac{1}{\mathrm{e}^{E_{\boldsymbol{k}}\beta} + 1} \right) \\ &= u_{\boldsymbol{k}} v_{\boldsymbol{k}} \tanh\left(\beta E_{\boldsymbol{k}}/2\right) \\ &= \frac{\Delta}{2E_{\boldsymbol{k}}} \tanh\left(\beta E_{\boldsymbol{k}}/2\right), \end{aligned} \tag{3.28}$$

where we have used that $2u_{\boldsymbol{k}}v_{\boldsymbol{k}} = \sin 2\theta = b/\sqrt{1+b^2}$. Using the fact that $\Delta \equiv -\frac{U}{2N}\sum_{\boldsymbol{k}} b_{\boldsymbol{k}}$, we arrive at the general form of the gap equation

$$1 = -\frac{U}{2N} \sum_{\boldsymbol{k}} \frac{\tanh(\beta E_{\boldsymbol{k}}/2)}{2E_{\boldsymbol{k}}}. \tag{3.29}$$

Equation (3.29), as it stands, is not possible to solve for Δ analytically. However, we can study the gap equation at temperature limits that simplifies it. We will consider two such limits: $T \to 0$ and $T \to T_c$.

3.2.4 The gap equation at $T \to 0$

The gap equation near $T = 0$ simplifies considerably,

$$1 = -\underbrace{\frac{U}{2N}}_{\equiv V} \sum_{\boldsymbol{k}} \frac{\tanh\left(\beta E_{\boldsymbol{k}}/2\right)}{2E_{\boldsymbol{k}}}\bigg|_{T\to 0} = \frac{V}{2} \sum_{\boldsymbol{k}} \frac{1}{\sqrt{(\epsilon_{\boldsymbol{k}} - \mu)^2 + \Delta^2}}. \tag{3.30}$$

Next, if we let

$$\sum_{\boldsymbol{k}} \to \int D(\epsilon)\mathrm{d}\epsilon \approx D(\epsilon_{\mathrm{F}}) \int \mathrm{d}\epsilon, \tag{3.31}$$

where $D(\epsilon_{\mathrm{F}})$ is the density of states (DOS) at the Fermi level[2], the gap equation simplifies further

$$1 = \overbrace{\frac{VD(\epsilon_{\mathrm{F}})}{2}}^{\equiv\lambda} \int_{-\omega_D}^{\omega_D} \frac{\mathrm{d}\epsilon}{\sqrt{\epsilon^2 + \Delta^2}} = \lambda \int_0^{\omega_D} \frac{\mathrm{d}\epsilon}{\sqrt{\epsilon^2 + \Delta^2}}, \tag{3.32}$$

where ω_D is a cut-off energy called the Debye frequency [40]. We introduce this cut-off energy because we wish to only consider energies close to the Fermi level. If we now make the substitution $x = \epsilon/\Delta$, the integral becomes easily solvable

$$\frac{1}{\lambda} = \int_0^{\omega_D/\Delta} \frac{\mathrm{d}x}{\sqrt{1+x^2}} = \sinh^{-1}\left(\frac{\omega_D}{\Delta}\right). \tag{3.33}$$

An important assumption of BCS theory is that $\lambda \equiv VD(\epsilon_{\mathrm{F}}) \ll 1$. Using this, the expression for Δ simplifies

$$\Delta \approx 2\omega_D \mathrm{e}^{-1/\lambda}. \tag{3.34}$$

We have now found an approximate solution to the gap equation at $T = 0$.

[2]The energy regime we are working in is within a thin shell around the Fermi level. Therefore, if we assume that the DOS varies slowly near the Fermi level, we may approximate $D(\epsilon)$ with $D(\epsilon_{\mathrm{F}})$.

3.2.5 The gap equation at $T \to T_c$

As we heat up a superconductor past its critical temperature T_c, the superconducting gap Δ vanishes because the superconductor transitions to a normal metal, i.e. the Cooper pairs break. Thus, at T_c the energy eigenvalue changes to $E_{\boldsymbol{k}} = |\epsilon_{\boldsymbol{k}} - \mu|$.

We will now find an expression for T_c by evaluating the gap equation at $T = T_c$. The calculations will follow [27]. At T_c, the gap equation has the form

$$1 = V \sum_{\boldsymbol{k}} \frac{\tanh(\beta|\epsilon_{\boldsymbol{k}} - \mu|/2)}{2|\epsilon_{\boldsymbol{k}} - \mu|}, \tag{3.35}$$

where $\beta = k_{\mathrm{B}} T_c$. Converting the sum to an integral, as in Eq. (3.31), gives

$$1 = V D(\epsilon_{\mathrm{F}}) \int_0^{\omega_D} \mathrm{d}\epsilon \; \frac{\tanh(\beta\epsilon/2)}{\epsilon}. \tag{3.36}$$

Substituting $x = \beta\epsilon/2$ and setting $\lambda = V D(\epsilon_{\mathrm{F}})$, we solve the integral by parts

$$\frac{1}{\lambda} = \int_0^{\beta\omega_D/2} \mathrm{d}x \; \frac{\tanh x}{x} = \tanh x \ln x \Big|_0^{\beta\omega_D/2} - \int_0^{\beta\omega_D/2} \mathrm{d}x \; \frac{\ln x}{\cosh^2 x}. \tag{3.37}$$

The first term is easily evaluated because the $\tanh x$ goes to 0 faster than the $\ln x$ goes to $-\infty$ at $x = 0$. The integral is a little trickier. Luckily though, the integrand converges quite fast and because typically $k_{\mathrm{B}} T_c \ll \omega_D$, we can safely set the upper limit of the integral to ∞ and still get a well approximated result. The resulting integral can be looked up, and we denote it $\ln C$ where $C = \pi \mathrm{e}^{-\gamma}/4$ and $\gamma \approx 0.5772...$ is Euler's constant. We are left with

$$\frac{1}{\lambda} = \tanh(\beta\omega_D/2) \ln(\beta\omega_D/2) - \ln C \approx \ln(\beta\omega_D/2C), \tag{3.38}$$

where we have approximated $\tanh(\beta\omega_D/2) \approx 1$ because $k_{\mathrm{B}} T_c \ll \omega_D$. The resulting expression for the critical temperature is

$$k_{\mathrm{B}} T_c \approx \frac{\omega_D}{2C} \mathrm{e}^{-1/\lambda}. \tag{3.39}$$

This is the temperature at which superconductivity occurs in BCS theory.

3.2.6 A universal number

Something worth noting is that both Eqs. (3.34) and (3.39) have the same dependency on λ. A consequence of this is the emergence of a surprising constant

$$\frac{\Delta(T = 0)}{k_{\mathrm{B}} T_c} = 4C \approx 1.76. \tag{3.40}$$

The ratio is a dimensionless, universal number completely independent of the physical properties of the system. In other words, this number should be constant for all superconductors. However, as one would expect, this only applies to weakly coupled superconductors. In fact, this ratio is often used as reference when comparing model accuracy between coupling regimes. Aluminium, for instance, has a ratio of approximately 1.77 with mercury at 2.30 [32]. Because $\lambda_{\mathrm{Hg}} > 1 > \lambda_{\mathrm{Al}}$, this discrepancy shows that BCS is inadequate for large λ.

3.3 Eliashberg theory

So far we have studied the Haldane-Hubbard model in the weak-coupling regime using the BCS theory of superconductivity. Aware of the limitations of this theory, we have still managed to show that superconductivity is theoretically possible in the graphene model. However, the strong-coupling regime is still of high interest. We stated earlier that many interesting phenomena in condensed matter theory exist due to electron correlation. The same is true for strongly correlated electrons. For example, it has been suggested that strongly interacting TIs could lose their edge states while still remaining insulating in the bulk [24]. Normally a TI has gapped states in the bulk with conducting edge states. However, if we introduce strongly coupled electrons, the system could lose the edge states while still being in a topologically non-trivial state. Even though the topic of this thesis is strictly focused on the bulk, studies like the one above is great motivation for studying behavior of topological states of matter at strong-coupling. This is what we will do next using the Eliashberg theory of superconductivity.

The diagonalized Hamiltonian is the same as Eq. (3.2)

$$H = \sum_{\boldsymbol{k}\sigma} (\epsilon_{\boldsymbol{k}} - \mu) c^\dagger_{\boldsymbol{k}\sigma} c_{\boldsymbol{k}\sigma} + \frac{U}{4N} \sum_{\boldsymbol{k}\boldsymbol{k}'\boldsymbol{q},\sigma} c^\dagger_{\boldsymbol{k}+\boldsymbol{q}\sigma} c^\dagger_{\boldsymbol{k}'-\boldsymbol{q},-\sigma} c_{\boldsymbol{k}',-\sigma} c_{\boldsymbol{k}\sigma}. \tag{3.41}$$

One could argue that our choice of interaction, on-site Hubbard, does not do Eliashberg theory justice in some sense because it is constant and frequency independent like Coulomb. Given that Eliashberg theory focuses on the frequency dependence of phonons, this is a reasonable statement. It seems plausible that simplifying the interaction to this extent could prove unproductive. However, our goal is not to study the frequency dependency of the Haldane-Hubbard model, but rather the consequences of including the electron self-energy S at strong-coupling. The Hubbard interaction is simply phenomenological, as it generalizes the combined effects of the Coulomb and electron-phonon forces. Usually one uses an effective description such as the Hubbard interaction to represent weakly coupled electron-phonon effects. The natural question is then: why can we do the same for strong-coupling? It has been shown that electrons hopping in the vicinity of phonons on a honeycomb lattice can exhibit coupling strengths λ of up to 0.6 [41]. Hence, studying the effect of S on the gap equation at strong-coupling using the Haldane-Hubbard model is an interesting topic of discussion.

3.3.1 Green's functions and the equation of motion

Traditionally, the formalism of BCS theory has taken a more straight forward approach combining statistical and quantum mechanics. Eliashberg theory, on the other hand, was presented using a Green's function [33] formalism. More specifically the Matsubara or imaginary-time Green's function formalism. Not only is this formalism handy but also somewhat necessary. The big difference between BCS and Eliashberg is that Eliashberg theory takes retardation effects and the electron self-energy into account. Since both retardation effects and self-energies are more easily captured in the Green's function formalism, it is the natural choice[3].

In Eliashberg theory, we define three single-particle Green's functions

$$G(\boldsymbol{k}, \tau) \equiv -\langle T_\tau c_{\boldsymbol{k}\sigma}(\tau) c^\dagger_{\boldsymbol{k}\sigma}(0) \rangle \tag{3.42}$$

$$F(\boldsymbol{k}, \tau) \equiv -\langle T_\tau c_{\boldsymbol{k}\uparrow}(\tau) c_{-\boldsymbol{k}\downarrow}(0) \rangle \tag{3.43}$$

$$F^\dagger(\boldsymbol{k}, \tau) \equiv -\langle T_\tau c^\dagger_{-\boldsymbol{k}\downarrow}(\tau) c^\dagger_{\boldsymbol{k}\uparrow}(0) \rangle. \tag{3.44}$$

[3] BCS theory was later described by the same formalism.

G is a normal Green's function while F and $F^\dagger$ are anomalous Green's functions[4], see Fig. 2.4 for their diagrammatic interpretation. The imaginary-time dependent operators above are defined in the Heisenberg picture

$$\begin{aligned} c_\nu(\tau) &\equiv \mathrm{e}^{\tau H} c_\nu \mathrm{e}^{-\tau H} \\ c_\nu^\dagger(\tau) &\equiv \mathrm{e}^{\tau H} c_\nu^\dagger \mathrm{e}^{-\tau H}. \end{aligned} \tag{3.45}$$

Our goal is similar to what we did in the previous section: we want to find the equations whose solution is the superconducting gap and a way to find the critical temperature of the system. The information needed to achieve this goal is all encoded in the following self-energies [34]

$$S(\boldsymbol{p}, i\omega_n) = -\frac{1}{\beta}\sum_{\boldsymbol{q},m} V_{\text{eff}}(\boldsymbol{q}, i\omega_m) G(\boldsymbol{p}+\boldsymbol{q}, i\omega_n + i\omega_m) \tag{3.46}$$

$$W(\boldsymbol{p}, i\omega_n) = -\frac{1}{\beta}\sum_{\boldsymbol{q},m} V_{\text{eff}}(\boldsymbol{q}, i\omega_m) F(\boldsymbol{p}+\boldsymbol{q}, i\omega_n + i\omega_m), \tag{3.47}$$

where $\omega_n \equiv (2n+1)\pi/\beta$ are called the Matsubara frequencies. We call S the electron self-energy (neglected in BCS theory) and W the superconducting gap (Δ in BCS theory). See Fig. 2.5 for a diagram of S and W. To be able to use these self-energies we first need to find expressions for G, F and $F^\dagger$ self-consistently. One way to do this is to use the equation of motion approach where we evaluate the time-derivative of the Green's functions to find a self-consistent expression [42]. The detailed derivation of the Eliashberg equations is present in Appendix D. However, we will summarize the results in the following starting by finding the equation of motion for G

$$\frac{\partial}{\partial\tau} G_\sigma(\boldsymbol{k}, \tau) = -\delta(\tau) - \left\langle T_\tau \left[H, c_{\boldsymbol{k}\sigma}(\tau)\right] c_{\boldsymbol{k}\sigma}^\dagger(0)\right\rangle. \tag{3.48}$$

Note that we can keep the spin index σ above to differentiate between the two spin states in case of spin-dependent forces. However, going forward we will assume no spin-dependent forces and thus drop the index. The commutator $[H, c_{\boldsymbol{k}\sigma}(\tau)]$ is straight forward and gives

$$\left(\frac{\partial}{\partial\tau} + \xi_{\boldsymbol{k}}\right) G(\boldsymbol{k}, \tau) = -\delta(\tau) + \frac{U}{2N}\sum_{\boldsymbol{pq}} \langle T_\tau c_{\boldsymbol{p}+\boldsymbol{q}\downarrow}^\dagger(\tau) c_{\boldsymbol{p}\downarrow}(\tau) c_{\boldsymbol{k}+\boldsymbol{q}\uparrow}(\tau) c_{\boldsymbol{k}\uparrow}^\dagger(0)\rangle \tag{3.49}$$

for spin up, where $\xi_{\boldsymbol{k}} \equiv \epsilon_{\boldsymbol{k}} - \mu$. To make sense of this equation we have to simplify the sum on the RHS. The summand is a statistical mean of a time-ordered product of four fermion operators, i.e. like a two-particle Green's function. However, because we would like to express our theory by only single-particle Green's functions like the ones in Eqs. (3.42) to (3.44), we need to decouple the time-ordered product of four operators into two time-ordered products of two operators. This is of course only an approximation, much like the mean-field approximation in BCS theory of Section 3.2.1. When decoupling our time-ordered mean, we have a few choices: we can decouple into $\langle cc\rangle$, $\langle c^\dagger c^\dagger\rangle$, and $\langle cc^\dagger\rangle$. These choices are highlighted because each corresponds to F, $F^\dagger$, and G respectively[5]. Hence, we decouple the summand in Eq. (3.49) into single-particle Green's functions by a method analogous to Wick's theorem[6]

$$\langle T_\tau c_{\boldsymbol{p}+\boldsymbol{q}\downarrow}^\dagger(\tau) c_{\boldsymbol{p}\downarrow}(\tau) c_{\boldsymbol{k}+\boldsymbol{q}\uparrow}(\tau) c_{\boldsymbol{k}\uparrow}^\dagger(0)\rangle \to -\delta_{\boldsymbol{q}0} G(\boldsymbol{p}, 0) G(\boldsymbol{k}, \tau) - \delta_{\boldsymbol{p},-\boldsymbol{k}-\boldsymbol{q}} F(\boldsymbol{k}+\boldsymbol{q}, 0) F^\dagger(\boldsymbol{k}, \tau). \tag{3.50}$$

This allows us to rewrite Eq. (3.49) as

[4] They are anomalous because they do not conserve particle number.

[5] Keeping the correct spin and momentum combinations in mind.

[6] Normal order the fermion operators and change sign for each permutation of the operators. For more details on this method see Appendix D and chapter 51 of [33] or 10.1.1 of [34].

$$\left(\frac{\partial}{\partial\tau}+\xi_{\boldsymbol{k}}\right)G(\boldsymbol{k},\tau)=-\delta(\tau)-\frac{U}{2N}\sum_{\boldsymbol{q}}\left[G(\boldsymbol{q},0)G(\boldsymbol{k},\tau)+F(\boldsymbol{q},0)F^{\dagger}(\boldsymbol{k},\tau)\right], \tag{3.51}$$

where we have changed the summation momentum $\boldsymbol{q}\to\boldsymbol{q}-\boldsymbol{k}$ in the F–term. This kind of decoupling is analogous to the mean-field approximation we did for BCS theory. Unlike the BCS theory, however, the G–term in Eq. (3.51) is present. This term describes the exchange self-energy of the electron due to the phonon-induced electron interaction. In weakly coupled superconductors, this term causes only a weak change in the electron's effective mass, resulting in it being ignored. The F–term is present in both theories because this term is what makes the superconducting state special.

Now that we have decoupled the equation of motion for G, we do the same for F and $F^{\dagger}$

$$\left(\frac{\partial}{\partial\tau}+\xi_{\boldsymbol{k}}\right)F(\boldsymbol{k},\tau)=\frac{U}{2N}\sum_{\boldsymbol{q}}\Big[-G(\boldsymbol{q},0)F(\boldsymbol{k},\tau)+F(\boldsymbol{q},0)G(-\boldsymbol{k},-\tau)\Big] \tag{3.52}$$

$$\left(\frac{\partial}{\partial\tau}-\xi_{\boldsymbol{k}}\right)F^{\dagger}(\boldsymbol{k},\tau)=\frac{U}{2N}\sum_{\boldsymbol{q}}\Big[G(\boldsymbol{q},0)F^{\dagger}(\boldsymbol{k},\tau)-F^{\dagger}(\boldsymbol{q},0)G(\boldsymbol{k},\tau)\Big]. \tag{3.53}$$

The next step is to Fourier transform Eqs. (3.51) to (3.53) to the frequency domain $i\omega_n$. Using the definition of the transform

$$G(\boldsymbol{k},\tau)=\frac{1}{\beta}\sum_{n\in\mathbb{Z}}\mathrm{e}^{-i\omega_n\tau}G(\boldsymbol{k},i\omega_n) \tag{3.54}$$

$$G(\boldsymbol{k},i\omega_n)=\int_0^{\beta}\mathrm{d}\tau\ \mathrm{e}^{i\omega_n\tau}G(\boldsymbol{k},\tau), \tag{3.55}$$

we get

$$(-i\omega_n+\xi_{\boldsymbol{k}})\,G(\boldsymbol{k},i\omega_n)=-1-\frac{1}{\beta}\frac{U}{2N}\sum_{\boldsymbol{q},m}\left[G(\boldsymbol{q},i\omega_m)G(\boldsymbol{k},i\omega_n)+F(\boldsymbol{q},i\omega_m)F^{\dagger}(\boldsymbol{k},i\omega_n)\right] \tag{3.56}$$

$$(-i\omega_n+\xi_{\boldsymbol{k}})\,F(\boldsymbol{k},i\omega_n)=\frac{1}{\beta}\frac{U}{2N}\sum_{\boldsymbol{q},m}\Big[-G(\boldsymbol{q},i\omega_m)F(\boldsymbol{k},i\omega_n)+F(\boldsymbol{q},i\omega_m)G(-\boldsymbol{k},-i\omega_n)\Big] \tag{3.57}$$

$$(-i\omega_n-\xi_{\boldsymbol{k}})\,F^{\dagger}(\boldsymbol{k},i\omega_n)=\frac{1}{\beta}\frac{U}{2N}\sum_{\boldsymbol{q},m}\Big[G(\boldsymbol{q},i\omega_m)F^{\dagger}(\boldsymbol{k},i\omega_n)-F^{\dagger}(\boldsymbol{q},i\omega_m)G(\boldsymbol{k},i\omega_n)\Big]. \tag{3.58}$$

Looking at the sums of the above equations, we notice the self-energies from Eqs. (3.46) and (3.47)

$$S=-\frac{1}{\beta}\frac{U}{2N}\sum_{\boldsymbol{q},m}G(\boldsymbol{q},i\omega_m) \tag{3.59}$$

$$W=-\frac{1}{\beta}\frac{U}{2N}\sum_{\boldsymbol{q},m}F(\boldsymbol{q},i\omega_m) \tag{3.60}$$

$$W^{\dagger}=-\frac{1}{\beta}\frac{U}{2N}\sum_{\boldsymbol{q},m}F^{\dagger}(\boldsymbol{q},i\omega_m) \tag{3.61}$$

with the distinct difference from Eqs. (3.46) and (3.47) being that $V_{\text{eff}} = U/2N$ and that our self-energy expressions are independent of both momentum and frequency. The latter remark is not surprising given that our Hubbard interaction strength is constant. We can re-write Eqs. (3.56) to (3.58) using S and W

$$(i\omega_n - \xi_{\boldsymbol{k}})\, G(\boldsymbol{k}, i\omega_n) = 1 - SG(\boldsymbol{k}, i\omega_n) - WF^\dagger(\boldsymbol{k}, i\omega_n) \tag{3.62}$$

$$(i\omega_n - \xi_{\boldsymbol{k}})\, F(\boldsymbol{k}, i\omega_n) = WG(-\boldsymbol{k}, -i\omega_n) - SF(\boldsymbol{k}, i\omega_n) \tag{3.63}$$

$$(-i\omega_n - \xi_{\boldsymbol{k}})\, F^\dagger(\boldsymbol{k}, i\omega_n) = W^\dagger G(\boldsymbol{k}, i\omega_n) - SF^\dagger(\boldsymbol{k}, i\omega_n). \tag{3.64}$$

Typically one writes the bare propagator as $G^{(0)}(\boldsymbol{k}, i\omega_n) = [i\omega_n - \xi_{\boldsymbol{k}}]^{-1}$. Using this we can see the resemblance between Eqs. (3.62) to (3.64) and the Dyson equations in Eq. (2.21).

Solving for G and F gives

$$G(\boldsymbol{k}, i\omega_n) = \frac{i\omega_n + \xi_{\boldsymbol{k}} - S}{(i\omega_n)^2 - (\xi_{\boldsymbol{k}} - S)^2 - |W|^2} \tag{3.65}$$

$$F(\boldsymbol{k}, i\omega_n) = \frac{-W}{(i\omega_n)^2 - (\xi_{\boldsymbol{k}} - S)^2 - |W|^2}. \tag{3.66}$$

We have now found expressions for G and F self-consistently and can insert them back into the self-energy equations to get the so-called Eliashberg equations

$$S = -\frac{1}{\beta}\frac{U}{2N}\sum_{\boldsymbol{k},n} \frac{i\omega_n + \xi_{\boldsymbol{k}} - S}{(i\omega_n)^2 - (\xi_{\boldsymbol{k}} - S)^2 - |W|^2} \tag{3.67}$$

$$W = -\frac{1}{\beta}\frac{U}{2N}\sum_{\boldsymbol{k},n} \frac{-W}{(i\omega_n)^2 - (\xi_{\boldsymbol{k}} - S)^2 - |W|^2}. \tag{3.68}$$

This set of equations can be solved for S and W, and even for T_c ($W \to 0$ in this limit). The only thing stopping us is the sums over frequency and momentum. We can simplify both equations by first carrying out the sum over frequency using a known method for evaluating so-called Matsubara sums, followed by evaluating the momentum sum by transforming it to an energy integral for two temperature limits.

3.3.2 Frequency and momentum sums

The method used to evaluate the frequency sum is described in detail in Appendix D.2 and we only state the results here

$$S = -\frac{U}{4} + \frac{U}{2N}\sum_{\boldsymbol{k}}(\xi_{\boldsymbol{k}} - S)\frac{\tanh\left(\frac{\beta}{2}\sqrt{(\xi_{\boldsymbol{k}} - S)^2 + |W|^2}\right)}{2\sqrt{(\xi_{\boldsymbol{k}} - S)^2 + |W|^2}} \tag{3.69}$$

$$1 = -\frac{U}{2N}\sum_{\boldsymbol{k}} \frac{\tanh\left(\frac{\beta}{2}\sqrt{(\xi_{\boldsymbol{k}} - S)^2 + |W|^2}\right)}{2\sqrt{(\xi_{\boldsymbol{k}} - S)^2 + |W|^2}}. \tag{3.70}$$

Equation (3.69) does not have a counterpart in BCS theory. The existence of this equation means that we are accounting for the electron's self-energy. A physical consequence of this is that the electron's effective mass is no longer the same as in BCS theory. Equation (3.70), on the other

hand, does have a counterpart, namely the gap equation from Eq. (3.29). There are only two differences between Eqs. (3.29) and (3.70): the existence of S and that the gap is called Δ in BCS and W in Eliashberg. In other words, if we set $S = 0$, Eqs. (3.29) and (3.70) are completely equivalent. Like in BCS theory, the momentum sum in Eq. (3.70) can not be analytically evaluated for all temperatures. Which is why we would need to study this equation in the two temperature limits $T \to 0$ and $T \to T_c$ like in BCS theory. To do so, we will convert the momentum sum to an energy integral:

$$\sum_{\boldsymbol{k}} \to \int \mathrm{d}\epsilon \; D(\epsilon) \approx D(\epsilon_{\mathrm{F}}) \int_{-\omega_D}^{\omega_D} \mathrm{d}\epsilon, \tag{3.71}$$

where ω_D is the Debye frequency (a cut-off energy) and $D(\epsilon_{\mathrm{F}})$ is the density of states at the Fermi level ϵ_{F}. This is all similar to the steps in BCS theory. The density of states near the Dirac point is derived in Appendix C and given as

$$D(\epsilon_{\mathrm{F}}) = \frac{2\pi N}{A_{\mathrm{BZ}}} \frac{|\epsilon_{\mathrm{F}}|}{\sqrt{b^2 - 4a(m^2 - \epsilon_{\mathrm{F}}^2)}} \Theta(|\epsilon_{\mathrm{F}}| - m), \tag{3.72}$$

see Appendix C for definition of symbols.

The sum in Eq. (3.69) can actually be evaluated analytically for all T by converting it to an energy integral. Converting both Eqs. (3.69) and (3.70) to energy integrals gives

$$S = -\frac{U}{4} - \frac{\lambda}{\beta} \ln \left[\frac{\cosh\left(\frac{\beta}{2}\sqrt{(\omega_D - S)^2 + |W|^2}\right)}{\cosh\left(\frac{\beta}{2}\sqrt{(\omega_D + S)^2 + |W|^2}\right)} \right] \tag{3.73}$$

$$1 = \lambda \int_{-\omega_D}^{\omega_D} \mathrm{d}\epsilon \; \frac{\tanh\left(\frac{\beta}{2}\sqrt{(\epsilon - S)^2 + |W|^2}\right)}{2\sqrt{(\epsilon - S)^2 + |W|^2}}, \tag{3.74}$$

where $\lambda \equiv -UD(\epsilon_{\mathrm{F}})/2N$. Note how similar Eq. (3.74) is to the BCS gap equation at Eq. (3.29). The only real difference is of course the existence of S and the additional equation which determines S. Its this shift in the electron energy which makes the Eliashberg treatment of the Haldane-Hubbard model interesting. We will study these equations in the two temperature limits $T \to 0$ and $T \to T_c$.

3.3.3 $T \to 0$

When $T \to 0$, the hyperbolic functions in Eqs. (3.73) and (3.74) go as $\cosh(\beta x) \to \exp(\beta x)/2$ and $\tanh(\beta x) \to 1$. With this in mind these equations become

$$S = \frac{N\lambda}{2D(\epsilon_{\mathrm{F}})} - \frac{\lambda}{2}\left(\sqrt{(\omega_D - S)^2 + |W|^2} - \sqrt{(\omega_D + S)^2 + |W|^2}\right) \tag{3.75}$$

and

$$1 = \frac{\lambda}{2}\left[\sinh^{-1}\left(\frac{\omega_D - S}{|W|}\right) + \sinh^{-1}\left(\frac{\omega_D + S}{|W|}\right)\right]. \tag{3.76}$$

These equations form a set of equations with S and W as the unknowns. They can be solved numerically as a function of e.g. λ.

Unsurprisingly, we see a likeness between the BCS gap from Eq. (3.34) and Eq. (3.76) above. The only difference is that $S = 0$ and $\lambda \ll 1$ in BCS theory, which is why $\Delta \sim \mathrm{e}^{-1/\lambda}$.

3.3.4 $T \to T_c$

When we approach the critical temperature of a superconductor from below, the superconducting gap approaches zero. This is because when we heat up a superconductor it stops being superconducting after surpassing T_c. Thus, $W \to 0$ in this limit. Setting $W = 0$ in Eqs. (3.73) and (3.74) and performing the energy integral gives

$$S = \frac{N\lambda}{2D(\epsilon_\mathrm{F})} - \frac{\lambda}{\beta} \ln \left[\frac{\cosh\left(\frac{\beta}{2}|\omega_D - S|\right)}{\cosh\left(\frac{\beta}{2}|\omega_D + S|\right)} \right] \tag{3.77}$$

and

$$\begin{aligned} \frac{2}{\lambda} = {} & \mathrm{sgn}(\omega_D - S) \tanh\left(\frac{\beta}{2}|\omega_D - S|\right) \ln\left(\frac{\beta}{2}|\omega_D - S|\right) + \\ & \tanh\left(\frac{\beta}{2}|\omega_D + S|\right) \ln\left(\frac{\beta}{2}|\omega_D + S|\right) - \\ & [\mathrm{sgn}(\omega_D - S) + 1] \ln C, \end{aligned} \tag{3.78}$$

where $C \approx 0.44$ and $\mathrm{sgn}(x)$ is the sign function. With S and $\beta \equiv 1/k_\mathrm{B}T_c$ as the unknowns, the above two equations form a set of equations numerically solvable as a function of e.g. λ.

Equation (3.78) is the Eliashberg counterpart to Eq. (3.38). Due to the inclusion of S, we need to consider the sign function. Like before, if we set $S = 0$ we get Eq. (3.38) because $\omega_D > 0$. Note that we have let $\tanh(\beta\omega_D/2) \to 1$ due to $\beta\omega_D \gg 1$ to arrive at the $k_\mathrm{B}T_c \sim \mathrm{e}^{-1/\lambda}$ expression in Eq. (3.39). In Eq. (3.78), however, we keep the factor to ensure numerical accuracy in cases where S is large enough to have an impact on the approximation.

3.4 Solving the Eliashberg equations

It is obvious that Eqs. (3.75) and (3.76), and Eqs. (3.77) and (3.78) are difficult to solve analytically for (S, W) and (S, T_c), respectively. Which is why we need to implement a numerical method of solving the non-linear sets of equations.

3.4.1 Numerical method

The method is straight-forward: starting with an initial guess for all unknowns, work iteratively toward a convergent solution approaching $f(\boldsymbol{x}) = g(\boldsymbol{x}) = 0$ by adjusting $\boldsymbol{x}$ for each iteration. Here, f and g are the two equations that make up the set either for $T = 0$ or $T = T_c$, and $\boldsymbol{x}$ are the guesses. The numerical tools we can use to achieve this is Python's root solver `scipy.optimize.root()`[7]. The function `root()` takes the following arguments (among others): the function to evaluate[8], an initial guess for each unknown, known parameters needed in the equations, and a specific root solving method. The default root solving method is MINPACK's hybrd and hybrj routines (modified Powell method), which is the one we will be using. The root solver returns an array of solutions (one element for each unknown), a success boolean and more.

We have used the following known parameters: $\omega_D = 1.7$, $\mu = -1$ and $t' = 0.9t'_c$, see Eq. (3.4) for the definition of t'_c, in addition to an array of λ–values ranging from 0 to 1.5. The motivation for this choice is simply trial and error. Depending on the equations we want to solve, we have also supplied the solver with initial guesses of S, W and T_c. In short, for each λ, we have solved the equation(s) given the list of known parameters and initial guesses, in the end producing $S(\lambda)$

[7] See documentation at docs.scipy.org

[8] If the equation set has multiple equations, the user must create a wrapper function that returns the output of e.g. both $f(\boldsymbol{x})$ and $g(\boldsymbol{x})$.

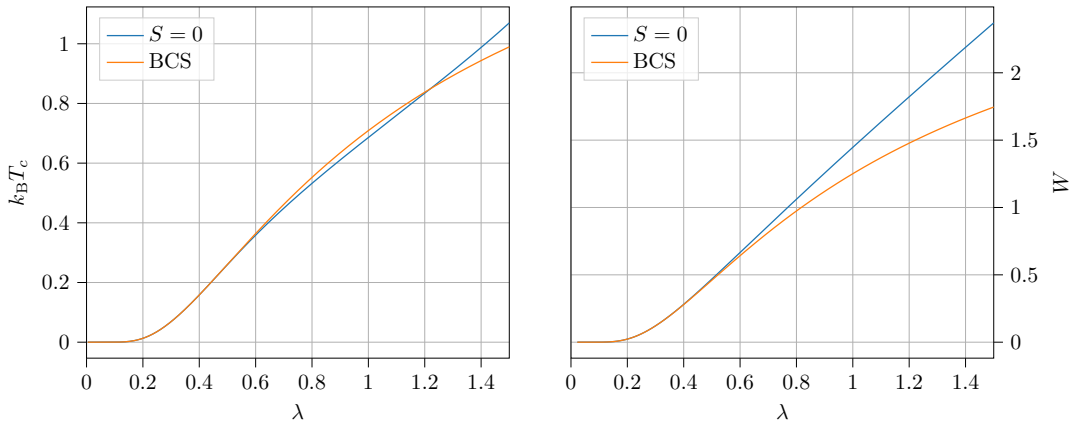


Figure 3.2: Plot of T_c (left panel) and W (right panel) using BCS and Eliashberg ($S = 0$) solutions. The left panel is the solution to Eqs. (3.39) and (3.78), while the right panel is the solution to Eqs. (3.34) and (3.76). We expect the two solutions to be the same. However, due to approximations made in the BCS case the two solutions deviate.

and $W(\lambda)$ or $T_c(\lambda)$. To get a decent resolution and range in λ, we depend on the set of equations to behave for our choice of parameter values.

3.4.2 Decoupled equations (fixed S)

The first step is to do some sanity checks to see if the method is able to solve a single equation at a time for a fixed S. We will thus solve Eqs. (3.76) and (3.78) (completely separately) for W and T_c, respectively, as a function of λ. Ignoring for the moment Eqs. (3.75) and (3.77) as they do not make sense when S is constant. One way of knowing if our solution is sensible is to compare it to the BCS results of Eqs. (3.34) and (3.39). This should be a decent test especially if we set $S = 0$ because then the equations obtained using Eliashberg and BCS theory should be exactly the same. Figure 3.2 shows the solution to Eqs. (3.76) and (3.78) for $S = 0$ plotted with the BCS solutions. We see clearly that the Eliashberg and BCS solutions are practically the same up until $\lambda \sim 0.6$. For $\lambda > 0.6$, however, the solutions start to grow further apart. At first glance this behavior is odd considering that both theories should produce the same results in the case of $S = 0$. However, the difference in solutions is due to an approximation we made in the BCS case. Recall that for $T = 0$, we assumed $\lambda \ll 1$ and thus approximated $\sinh(1/\lambda) \to \exp(1/\lambda)/2$, see Eq. (3.34). Moreover, for $T = T_c$, we assumed $\beta\omega_D \gg 1$ and let $\tanh(\beta\omega_D/2) \to 1$, see Eq. (3.38). If we had not made these approximations the solutions to the Eliashberg equations would be identical to BCS. Going forward we will not use the approximated BCS solutions when comparing results but rather treat the Eliashberg solutions for $S = 0$ as the BCS solutions.

Next we solve the Eliashberg equations for fixed but non-zero S to study how the solutions change compared to the BCS case ($S = 0$). Figure 3.3 shows a plot of both T_c and W for increasing values of S as a function of λ. We see T_c for $S > 0$ start off lower than $S = 0$ but always end up surpassing $S = 0$ for larger λ. Moreover, the larger S is, the sooner it surpasses $S = 0$ growing more rapidly. W on the other hand, does not grow more rapidly for larger S. In fact, as S increases, W grows slower than $S = 0$.

We see that as S approaches $\omega_D = 1.7$ in Fig. 3.3, T_c grows far more rapidly as a function of λ than when $S \ll \omega_D$. This increase can be understood by studying each term of Eq. (3.78)

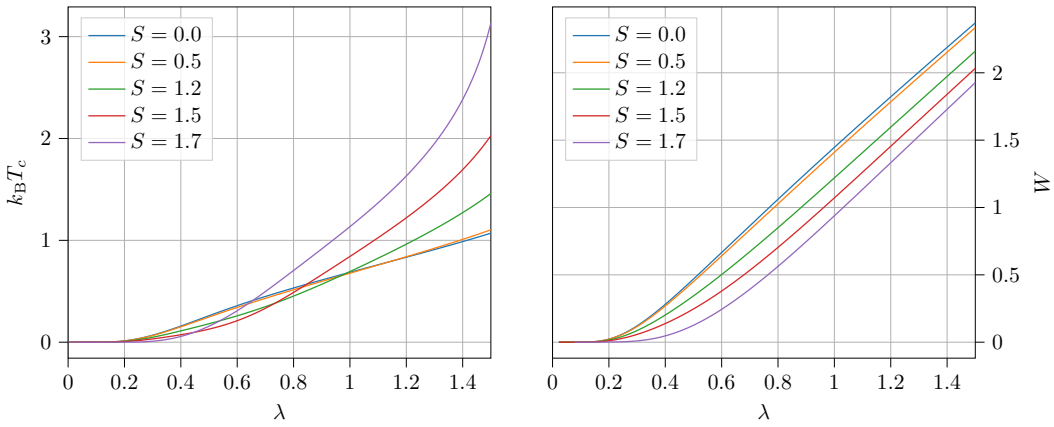


Figure 3.3: Plot of T_c (left panel) and W (right panel) for multiple values of fixed S. The left panel is the solution to Eq. (3.78) and the right panel is the solution to Eq. (3.76). The figure shows the S dependence of these equations.

$$\frac{2}{\lambda} = \operatorname{sgn}(\omega_D - S)\tanh\left(\frac{\beta}{2}|\omega_D - S|\right)\ln\left(\frac{\beta}{2}|\omega_D - S|\right) + \tanh\left(\frac{\beta}{2}|\omega_D + S|\right)\ln\left(\frac{\beta}{2}|\omega_D + S|\right) - [\operatorname{sgn}(\omega_D - S) + 1]\ln C.$$

When $S \to \omega_D$ for large λ, the first term on the RHS falls to zero due to the $\tanh(\beta|\omega_D - S|/2)$ factor. Meanwhile, the second term keeps growing due to the logarithm. With the LHS decreasing due to large λ and the RHS growing because of the logarithm, the only way the equation can have a solution is if the RHS's growth is decreased. This is achieved by decreasing β in the argument of the logarithm, i.e. increasing T_c.

3.4.3 Coupled equations

Now that we have studied the effect of fixed S on the BCS equations, we should try to solve the actual Eliashberg equations. Recall that a major difference between BCS and Eliashberg is the presence of an equation for S in the latter theory. Which is why we will solve the set of non-linear equations that determines S and W, or S and T_c, simultaneously instead of keeping S fixed like in the previous subsection. We will solve the coupled equations by finding the values of S and W (or T_c) that satisfy the equations for a given λ keeping all other parameters constant. Repeating this process for varying λ gives us the solutions for T_c and W as a function of λ shown in Fig. 3.4, denoted $T_c(\lambda, S(\lambda))$ and $W(\lambda, S(\lambda))$ respectively. We have also plotted the solution of S and the BCS equations ($S = 0$) in Fig. 3.4. The right hand panel is produced from solving Eqs. (3.75) and (3.76). While the left hand panel is produced from solving Eqs. (3.77) and (3.78).

In Fig. 3.4, the left hand panel is for $T = T_c$ and the right hand panel is for $T = 0$. The BCS curve (blue solid) for both $T = 0$ and $T = T_c$ is meant as a reference with $S = 0$. The $S(\lambda)$ curve (green dashed) in both panels has its y-axis on the right side and is one of the solutions of the coupled equations. The other solutions are $T_c(\lambda)$ and $W(\lambda)$ for $T = T_c$ and $T = 0$, respectively (orange solid). Their y-axis is on the left in both panels. We note that, compared to Fig. 3.3, Fig. 3.4 does not show solutions for a wide range of λ-values. For $T = T_c$, the solutions for both S and T_c exist in the interval $0.13 \lesssim \lambda < 0.65$. While for $T = 0$, solutions for S and W exist in $0 \lesssim \lambda < 0.65$. The reason for this is the difficulty of solving the coupled pair of equations for all λ

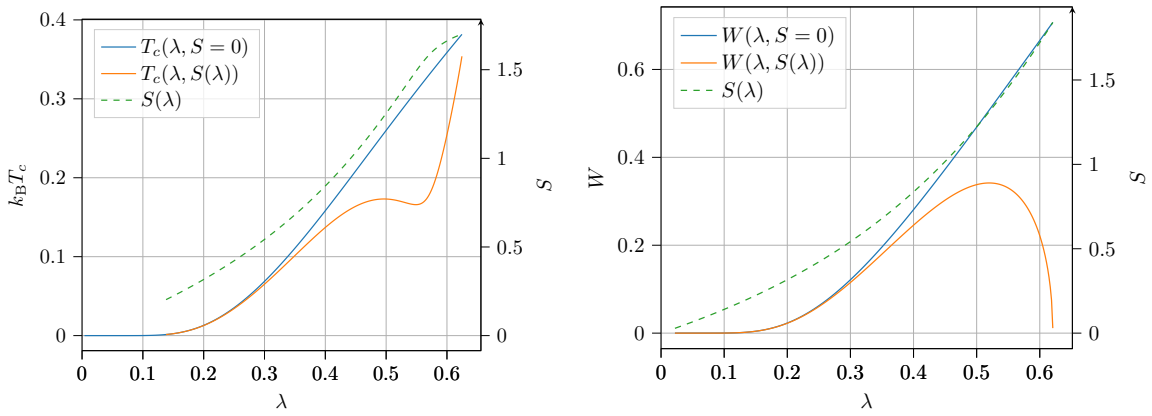


Figure 3.4: Plot of T_c (left panel) and W (right panel). The left panel shows the solutions S (green dashed) and T_c (orange solid) of Eqs. (3.77) and (3.78), while the right panel shows the solutions S (green dashed) and W (orange solid) of Eqs. (3.75) and (3.76). In both panels, the BCS solution (blue solid), i.e. for $S = 0$, is also drawn for comparison. Note that the T_c and W y-axes are on the left side of each panel, while the y-axis of S is on the right side of each panel.

for the given parameter values.

Even though we are unable to solve the equations for small λ, we can still see the trend of each curve. Both T_c and W decay to 0 as λ approaches 0 just like the BCS solutions. And by studying Eqs. (3.75) and (3.77), we can clearly see that $S(\lambda)$ should approach 0 linearly for small λ. This behavior is apparent in both panels of Fig. 3.4.

The upper limit of $\lambda \lesssim 0.65$ in Fig. 3.4 is caused by the fact that $S(\lambda \approx 0.65) \approx \omega_D = 1.7$. It turns out that S cannot be larger than ω_D because of Eq. (3.74)

$$1 = \lambda \int_{-\omega_D}^{\omega_D} \mathrm{d}\epsilon \, \frac{\tanh\left(\frac{\beta}{2}\sqrt{(\epsilon - S)^2 + |W|^2}\right)}{2\sqrt{(\epsilon - S)^2 + |W|^2}}.$$

To understand this we first have to realize that the shape of the integrand as a function of ϵ, for a given β and W, looks like a Gaussian about S. In fact, if we increase β or decrease W, the Gaussian-like shape looks more like a Dirac delta distribution. For the integral to be sizable, the integral limits should include the peak of the integrand. However, if $S > \omega_D$, the peak will not be included and thus the integral might not be large enough to satisfy the LHS of Eq. (3.74). This behavior limits $S < \omega_D$, and because $S(\lambda)$ only seems to increase with λ, we do not find any solutions for S, T_c or W for $\lambda > 0.65$ in Fig. 3.4.

T_c and W for $S \neq 0$ in Fig. 3.4 have a somewhat surprising λ dependency. Still, we can make sense of it if we compare it to Fig. 3.3. Lets begin with T_c in the left panel of Fig. 3.4 and its change of curvature before and after $\lambda \approx 0.5$. We can convince ourselves that the two stationary points at $\lambda \approx 0.5$ and $\lambda \approx 0.55$ make sense by comparing $T_c(\lambda)$ and $S(\lambda)$ from Fig. 3.4 with Fig. 3.3. From Fig. 3.4 we see that $S(\lambda \approx 0.5) \approx 1.25$ and that $S(\lambda \approx 0.55) \approx 1.5$. Comparing these two points to the left panel of Fig. 3.3, we see that $T_c(\lambda \approx 0.5)$ for $S = 1.2$ (green curve) is larger than $T_c(\lambda \approx 0.55)$ for $S = 1.5$ (red curve). This is why T_c in Fig. 3.4 decreases slightly. We can explain the rapid increase of T_c in Fig. 3.4 at $\lambda \approx 0.6$ similarly. We see that $S(\lambda \approx 0.6) \approx 1.7$ in the left panel of Fig. 3.4. And in Fig. 3.3, $T_c(\lambda \approx 0.6)$ for $S = 1.7$ (purple curve) is larger than the two previous points we just discussed, and almost as large as the BCS case (blue curve).

A similar analysis for W in Fig. 3.4 can be made to explain its shape. As S increases in the right panel of Fig. 3.3, $W(\lambda)$ gets flatter. This results in $W(\lambda)$ in Fig. 3.4 to keep decreasing when

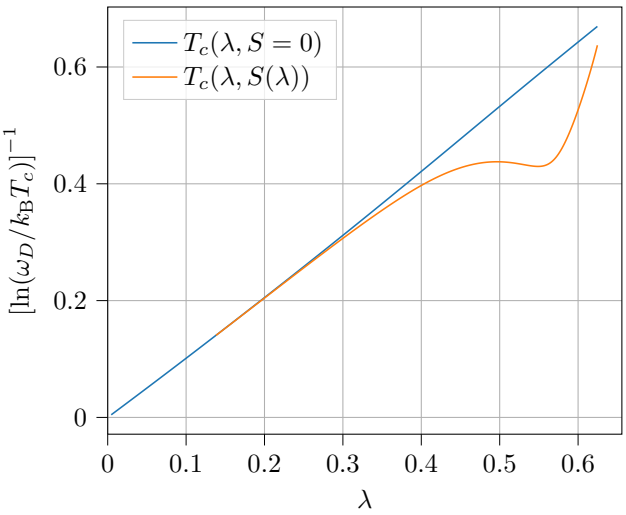


Figure 3.5: Plot comparing the pre-factor of T_c in BCS and Eliashberg theory (based on the same data as left panel of Fig. 3.4). The figure is meant to be compared to Fig. 1 in [44].

$S(\lambda)$ gets large enough.

The presupposition has been that BCS and Eliashberg theory converge in the weak-coupling limit. This is at least true in our case as both solutions for T_c and W in Fig. 3.4 coincide for $\lambda < 0.3$. We can then argue that the strong-coupling regime begins when the BCS and Eliashberg solutions start to deviate, e.g. $\lambda > 0.3$ or even as high as $\lambda > 0.5$. The most striking difference between the BCS and Eliashberg solutions for large λ is the aggressive decline of W to almost zero. We can interpret this as the upper limit of $S \approx \omega_D$ being detrimental to superconductivity, because recall that $W \ll 1$ means that almost zero energy is required to break Cooper pairs. To summarize, this means that including a shift S in the quasi-particle spectrum can affect the superconductive properties of the Haldane-Hubbard model significantly.

Although it is common to think that BCS and Eliashberg theory converge in the weak-coupling limit, Karakozov et al. [43] showed that the BCS pre-factor of T_c gets a correction from Eliashberg theory for small λ as well. Marsiglio [44] has re-derived this result for T_c and investigated the weak-coupling limit further. The correction is a reduction of the pre-factor in Eq. (3.39) arising within Eliashberg theory. We can compare our result of T_c with Marsiglio's by recreating Fig. 1 in [44]. Our recreated figure of Fig. 1 in [44] is shown in Fig. 3.5. Figure 1 in [44] shows a larger difference between the BCS and Eliashberg curves than Fig. 3.5. As $\lambda \to 0$, the BCS and Eliashberg solution converge almost exactly in Fig. 3.5. This result is not surprising because S goes to zero linearly as a function of λ. And as we know, when $S = 0$ our Eliashberg equations are equivalent to the BCS equations. Thus, the differences in the two figures are probably caused by the fact that Marsiglio's Hamiltonian includes a more explicit electron-phonon coupling term as well, i.e. accounting for retardation effects. The consequence of this is that the gap and self-energy become frequency dependent. Our model, on the other hand, has a constant interaction term which only gives us an equation to determine S independent of frequency. However, the inclusion of S seems to have an effect in the correct direction: the Eliashberg curve in Fig. 3.5 stays beneath the BCS curve like in Fig. 1 in [44].

We have learned that the gap equation can change its behavior dramatically by simply accounting for the electron self-energy S. The most dramatic behavior can be observed near $\lambda \sim 0.6$ which also seems to be a break point for the validity of the theory. Recall that Migdal's "theorem" states that $\lambda\omega_D/\epsilon_F \ll 1$ for Eliashberg theory to be valid. In our case, with $\omega_D = 1.7$, $|\mu| = \epsilon_F = 1$, and $\lambda < 0.65$, the ratio becomes $\lambda\omega_D/\epsilon_F \lesssim 1.1$. If we hope to abide by Migdal's "theorem" for

$\lambda > 0.65$, we must satisfy $\omega_D \ll \epsilon_\mathrm{F}$. However, attempting this results in the solver failing to find any solutions. There can be multiple reasons for this. First, as we increase $|\mu|$, the density of states $D(\epsilon_\mathrm{F})$ also increases, see Fig. C.1. This causes imbalance in the equations for S, W, and T_c resulting in convergence issues. Second, if we simply lower ω_D, S will approach ω_D for even lower λ than before. Which is the opposite of what we want to study, but it does keep the Migdal ratio under 1. Regardless, we are not free to set ω_D arbitrarily low as the solver is unable to converge for $\omega_D \ll 1$. The point to make is that with the limit of $\lambda < 0.65$ in Fig. 3.4, we are also reaching a limit where we can be confident in the validity of the results with the current parameter values[9]. In other words, if we want to study the current model for larger λ, the parameters need to be tuned and a better scheme for numerical solving needs to be implemented.

[9]The parameter values used are summarized in Section 3.4.1.

Chapter 4

Summary and outlook

In this thesis, we have studied the superconductive properties of an attractive Haldane-Hubbard model. Our goal has been to study the Haldane-Hubbard model in both a weak and strong-coupling regime using BCS and Eliashberg theory, respectively. With interacting TIs becoming a hotter topic in recent years, and research on attractively interacting Haldane-like models being scarce, makes this an interesting model to study. We began by introducing some key concepts like useful lattice models, BCS and Eliashberg theory and a primer on topological insulators. After which, we derived the gap equation for the attractive Haldane-Hubbard model from BCS theory in the weak-coupling limit, and evaluated it for two temperature limits: $T \to 0$ and $T \to T_c$. The limits gave us an expression for the BCS gap Δ and critical temperature T_c, respectively.

We followed up by focusing on Eliashberg theory which allows for considering the strong coupling limit. The main difference between the BCS and Eliashberg approach was the use of the Green's function formalism. Additionally, in Eliashberg theory, we included the electron self-energy S. In principle, the self-energy S exists in BCS theory but is always neglected. The inclusion of S in Eliashberg theory means that we are accounting for a shift in the quasi-particle spectrum, which usually has an impact on superconductivity. In practice, however, this means we have to solve another equation in addition to the gap equation to determine S.

The task at hand was to solve the Eliashberg equations for S, W and T_c in the usual temperature limits $T \to 0$ and $T \to T_c$. We did this using a root solver in Python and obtained interesting results. The Eliashberg solutions for T_c and W showed that they were consistently lower than the BCS solutions. In fact, as λ increased towards 1, W kept decreasing almost to 0. From this, we learned that as λ increases so does S, which means that S reaches a maximum value ω_D at some point which turns out to be detrimental to superconductivity at $T = 0$.

The upper limit of $S < \omega_D$ is introduced when we set the integration limits of the energy integral in Eq. (3.71) close to the Fermi level. This approximation is made because of the assumption that almost all of the interesting physics happens near the Fermi level. An alternative approach to this approximation is to do the momentum sum (lattice sum) directly instead of converting it to an energy integral. This way, we might be able to solve the coupled equations for larger λ and thus expand our strong-coupling regime.

If we wish to increase λ using the root solving method in this thesis, we have to pay attention to our choice of parameter values. Especially the parameters that make up the Migdal ratio $\lambda\omega_D/\epsilon_{\mathrm{F}}$. Since this ratio should be lower than 1 for the Eliashberg theory to remain valid, we have to ensure that $\omega_D < \epsilon_{\mathrm{F}}$ when increasing λ beyond e.g. 1. However, trying this with the current numerical method results in unstable solutions. It is therefore advised to implement a better numerical scheme to solve the coupled equations or, better yet, do the lattice sum directly.

Appendix A

Chern number of the Haldane model

The goal of this appendix is to show that the Haldane model is topologically non-trivial by calculating a topological index called the Chern number. The Chern number C is proportional to the total Berry curvature

$$C = \frac{1}{2\pi} \oiint_{\mathrm{BZ}} \mathrm{d}\boldsymbol{k} \cdot \boldsymbol{\Omega} \tag{A.1}$$

where $\boldsymbol{\Omega} \equiv \nabla \times \boldsymbol{A}$ is the Berry curvature and $\boldsymbol{A} \equiv -i\langle u(\boldsymbol{k})|\partial_{\boldsymbol{k}}|u(\boldsymbol{k})\rangle$ is the Berry connection. $|u(\boldsymbol{k})\rangle$ is one of the eigenvectors of the Hamiltonian in Eq. (2.23) in reciprocal space. BZ denotes that we have to integrate over the Brillouin zone, see Fig. A.1.

The Chern number is an integer, and if $C \neq 0$, we will have shown that the model is topologically non-trivial. To understand why $C \neq 0$ is significant, we should try to interpret the Chern number as the number of singularities/poles inside the BZ. If there are no singularities inside the BZ, then the system is topologically trivial, i.e. there is nothing special about the system topologically speaking. However, if there are singularities, then the system is non-trivial. Which means that a property of the system is protected by its topology.

A.1 Calculating the Chern number

To show that the Haldane model given by Eq. (2.23) is topologically non-trivial, we first have to Fourier transform it. Using the results of Appendices B.1.1 and B.1.2, i.e. Eqs. (B.5) and (B.7), we rewrite H

$$H = h_0(\boldsymbol{k})I + \boldsymbol{h}(\boldsymbol{k}) \cdot \boldsymbol{\sigma}. \tag{A.2}$$

We use this notation because we will follow [45] in calculating the Chern number. Above, $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ are the Pauli matrices, I is the 2×2 identity matrix, and $\boldsymbol{h} = (h_x, h_y, h_z)$ where

$$\begin{aligned} h_0 &= -2t' \cos\phi \sum_j \cos(\boldsymbol{k} \cdot \boldsymbol{\delta}_j), \quad & h_x &= -t \sum_j \cos(\boldsymbol{k} \cdot \boldsymbol{e}_j) \\ h_y &= t \sum_j \sin(\boldsymbol{k} \cdot \boldsymbol{e}_j), & h_z &= 2t' \sin\phi \sum_j \sin(\boldsymbol{k} \cdot \boldsymbol{\delta}_j). \end{aligned} \tag{A.3}$$

Now that H is expressed in reciprocal space, we can calculate the Berry connection $\boldsymbol{A}$. The eigenvectors of H in a specific gauge are

$$|u_{\pm}^{(1)}(\boldsymbol{k})\rangle = \frac{1}{N^{(1)}(\boldsymbol{k})} \begin{pmatrix} h_z \pm |\boldsymbol{h}| \\ h_x + ih_y \end{pmatrix}, \tag{A.4}$$

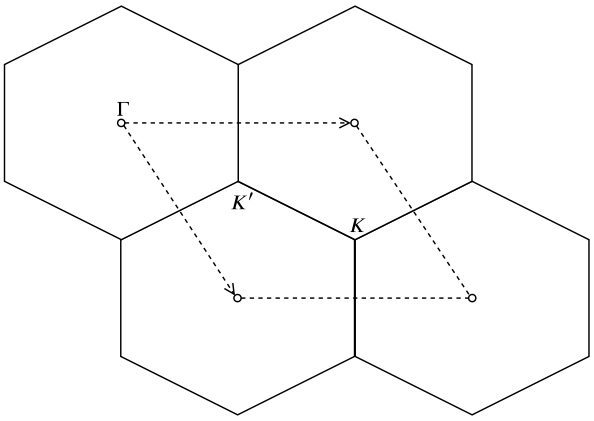


Figure A.1: Brillouin zone (enclosed by dashed lines) of the honeycomb lattice with symmetry points Γ, K and K'.

where $N^{(1)}(\boldsymbol{k})$ is the normalization constant in this gauge. We will get a Chern number $C_\pm$ for each of the eigenvectors. However, we will focus on the lower band because we assume the upper band to be unoccupied.

There are two points in the BZ where the eigenvector $|u_-^{(1)}(\boldsymbol{k})\rangle$ can be singular, i.e. both elements are zero. These points are the Dirac points $\boldsymbol{K} = (4\pi/3\sqrt{3}, 0)$ and $\boldsymbol{K'} = -\boldsymbol{K}$. However, whether or not $|u_-^{(1)}\rangle$ is singular in any of these points depends on whether $0 < \phi < \pi$ or $\pi < \phi < 2\pi$. Thus, we will consider these two cases separately.

A.1.1 Case 1: $0 < \phi < \pi$

In this case we see that $|u_-^{(1)}(\boldsymbol{k})\rangle$ is singular at $\boldsymbol{k} = \boldsymbol{K'}$. At $\boldsymbol{K'}$ both $h_x = h_y = 0$ and $h_z = 3\sqrt{3}t' \sin\phi > 0$. Which means we cannot use $|u_-^{(1)}(\boldsymbol{k})\rangle$ in Eq. (A.1) when integrating over the area which contains $\boldsymbol{K'}$. Thus, we gauge transform $|u_-^{(1)}\rangle \to |u_-^{(2)}\rangle$ where $|u_-^{(2)}(\boldsymbol{k})\rangle = |u_-^{(1)}(\boldsymbol{k})\rangle \exp(i\varphi(\boldsymbol{k}))$ such that it is well-defined at $\boldsymbol{k} = \boldsymbol{K'}$

$$\begin{aligned} |u_-^{(2)}(\boldsymbol{k})\rangle &= |u_-^{(1)}(\boldsymbol{k})\rangle \left(\frac{h_z + |\boldsymbol{h}|}{h_x + ih_y} \cdot \left| \frac{h_z + |\boldsymbol{h}|}{h_x + ih_y} \right|^{-1} \right) \\ &= \frac{1}{N^{(2)}(\boldsymbol{k})} \begin{pmatrix} -h_x + ih_y \\ h_z + |\boldsymbol{h}| \end{pmatrix}. \end{aligned} \tag{A.5}$$

This transformation transforms the Berry connection as $\boldsymbol{A}_-^{(2)}(\boldsymbol{k}) = \boldsymbol{A}_-^{(1)}(\boldsymbol{k}) + \nabla_{\boldsymbol{k}}\varphi(\boldsymbol{k})$. $|u_-^{(2)}(\boldsymbol{k})\rangle$ is not singular at $\boldsymbol{K'}$ but rather at $\boldsymbol{K}$. Thus, we can cut the BZ into two parts and use $|u_-^{(1)}\rangle$ in the part that includes $\boldsymbol{K}$ and $|u_-^{(2)}\rangle$ in the part with $\boldsymbol{K'}$.

Lets draw a circle around $\boldsymbol{K'}$ and call the enclosed area $\mathcal{D}^{(1)}$. We call the rest of the BZ $\mathcal{D}^{(2)}$, see Fig. A.2. The Chern number can be rewritten using Stokes theorem such that we only need to consider the boundaries of the two areas, denoted $\partial\mathcal{D}^{(1)}$ and $\partial\mathcal{D}^{(2)}$

$$C_- = \frac{1}{2\pi} \left(\oint_{\partial\mathcal{D}^{(1)}} \mathrm{d}\boldsymbol{k} \cdot \boldsymbol{A}_-^{(1)} + \oint_{\partial\mathcal{D}^{(2)}} \mathrm{d}\boldsymbol{k} \cdot \boldsymbol{A}_-^{(2)} \right). \tag{A.6}$$

Next we point out that because $\mathcal{D}^{(1)}$ is enclosed by $\mathcal{D}^{(2)}$, a consequence of Stokes theorem is that integrating along the boundary $\partial\mathcal{D}^{(1)}$ is the same as integrating along $\partial\mathcal{D}^{(2)}$ with a minus sign, see Fig. A.3. Using this fact, Eq. (A.6) simplifies

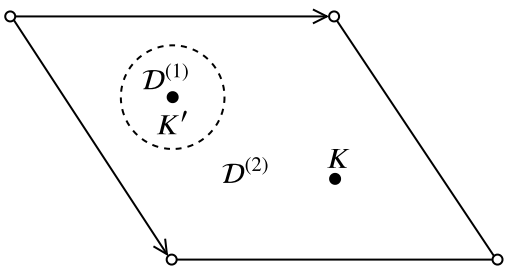


Figure A.2: The two areas of integration inside the BZ, namely $\mathcal{D}^{(1)}$ and $\mathcal{D}^{(2)}$. The radius of the circle enclosing $\mathcal{D}^{(1)}$ can be arbitrarily small.

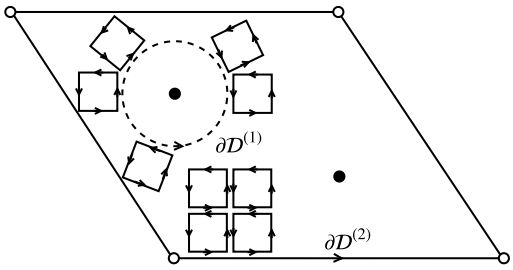


Figure A.3: We can draw small loops inside the BZ to visualize Stokes theorem. The direction of $\partial\mathcal{D}^{(1)}$ is opposite the inner loops of $\mathcal{D}^{(2)}$.

$$\begin{aligned} C_- &= \frac{1}{2\pi}\oint_{\partial\mathcal{D}^{(1)}} \mathrm{d}\boldsymbol{k}\cdot\left(\boldsymbol{A}_-^{(1)}-\boldsymbol{A}_-^{(2)}\right) \\ &= -\frac{1}{2\pi}\oint_{\partial\mathcal{D}^{(1)}} \mathrm{d}\boldsymbol{k}\cdot\nabla_{\boldsymbol{k}}\varphi(\boldsymbol{k}) \\ &= -\frac{1}{2\pi}\int_0^{2\pi}\mathrm{d}\theta\;\partial_\theta\varphi(\boldsymbol{k}) \\ &= -\frac{1}{2\pi}\Big[\varphi(\theta=2\pi)-\varphi(\theta=0)\Big]. \end{aligned} \tag{A.7}$$

All that remains now is to derive an expression for the phase $\varphi(\boldsymbol{k})$. The transformation $\exp(i\varphi(\boldsymbol{k}))$ is defined as

$$\mathrm{e}^{i\varphi(\boldsymbol{k})} = \frac{h_z+|\boldsymbol{h}|}{h_x+ih_y}\cdot\left|\frac{h_z+|\boldsymbol{h}|}{h_x+ih_y}\right|^{-1} \tag{A.8}$$

The right hand side (RHS) of Eq. (A.8) is too difficult to solve as is. To simplify, we decrease the radius of $\partial\mathcal{D}^{(1)}$ and Taylor expand the RHS about $\boldsymbol{k}=\boldsymbol{K}'+\boldsymbol{q}$ where $|\boldsymbol{q}|\ll 1$. Writing out the elements of $\boldsymbol{h}(\boldsymbol{K}'+\boldsymbol{q})$

$$\begin{aligned} h_x(\boldsymbol{K}'+\boldsymbol{q}) &= -t\sum_j\Big[\cos\big(\boldsymbol{K}'\cdot\boldsymbol{e}_j\big)\cos(\boldsymbol{q}\cdot\boldsymbol{e}_j)-\sin\big(\boldsymbol{K}'\cdot\boldsymbol{e}_j\big)\sin(\boldsymbol{q}\cdot\boldsymbol{e}_j)\Big] \\ h_y(\boldsymbol{K}'+\boldsymbol{q}) &= t\sum_j\Big[\sin\big(\boldsymbol{K}'\cdot\boldsymbol{e}_j\big)\cos(\boldsymbol{q}\cdot\boldsymbol{e}_j)+\cos\big(\boldsymbol{K}'\cdot\boldsymbol{e}_j\big)\sin(\boldsymbol{q}\cdot\boldsymbol{e}_j)\Big] \\ h_z(\boldsymbol{K}'+\boldsymbol{q}) &= 2t'\sin\phi\sum_j\Big[\sin\big(\boldsymbol{K}'\cdot\boldsymbol{\delta}_j\big)\cos(\boldsymbol{q}\cdot\boldsymbol{\delta}_j)+\cos\big(\boldsymbol{K}'\cdot\boldsymbol{\delta}_j\big)\sin(\boldsymbol{q}\cdot\boldsymbol{\delta}_j)\Big]. \end{aligned} \tag{A.9}$$

Now we Taylor expand each function to first order in $\boldsymbol{q}$, and perform the sums using the definition of the NN and 2NN vectors $\boldsymbol{e}_j$ and $\boldsymbol{\delta}_j$ from Eq. (2.1).

$$
\begin{aligned}
h_x(\boldsymbol{K}'+\boldsymbol{q}) &= -t\sum_j \Big[\cos\big(\boldsymbol{K}'\cdot\boldsymbol{e}_j\big) - (\boldsymbol{q}\cdot\boldsymbol{e}_j)\sin\big(\boldsymbol{K}'\cdot\boldsymbol{e}_j\big)\Big] = -\frac{3}{2}tq_x \\
h_y(\boldsymbol{K}'+\boldsymbol{q}) &= t\sum_j \Big[\sin\big(\boldsymbol{K}'\cdot\boldsymbol{e}_j\big) + (\boldsymbol{q}\cdot\boldsymbol{e}_j)\cos\big(\boldsymbol{K}'\cdot\boldsymbol{e}_j\big)\Big] = \frac{3}{2}tq_y \\
h_z(\boldsymbol{K}'+\boldsymbol{q}) &= 2t'\sin\phi\sum_j \Big[\sin\big(\boldsymbol{K}'\cdot\boldsymbol{\delta}_j\big) + (\boldsymbol{q}\cdot\boldsymbol{\delta}_j)\cos\big(\boldsymbol{K}'\cdot\boldsymbol{\delta}_j\big)\Big] = 3\sqrt{3}t'\sin\phi.
\end{aligned}
\tag{A.10}
$$

The RHS of Eq. (A.8) can now be evaluated with ease

$$
\mathrm{e}^{i\varphi(\boldsymbol{k})} = \frac{h_z+|\boldsymbol{h}|}{h_x+ih_y}\cdot\left|\frac{h_z+|\boldsymbol{h}|}{h_x+ih_y}\right|^{-1} = \frac{|h_x+ih_y|}{h_x+ih_y}. \tag{A.11}
$$

The last transition is allowed because $h_z(\boldsymbol{K}') > 0$ in the current case. If we insert the simplified h_x and h_y into Eq. (A.11), we get

$$
\mathrm{e}^{i\varphi(\boldsymbol{q})} = -\frac{|q_x-iq_y|}{q_x-iq_y} = -\frac{\left|q\mathrm{e}^{-i\theta}\right|}{qe^{-i\theta}} = \mathrm{e}^{i(\theta+\pi/2)}, \tag{A.12}
$$

where $q = |q_x+iq_y| = |\boldsymbol{q}|$. From Eq. (A.12) we see that $\varphi(\theta) = \theta+\pi/2$, and inserting this result into Eq. (A.7) gives us the final Chern number

$$
C_- = \frac{1}{2\pi}(\pi/2 - 2\pi - \pi/2) = -1. \tag{A.13}
$$

The Chern number of the lower band for $0 < \phi < \pi$ is $C_- = -1$. We have thus shown that because $C_- \neq 0$, the Haldane model is indeed topologically non-trivial. Furthermore, because the sum of Chern numbers over all bands must be zero, we know $C_+ = 1$.

A.1.2 Case 2: $\pi < \phi < 2\pi$

In this case we have to focus on $\boldsymbol{K}$ rather than $\boldsymbol{K}'$ because $h_z(\boldsymbol{K}) = -3\sqrt{3}t'\sin\phi > 0$ when $\pi < \phi < 2\pi$. The method of finding $C_\pm$ in this case is similar to the first case, and the equivalent of Eq. (A.10) is

$$
\begin{aligned}
h_x(\boldsymbol{q}+\boldsymbol{K}) &= \frac{3}{2}tq_x \\
h_y(\boldsymbol{q}+\boldsymbol{K}) &= \frac{3}{2}tq_y \\
h_z(\boldsymbol{q}+\boldsymbol{K}) &= -3\sqrt{3}t'\sin\phi.
\end{aligned}
\tag{A.14}
$$

Which gives $\varphi(\theta) = -\theta$, and in turn $C_- = +1$ and $C_+ = -1$.

Appendix B

Diagonalizing the Hamiltonian

The total tight-binding Hamiltonian of electrons on a honeycomb lattice with an attractive Hubbard potential has the form

$$H = H_K + H_U, \tag{B.1}$$

where H_K represents the kinetic term and H_U the Hubbard potential term. The kinetic term could include nearest and second-nearest neighbor hopping, in which case

$$H_K = H_t + H_{t'}. \tag{B.2}$$

Above, the first term represents nearest neighbor and the last second-nearest neighbor hopping. In the following, we will first diagonalize Eq. (B.1) to the $\boldsymbol{k}$-basis and then diagonalize the emerging k-space Hamiltonian such that there is no mixing of fermion operators.

B.1 Fourier transforms

B.1.1 Nearest neighbor hopping term

The nearest neighbor (NN) hopping term H_t from the tight-binding model is defined as

$$H_t = -t \sum_{\langle i,j\rangle\sigma} (c^\dagger_{i\sigma} d_{j\sigma} + d^\dagger_{j\sigma} c_{i\sigma}), \tag{B.3}$$

where t is the NN hopping parameter, $c_{i\sigma}$ and $c^\dagger_{i\sigma}$ are the fermionic annihilation and creation operators at lattice site i with spin σ on sublattice A, and the d-operators are equivalent to the c-operators but for sublattice B. The sum is taken over all NN lattice sites $\langle i, j\rangle$. The neighbor vectors in the honeycomb lattice are shown in Fig. 2.2.

Equation (B.3) can be Fourier transformed to reciprocal space using the definition

$$c_{i\sigma} = \frac{1}{\sqrt{N}} \sum_i c_{\boldsymbol{k}\sigma} \mathrm{e}^{i\boldsymbol{k}\cdot\boldsymbol{r}_i}, \tag{B.4}$$

where N is the number of unit cells, $\boldsymbol{k}$ is the reciprocal lattice vector, and $\boldsymbol{r}_i$ is the real-space position of lattice site i. Using this definition in Eq. (B.3) gives the Fourier transformed Hamiltonian

$$\begin{aligned} H_t &= -t \sum_{\boldsymbol{k}\boldsymbol{k}'\sigma} c^\dagger_{\boldsymbol{k}\sigma} d_{\boldsymbol{k}'\sigma} \underbrace{\left(\frac{1}{N} \sum_i \mathrm{e}^{-i(\boldsymbol{k}-\boldsymbol{k}')\cdot \boldsymbol{r}_i} \right)}_{\equiv \delta_{\boldsymbol{k}\boldsymbol{k}'}} \sum_{j=1}^{3} \mathrm{e}^{i\boldsymbol{k}'\cdot \boldsymbol{e}_j} + \text{h.c.} \\ &= -t \sum_{\boldsymbol{k}\sigma} \Big[S(\boldsymbol{k}) c^\dagger_{\boldsymbol{k}\sigma} d_{\boldsymbol{k}\sigma} + S^*(\boldsymbol{k}) d^\dagger_{\boldsymbol{k}\sigma} c_{\boldsymbol{k}\sigma} \Big]. \end{aligned} \tag{B.5}$$

Above, $\boldsymbol{e}_j$ are the NN lattice vectors shown in Fig. 2.2. We have further defined the neighbor vector $\boldsymbol{r}_j = \boldsymbol{r}_i + \boldsymbol{e}_j$ and $S(\boldsymbol{k}) = \sum_j \exp(i\boldsymbol{k}\cdot\boldsymbol{e}_j)$.

Equation (B.5) gives us the final form of the Fourier transformed NN term of the Hamiltonian.

B.1.2 Second-nearest neighbor hopping term

We can transform the second-nearest neighbor (2NN) term $H_{t'}$ the same way as in Appendix B.1.1. The 2NN term is defined as

$$H_{t'} = -t' \sum_{\langle\langle i,j\rangle\rangle\sigma} \Big(c^\dagger_{i\sigma} c_{j\sigma} \mathrm{e}^{i\phi} + c^\dagger_{j\sigma} c_{i\sigma} \mathrm{e}^{-i\phi} + d^\dagger_{i\sigma} d_{j\sigma} \mathrm{e}^{-i\phi} + d^\dagger_{j\sigma} d_{i\sigma} \mathrm{e}^{i\phi} \Big). \tag{B.6}$$

Here, the sum is taken over 2NN lattice sites $\langle\langle i,j\rangle\rangle$. Moreover, the 2NN hopping parameter is complex with absolute value t' and phase $\pm\phi$. Equation (B.6) can be Fourier transformed like Eq. (B.3), which gives

$$H_{t'} = -t' \sum_{\boldsymbol{k}\sigma} \Big[Q_+(\boldsymbol{k}) c^\dagger_{\boldsymbol{k}\sigma} c_{\boldsymbol{k}\sigma} + Q_-(\boldsymbol{k}) d^\dagger_{\boldsymbol{k}\sigma} d_{\boldsymbol{k}\sigma} \Big], \tag{B.7}$$

where we have defined $Q_\pm(\boldsymbol{k}) = 2\sum_j \cos(\boldsymbol{k}\cdot\boldsymbol{\delta}_j \pm \phi)$ with $\boldsymbol{\delta}_j$ as the three 2NN vectors shown in Fig. 2.2. Equation (B.7) is the final form of the Fourier transform of Eq. (B.6).

B.1.3 Hubbard interaction term

The on-site Hubbard interaction term is defined as [26]

$$H_U = U \sum_{i\in A} c^\dagger_{i\uparrow} c_{i\uparrow} c^\dagger_{i\downarrow} c_{i\downarrow} + U \sum_{j\in B} d^\dagger_{j\uparrow} d_{j\uparrow} d^\dagger_{j\downarrow} d_{j\downarrow}, \tag{B.8}$$

where $U < 0$. Equation (B.8) adds up the total potential energy of electrons on each fully occupied lattice site. The two terms account for the two sublattices. However, because there is exactly one site from each sublattice in a unit cell, they both Fourier transform the same way

$$\begin{aligned} H_U &= \frac{U}{N} \sum_{\boldsymbol{k}_1\boldsymbol{k}_2\boldsymbol{k}_3\boldsymbol{k}_4} c^\dagger_{\boldsymbol{k}_1\uparrow} c_{\boldsymbol{k}_2\uparrow} c^\dagger_{\boldsymbol{k}_3\downarrow} c_{\boldsymbol{k}_4\downarrow} \underbrace{\left(\frac{1}{N} \sum_{i\in A} \mathrm{e}^{-i(\boldsymbol{k}_1-\boldsymbol{k}_2+\boldsymbol{k}_3-\boldsymbol{k}_4)\cdot r_i} \right)}_{=\delta_{\boldsymbol{k}_1,\boldsymbol{k}_2-\boldsymbol{k}_3+\boldsymbol{k}_4}} + (c\to d) \\ &= \frac{U}{N} \sum_{\boldsymbol{k}_2\boldsymbol{k}_3\boldsymbol{k}_4} c^\dagger_{\boldsymbol{k}_2-\boldsymbol{k}_3+\boldsymbol{k}_4,\uparrow} c_{\boldsymbol{k}_2\uparrow} c^\dagger_{\boldsymbol{k}_3\downarrow} c_{\boldsymbol{k}_4\downarrow} + (c\to d). \end{aligned} \tag{B.9}$$

For simplicity, we rename $\boldsymbol{k}_2 = \boldsymbol{k}$, $\boldsymbol{k}_4 = \boldsymbol{k}'$ and $\boldsymbol{k}_3 = \boldsymbol{k}' - \boldsymbol{q}$. Which gives us the final form of the transformed H_U expressed by a momentum transfer $\boldsymbol{q}$

$$H_U = \frac{U}{N} \sum_{\boldsymbol{k}\boldsymbol{k}'\boldsymbol{q}} \Big(c^\dagger_{\boldsymbol{k}+\boldsymbol{q}\uparrow} c_{\boldsymbol{k}\uparrow} c^\dagger_{\boldsymbol{k}'-\boldsymbol{q}\downarrow} c_{\boldsymbol{k}'\downarrow} + d^\dagger_{\boldsymbol{k}+\boldsymbol{q}\uparrow} d_{\boldsymbol{k}\uparrow} d^\dagger_{\boldsymbol{k}'-\boldsymbol{q}\downarrow} d_{\boldsymbol{k}'\downarrow} \Big). \tag{B.10}$$

B.2 Diagonalizing

We will now diagonalize the Hamiltonian $H = H_t + H_U$ to a fermion basis without mixing of operators. Although we restrict ourselves to NN hopping here, the following derivation can be done with 2NN as well. Note that the contribution from the 2NN term has no mixing of (c, d) operators, which means the matrix representation of H would have non-zero diagonal elements.

B.2.1 Hopping term

The Fourier transformed NN hopping term is derived in Appendix B.1.1 and given as

$$H_t = -t \sum_{\boldsymbol{k}\sigma} \Big[S(\boldsymbol{k}) c^\dagger_{\boldsymbol{k}\sigma} d_{\boldsymbol{k}\sigma} + S^*(\boldsymbol{k}) d^\dagger_{\boldsymbol{k}\sigma} c_{\boldsymbol{k}\sigma} \Big]. \tag{B.11}$$

Above, $S(\boldsymbol{k}) = \sum_j \exp(i\boldsymbol{k} \cdot \boldsymbol{e_j})$ where $\boldsymbol{e}_j$ are the NN vectors shown in Fig. 2.2. Moreover, t is the NN hopping parameter and $(c^\dagger, c)$ and $(d^\dagger, d)$ are the fermionic creation and annihilation operators for sub-lattice A and B, respectively.

Equation (B.11) is non-diagonal due to the mixing of c and d operators. Writing it out on matrix form makes it clearer

$$H_t = \sum_{\boldsymbol{k}\sigma} \begin{pmatrix} c^\dagger_{\boldsymbol{k}\sigma} & d^\dagger_{\boldsymbol{k}\sigma} \end{pmatrix} \begin{pmatrix} 0 & -tS(\boldsymbol{k}) \\ -tS^*(\boldsymbol{k}) & 0 \end{pmatrix} \begin{pmatrix} c_{\boldsymbol{k}\sigma} \\ d_{\boldsymbol{k}\sigma} \end{pmatrix}. \tag{B.12}$$

We wish to diagonalize the Hamiltonian such that we can easily read off the dispersion relation. Lets rewrite Eq. (B.12) from the current basis $\phi^\dagger_{\boldsymbol{k}\sigma} = \begin{pmatrix} c^\dagger_{\boldsymbol{k}\sigma} & d^\dagger_{\boldsymbol{k}\sigma} \end{pmatrix}$ to the new, diagonal basis $\tilde{\phi}^\dagger_{\boldsymbol{k}\sigma} = \begin{pmatrix} \alpha^\dagger_{\boldsymbol{k}\sigma} & \beta^\dagger_{\boldsymbol{k}\sigma} \end{pmatrix}$

$$\begin{aligned} H_t &= \sum_{\boldsymbol{k}\sigma} \Big(\phi^\dagger_{\boldsymbol{k}\sigma} M^{-1} \Big) \Big[M \begin{pmatrix} 0 & -tS(\boldsymbol{k}) \\ -tS^*(\boldsymbol{k}) & 0 \end{pmatrix} M^{-1} \Big] \Big(M \phi_{\boldsymbol{k}\sigma} \Big) \\ &= \sum_{\boldsymbol{k}\sigma} \tilde{\phi}^\dagger_{\boldsymbol{k}\sigma} \begin{pmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{pmatrix} \tilde{\phi}_{\boldsymbol{k}\sigma}, \end{aligned} \tag{B.13}$$

where $\lambda_\pm = \pm t|S(\boldsymbol{k})|$ are the eigenvalues of the Hamiltonian, i.e. upper and lower band of the dispersion relation. A matrix M that can diagonalize Eq. (B.12) is

$$M = \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{\frac{S(\boldsymbol{k})}{S^*(\boldsymbol{k})}} & -\sqrt{\frac{S(\boldsymbol{k})}{S^*(\boldsymbol{k})}} \\ 1 & 1 \end{pmatrix}, \tag{B.14}$$

from which we see how the new fermionic operators α, β relate to the old c, d

$$\begin{aligned} \alpha_{\boldsymbol{k}\sigma} &= \frac{1}{\sqrt{2}} (u_{\boldsymbol{k}} c_{\boldsymbol{k}\sigma} - u_{\boldsymbol{k}} d_{\boldsymbol{k}\sigma}) \\ \beta_{\boldsymbol{k}\sigma} &= \frac{1}{\sqrt{2}} (c_{\boldsymbol{k}\sigma} + d_{\boldsymbol{k}\sigma}), \end{aligned} \tag{B.15}$$

where we have rewritten $u_{\boldsymbol{k}} \equiv \sqrt{S(\boldsymbol{k})/S^*(\boldsymbol{k})}$.

To summarize, we have now expressed the NN term in the Hamiltonian using a new basis

$$H_t = \sum_{\boldsymbol{k}\sigma} \Big(\lambda_+ \alpha^\dagger_{\boldsymbol{k}\sigma} \alpha_{\boldsymbol{k}\sigma} + \lambda_- \beta^\dagger_{\boldsymbol{k}\sigma} \beta_{\boldsymbol{k}\sigma} \Big). \tag{B.16}$$

H_t is diagonal and we can easily read off the eigenvalues $\lambda_\pm$. A plot of the eigenvalues is shown in Fig. B.1 near the so called Dirac point $\boldsymbol{K} = (4\pi/3\sqrt{3}, 0)$. This is the famous Dirac cone.

We can invert Eq. (B.15) and express the Hubbard interaction term in terms of these new operators. That will be our next task.

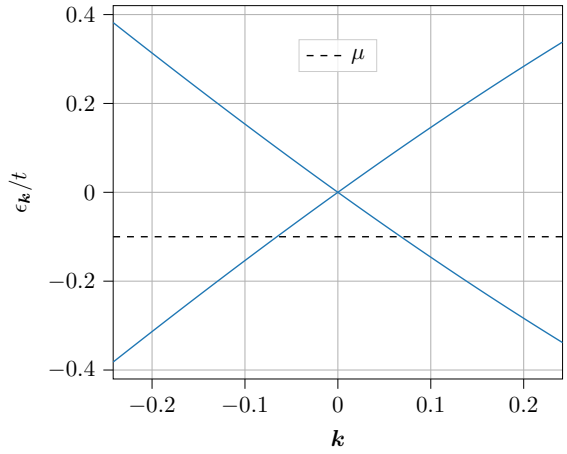


Figure B.1: The dispersion relation of the NN tight-binding Hamiltonian near the $\boldsymbol{K}$-point. The bands form a cone in two dimensions called a Dirac Cone. An arbitrary Fermi level μ in units of t is drawn together with the energy bands.

B.2.2 Hubbard interaction term

We want to re-express the interaction term H_U in terms of the new operators. We begin with the Fourier transform of H_U derived in Appendix B.1.3

$$H_U = \frac{U}{N}\sum_{\boldsymbol{k}\boldsymbol{k}'\boldsymbol{q}}\left(c^\dagger_{\boldsymbol{k}+\boldsymbol{q}\uparrow}c^\dagger_{\boldsymbol{k}'-\boldsymbol{q}\downarrow}c_{\boldsymbol{k}'\downarrow}c_{\boldsymbol{k}\uparrow} + d^\dagger_{\boldsymbol{k}+\boldsymbol{q}\uparrow}d^\dagger_{\boldsymbol{k}'-\boldsymbol{q}\downarrow}d_{\boldsymbol{k}'\downarrow}d_{\boldsymbol{k}\uparrow}\right). \tag{B.17}$$

Here, $U < 0$ is the attractive interaction strength, N is the number of unit cells, and $\boldsymbol{q}$ can be interpreted as a momentum transfer in a scattering process.

We will now express Eq. (B.17) in terms of α and β operators, see Eq. (B.15). Inserting Eq. (B.15) in Eq. (B.17) gives

$$H_U = \frac{U}{N}\sum_{\boldsymbol{k}\boldsymbol{k}'\boldsymbol{q}}(\tilde{\alpha}^\dagger_{\boldsymbol{k}+\boldsymbol{q}\uparrow} + \tilde{\beta}^\dagger_{\boldsymbol{k}+\boldsymbol{q}\uparrow})(\tilde{\alpha}^\dagger_{\boldsymbol{k}'-\boldsymbol{q}\downarrow} + \tilde{\beta}^\dagger_{\boldsymbol{k}'-\boldsymbol{q}\downarrow})(\tilde{\alpha}_{\boldsymbol{k}'\downarrow} + \tilde{\beta}_{\boldsymbol{k}'\downarrow})(\tilde{\alpha}_{\boldsymbol{k}\uparrow} + \tilde{\beta}_{\boldsymbol{k}\uparrow}) + (\tilde{\alpha} \to -\tilde{\alpha}), \tag{B.18}$$

where we have defined $\tilde{\alpha}_{\boldsymbol{k}\sigma} \equiv \alpha_{\boldsymbol{k}\sigma}/\sqrt{2}u_{\boldsymbol{k}}$ and $\tilde{\beta}_{\boldsymbol{k}\sigma} \equiv \beta_{\boldsymbol{k}\sigma}/\sqrt{2}$. At first glance this expression looks worse. However, by introducing a Fermi energy μ and choosing it such that e.g. only the lower band λ_- is occupied (with zero states in the upper band λ_+), we can disregard all terms with α in Eq. (B.18)[1]. We can do this because all terms including an α-operator will give exactly zero contribution when the Fermi energy lies inside the lower band. This simplification leaves us with

$$H_U = \frac{U}{2N}\sum_{\boldsymbol{k}\boldsymbol{k}'\boldsymbol{q}}\beta^\dagger_{\boldsymbol{k}+\boldsymbol{q}\uparrow}\beta^\dagger_{\boldsymbol{k}'-\boldsymbol{q}\downarrow}\beta_{\boldsymbol{k}'\downarrow}\beta_{\boldsymbol{k}\uparrow}. \tag{B.19}$$

In doing this, we have introduced a somewhat fixed Fermi level, which means we will include another term H_μ in the Hamiltonian H

$$H_\mu = -\mu\sum_{\boldsymbol{k}\sigma}\left(c^\dagger_{\boldsymbol{k}\sigma}c_{\boldsymbol{k}\sigma} + d^\dagger_{\boldsymbol{k}\sigma}d_{\boldsymbol{k}\sigma}\right). \tag{B.20}$$

[1]The Fermi level μ is drawn as a horizontal line in Fig. B.1.

The total Hamiltonian H to this point is

$$\begin{aligned} H &= H_t + H_\mu + H_U \\ &= \sum_{\boldsymbol{k}\sigma} (\epsilon_{\boldsymbol{k}} - \mu) c^\dagger_{\boldsymbol{k}\sigma} c_{\boldsymbol{k}\sigma} + \frac{U}{4N} \sum_{\boldsymbol{k}\boldsymbol{k}'\boldsymbol{q},\sigma} c^\dagger_{\boldsymbol{k}+\boldsymbol{q}\sigma} c^\dagger_{\boldsymbol{k}'-\boldsymbol{q},-\sigma} c_{\boldsymbol{k}',-\sigma} c_{\boldsymbol{k}\sigma}, \end{aligned} \tag{B.21}$$

where $\epsilon_{\boldsymbol{k}} \equiv -t|S(\boldsymbol{k})|$ is the lower band of the dispersion relation. We have renamed all β-operators to c-operators for familiarity's sake. The last term has also been made slightly more general by introducing a spin-sum. H would look the same had we included 2NN hopping. The only visible difference would have been in $\epsilon_{\boldsymbol{k}}$ which for 2NN is defined as

$$\epsilon_{\boldsymbol{k}} = -2t' \cos\phi \sum_j \cos(\boldsymbol{k}\cdot\boldsymbol{\delta}_j) \pm \left[t^2 |S(\boldsymbol{k})|^2 + 4t'^2 \sin^2\phi \Big(\sum_j \sin(\boldsymbol{k}\cdot\boldsymbol{\delta}_j) \Big)^2 \right]^{1/2}, \tag{B.22}$$

where $\boldsymbol{\delta}_j$ are the 2NN vectors and the 2NN hopping parameter is defined as $t'\mathrm{e}^{\pm i\phi}$.

Appendix C

Density of states

In this appendix we will derive the density of states (DOS) at the Fermi level near the Dirac point. The DOS at the Fermi level $\epsilon_{\mathrm{F}} = \mu < 0$ is defined as

$$D(\epsilon_{\mathrm{F}}) = \sum_{\boldsymbol{k}} \delta(\epsilon_{\mathrm{F}} - \epsilon_{\boldsymbol{k}}), \tag{C.1}$$

where $\delta(x)$ is the Dirac delta distribution. Equation (C.1) is too difficult to solve because the dispersion relation $\epsilon_{\boldsymbol{k}}$ is too complicated. We begin simplifying $\epsilon_{\boldsymbol{k}}$ from Eq. (3.3) by setting $\phi = \pi/2$ and introducing $v \equiv t'/t$

$$\epsilon_{\boldsymbol{k}} = -\Big[|S(\boldsymbol{k})|^2 + 4v^2\Big(\sum_j \sin(\boldsymbol{k}\cdot\boldsymbol{\delta}_j)\Big)^2\Big]^{1/2}. \tag{C.2}$$

Recall the definition $S(\boldsymbol{k}) \equiv \sum_j \exp(i\boldsymbol{k}\cdot\boldsymbol{e}_j)$. From this, we can easily derive its absolute square

$$|S(\boldsymbol{k})|^2 = 3 + 2\sum_j \cos(\boldsymbol{k}\cdot\boldsymbol{\delta}_j). \tag{C.3}$$

The dispersion relation is still too complicated. One way to simplify it is to write it as a power series about the Dirac point $\boldsymbol{K} = (4\pi/3\sqrt{3}, 0)$. This choice is reasonable because we are really only interested in the DOS near the Dirac point, where the gap is smallest. We begin by expanding Eq. (C.3) letting $\boldsymbol{k} \to \boldsymbol{k} + \boldsymbol{K}$ where $|\boldsymbol{k}| \ll 1$

$$|S(\boldsymbol{k}+\boldsymbol{K})|^2 = 3 + 2\sum_j \Big[\cos(\boldsymbol{K}\cdot\boldsymbol{\delta}_j)\cos(\boldsymbol{k}\cdot\boldsymbol{\delta}_j) - \sin(\boldsymbol{K}\cdot\boldsymbol{\delta}_j)\sin(\boldsymbol{k}\cdot\boldsymbol{\delta}_j)\Big]. \tag{C.4}$$

Taylor expanding the terms with $\boldsymbol{k}$ to $\mathcal{O}(\boldsymbol{k}^4)$ and performing the sums leaves us with

$$|S(\boldsymbol{k})|^2 = \frac{9}{4}k^2 - \frac{27}{64}k^4. \tag{C.5}$$

If we expand $\sum_j \sin(\boldsymbol{k}\cdot\boldsymbol{\delta}_j)$ in Eq. (C.2) similarly, we end up with a less complicated dispersion relation

$$\epsilon_{\boldsymbol{k}} = -\sqrt{\frac{9}{4}k^2 - \frac{27}{64}k^4 + 4v^2\Big(-\frac{3\sqrt{3}}{2} + \frac{9\sqrt{3}}{8}k^2\Big)^2} = -\sqrt{ak^4 + bk^2 + m^2}, \tag{C.6}$$

where

$$\begin{aligned} a &= \frac{243}{16}v^2 - \frac{27}{64} \\ b &= -\frac{81}{2}v^2 + \frac{9}{4} \\ m^2 &= 27v^2. \end{aligned} \tag{C.7}$$

The DOS is now much easier to derive

$$D(\epsilon_{\mathrm{F}}) = \sum_{\boldsymbol{k}} \delta\Big(- |\epsilon_{\mathrm{F}}| + \sqrt{ak^4 + bk^2 + m^2}\Big). \tag{C.8}$$

If we let

$$\frac{1}{N}\sum_{\boldsymbol{k}} \to \frac{1}{A_{\mathrm{BZ}}}\int \mathrm{d}^2k, \tag{C.9}$$

where $A_{\mathrm{BZ}} = \frac{8\pi^2}{3\sqrt{3}a^2}$ is the Brillouin zone area and N is the number of unit cells, we can write Eq. (C.8) as an integral

$$D(\epsilon_{\mathrm{F}}) = \frac{N}{A_{\mathrm{BZ}}}\int_0^{2\pi} \mathrm{d}\theta \int_0^{\infty} \mathrm{d}k\ k\delta\Big(|\epsilon_{\mathrm{F}}| - \sqrt{ak^4 + bk^2 + m^2}\Big). \tag{C.10}$$

Next, we substitute $u = \sqrt{ak^4 + bk^2 + m^2}$. Differentiating it gives

$$\mathrm{d}u = \frac{2ak^2 + b}{u} k\mathrm{d}k. \tag{C.11}$$

We find k^2 expressed by u by solving $ak^4 + bk^2 + m^2 - u^2 = 0$, whose solutions are

$$k^2 = \frac{-b \pm \sqrt{b^2 - 4a(m^2 - u^2)}}{2a}. \tag{C.12}$$

We choose the positive solution because the negative solution would cause the DOS to be negative. Furthermore, as a function of v the positive solution is positive up until $v_c = t_c'/t$. The DOS integral becomes

$$D(\epsilon_{\mathrm{F}}) = \frac{2\pi N}{A_{\mathrm{BZ}}}\int_m^{\infty} \mathrm{d}u\ \frac{u}{\sqrt{b^2 - 4a(m^2 - u^2)}}\ \delta(|\epsilon_{\mathrm{F}}| - u). \tag{C.13}$$

The delta function and integral limits put a constraint on the integral, which is why we introduce the Heaviside step function $\Theta(x)$, and arrive at the final expression for $D(\epsilon_{\mathrm{F}})$

$$D(\epsilon_{\mathrm{F}}) = \frac{2\pi N}{A_{\mathrm{BZ}}}\ \frac{|\epsilon_{\mathrm{F}}|}{\sqrt{b^2 - 4a(m^2 - \epsilon_{\mathrm{F}}^2)}}\ \Theta(|\epsilon_{\mathrm{F}}| - m). \tag{C.14}$$

A plot of the DOS in Eq. (C.14) as a function of μ and v is shown in Fig. C.1. The DOS as a function of μ is almost linear for $v = 0$. And increasing v shows that the DOS becomes non-zero only after passing a critical $\mu_c = \sqrt{27}v_c$, after which the DOS grows approximately linearly. This behavior is what we would expect. The DOS jumps at μ_c for a given v because that is the moment the Fermi level moves inside the band.

The DOS as a function of v is slightly trickier to understand. We see that the DOS increases as a function of v until it collapses to 0 at v_c. Which means that the DOS is at its highest when the Fermi level is on the brink of moving inside the gap due to the gap growing larger. To understand this effect, we should study the dispersion relation for increasing v, see Fig. 3.1. The figure shows that increasing v not only increases the gap between the bands but also alters the shape of the

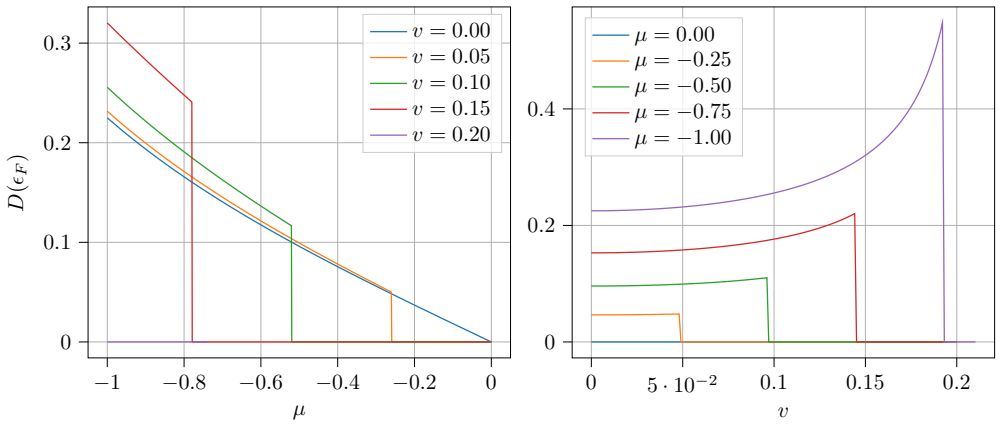


Figure C.1: Density of states at Fermi level near the Dirac point as a function of μ and $v \equiv t'/t$.

bands. Changes significant to the DOS are (1) the Fermi surface for $v \lesssim v_c$ is smaller than for $v = 0$ and (2) the slope of the band at the Fermi surface is smaller, i.e. the band flattens. Of the two effects, (1) contributes towards a decrease in the DOS, while (2) increases it. It is easier to see why (2) increases the DOS if we rewrite it as [40]

$$D(\epsilon_{\mathrm{F}}) = \frac{\Omega}{(2\pi)^2} \int\limits_{\epsilon=\epsilon_{\mathrm{F}}} \frac{\mathrm{d}S}{|\nabla_{\boldsymbol{k}} \epsilon_{\boldsymbol{k}}|}, \tag{C.15}$$

where we integrate over the constant Fermi surface. Equation (C.15) is inversely "proportional" to the gradient of the dispersion relation at the Fermi surface. Hence, a decrease of the gradient, or flattening of the band, should increase the DOS.

In summary, we observe that effect (2) dominates. Resulting in an increase of $D(\epsilon_{\mathrm{F}})$ as a function of v up until v_c.

Appendix D

Eliashberg equations

In the following appendix we will derive the Eliashberg equations using the equations of motion approach shown in the appendix of [42] together with chapter 10 of [34].

The diagonalized Hamiltonian is given by Eq. (B.21)

$$H = \sum_{\boldsymbol{k}\sigma} (\epsilon_{\boldsymbol{k}} - \mu) c^\dagger_{\boldsymbol{k}\sigma} c_{\boldsymbol{k}\sigma} + \frac{U}{4N} \sum_{\boldsymbol{k}\boldsymbol{k}'\boldsymbol{q},\sigma} c^\dagger_{\boldsymbol{k}+\boldsymbol{q}\sigma} c^\dagger_{\boldsymbol{k}'-\boldsymbol{q},-\sigma} c_{\boldsymbol{k}',-\sigma} c_{\boldsymbol{k}\sigma}, \tag{D.1}$$

where c ($c^\dagger$) is the pseudo-particle fermionic annihilation (creation) operator. The single-particle $\boldsymbol{k}$-state energy $\epsilon_{\boldsymbol{k}}$ (for NN and 2NN) is given by

$$\epsilon_{\boldsymbol{k}} = -2t' \cos\phi \sum_j \cos(\boldsymbol{k}\cdot\boldsymbol{\delta}_j) \pm \left[t^2 |S(\boldsymbol{k})|^2 + 4t'^2 \sin^2\phi \Big(\sum_j \sin(\boldsymbol{k}\cdot\boldsymbol{\delta}_j) \Big)^2 \right]^{1/2}. \tag{D.2}$$

D.1 Imaginary-time Green's functions

The definition of the single-particle Green's function, defined in momentum space, as a function of imaginary-time τ is

$$G(\boldsymbol{k},\tau) \equiv -\langle T_\tau c_{\boldsymbol{k}\sigma}(\tau) c^\dagger_{\boldsymbol{k}\sigma}(0) \rangle. \tag{D.3}$$

Recall that, in the Heisenberg picture, the imaginary-time dependency of operators is defined by

$$\begin{aligned} c_\nu(\tau) &\equiv \mathrm{e}^{\tau H} c_\nu \mathrm{e}^{-\tau H} \\ c^\dagger_\nu(\tau) &\equiv \mathrm{e}^{\tau H} c^\dagger_\nu \mathrm{e}^{-\tau H}. \end{aligned} \tag{D.4}$$

Note that both the equation for $c_\nu(\tau)$ and $c^\dagger_\nu(\tau)$ is needed because $\mathrm{e}^{\tau H}$ is not a unitary operator.

We use the equation of motion approach to derive the Eliashberg equations for our system given by Eq. (D.1). We start by taking the imaginary-time derivative of G

$$\begin{aligned} \frac{\partial}{\partial\tau} G(\boldsymbol{k},\tau) &= -\delta(\tau) \left\langle \left\{ c_{\boldsymbol{k}\sigma}(\tau=0), c^\dagger_{\boldsymbol{k}\sigma}(0) \right\} \right\rangle - \left\langle T_\tau \left[\frac{\partial}{\partial\tau} c_{\boldsymbol{k}\sigma}(\tau) \right] c^\dagger_{\boldsymbol{k}\sigma}(0) \right\rangle \\ &= -\delta(\tau) - \left\langle T_\tau \left[H, c_{\boldsymbol{k}\sigma}(\tau) \right] c^\dagger_{\boldsymbol{k}\sigma}(0) \right\rangle. \end{aligned} \tag{D.5}$$

To proceed, we need to derive an expression for the commutator $[H, c_{\boldsymbol{k}\sigma}(\tau)]$. From the definition of the operators in Eq. (D.4), we see that $[H, c_{\boldsymbol{k}\sigma}(\tau)] = \mathrm{e}^{\tau H} [H, c_{\boldsymbol{k}\sigma}] \mathrm{e}^{-\tau H}$ because H commutes with $\mathrm{e}^{\pm\tau H}$ [1]. This leaves us with the task of finding the time-independent commutator

[1] Recall that $\mathrm{e}^{\pm\tau H}$ is simply the series expansion of the exponential function, i.e. integer powers of H.

$$[H, c_{\boldsymbol{k}\sigma}] = \sum_{\nu\sigma'} \xi_\nu \left[c^\dagger_{\nu\sigma'} c_{\nu\sigma'}, c_{\boldsymbol{k}\sigma} \right] + \frac{U}{4N} \sum_{\nu\nu' q,s} \left[c^\dagger_{\nu+q,s} c^\dagger_{\nu'-q,-s} c_{\nu',-s} c_{\nu,s}, c_{\boldsymbol{k}\sigma} \right]. \tag{D.6}$$

The first commutator is easily evaluated by using the relation

$$[AB, C] = A\,[B, C] + [A, C]\,B = A\,\{B, C\} - \{A, C\}\,B. \tag{D.7}$$

This gives

$$\begin{aligned} \left[c^\dagger_{\nu\sigma'} c_{\nu\sigma'}, c_{\boldsymbol{k}\sigma} \right] &= c^\dagger_{\nu\sigma'} \left\{ c_{\nu\sigma'}, c_{\boldsymbol{k}\sigma} \right\} - \left\{ c^\dagger_{\nu\sigma'}, c_{\boldsymbol{k}\sigma} \right\} c_{\nu\sigma'} \\ &= 0 - \delta_{\nu\boldsymbol{k}} \delta_{\sigma\sigma'} c_{\nu\sigma'}. \end{aligned} \tag{D.8}$$

The second commutator in Eq. (D.6) can be evaluated similarly by using Eq. (D.7) twice

$$\begin{aligned} \left[c^\dagger_{\nu+q,s} c^\dagger_{\nu'-q,-s} c_{\nu',-s} c_{\nu,s}, c_{\boldsymbol{k}\sigma} \right] &= c^\dagger_{\nu+q,s} c^\dagger_{\nu'-q,-s} \left[c_{\nu',-s} c_{\nu,s}, c_{\boldsymbol{k}\sigma} \right] + \left[c^\dagger_{\nu+q,s} c^\dagger_{\nu'-q,-s}, c_{\boldsymbol{k}\sigma} \right] c_{\nu',-s} c_{\nu,s} \\ &= 0 + \left(c^\dagger_{\nu+q,s} \left\{ c^\dagger_{\nu'-q,-s}, c_{\boldsymbol{k}\sigma} \right\} - \left\{ c^\dagger_{\nu+q,s}, c_{\boldsymbol{k}\sigma} \right\} c^\dagger_{\nu'-q,-s} \right) c_{\nu',-s} c_{\nu,s} \\ &= \left(\delta_{\nu',\boldsymbol{k}+q} \delta_{\sigma,-s} c^\dagger_{\nu+q,s} - \delta_{\nu,\boldsymbol{k}-q} \delta_{\sigma,s} c^\dagger_{\nu'-q,-s} \right) c_{\nu',-s} c_{\nu,s}. \end{aligned} \tag{D.9}$$

Performing the sums over ν, ν' and s in Eq. (D.6) gives

$$\begin{aligned} [H, c_{\boldsymbol{k}\sigma}] &= -\xi_{\boldsymbol{k}} c_{\boldsymbol{k}\sigma} + \frac{U}{4N} \sum_{\boldsymbol{pq}} \left(c^\dagger_{\boldsymbol{p}+\boldsymbol{q},-\sigma} c_{\boldsymbol{k}+\boldsymbol{q},\sigma} c_{\boldsymbol{p},-\sigma} - c^\dagger_{\boldsymbol{p}-\boldsymbol{q},-\sigma} c_{\boldsymbol{p},-\sigma} c_{\boldsymbol{k}-\boldsymbol{q},\sigma} \right) \\ &= -\xi_{\boldsymbol{k}} c_{\boldsymbol{k}\sigma} - \frac{U}{2N} \sum_{\boldsymbol{pq}} c^\dagger_{\boldsymbol{p}+\boldsymbol{q},-\sigma} c_{\boldsymbol{p},-\sigma} c_{\boldsymbol{k}+\boldsymbol{q},\sigma}, \end{aligned} \tag{D.10}$$

where we have let $\boldsymbol{q} \to -\boldsymbol{q}$ in the second term and anti-commuted the annihilation operators in the first term. Inserting this result back into Eq. (D.5) gives

$$\left(\frac{\partial}{\partial\tau} + \xi_{\boldsymbol{k}} \right) G_\sigma(\boldsymbol{k}, \tau) = -\delta(\tau) + \frac{U}{2N} \sum_{\boldsymbol{pq}} \langle T_\tau c^\dagger_{\boldsymbol{p}+\boldsymbol{q},-\sigma}(\tau) c_{\boldsymbol{p},-\sigma}(\tau) c_{\boldsymbol{k}+\boldsymbol{q},\sigma}(\tau) c^\dagger_{\boldsymbol{k},\sigma}(0) \rangle. \tag{D.11}$$

The summand on the RHS can be approximated by decoupling it into two $\langle cc^\dagger \rangle$ terms to give two Green's functions G, analogous to Wick's theorem. However, these are not the only significant combinations of fermion operators. We may also allow $\langle cc \rangle$ and $\langle c^\dagger c^\dagger \rangle$ when decoupling our time-ordered product. Such combinations give rise to so-called Gor'kov anomalous amplitudes F and $F^\dagger$, which are defined as

$$F(\boldsymbol{k}, \tau) \equiv -\langle T_\tau c_{\boldsymbol{k}\uparrow}(\tau) c_{-\boldsymbol{k}\downarrow}(0) \rangle \tag{D.12}$$

and

$$F^\dagger(\boldsymbol{k}, \tau) \equiv -\langle T_\tau c^\dagger_{-\boldsymbol{k}\downarrow}(\tau) c^\dagger_{\boldsymbol{k}\uparrow}(0) \rangle. \tag{D.13}$$

Thus, we get two contributions instead of one

$$\begin{aligned} \langle T_\tau c^\dagger_{\boldsymbol{p}+\boldsymbol{q},-\sigma}(\tau) c_{\boldsymbol{p},-\sigma}(\tau) c_{\boldsymbol{k}+\boldsymbol{q},\sigma}(\tau) c^\dagger_{\boldsymbol{k},\sigma}(0) \rangle &\to -\langle T_\tau c_{\boldsymbol{p},-\sigma}(\tau) c^\dagger_{\boldsymbol{p}+\boldsymbol{q},-\sigma}(\tau) \rangle \langle T_\tau c_{\boldsymbol{k}+\boldsymbol{q},\sigma}(\tau) c^\dagger_{\boldsymbol{k},\sigma}(0) \rangle \\ &\to -\delta_{\boldsymbol{q}\boldsymbol{0}} G_{-\sigma}(\boldsymbol{p}, 0) G_\sigma(\boldsymbol{k}, \tau) \end{aligned} \tag{D.14}$$

and

$$\begin{aligned}\langle T_\tau c^\dagger_{\boldsymbol{p}+\boldsymbol{q},-\sigma}(\tau)c_{\boldsymbol{p},-\sigma}(\tau)c_{\boldsymbol{k}+\boldsymbol{q},\sigma}(\tau)c^\dagger_{\boldsymbol{k},\sigma}(0)\rangle &\to -\langle T_\tau c_{\boldsymbol{k}+\boldsymbol{q},\sigma}(\tau)c_{\boldsymbol{p},-\sigma}(\tau)\rangle\langle T_\tau c^\dagger_{\boldsymbol{p}+\boldsymbol{q},-\sigma}(\tau)c^\dagger_{\boldsymbol{k},\sigma}(0)\rangle \\ &\to -\delta_{\boldsymbol{p},-\boldsymbol{k}-\boldsymbol{q}}F(\sigma[\boldsymbol{k}+\boldsymbol{q}],0)F^\dagger(\sigma\boldsymbol{k},\sigma\tau).\end{aligned} \tag{D.15}$$

Combining both contributions to Eq. (D.5) we get

$$\left(\frac{\partial}{\partial\tau}+\xi_{\boldsymbol{k}}\right)G_\sigma(\boldsymbol{k},\tau) = -\delta(\tau) - \frac{U}{2N}\sum_{\boldsymbol{q}}\left[G_{-\sigma}(\boldsymbol{q},0)G_\sigma(\boldsymbol{k},\tau)+F(\boldsymbol{q},0)F^\dagger(\sigma\boldsymbol{k},\sigma\tau)\right], \tag{D.16}$$

where we have let $\boldsymbol{p}\to\boldsymbol{q}$ in the G-term and let $\boldsymbol{q}\to\boldsymbol{q}-\boldsymbol{k}$ in the F-term. Moreover, because we can let $\boldsymbol{q}\to-\boldsymbol{q}$, we are free to drop the σ inside F.

Next we derive equations of motion for F and $F^\dagger$. Differentiating them w.r.t. imaginary-time gives

$$\frac{\partial}{\partial\tau}F(\boldsymbol{k},\tau) = -\langle T_\tau\left[H,c_{\boldsymbol{k}\uparrow}(\tau)\right]c_{-\boldsymbol{k}\downarrow}(0)\rangle \tag{D.17}$$

and

$$\frac{\partial}{\partial\tau}F^\dagger(\boldsymbol{k},\tau) = -\langle T_\tau\left[H,c^\dagger_{-\boldsymbol{k}\downarrow}(\tau)\right]c^\dagger_{\boldsymbol{k}\uparrow}(0)\rangle. \tag{D.18}$$

The commutators can be evaluated same as before by using Eq. (D.7). Skipping the intermediate steps gives

$$\left(\frac{\partial}{\partial\tau}+\xi_{\boldsymbol{k}}\right)F(\boldsymbol{k},\tau) = \frac{U}{2N}\sum_{\boldsymbol{pq}}\langle T_\tau c^\dagger_{\boldsymbol{p}+\boldsymbol{q}\downarrow}(\tau)c_{\boldsymbol{p}\downarrow}(\tau)c_{\boldsymbol{k}+\boldsymbol{q}\uparrow}(\tau)c_{-\boldsymbol{k}\downarrow}(0)\rangle \tag{D.19}$$

and

$$\left(\frac{\partial}{\partial\tau}-\xi_{\boldsymbol{k}}\right)F^\dagger(\boldsymbol{k},\tau) = \frac{U}{2N}\sum_{\boldsymbol{pq}}\langle T_\tau c^\dagger_{\boldsymbol{p}+\boldsymbol{q}\uparrow}(\tau)c^\dagger_{-\boldsymbol{k}-\boldsymbol{q}\downarrow}(\tau)c_{\boldsymbol{p}\uparrow}(\tau)c^\dagger_{\boldsymbol{k}\uparrow}(0)\rangle. \tag{D.20}$$

We can decouple the summand in the same way as before. For F we get

$$\begin{aligned}\langle T_\tau c^\dagger_{\boldsymbol{p}+\boldsymbol{q}\downarrow}(\tau)c_{\boldsymbol{p}\downarrow}(\tau)c_{\boldsymbol{k}+\boldsymbol{q}\uparrow}(\tau)c_{-\boldsymbol{k}\downarrow}(0)\rangle &\to -\langle T_\tau c_{\boldsymbol{p}\downarrow}(\tau)c^\dagger_{\boldsymbol{p}+\boldsymbol{q}\downarrow}(\tau)\rangle\langle T_\tau c_{\boldsymbol{k}+\boldsymbol{q}\uparrow}(\tau)c_{-\boldsymbol{k}\downarrow}(0)\rangle \\ &\to -\delta_{\boldsymbol{q}\boldsymbol{0}}G_\downarrow(\boldsymbol{p},0)F(\boldsymbol{k},\tau)\end{aligned} \tag{D.21}$$

and

$$\begin{aligned}\langle T_\tau c^\dagger_{\boldsymbol{p}+\boldsymbol{q}\downarrow}(\tau)c_{\boldsymbol{p}\downarrow}(\tau)c_{\boldsymbol{k}+\boldsymbol{q}\uparrow}(\tau)c_{-\boldsymbol{k}\downarrow}(0)\rangle &\to \langle T_\tau c_{\boldsymbol{k}+\boldsymbol{q}\uparrow}(\tau)c_{\boldsymbol{p}\downarrow}(\tau)\rangle\langle T_\tau c_{-\boldsymbol{k}\downarrow}(0)c^\dagger_{\boldsymbol{p}+\boldsymbol{q}\downarrow}(\tau)\rangle \\ &\to \delta_{\boldsymbol{p},-\boldsymbol{k}-\boldsymbol{q}}F(\boldsymbol{k}+\boldsymbol{q},0)G_\downarrow(-\boldsymbol{k},-\tau),\end{aligned} \tag{D.22}$$

and the following for $F^\dagger$

$$\begin{aligned}\langle T_\tau c^\dagger_{\boldsymbol{p}+\boldsymbol{q}\uparrow}(\tau)c^\dagger_{-\boldsymbol{k}-\boldsymbol{q}\downarrow}(\tau)c_{\boldsymbol{p}\uparrow}(\tau)c^\dagger_{\boldsymbol{k}\uparrow}(0)\rangle &\to -\langle T_\tau c^\dagger_{-\boldsymbol{k}-\boldsymbol{q}\downarrow}(\tau)c^\dagger_{\boldsymbol{p}+\boldsymbol{q}\uparrow}(\tau)\rangle\langle T_\tau c_{\boldsymbol{p}\uparrow}(\tau)c^\dagger_{\boldsymbol{k}\uparrow}(0)\rangle \\ &\to -\delta_{\boldsymbol{p}\boldsymbol{k}}F^\dagger(\boldsymbol{k}+\boldsymbol{q},0)G_\uparrow(\boldsymbol{k},\tau)\end{aligned} \tag{D.23}$$

and

$$\begin{aligned}\langle T_\tau c^\dagger_{\boldsymbol{p}+\boldsymbol{q}\uparrow}(\tau)c^\dagger_{-\boldsymbol{k}-\boldsymbol{q}\downarrow}(\tau)c_{\boldsymbol{p}\uparrow}(\tau)c^\dagger_{\boldsymbol{k}\uparrow}(0)\rangle &\to \langle T_\tau c_{\boldsymbol{p}\uparrow}(\tau)c^\dagger_{\boldsymbol{p}+\boldsymbol{q}\uparrow}(\tau)\rangle\langle T_\tau c^\dagger_{-\boldsymbol{k}-\boldsymbol{q}\downarrow}(\tau)c^\dagger_{\boldsymbol{k}\uparrow}(0)\rangle \\ &\to \delta_{\boldsymbol{q}\boldsymbol{0}}G_\uparrow(\boldsymbol{p},0)F^\dagger(\boldsymbol{k},\tau).\end{aligned} \tag{D.24}$$

Combining these contributions into Eqs. (D.19) and (D.20) gives

$$\left(\frac{\partial}{\partial\tau}+\xi_{\boldsymbol{k}}\right)F(\boldsymbol{k},\tau)=\frac{U}{2N}\sum_{\boldsymbol{q}}\Big[-G_\downarrow(\boldsymbol{q},0)F(\boldsymbol{k},\tau)+F(\boldsymbol{q},0)G_\downarrow(-\boldsymbol{k},-\tau)\Big] \tag{D.25}$$

and

$$\left(\frac{\partial}{\partial\tau}-\xi_{\boldsymbol{k}}\right)F^\dagger(\boldsymbol{k},\tau)=\frac{U}{2N}\sum_{\boldsymbol{q}}\Big[G_\uparrow(\boldsymbol{q},0)F^\dagger(\boldsymbol{k},\tau)-F^\dagger(\boldsymbol{q},0)G_\uparrow(\boldsymbol{k},\tau)\Big], \tag{D.26}$$

where we have let $\boldsymbol{q}\to\boldsymbol{q}-\boldsymbol{k}$ in the last term in both equations.

Next we Fourier transform Eqs. (D.16), (D.25) and (D.26). The transform is defined by the following equations

$$G(\boldsymbol{k},\tau)=\frac{1}{\beta}\sum_{m\in\mathbb{Z}}\mathrm{e}^{-i\omega_m\tau}G(\boldsymbol{k},i\omega_m) \tag{D.27}$$

$$G(\boldsymbol{k},i\omega_m)=\int_0^\beta \mathrm{d}\tau G(\boldsymbol{k},\tau)\mathrm{e}^{i\omega_m\tau}, \tag{D.28}$$

and similarly for F and $F^\dagger$. The Matsubara frequency ω_n is defined as $\omega_n=(2n+1)\pi/\beta$ for fermions. Fourier transforming Eqs. (D.16), (D.25) and (D.26) gives the following equations

$$(-i\omega_n+\xi_{\boldsymbol{k}})\,G_\sigma(\boldsymbol{k},i\omega_n)=-1-\frac{1}{\beta}\frac{U}{2N}\sum_{\boldsymbol{q},m}\left[G_{-\sigma}(\boldsymbol{q},i\omega_m)G_\sigma(\boldsymbol{k},i\omega_n)+F(\boldsymbol{q},i\omega_m)F^\dagger(\sigma\boldsymbol{k},\sigma i\omega_n)\right], \tag{D.29}$$

$$(-i\omega_n+\xi_{\boldsymbol{k}})\,F(\boldsymbol{k},i\omega_n)=\frac{1}{\beta}\frac{U}{2N}\sum_{\boldsymbol{q},m}\Big[-G_\downarrow(\boldsymbol{q},i\omega_m)F(\boldsymbol{k},i\omega_n)+F(\boldsymbol{q},i\omega_m)G_\downarrow(-\boldsymbol{k},-i\omega_n)\Big], \tag{D.30}$$

and

$$(-i\omega_n-\xi_{\boldsymbol{k}})\,F^\dagger(\boldsymbol{k},i\omega_n)=\frac{1}{\beta}\frac{U}{2N}\sum_{\boldsymbol{q},m}\Big[G_\uparrow(\boldsymbol{q},i\omega_m)F^\dagger(\boldsymbol{k},i\omega_n)-F^\dagger(\boldsymbol{q},i\omega_m)G_\uparrow(\boldsymbol{k},i\omega_n)\Big]. \tag{D.31}$$

Looking at the above equations we see a few reoccurring sums

$$S_\sigma\equiv-\frac{1}{\beta}\frac{U}{2N}\sum_{\boldsymbol{q},m}G_\sigma(\boldsymbol{q},i\omega_m) \tag{D.32}$$

$$W\equiv-\frac{1}{\beta}\frac{U}{2N}\sum_{\boldsymbol{q},m}F(\boldsymbol{q},i\omega_m) \tag{D.33}$$

$$W^\dagger\equiv-\frac{1}{\beta}\frac{U}{2N}\sum_{\boldsymbol{q},m}F^\dagger(\boldsymbol{q},i\omega_m). \tag{D.34}$$

We can thus re-express Eqs. (D.29) to (D.31) using S, W and $W^\dagger$

$$(i\omega_n - \xi_{\boldsymbol{k}} + S_{-\sigma})\, G_\sigma(\boldsymbol{k}, i\omega_n) + W F^\dagger(\sigma\boldsymbol{k}, \sigma i\omega_n) = 1, \tag{D.35}$$

$$(i\omega_n - \xi_{\boldsymbol{k}} + S_\downarrow)\, F(\boldsymbol{k}, i\omega_n) - W G_\downarrow(-\boldsymbol{k}, -i\omega_n) = 0, \tag{D.36}$$

and

$$(-i\omega_n - \xi_{\boldsymbol{k}} + S_\uparrow)\, F^\dagger(\boldsymbol{k}, i\omega_n) - W^\dagger G_\uparrow(\boldsymbol{k}, i\omega_n) = 0. \tag{D.37}$$

We can find expressions for G and $F^\dagger$ by solving Eqs. (D.35) and (D.37). Doing so gives

$$G_\uparrow(\boldsymbol{k}, i\omega_n) = \frac{i\omega_n + \xi_{\boldsymbol{k}} - S_\uparrow}{(i\omega_n)^2 - i\omega_n(S_\uparrow - S_\downarrow) - \xi_{\boldsymbol{k}}^2 + \xi_{\boldsymbol{k}}(S_\uparrow + S_\downarrow) - S_\uparrow S_\downarrow - |W|^2} \tag{D.38}$$

and

$$F^\dagger(\boldsymbol{k}, i\omega_n) = \frac{-W^\dagger}{(i\omega_n)^2 - i\omega_n(S_\uparrow - S_\downarrow) - \xi_{\boldsymbol{k}}^2 + \xi_{\boldsymbol{k}}(S_\uparrow + S_\downarrow) - S_\uparrow S_\downarrow - |W|^2}. \tag{D.39}$$

If we assume no spin dependent forces, we can set $G_\sigma = G$ and $S_\sigma = S$

$$G(\boldsymbol{k}, i\omega_n) = \frac{i\omega_n + \xi_{\boldsymbol{k}} - S}{(i\omega_n)^2 - (\xi_{\boldsymbol{k}} - S)^2 - |W|^2} \tag{D.40}$$

$$F^\dagger(\boldsymbol{k}, i\omega_n) = \frac{-W^\dagger}{(i\omega_n)^2 - (\xi_{\boldsymbol{k}} - S)^2 - |W|^2}. \tag{D.41}$$

With these equations for G and $F^\dagger$, and with $F^\dagger$ the complex conjugate of F, we are in a position to set up equations for the self-energies S and W

$$S = -\frac{1}{\beta}\frac{U}{2N}\sum_{\boldsymbol{k},n} \frac{i\omega_n + \xi_{\boldsymbol{k}} - S}{(i\omega_n)^2 - (\xi_{\boldsymbol{k}} - S)^2 - |W|^2} \tag{D.42}$$

$$W = -\frac{1}{\beta}\frac{U}{2N}\sum_{\boldsymbol{k},n} \frac{-W}{(i\omega_n)^2 - (\xi_{\boldsymbol{k}} - S)^2 - |W|^2}. \tag{D.43}$$

D.2 Matsubara sum

The next step is to carry out the sums over the Matsubara frequencies $\omega_n = (2n+1)\pi/\beta$. A method of evaluating Matsubara sums can be found in section 25 of [33] and in section 4.2 of [46]. Let us identify two kinds of sums in Eqs. (D.42) and (D.43)

$$S_1 \equiv \sum_n \frac{i\omega_n}{(i\omega_n)^2 - (\xi_{\boldsymbol{q}} - S)^2 - |W|^2} \tag{D.44}$$

and

$$S_2 \equiv \sum_n \frac{1}{(i\omega_n)^2 - (\xi_{\boldsymbol{q}} - S)^2 - |W|^2}. \tag{D.45}$$

We will evaluate S_1 in detail[2], which we rewrite as

[2] Evaluation of S_2 is similar to S_1.

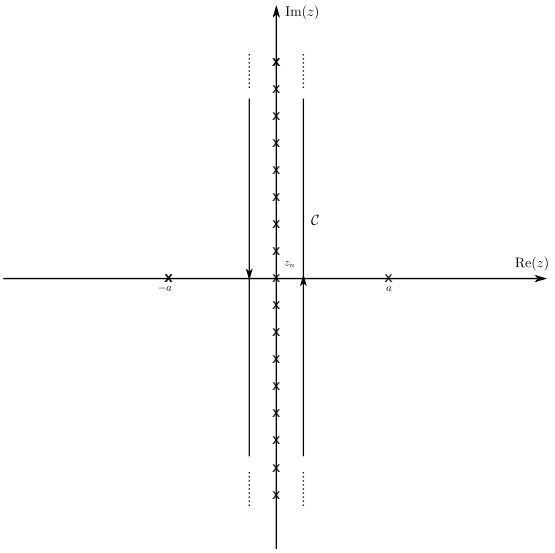


Figure D.1: Contour $\mathcal{C}$ enclosing all poles z_n (denoted by crosses) of the counting function $-\beta(\mathrm{e}^{\beta z}+1)^{-1}$ along the imaginary axis. $z=\pm a$ are the poles of the original summand.

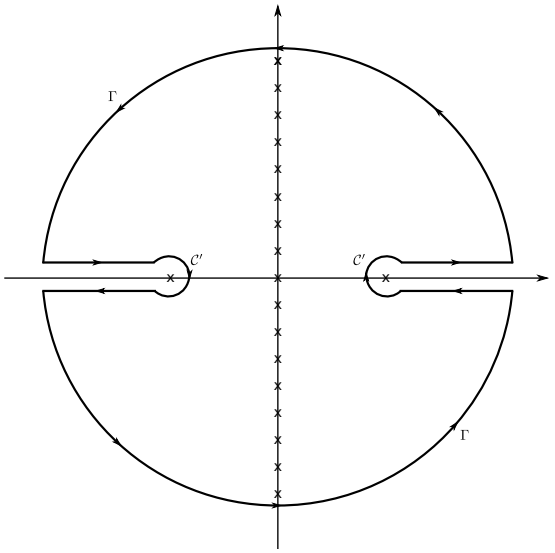


Figure D.2: $\mathcal{C}$ inflated to Γ and $\mathcal{C}'$. This contour encloses the same poles as $\mathcal{C}$ when blown up to infinity. In the limit $|z|\to\infty$, the integral over Γ decays and only a clock-wise integration along $\mathcal{C}'$ contributes.

$$S_1=\sum_n \frac{i\omega_n}{(i\omega_n)^2-a^2} \tag{D.46}$$

where $a\equiv\sqrt{(\xi_{\boldsymbol{q}}-S)^2+|W|^2}$. The method involves rewriting the sum as a contour integral of a product between the summand and a "counting function" which has poles $z_n=i\omega_n$ along the entire imaginary axis. One such function, for fermions, is $g(z)=-\beta(\mathrm{e}^{\beta z}+1)^{-1}$. This function is essentially the Fermi-Dirac distribution function and has unit residue at each pole. Thus, we can write S_1 as

$$S_1=\sum_n \frac{i\omega_n}{(i\omega_n)^2-a^2}=\frac{-\beta}{2\pi i}\oint_{\mathcal{C}}\frac{\mathrm{d}z}{\mathrm{e}^{\beta z}+1}\frac{z}{z^2-a^2} \tag{D.47}$$

where $z=i\omega_n$. The contour $\mathcal{C}$ is illustrated in Fig. D.1. Since the contour encloses the poles of g, the integral reduces to a sum over residues. g is constructed to have unit residue at each pole, which means the sum over residues becomes a sum over ω_n, and thus we get S_1. We note that S_1 is on the edge of convergence. To push it over, we introduce a convergence factor $\mathrm{e}^{i\omega_n\eta}$ where $\eta=0^+$. With this factor, S_1 takes the form

$$S_1=\lim_{\eta\to0^+}\sum_n \frac{i\omega_n\mathrm{e}^{i\omega_n\eta}}{(i\omega_n)^2-a^2}=\lim_{\eta\to0^+}\frac{-\beta}{2\pi i}\oint_{\mathcal{C}}\frac{\mathrm{d}z}{\mathrm{e}^{\beta z}+1}\frac{z\mathrm{e}^{z\eta}}{z^2-a^2}. \tag{D.48}$$

Next we inflate $\mathcal{C}$ to Γ and $\mathcal{C}'$ as shown in Fig. D.2. This contour encloses the same poles as $\mathcal{C}$. We note that in the limit $|z|\to\infty$, the contribution from Γ vanishes. We see this by studying the order of the integrand for both cases $\mathrm{Re}(z)>0$ and $\mathrm{Re}(z)<0$[3]. For $\mathrm{Re}(z)>0$ the integrand is of order $|z|^{-1}\exp[-\mathrm{Re}(z)(\beta-\eta)]$, and for $\mathrm{Re}(z)<0$ it is of the order $|z|^{-1}\exp[\mathrm{Re}(z)\eta]$. In both cases, the integrand decays as $|z|\to\infty$, i.e. we are left with an integral along only $\mathcal{C}'$

[3]The sign of $\mathrm{Im}(z)$ is not of interest because the imaginary part only contributes to oscillations.

$$S_1 = \lim_{\eta\to 0^+} \frac{-\beta}{2\pi i} \oint_{\mathcal{C}'} \frac{\mathrm{d}z}{\mathrm{e}^{\beta z}+1} \frac{z\mathrm{e}^{z\eta}}{z^2-a^2}. \tag{D.49}$$

This has reduced the problem from a sum over infinite poles to a sum over only two simple poles enclosed by $\mathcal{C}'$, namely $z = \pm a$. By use of the residue theorem, S_1 becomes[4]

$$S_1 = \lim_{\eta\to 0^+} \sum_{z_k=\pm a} \mathrm{Res}\left[\frac{\beta}{\mathrm{e}^{\beta z_k}+1} \frac{z_k \mathrm{e}^{z_k\eta}}{(z_k-a)(z_k+a)}\right] = \frac{\beta}{2}. \tag{D.50}$$

A similar derivation for S_2 yields

$$S_2 = \lim_{\eta\to 0^+} \sum_{z_k=\pm a} \mathrm{Res}\left[\frac{\beta}{\mathrm{e}^{\beta z_k}+1} \frac{\mathrm{e}^{z_k\eta}}{(z_k-a)(z_k+a)}\right] = -\beta \frac{\tanh(\beta a/2)}{2a}. \tag{D.51}$$

Using these results for S_1 and S_2, we can rewrite Eqs. (D.42) and (D.43)

$$S = -\frac{U}{4} + \frac{U}{2N}\sum_{\boldsymbol{k}} (\xi_{\boldsymbol{k}} - S) \frac{\tanh\left(\frac{\beta}{2}\sqrt{(\xi_{\boldsymbol{k}}-S)^2+|W|^2}\right)}{2\sqrt{(\xi_{\boldsymbol{k}}-S)^2+|W|^2}} \tag{D.52}$$

$$1 = -\frac{U}{2N}\sum_{\boldsymbol{k}} \frac{\tanh\left(\frac{\beta}{2}\sqrt{(\xi_{\boldsymbol{k}}-S)^2+|W|^2}\right)}{2\sqrt{(\xi_{\boldsymbol{k}}-S)^2+|W|^2}}. \tag{D.53}$$

These are the equations we need to solve for S and W as a function of T and U. However, they need to be simplified by evaluating the $\boldsymbol{k}$-sums first.

D.3 Energy integral

The sum in Eq. (D.52) can be evaluated analytically if converted to an energy integral. We will assume an isotropic Fermi surface and rewrite the sum as

$$\sum_{\boldsymbol{k}} \to \int \mathrm{d}\epsilon D(\epsilon) \approx D(\epsilon_{\mathrm{F}}) \int_{-\omega_D}^{\omega_D} \mathrm{d}\epsilon, \tag{D.54}$$

where ω_D is the Debye frequency and $D(\epsilon)$ is the density of states. The last approximation is valid if we stay close to the Fermi surface ϵ_{F}. Using Eq. (D.54), Eq. (D.52) becomes

$$S = -\frac{U}{4} + \frac{UD(\epsilon_{\mathrm{F}})}{2N} \int_{-\omega_D}^{\omega_D} \mathrm{d}\epsilon\ (\epsilon - S) \frac{\tanh\left(\frac{\beta}{2}\sqrt{(\epsilon-S)^2+|W|^2}\right)}{2\sqrt{(\epsilon-S)^2+|W|^2}}. \tag{D.55}$$

A change of variables $x(\epsilon) \equiv \frac{\beta}{2}\sqrt{(\epsilon-S)^2+|W|^2}$ gives

$$S = -\frac{U}{4} - \frac{\lambda}{\beta} \int_{x(-\omega_D)}^{x(\omega_D)} \mathrm{d}x\ \tanh(x), \tag{D.56}$$

where we have defined the positive coupling constant $\lambda \equiv -UD(\epsilon_{\mathrm{F}})/2N$. Performing the integral gives us the final form of Eq. (D.52)

[4] Change of sign due to the negative nature of $\mathcal{C}'$.

$$S = -\frac{U}{4} - \frac{\lambda}{\beta} \ln \left[\frac{\cosh\left(\frac{\beta}{2}\sqrt{(\omega_D - S)^2 + |W|^2}\right)}{\cosh\left(\frac{\beta}{2}\sqrt{(\omega_D + S)^2 + |W|^2}\right)} \right]. \tag{D.57}$$

Looking back on Eq. (D.53) we see its resemblance to the BCS gap equation. In fact, if we set $S = 0$ the BCS gap Δ and W are equivalent. Just like in BCS, the sum in Eq. (D.53) cannot be evaluated analytically for all β. Hence, we need to consider two temperature limits, namely $T \to 0$ and $T \to T_c$.

D.3.1 $T \to 0$

In the limit $T \to 0$, β grows to infinity. For $\beta x \gg 1$, $\cosh(\beta x) \to \exp(\beta x)/2$ and $\tanh(\beta x) \to 1$. With this in mind Eqs. (D.53) and (D.57) become

$$1 = \frac{\lambda}{2} \int_{-\omega_D}^{\omega_D} \mathrm{d}\epsilon \, \frac{1}{\sqrt{(\epsilon - S)^2 + |W|^2}}, \tag{D.58}$$

and

$$S = -\frac{U}{4} - \frac{\lambda}{2}\left(\sqrt{(\omega_D - S)^2 + |W|^2} - \sqrt{(\omega_D + S)^2 + |W|^2}\right) \tag{D.59}$$

respectively. The integral in Eq. (D.58) can be performed analytically by defining the variable $x(\epsilon) \equiv (\epsilon - S)/|W|$

$$\begin{aligned} 1 &= \frac{\lambda}{2} \int_{x(-\omega_D)}^{x(\omega_D)} \mathrm{d}x \, \frac{1}{\sqrt{1 + x^2}} \\ &= \frac{\lambda}{2} \left[\sinh^{-1}\left(\frac{\omega_D - S}{|W|}\right) + \sinh^{-1}\left(\frac{\omega_D + S}{|W|}\right) \right]. \end{aligned} \tag{D.60}$$

Summarizing both equations for completeness gives

$$S = -\frac{U}{4} - \frac{\lambda}{2}\left(\sqrt{(\omega_D - S)^2 + |W|^2} - \sqrt{(\omega_D + S)^2 + |W|^2}\right) \tag{D.61}$$

and

$$1 = \frac{\lambda}{2} \left[\sinh^{-1}\left(\frac{\omega_D - S}{|W|}\right) + \sinh^{-1}\left(\frac{\omega_D + S}{|W|}\right) \right]. \tag{D.62}$$

These equations can be solved numerically to find both S and W at $T = 0$ as a function of e.g. U.

D.3.2 $T \to T_c$

A superconductor transitions to a metal when $T > T_c$. Which means that at the limit $T = T_c$, the superconducting gap W must go to zero. We can thus rewrite Eqs. (D.53) and (D.57) to

$$1 = \frac{\lambda}{2} \int_{-\omega_D}^{\omega_D} \mathrm{d}\epsilon \, \frac{\tanh\left(\frac{\beta}{2}|\epsilon - S|\right)}{|\epsilon - S|}, \tag{D.63}$$

and

$$S = -\frac{U}{4} - \frac{\lambda}{\beta} \ln \left[\frac{\cosh\left(\frac{\beta}{2}|\omega_D - S|\right)}{\cosh\left(\frac{\beta}{2}|\omega_D + S|\right)} \right]. \tag{D.64}$$

where $\beta = 1/k_{\mathrm{B}}T_c$. The strategy to solve Eq. (D.63) is to first do a change of variables, then split up the integral limits, and finally integrate by parts. We begin by defining the new variable $x(\epsilon) \equiv \frac{\beta}{2}|\epsilon - S|$ with $\mathrm{d}x/\mathrm{d}\epsilon = \frac{\beta}{2}\mathrm{sgn}(\epsilon - S)$, where $\mathrm{sgn}(a)$ gives the sign of a. Splitting the integration limits gives

$$1 = \frac{\lambda}{2} \left[\mathrm{sgn}(-\omega_D - S) \int_{\frac{\beta}{2}|-\omega_D - S|}^{0} \mathrm{d}x \ \frac{\tanh(x)}{x} + \mathrm{sgn}(\omega_D - S) \int_{0}^{\frac{\beta}{2}|\omega_D - S|} \mathrm{d}x \ \frac{\tanh(x)}{x} \right]. \tag{D.65}$$

Both integrals can be evaluated in the same manner. Let us focus on the last integral. We continue by first integrating it by parts

$$\begin{aligned} \int_{0}^{\frac{\beta}{2}|\omega_D - S|} \mathrm{d}x \ \frac{\tanh(x)}{x} &= \tanh(x)\ln(x)\Big|_{0}^{\frac{\beta}{2}|\omega_D - S|} - \int_{0}^{\frac{\beta}{2}|\omega_D - S|} \frac{\ln(x)}{\cosh^2(x)} \\ &= \tanh\left(\frac{\beta}{2}|\omega_D - S|\right) \ln\left(\frac{\beta}{2}|\omega_D - S|\right) - \ln C. \end{aligned} \tag{D.66}$$

In the last integral we have assumed that $\frac{\beta}{2}|\omega_D - S| \gg 1$ and thus set the upper limit of integration to infinity. This is a reasonable decision because the integrand $\ln x/\cosh^2 x$ is rapidly convergent. With ∞ as the upper limit, the last integral is well known with $C = \pi \mathrm{e}^{-\gamma}/4$.[5] Note that we have not let $\frac{\beta}{2}|\omega_D - S| \to \infty$ in the first term as it is logarithmically divergent. Using this result, we arrive at the final form of Eq. (D.63)

$$\begin{aligned} \frac{2}{\lambda} = \mathrm{sgn}(\omega_D - S) &\tanh\left(\frac{\beta}{2}|\omega_D - S|\right) \ln\left(\frac{\beta}{2}|\omega_D - S|\right) + \\ &\tanh\left(\frac{\beta}{2}|\omega_D + S|\right) \ln\left(\frac{\beta}{2}|\omega_D + S|\right) - \\ &[\mathrm{sgn}(\omega_D - S) + 1] \ln C. \end{aligned} \tag{D.67}$$

Equation (D.67) and

$$S = -\frac{U}{4} - \frac{\lambda}{\beta} \ln \left[\frac{\cosh\left(\frac{\beta}{2}|\omega_D - S|\right)}{\cosh\left(\frac{\beta}{2}|\omega_D + S|\right)} \right] \tag{D.68}$$

makes up the set of equations to solve for S and T_c as a function of e.g. U.

[5] γ is called Euler's constant and is $\gamma \approx 0.577215$.

Bibliography

[1] H. K. Onnes, Comm. Phys. Lab. Univ. Leiden **120b, 122b, 124c** (1911).

[2] W. Meissner and R. Ochsenfeld, Naturwissenschaften **21**, 787 (1933).

[3] V. L. Ginzburg and L. D. Landau, J. Exp. Theor. Phys. USSR **20**, 1064 (1950).

[4] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, Phys. Rev. **108**, 1175 (1957).

[5] I. Giæver, Phys. Rev. Lett. **5**, 147 (1960).

[6] I. Giæver, Phys. Rev. Lett. **5**, 464 (1960).

[7] I. Giæver and K. Megerle, Phys. Rev. **122**, 1101 (1961).

[8] B. D. Josephson, Physics Letters **1**, 251 (1962).

[9] L. Gor'Kov, Sov. Phys. JETP **7**, 158 (1958).

[10] G. Eliashberg, Sov. Phys. JETP **11**, 696 (1960).

[11] G. Eliashberg, Sov. Phys. JETP **12**, 1000 (1961).

[12] J. G. Bednorz and K. A. Müller, Zeitschrift für Physik B Condensed Matter **64**, 189 (1986).

[13] G. R. Stewart, Advances in Physics **66**, 75–196 (2017).

[14] J. G. Bednorz and K. A. Müller, "Perovskite-Type Oxides — The New Approach to High-Tc Superconductivity," (1987), Nobel lecture.

[15] L. D. Landau, Ukr. J. Phys. **11**, 19 (1937).

[16] S. Elitzur, Phys. Rev. D **12**, 3978 (1975).

[17] K. v. Klitzing, G. Dorda, and M. Pepper, Phys. Rev. Lett. **45**, 494 (1980).

[18] E. H. Hall, American Journal of Mathematics **2**, 287 (1879).

[19] D. J. Thouless, M. Kohmoto, M. P. Nightingale, and M. den Nijs, Phys. Rev. Lett. **49**, 405 (1982).

[20] F. D. M. Haldane, Phys. Rev. Lett. **61**, 2015 (1988).

[21] The Royal Swedish Academy of Sciences, "Topological phase transitions and topological phases of matter," (2016), Scientific Background on the Nobel Prize in Physics 2016.

[22] M. Z. Hasan and C. L. Kane, Rev. Mod. Phys. **82**, 3045 (2010).

[23] S. Rachel, Reports on Progress in Physics **81**, 116501 (2018).

[24] V. Gurarie, Physical Review B **83** (2011).

[25] N. W. Ashcroft and N. D. Mermin, *Solid state physics* (Harcourt College, 1976).

[26] J. Hubbard and B. H. Flowers, Proc. Roy. Soc. London. Series A. Math. Phys. Sci. **276**, 238 (1963).

[27] K. Fossheim and A. Sudbø, *Superconductivity: physics and applications* (John Wiley & Sons, 2004).

[28] L. N. Cooper, Phys. Rev. **104**, 1189 (1956).

[29] S. N. Bose, Zeitschrift für Physik **26**, 178 (1924).

[30] A. Einstein, Königliche Preußische Akademie der Wissenschaften **261-267** (1924).

[31] P. J. Ford and G. A. Saunders, *The Rise of the Superconductors* (CRC Press, 2005).

[32] G. A. Ummarino, "Eliashberg theory — Autumn school on correlated electrons: Emergent phenomena in correlated matter," (2013).

[33] A. L. Fetter and J. D. Walecka, *Quantum theory of many-particle systems* (Dover Publications, 2003).

[34] G. D. Mahan, *Many-particle physics*, 3rd ed., Physics of solids and liquids (Kluwer Academic/Plenum, New York, 2000).

[35] M. E. Peskin and D. V. Schroeder, *An introduction to quantum field theory* (Addison-Wesley, 1995).

[36] A. Migdal, Sov. Phys. JETP **7**, 996 (1958).

[37] J. E. Moore, Nature **464**, 194 (2010).

[38] M. V. Berry, Proc. R. Soc. Lond. A. **392** (1984).

[39] B. R. Holstein, American Journal of Physics **57**, 1079 (1989).

[40] C. Kittel, *Introduction to solid state physics* (John Wiley & Sons, 2005).

[41] E. Thingstad, A. Kamra, J. W. Wells, and A. Sudbø, Phys. Rev. B **101**, 214513 (2020).

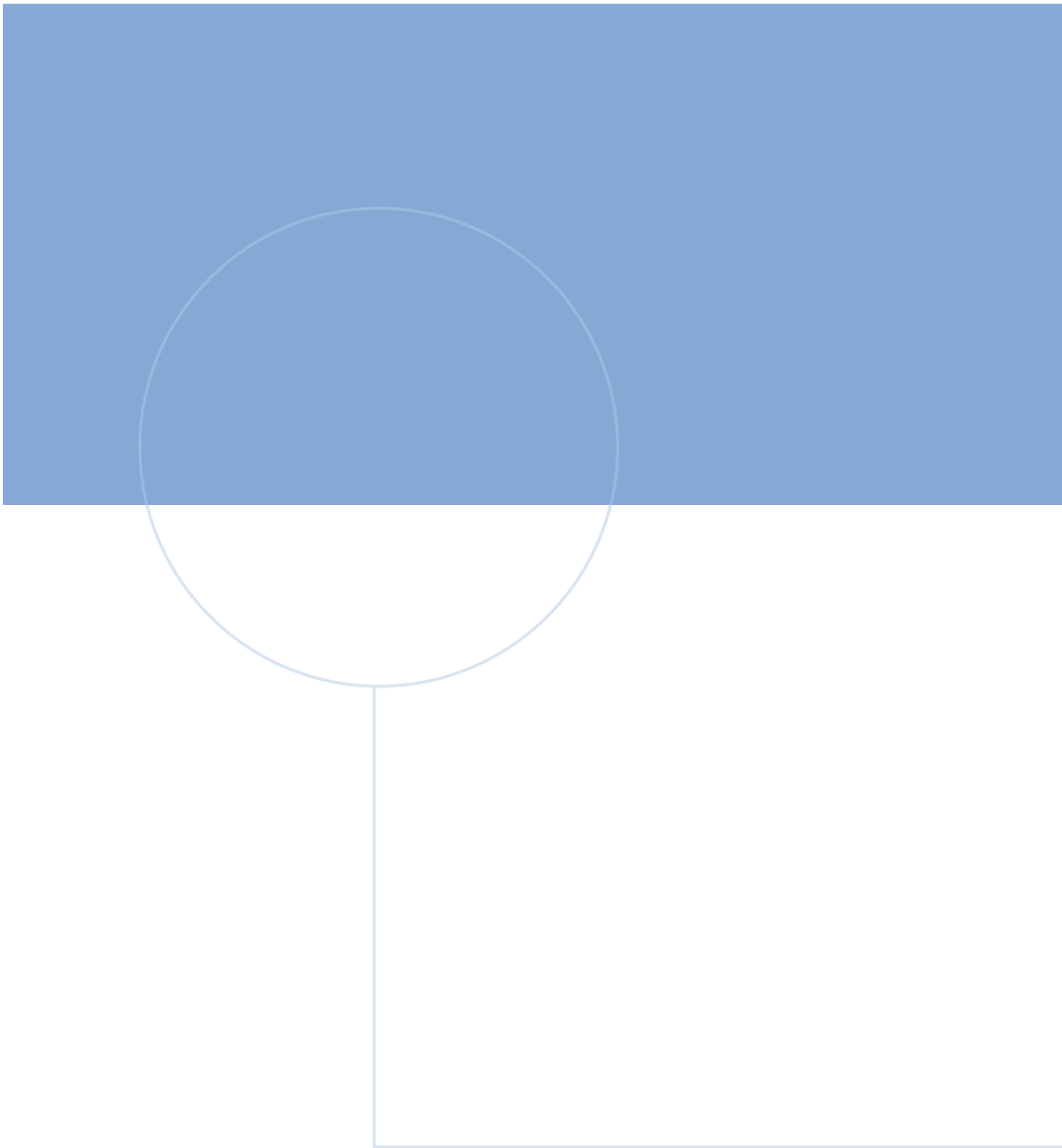
[42] F. Marsiglio, Annals of Physics , 168102 (2020).

[43] A. Karakozov, E. Maksimov, and S. Mashkov, Zh. Eksp. Teor. Fiz **68**, 1937 (1975).

[44] F. Marsiglio, Phys. Rev. B **98**, 024523 (2018).

[45] K. Sun, "Topological insulator (band structure theory) — lecture notes for physics 620," (2012).

[46] A. Altland and B. Simons, *Condensed Matter Field Theory* (Cambridge University Press, 2010).



NTNU

Norwegian University of
Science and Technology