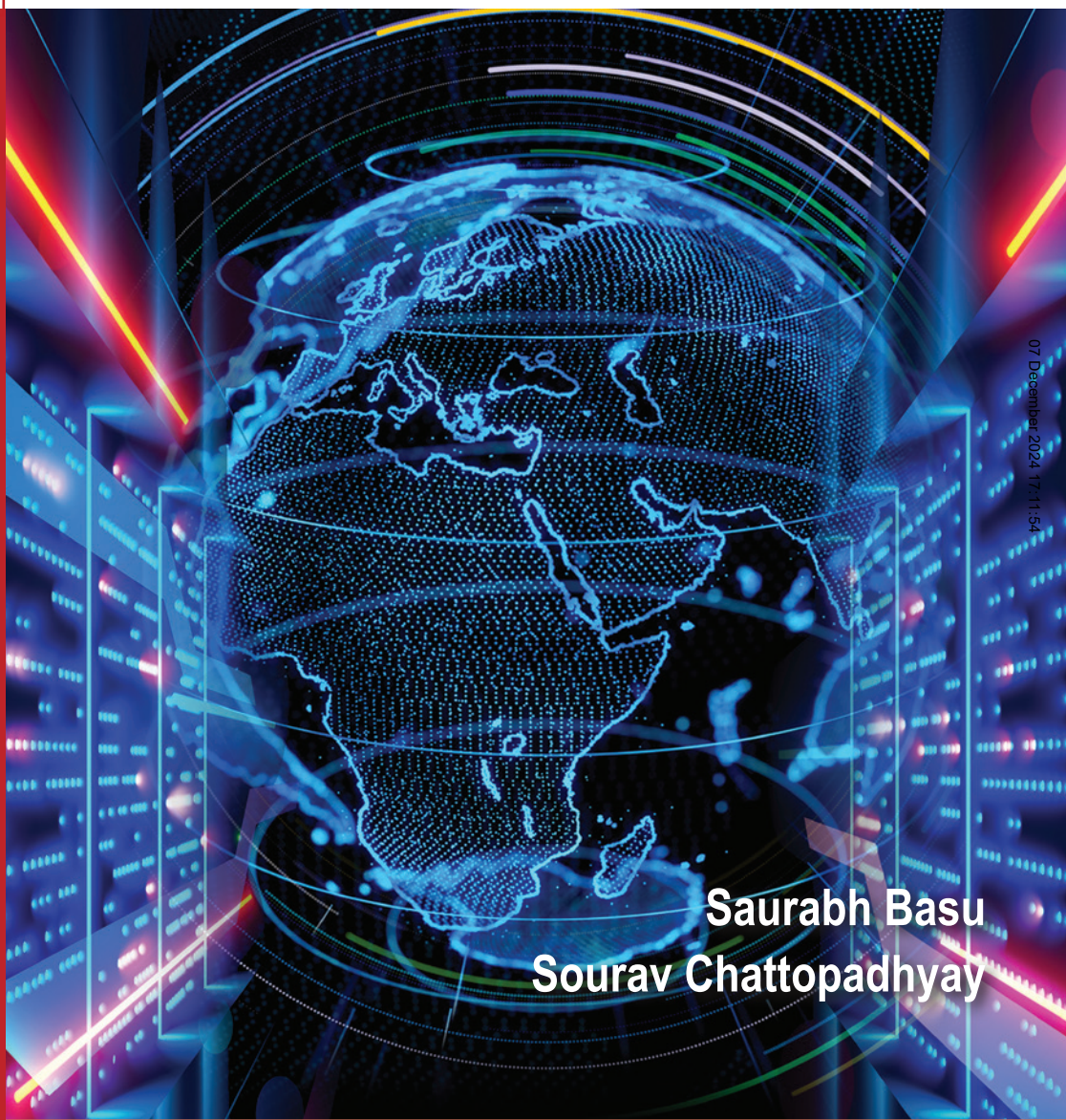
Principles

Modern Perspectives in the Study of Electronic Systems



Saurabh Basu
Sourav Chattopadhyay

AIP Publishing

scitation.org/books

Modern Perspectives in the Study of Electronic Systems

Modern Perspectives in the Study of Electronic Systems

Saurabh Basu
Department of Physics, IIT Guwahati,
Guwahati, Assam, Pin 781039, India

Sourav Chattopadhyay
Department of Physics, The ICFAI University of Tripura,
Agartala, Pin 799210, India

AIP Publishing Books
A publication of AIP Publishing
Melville, New York

First edition, published 2022.

Library of Congress Control Number: 2022944651

ISBN: 978-0-7354-2251-3 (Softcover)
ISBN: 978-0-7354-2253-7 (Online)
ISBN: 978-0-7354-2250-6 (ePub)
ISBN: 978-0-7354-2252-0 (ePDF)

Set in 10/13pt MinionPro by Nova Techset Private Limited, Bengaluru & Chennai, India, and bound by The Sheridan Group, USA

Published by AIP Publishing
1305 Walt Whitman Road, Suite 110, Melville, NY 11747-4300, USA

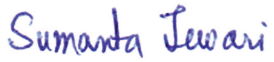
To our parents....

FOREWORD

The book titled *Modern Perspectives in the Study of Electronic Systems* by Professor Saurabh Basu and Dr. Sourav Chattopadhyay consists of a set of rather important topics that are at the heart of modern Condensed Matter Physics. It presents a detailed and pedagogical analysis of several important subjects, such as the Quantum Hall effect, Topology, and systems where interparticle interactions play an important role, such as Magnetism and Superconductivity. The choice of topics and the discussions therein seem perfect in bringing the key elements together that the students should learn amidst the vast literature that exists in today's world. The style of presentation is also very clear and should be helpful for self-study by the students.

The book is expected to be useful for advanced undergraduate students and also the beginners in graduate studies who aim to take up condensed matter for research in the future. The topics are lucidly described and each chapter provides a basic introduction to the subject, which should be useful while delving deeper into it during graduate study. There is always a thin gap that exists between undergraduate training deployed in the classrooms and graduate programs in a university or research institute, which the students have the responsibility to bridge on their own. I feel that this book will aid them in achieving this task in a fluent and successful manner.

Overall, I find this book to be a useful addition to the subject of Condensed Matter Physics and recommend it without hesitation to the advanced undergraduate and graduate students.



Dr. Sumanta Tewari
Professor
Department of Physics and Astronomy
Clemson University
Clemson, South Carolina, USA

PREFACE

It is somewhat implicit that the readers are familiar with the first course on quantum mechanics, which mainly deals with the properties of a single and non-relativistic particle in the presence of a given potential that usually has a simple form. We give a brief recap of some of these problems below. Readers are encouraged to look at the classic texts on quantum physics. For example, the eigensolutions (ϵ_n, ψ_n) of a free particle confined in an infinite potential well are found to be $\epsilon_n \sim n^2$, and $\phi_n \sim \sin \frac{n\pi x}{L}/\cos \frac{n\pi x}{L}$ or, in an infinite space, $\phi_n \sim e^{in\pi x/L}$ (n being the quantum number for the problem). Whereas, a parabolic potential (simple harmonic oscillator) yields an equidistant energy spectrum of the form, $\epsilon_n = (n + \frac{1}{2})\hbar\omega$, and the eigenfunctions are denoted by a Gaussian ($\sim e^{-x^2}$) multiplied by a polynomial (Hermite polynomial) which possesses an even or odd parity depending on whether n is even or odd. Such an even-odd nature of the eigenfunctions is an artefact of the symmetric potential, which allows both even and odd solutions. Hence, the Hilbert space gets fragmented into one half for even n, and the other half for odd n. Further, in a three-dimensional case, which is more complicated than its one dimensional counterpart, in the presence of a Coulomb potential, appropriate for a hydrogen (H) atom which has only one electron, the energy spectrum of an electron takes a form, $\epsilon_n = -\frac{13.6}{n^2}$ eV, where n denotes the principal quantum number.

Furthermore, the barrier transmission problems in quantum mechanics may find potential applications in the transport properties of semiconducting heterostructures. There are a variety of barriers, such as a finite step, finite well, or an infinite potential discontinuity (δ-function) etc. where the formulae of finding the reflection and the transmission coefficients are prescribed by matching the boundary conditions of the wavefunctions and their derivatives across the boundary. Interesting consequences occur when the energy of the particle is less than the barrier height, etc. An infinite sequence of potential profiles results in the Kronig-Penny model for solids, which describes the behavior of the electrons in a crystalline solid, and yields the band structure which is indispensable for describing the electronic properties of materials. Importantly, the results yield a classification of metals, semiconductors and insulators.

In addition, the readers at the advanced level are encouraged to read the solution of the Dirac equation both in the absence and in the presence of a variety of potential functions, for example, a Coulomb potential, etc. In condensed matter physics, the relativistic nature of particles finds an application in graphene, which shows a relativistic dispersion for the valence electrons close to the Fermi energy. This gives rise to a vast field of two dimensional (and even three dimensional generalizations exist) Dirac materials, where the electronic dispersion is linear in the wavevector, just like that of a photon. However, the electrons have a much lower velocity (of the order of $10^5 - 10^6$ m/s) than that of light, as thus earns the name, *pseudo relativistic* dispersion. This linear dispersion has important ramifications for the transport properties of Dirac materials.

Technically speaking, physics deals only with one body and many body problems. Because a two body problems reduces to a one body problem and a three body problems is unsolvable. However, physicists and chemists routinely worry about $\sim 10^{23}$ number of particles. With this many of them, the density is such that the particles spend enough time within a few de Broglie wavelengths from one another, and hence we need to go beyond the single particle description, that is, there is necessity to develop a quantum many body theory. The basic idea behind this approach is that instead of keeping track of a large number of strongly interacting particles, can we get away with a relatively smaller number of weakly interacting particles, called the quasiparticles, or the elementary excitations.

The study of condensed matter physics and its applications to the physical properties of various materials has found a place in the undergraduate curricula for a century or even more. The perspective on teaching of condensed matter has remained unchanged for most of this period. However, the developments in condensed matter over the last few decades require a new perspective of teaching and learning of the subject. Quantum Hall effect is one such discovery that has influenced the way condensed matter physics is taught to undergraduate students. The role of topology in condensed matter systems and the fashion in which it is interwoven with physical observables needs to be understood by a student for a deeper appreciation of the subject. However, the role of inter-particle interactions in shaping up the properties of materials cannot be ignored. Thus, to make a quintessential presentation for the undergraduate students, in this book, we have addressed selected topics that comprehensively contribute to the learning of condensed matter physics that emerged in not-so-distant past, as well as those topics that have firmly laid the foundation of conventional condensed matter physics.

A slight elucidation of the content will aid to a better understanding of the spirit of this text. In Chap. 1 a follow-up of material properties is taken up in the form of studying magnetism. Different types of magnetic order and materials were introduced to the students. The upshot of the discussion is that the electronic interactions drive the magnetic order, and hence they need to be incorporated for a comprehensive understanding of magnetic properties. In this connection, we introduce spin models, such as the Ising model, the XY model, Heisenberg model, etc. We solve these via controlled approximations. Furthermore, magnetism is shown to originate from itinerate electronic models, such as the Hubbard model. A self-consistent Hartree-Fock solution of the Hamiltonian yields a reasonable description for both the ferromagnetic and antiferromagnetic correlations.

Next, in Chap. 2, we embark on the transport properties of 2D electronic systems and focus solely on the role of a constant magnetic field therein. This brings us to the topic of the quantum Hall effect which is one of the main verticals of the book. The origin of the Landau levels and the passage of the Hall current through edge modes are discussed. The latter establishes a quantum Hall sample to be the first example of a topological insulator. Having discussed 2D electron gas, it is of topical interest to discuss the corresponding scenario in graphene. Thereafter, a crisp introduction of the fractional quantum Hall effect is included. It comprises a discussion of the Laughlin states, composite fermions, and the hierarchy scenario which will benefit the students in understanding the role of electronic interactions resulting in fractionally quantized Hall plateaus.

Our subsequent focus in Chap. 3 is the subject topology and its application to condensed matter physics. Introducing the subject from a formal standpoint, we discuss the band structure and topological invariants in 1D and 2D. In particular, we talk about the SSH and Kitaev models, which, apart from being a possible realization for a polyacetylene chain, have emerged as a paradigmatic tool to study topology in 1D. In 2D, the usual hobbyhorse, namely, graphene, is taken. We dwell upon the possibility pointed out by Haldane, whether graphene can become a topological insulator. Addition of the spin of electrons to the ongoing discussion emerged as a unique possibility to yield another version of the topological insulator, namely, the quantum spin Hall insulator, which may lie at the heart if the next generation spintronic devices.

We go on to discuss superconductivity in Chap. 4. We sequentially follow the historical developments of the field, phenomenological understanding of different phases, magnetic properties of superconductors, etc. Importantly, we present the BCS theory, in as much detail as possible, in an effort to provide a microscopic description of the thermodynamics and the electromagnetic phenomena. The finite momentum pairing and the FFLO states are hence described to compare and contrast between these and the BCS superconductors. Ginzburg-Landau theory, a phenomenological description of a superconductor to a metal transition, has been introduced in a nutshell. We finally wind up with a description of the experiments that determine the superconducting energy gap, and a very brief note of unconventional cuprate superconductors, and the pnictides or the chalcogenides. Finally, we focus on the application aspects and discuss Josephson junctions and SQUIDs.

All the while during the course of the book, we have included rigorous mathematical derivations wherever required, presented experimental details to connect with the ongoing discussions and tried to be as lucid as possible in our presentation of topics and concepts. A whole lot of schematic diagrams are presented for clarity as well. We hope that the students gain from the essence of this book, and it aids their understanding of both the topical and the traditional condensed matter physics. We shall be available and happy to answer queries, clarifications by students and researchers, and welcome comments for improvement.

Dr. Saurabh Basu
Department of Physics,
IIT Guwahati

Dr. Sourav Chattopadhyay
Department of Physics,
ICFAI University Tripura

AUTHORS' BIOGRAPHIES

Saurabh Basu was born and raised in Kolkata (the then Calcutta), India. His school and undergraduate colleges were in St. Lawrence High School and St. Xavier's College, respectively. He has completed Masters in Physics from IIT Bombay, PhD from IIT Kanpur, and after a couple of post-doctoral studies at TIFR, Mumbai and Queens University, Kingston, Canada, Saurabh had joined the department of Physics, IIT Guwahati, India in 2003 and is currently a professor there. His broad research interests are different fields of theoretical Condensed Matter Physics, with a focus on topological materials, higher order topological insulators, ultracold physics, non-Hermitian systems, charge and thermal transport in mesoscale and nanoscale devices, Floquet dynamics, critical phenomena, etc. Dr. Basu has more than 100 research publications in different peer-reviewed journals and a few book chapters. 12 PhD students have so far graduated on his guidance, and currently there are 6 students working for their PhD degree. Several Bachelors, Masters students, and intern students from different institutes have received his guidance. Apart from research and teaching at IIT Guwahati, he is passionate about improving the status of school and college education in the country. He has visited several institutions to upgrade a hands on training program there, and is committed to make science education more enjoyable for young students.

Dr. Sourav Chattopadhyay did his undergraduate degree in Physics (Hons.) from the Scottish Church College, Kolkata. He has completed both Masters in Physics and PhD from IIT Guwahati. Sourav had joined the department of Physics, ICFAI University Tripura, India in 2022 and is currently working as an assistant professor there. His broad research interests are computational statistical mechanics, different models of phase transition. Several Masters students, and intern students have worked under his guidance. He has written a book "Computer in the fingertip" for the upliftment of computer literacy for the less privileged children under the aegis of the "Outreach Program" of IIT Guwahati. He has delivered several invited talks in different institutes and has a keen interest in spreading science education via computer visuals.

ACKNOWLEDGMENTS

It is a proud privilege to acknowledge a lot of people who have actively or passively contributed in bringing up the book in the current form. First and foremost, we owe a lot to a number of current and past graduate students of the department of Physics, IIT Guwahati, such as Ms. Shilpi Roy, Mr. Sayan Mandal, Mr. Dipendu Halder, Mr. Koustav Roy, and Ms. Srijata Lahiri, Dr. Sudin Ganguly, Dr. Sk. Noor Nabi, Dr. Priyadarshini Kapri, and Dr. Priyanka Sinha. It is also a privilege to acknowledge a number of people who have contributed directly or indirectly to the preparation of the book. They are Professors Gaurav Dar (BITS Pilani, Goa campus), Avinash Singh (IIT Kanpur), S. Murakami (TokyoTech), B. Tanatar (Bilkent), A Perumal (HoD, Physics, IIT Guwahati), Dr. Sovan Ghosh (HOD Physics, ICFAI University Tripura), and apologies to the names that we forgot to mention.

It would be unfair to not mention the free and open source communities under *Stackexchange*. The authors of several posts in the platform have put great thought and effort in addressing a number of issues that are used during the preparation of the manuscript. We have learned and improved several things, which really came out to be helpful in our endeavor.

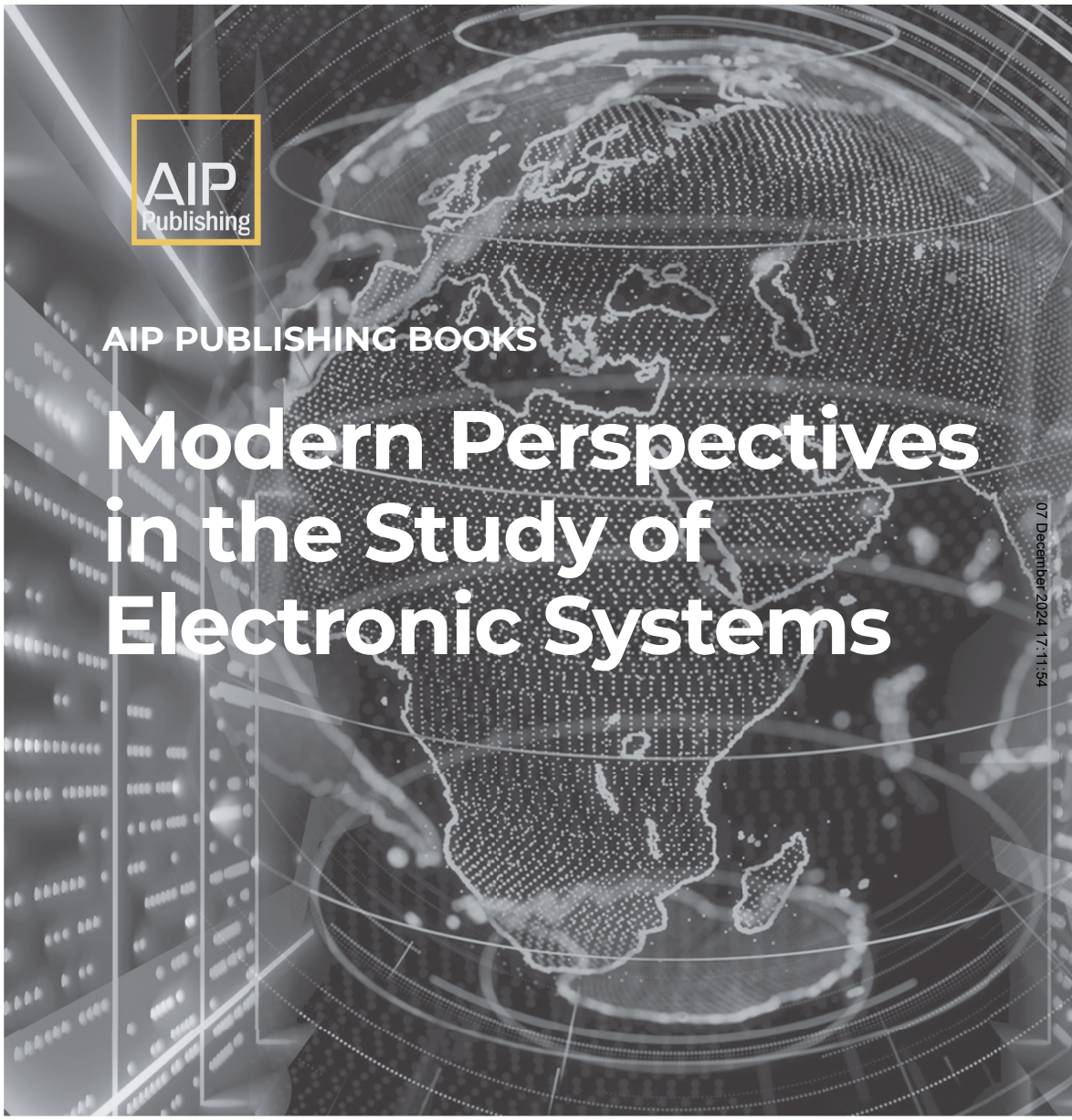

AIP
Publishing
AIP PUBLISHING BOOKS
Modern Perspectives in the Study of Electronic Systems
07 December 2024 17:11:54

TABLE OF CONTENTS

CHAPTER

1 MAGNETISM

1.1 INTRODUCTION

Magnetic phenomena were first observed in lodestones in ancient times. A lodestone is a naturally magnetized magnetite, namely, Ferric Oxide. The earliest mention of a lodestone attracting a needle was found in the first century AD in China. In fact, the Greek philosophers have also discussed the magnetic properties of materials during the period 600–500 BC. Even Indian medical texts, such as Sushruta Samhita, had prescribed magnetized materials to remove metallic objects from human bodies.

A large variety of metals, insulators, and even superconductors demonstrate magnetic properties. The importance of the magnetic phenomena gets enhanced due to the vast technological applications, which include equipment in our surroundings on a daily basis. Some of them are transformers, storage in computers, permanent magnets in motors, etc.

In the last couple of decades or so, magnetic materials have inroads into several intriguing discoveries, such as dilute magnetic semiconductors, iron-based superconductors, very large resistance in the presence of magnetic fields [colossal magnetoresistance (CMR) and giant magnetoresistance (GMR)], and more interestingly, applications in the field of spintronics, where the spin degrees of freedom (as opposed to the charge) are employed to carry information.

In the first course on solid state physics, we have seen that independent electron approximation is suitable for understanding certain properties of materials, such as metallicity. Even the insulating properties are well explained by the free electrons subjected to an effective potential due to the presence of ionic cores (remember the Kronig Penny model learned in the first course on solid state physics) via computation of the band structure within certain approximations, the most familiar of them is the tight binding model. Thus, within an independent electron approximation, it is possible to conceptually understand the energy spectrum resulting from an effective potential, phonons, etc.

However, there are more interesting phenomena, such as magnetic phenomena in solids, specifically ferromagnetism and antiferromagnetism where the many-electron aspects show up in a way that necessitates going beyond the single-particle picture. For example, consider spin waves, where the spin of one electron is flipped, while all the (valance) electrons take part, resulting in a collective behavior. The collective and the local aspects of the electron correlations are intertwined in a complicated way in giving rise to magnetic correlations. The most familiar type of magnetism that is widely discussed

https://doi.org/10.1063/9780735422537_001 1-1
Modern Perspectives in the Study of Electronic Systems, Saurabh Basu and Sourav Chattopadhyay

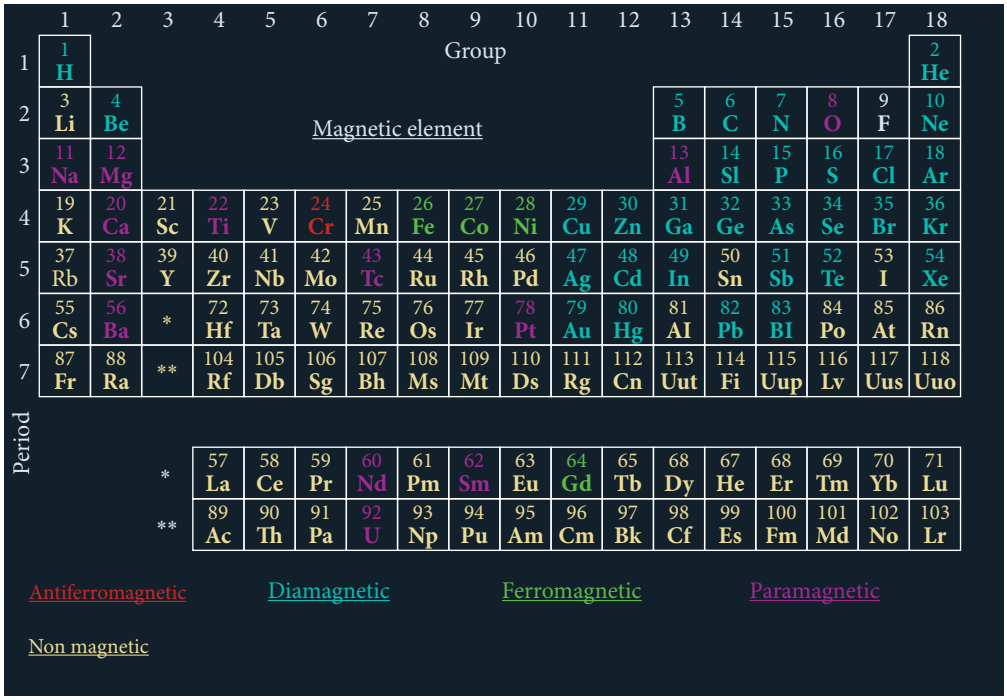


FIG. 1.1
The periodic table showing different magnetic elements.

in the literature is the ferromagnetism in 3d metals, such as Fe, Co, Ni, etc., where the exchange interaction between the largely delocalized 3d electrons facilitates ferromagnetism (see Fig. 1.1). However, in the $4f$ transition metals and their compounds (the Lanthanides) require a localized description. Even explanation of the antiferromagnetic arrangement of spins requires the exchange interaction to be invoked among the localized electrons.

Before we start discussing magnetic phenomena in solids, let us list only a few known magnetic materials, and their molar susceptibilities, χ_m defined by

$$\chi_m = \kappa \mathcal{V}_m = \kappa \frac{M}{\rho} \tag{1.1}$$

($\mathcal{V}_m$: molar volume, M: Molar mass, ρ: mass density) where κ is the volume susceptibility which appears in the proportionality of $\mathbf{M}$ (magnetization) and $\mathbf{B}$ (magnetic field) as $\mathbf{M} = \kappa \mathbf{H}$. The magnetic susceptibilities of a few common magnetic elements are listed in Table 1.1.

Table 1.1
Magnetic susceptibility of some common magnetic elements.

Name	χ_m ($\times 10^{-6}$cm^3/mol)	Name	χ_m ($\times 10^{-6}$cm^3/mol)
Aluminum (Al)	+16.5	Nickel (Ni)	Ferro
Antimony (Sb)	−37.0	Molybdenum (Mo)	+72.0
Bismuth (Bi)	−280.1	Phosphorus, Black (P)	−26.6
Boron (B)	−6.7	Platinum (Pt)	+201.9
Calcium (Ca)	+40.0	Potassium (K)	+20.8
Carbon (C)	−6.0	Rhodium (Rh)	+117.0
Cesium (Cs)	+29.0	Rubidium (Ru)	+17.0
Chromium (Cr)	+16.7	Selenium (Se)	−25.0
Copper (Cu)	−5.46	Silicon (Si)	−3.9
Cobalt (Co)	Ferro	Silver (Ag)	−19.5
Gallium (Ga)	−21.6	Strontium (Sr)	+92.0
Gold (Au)	−28	Sulfur α (S)	−15.5
Indium (In)	−64.0	Sodium (Na)	+16.0
Iridium (Tr)	+32.1	Tantalum (Ta)	+154.0
Iron (Fe)	Ferro	Thorium (Th)	+132.0
Lanthanum (La)	+118.0	Thallium (Tl)	−50.0
Lead (Pb)	−23.0	Tin Gray (Sn)	−37.0
Lithium (Li)	+14.2	Titanium (Ti)	+151.5
Manganese-α (Mn)	+529.0	Tungsten (W)	+59.0
Molybdenum (Mo)	−96.5	Vanadium (V)	+255.0
Niobium (Nb)	+195.0	Zinc (Zn)	−11.4

1.2 DIAMAGNETISM AND PARAMAGNETISM

To begin with, we shall review atomic magnetism with a view to understanding magnetic phenomena in insulators. It may be noted that these properties can quantitatively be understood by the independent electron approximation. In the following, we discuss how the atomic susceptibilities are computed. In the presence of an external magnetic field **B**, the kinetic energy operator assumes the form Ashcroft and Mermin (1976),

$$\mathcal{K} = \frac{1}{2m}\sum_i (\mathbf{p_i} - e\mathbf{A}_i)^2 \tag{1.2}$$

where $\mathbf{A}_i$ is the local vector potential derivable from the magnetic field $\mathbf{B}$. Choosing a symmetric gauge $\mathbf{A}_i = \frac{1}{2}(\mathbf{r}_i \times \mathbf{B})$, one gets,

$$\mathcal{K} = \frac{1}{2m}\sum_i \left(\mathbf{p}_i - \frac{e}{2}\mathbf{r}_i \times \mathbf{B}\right)^2$$
$$= \frac{1}{2m}\sum_i \left[\mathbf{p}_i^2 + \frac{e^2}{4}(\mathbf{r}_i \times \mathbf{B})^2 - \frac{e}{2}\{\mathbf{p}_i \cdot (\mathbf{r}_i \times \mathbf{B}) + (\mathbf{r}_i \times \mathbf{B}) \cdot \mathbf{p}_i\}\right]. \tag{1.3}$$

One can split the above expression as

$$\mathcal{K} = \mathcal{K}_0 + \frac{e^2}{8m}B^2 \sum_i \left[r_i^2 - 2e\mathbf{p}_i \cdot (\mathbf{r}_i \times \mathbf{B})\right] \tag{1.4}$$

where $\mathcal{K}_0$ is the kinetic energy without an external field. Using the vector identity,

$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \mathbf{b} \cdot (\mathbf{c} \times \mathbf{a}) = \mathbf{c} \cdot (\mathbf{a} \times \mathbf{b})$.

That is, writing,

$$\sum_i \mathbf{p}_i \cdot (\mathbf{r}_i \times \mathbf{B}) = \sum_i \mathbf{B} \cdot (\mathbf{r}_i \times \mathbf{p}_i) = \sum_i \mathbf{B} \cdot \mathbf{L}_i = \mathbf{B} \cdot \mathbf{L} \tag{1.5}$$

where $\mathbf{L}$ $(= \sum_i \mathbf{L}_i)$ is the total electronic orbital angular momentum. Thus, the kinetic energy (and also the Hamiltonian $\mathcal{H}$) (adding the spin of the electrons) becomes,

$$\mathcal{H} = \mathcal{K} = \mathcal{K}_0 + \mu_B \mathbf{L} \cdot \mathbf{B} + \frac{e^2}{8m}B^2 \sum_i (x_i^2 + y_i^2). \tag{1.6}$$

Treating the last two terms on the right hand side (RHS) as perturbation, that is, treating $\mathcal{H} = \mathcal{H}_0 + \mathcal{H}'$, we have for $\mathcal{H}'$,

$$\mathcal{H}' = \mu_B(\mathbf{L} + g\mathbf{S}) \cdot \mathbf{B} + \frac{e^2}{8m}B^2 \sum_i (x_i^2 + y_i^2) \tag{1.7}$$

where $\mu_B = \frac{eh}{2m}$ is the Bohr magneton with a value 0.579×10^{-8} eV/G and the landé g factor is given by $g \approx 2$. The energy correction can be computed via a perturbation theory using the unperturbed states of the Hamiltonian $\mathcal{H}_0$.

It is instructive to note that the magnetization, M is defined as the first derivative of the free energy, F with respect to the fixed B. That is,

$$M = -\frac{1}{V}\frac{\partial F}{\partial B}, \tag{1.8}$$

where V denotes the volume.

Thus, the magnetic susceptibility, χ is obtained from the magnetization as

$$\chi = lim_{B \to 0}\frac{\partial M}{\partial B} \tag{1.9}$$

thereby implying that the susceptibility is the second derivative of the free energy with respect to B. Therefore, we need to evaluate both the terms in $\mathcal{H}'$ up to second order in B, that is, B^2.

Hence, performing a second-order perturbation theory one gets (Ashcroft and Mermin, 1976),

$$\Delta E_n = \sum_{n \neq m} \frac{|\langle \phi_n | \mu_B \mathbf{B} \cdot (\mathbf{L} + g\mathbf{S}) | \phi_m \rangle|^2}{E_n - E_m} + \frac{e^2}{8m} B^2 \langle \phi_n | \sum_i (x_i^2 + y_i^2) | \phi_n \rangle. \tag{1.10}$$

1.3 MAGNETIC PROPERTIES OF FILLED AND PARTIALLY FILLED SHELL MATERIALS

Here, we shall discuss the magnetic properties of insulating materials, that is, in which all the valence electronic shells are filled. Hence, the total orbital and spin angular momentum is zero, that is the only surviving term is given by

$$\Delta E_n = \frac{e^2}{8m} B^2 \langle \phi_n | \sum_i (x_i^2 + y_i^2) | \phi_n \rangle. \tag{1.11}$$

Let us evaluate the change in the ground state energy (we call the ground state wavefunction as $|\phi_0\rangle$),

$$\begin{aligned} \Delta E_0 &= \frac{e^2}{8m} B^2 \langle \phi_0 | \sum_i (x_i^2 + y_i^2) | \phi_0 \rangle \\ &= \left(\frac{2}{3}\right) \frac{e^2}{8m} B^2 \langle \phi_0 | \sum_i (r_i^2) | \phi_0 \rangle \\ &= \frac{e^2}{12m} B^2 \langle \phi_0 | \sum_i (r_i^2) | \phi_0 \rangle. \end{aligned} \tag{1.12}$$

This should suffice to calculate the susceptibility of a material using,

$$\begin{aligned} \chi &= -\frac{\mathcal{N}}{\mathcal{V}} \frac{\partial^2 (\Delta E_0)}{\partial B^2} \\ &= -\frac{e^2}{6m} \frac{\mathcal{N}}{\mathcal{V}} \langle \phi_0 | \sum_i (r_i^2) | \phi_0 \rangle \\ &= -\frac{e^2}{6m} n \langle \phi_0 | \sum_i (r_i^2) | \phi_0 \rangle \end{aligned} \tag{1.13}$$

where $\mathcal{N}$ denotes the number of ions and n $(= \frac{\mathcal{N}}{\mathcal{V}})$ is the density. The negative sign in front of the above expression indicates diamagnetic properties, where the moment is induced opposite to the applied field. χ in the above equation is known as the Larmor diamagnetic susceptibility. One may note that in the sum, electrons in the outermost shells contribute maximally owing to their large mean square distance from the nucleus. Consider Z_{out} to be the number of electrons in the outermost shell and r_{out} to be the corresponding distances, then the largest term in the sum yields the susceptibility to having a form,

$$\chi = -\frac{e^2}{6m} n Z_{out} r_{out}^2. \tag{1.14}$$

Above formula [Eq. (1.14)] correctly explains the magnetic behavior of the alkali halides, such as He, Ne, Ar, Kr, Xe and their ionic configurations.

Exploration of the magnetic properties of materials with partially filled shells is easier and often taught in the first course on statistical mechanics (Pathria and Beale, 2011). A brief recap is presented in the following.

Consider $\mathcal{N}$ identical non-interacting spin-S particles in presence of an external magnetic field B. The corresponding Zeeman Hamiltonian can be written as

$$\mathcal{H} = -\mu_B \mathbf{S} \cdot \mathbf{B}. \tag{1.15}$$

It may be noted that electronic degrees of freedom are not important for this discussion. The canonical partition function is written as

$$\mathcal{Z} = \sum_{\{S\}} e^{-\beta \mathcal{H}\{S\}} = \sum_{|\mathbf{S}|=-S}^{+S} e^{-\beta \mu_B S B} \tag{1.16}$$

where it is assumed that the lowest $(2S+1)$ states ($-S$ to $+S$) are thermally excited at a temperature, T with appreciable probability. The above sum in Eq. (1.16) is computed in the form of a geometric progression (GP) series,

$$e^{-\beta \mathcal{H}} = \frac{e^{\beta \mu_B B(S+\frac{1}{2})} - e^{-\beta \mu_B B(S+\frac{1}{2})}}{e^{\beta \mu_B B/2} - e^{-\beta \mu_B B/2}}. \tag{1.17}$$

As earlier, the magnetization, M can be calculated using,

$$M = -\frac{\partial F}{\partial B} = \mu_B S B_S\,(\beta \mu_B S B) \tag{1.18}$$

$B_S(x)$ is called the Brillouin function, and is defined by

$$B_S(x) = \frac{2S+1}{2S} \coth\left(\frac{2S+1}{2S} x\right) - \frac{1}{2S} \coth\left(\frac{1}{2S} x\right) \tag{1.19}$$

where $x = \mu_B B/k_B T$. The Brillouin function $B_S(x)$ as a function of x for a number of S values are shown in Fig. 1.2.

As can be seen from the plot, for $x \gg 1$, that is, $\mu_B B \gg k_B T$, $B_S(x) \to 1$. Thus, at lower temperatures and large values of external fields, the magnetization reaches its saturation value. While in the other limit, that is for $\mu_B B \ll k_B T$, one may do a small x-expansion which yields,

$$\coth(x) \approx \frac{1}{x} + \frac{1}{3} x + \mathcal{O}(x^3). \tag{1.20}$$

Thus,

$$B_S(x) \approx \frac{S+1}{3S} x + \mathcal{O}(x^3). \tag{1.21}$$

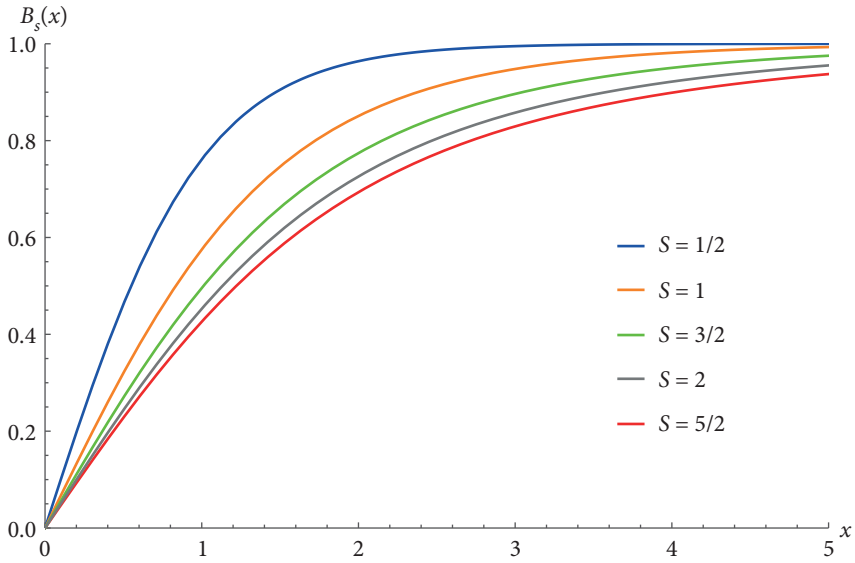


FIG. 1.2
Plot of different Brillouin functions $B_S(x)$ for different spin S values.

Hence, the susceptibility is computed as

$$\chi = \frac{\mu_B^2}{3}\frac{S(S+1)}{k_B T} = \frac{C_0}{T}. \tag{1.22}$$

This is called Curie's law. We may denote the above expression by χ_{Curie} [that is, replace χ by χ_{Curie} in Eq. (1.22)] and C_0 is called the Curie constant. Thus, the susceptibility to behaving inversely with temperature is a feature of 'paramagnetic solids' whose ordering of the magnetic moments is facilitated by a magnetic field, which, however may be hindered by thermal effects. The law is found to be valid at high temperatures, even when considerable magnetic interactions exist among the magnetic moments (Fig. 1.3).

The magnetic properties of the rare earth materials, such as La, Lu, Nd, Ce, Dy, etc. are adequately described by Curie's law. However, in transition metals, such as Fe, Zn, Cr, etc., there are additional phenomena, such as crystal field splitting which are dominant for partially filled *d*-shells. In order to understand the crystal field splitting in partially filled shells (say *d* shells) one needs to understand the Russel–Saunders (RS) coupling and Hund's rules (three of them). We postpone their discussion here and include them in the appendix.[1]

[1] Classic texts such as Ashcroft-Mermin, Kittel, etc. exist which treat these topics quite elaborately, and second, the focus of this text is to understand the magnetic properties from an electronic perspective.

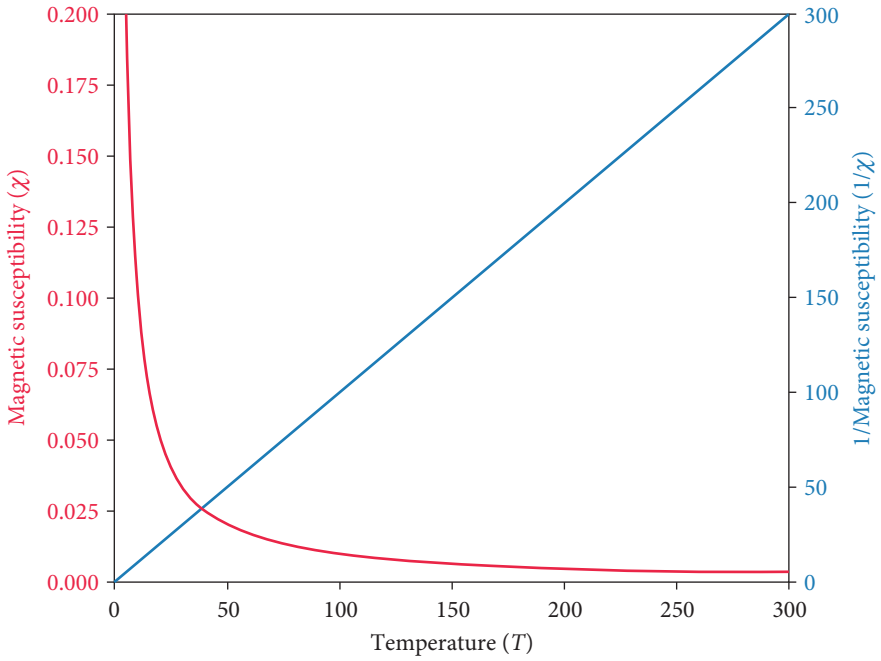


FIG. 1.3
χ_{Curie} (in red) and its inverse (in blue) are plotted as a function of temperature (T) and $1/T$ respectively. The slope of the linear plot yields C_0.

Having discussed Curie's law, one should be aware that there is another type of paramagnetic phenomenon applicable to metals, namely, Pauli paramagnetism which refers to the magnetic moments of conduction electrons. These electrons behave in a way that is distinct with respect to the localized electrons in partially filled ionic shells. Standard techniques of statistical mechanics can be applied (for example, see Pathria and Beale, 2011) to obtain the magnetic susceptibility, χ_{Pauli} which is a constant (as opposed to having an inverse temperature dependence as that in χ_{Curie}). In fact, the expression is

$$\chi_{\text{Pauli}} = \mu_B^2 N(\varepsilon_F), \tag{1.23}$$

where $N(\varepsilon_F)$ refers to the density of states (DOS) at the Fermi level. Thus, the susceptibility, χ of free electrons (Pauli) is a constant (independent of temperature), while for electrons bound to atoms, χ depends inversely on temperature (Curie).

A deeper introspection reveals that the Pauli susceptibility can be cast into Curie's law with the temperature, T replaced with T_F, the Fermi temperature. Since $T_F \approx 50\,000$ K for typical metals, χ_{Pauli} is typically three orders of magnitude lower than χ_{Curie} even at room temperature.

Furthermore, there is a finite magnetization of the conduction electrons that is diamagnetic in nature, giving rise to a phenomenon called Landau diamagnetism. The susceptibility, let us call it χ_{Landau}, which can be shown to be (Pathria and Beale, 2011)

$$\chi_{\text{Landau}} = -\frac{1}{3}\chi_{\text{Pauli}}. \tag{1.24}$$

Thus, the total susceptibility is a combination of the Pauli term, Landau susceptibility, Larmor susceptibility, etc. It is prudent to say that it is quite complicated to extricate a particular contribution from a combination of these effects.

1.4 FERROMAGNETISM AND ANTIFERROMAGNETISM

From the discussion that we had so far, it is clear that electronic interactions cannot be neglected in order to explain certain types of magnetic ordering, such as ferromagnetism and antiferromagnetism. Without such interactions (which are called magnetic interactions; however, the origin is electronic in nature), thermal effects would have randomized the magnetic moments (or the spins of the electrons) resulting in the absence of any sort of magnetic order. Even in antiferromagnetic materials where there is no net magnetization, there is still antiparallel order among the neighboring spins.

Before we bring about the electronic interactions to explain magnetism, let us discuss how one may get the magnetic Hamiltonian involving only the spins (or the magnetic moments) of the charge carriers. From a classical perspective, it is natural to expect that the dipolar force between moments, $\mathbf{m}_1$ and $\mathbf{m}_2$ separated by a distance $|\mathbf{r}|$ described by a potential of the form,

$$V(\mathbf{r}) = \frac{\mathbf{m_1}\cdot\mathbf{m_2} - 3(\mathbf{m_1}\cdot\hat{\mathbf{r}})(\mathbf{m_2}\cdot\hat{\mathbf{r}})}{\mathbf{r^3}} \tag{1.25}$$

will be operative. However, putting $\mathbf{m}_i$ to be of the order of atomic moments, that is the Bohr magneton (μ_B), and $|\mathbf{r}|$ to be of the order of Bohr radius,[2] a_0, V comes in the range 10^{-4} eV, which is far smaller than the typical atomic energies and hence cannot account for magnetic ordering, that typically involves an energy scale of the order of an eV. Thus, quantum mechanical exchange is really the dominant player, and dipolar interactions can safely be ignored.

An immediate motive is to arrive at a Heisenberg-like model[3], although here we restrict ourselves to an Ising-like Hamiltonian. To keep the matter even simpler, let us talk about two Ising-like spins [spins only have discrete degrees of freedom, for example, two degrees of freedom, pointing up ($\uparrow$) and down ($\downarrow$)]. Consider S_1 and S_2 both have values $s = 1/2$, which means that the spins will have two possible orientations where they can either be pointing up or down. Now consider the addition of these two $s = \frac{1}{2}$ particles,

$$S = S_1 + S_2. \tag{1.26}$$

[2] In quantum mechanical systems, a length scale is usually taken as the Bohr radius.
[3] Heisenberg model is discussed later.

Writing the problem in the (s, m_s) basis [S^2 has the eigenvalue $s(s+1)$ and S_z has the eigenvalue m_s, with $\hbar = 1$], the direct product space $(s_1, m_{s_1}) \otimes (s_2, m_{s_2})$ becomes four dimensional. Now $m_s = \pm\frac{1}{2}\hbar$ denote the $|\uparrow\rangle$-spin and the $|\downarrow\rangle$-spin states with eigenvalues $+\frac{\hbar}{2}$ and $-\frac{\hbar}{2}$ respectively. Thus, the basis states are spanned by $|\uparrow\uparrow\rangle$, $|\uparrow\downarrow\rangle$, $|\downarrow\uparrow\rangle$ and $|\downarrow\downarrow\rangle$. These are eigenstates of S_1^2, S_2^2, S_{1z} and S_{2z}. The eigenvalues for the total spin, S, that is, $s\ (= s_1 + s_2)$ are 0 and 1. For the total z-component of the spin, S_z, that is, $m_s\ (= m_{s_1} + m_{s_2})$ are 1, 0, 0, -1. One can check that,

$$S_z|\uparrow\uparrow\rangle = \hbar|\uparrow\uparrow\rangle; \quad S_z|\uparrow\downarrow\rangle = 0$$
$$S_z|\downarrow\downarrow\rangle = -\hbar|\downarrow\downarrow\rangle; \quad S_z|\uparrow\downarrow\rangle = 0. \tag{1.27}$$

Thus, we can take a linear combination of the above basis states to write the wavefunctions as

$$\chi_{00} = \frac{1}{\sqrt{2}}\left[|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle\right] \quad : \quad s = 0 \text{ singlet} \tag{1.28}$$

$$\left.\begin{aligned} \chi_{11} &= \frac{1}{\sqrt{2}}\left[|\uparrow\uparrow\rangle\right] \\ \chi_{10} &= \frac{1}{\sqrt{2}}\left[|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle\right] \\ \chi_{1-1} &= \frac{1}{\sqrt{2}}\left[|\downarrow\downarrow\rangle\right] \end{aligned}\right\} \quad : \quad \text{s = 1 triplets .} \tag{1.29}$$

The first one denotes a singlet wavefunction, χ_s, while the last three denote triplet states, χ_t. The singlet state is odd (changes sign) under the exchange of spins, while the triplet states are even (no change in sign). These are the eigenstates of the total spin operators, namely, S^2 and S_z. It can be checked that the S^2 operator has an eigenvalue 0 for the singlet state, and 1 for the triplet states. Also, the total spin, S satisfies

$$\begin{aligned} S^2 &= (S_1 + S_2)^2 = S_1^2 + S_2^2 + 2S_1 \cdot S_2 \\ &= \frac{3}{4}\hbar^2 + \frac{3}{4}\hbar^2 + 2S_1 \cdot S_2 \\ &= \frac{3}{2}\hbar^2 + 2S_1 \cdot S_2. \end{aligned} \tag{1.30}$$

Thus, for the singlet state, the operator $S_1 \cdot S_2$ has an eigenvalue $-\frac{3}{4}\hbar^2$, and $\frac{1}{4}\hbar^2$ for the triplet states. Denoting these eigenvalues by E_s and E_t respectively, we can write down a spin-only Hamiltonian, namely (here we have assumed $\hbar = 1$),

$$\mathcal{H} = \frac{1}{4}(E_s + 3E_t) - (E_s - E_t)S_1 \cdot S_2. \tag{1.31}$$

Readers can check whether the Hamiltonian has energies E_s for the singlet state and E_t for the triplet states by operating the Hamiltonian in Eq. (1.31) on the states in Eqs. (1.28) and (1.29). We may ignore the constant term, $(\frac{E_s+3E_t}{4})$ (or re-define the zero energy which is common to all states), and re-write the Hamiltonian as

$$\mathcal{H} = -JS_1 \cdot S_2, \tag{1.32}$$

where $J = E_s - E_t$. If J is positive, the system favors parallel alignment of spins and if J is negative, it favors antiparallel alignment.

One may wish to extend the above scenario to an array of spins with full spin-rotational symmetry and interacting via nearest neighbor exchange coupling to arrive at the Heisenberg model,

$$\mathcal{H} = -J \sum_{i,\delta} \mathbf{S}_i \cdot \mathbf{S}_{i+\delta}, \tag{1.33}$$

where δ refers to the neighbors of site i. This model was solved exactly by Bethe (1931) and later on by the others in one dimension.

Now, if one includes the orbital wavefunctions, in addition to the spin states, one can write the total eigenfunction as

$$\begin{aligned}\psi_{1,2}(\mathbf{r}_1, \mathbf{r}_2) &= \psi_{sym}(\mathbf{r}_1, \mathbf{r}_2)\chi_s(1,2) \\ &= \psi_{antisym}(\mathbf{r}_1, \mathbf{r}_2)\chi_t(1,2),\end{aligned} \tag{1.34}$$

where ψ_{sym} and $\psi_{antisym}$ denote the symmetric and antisymmetric states which are respectively, even and odd under exchange of $\mathbf{r}_1$ and $\mathbf{r}_2$. This takes care of the total fermionic wavefunction being antisymmetric. Furthermore, the total Hamiltonian can be written as

$$\mathcal{H} = \mathcal{H}_1 + \mathcal{H}_2 + \mathcal{H}_{12}, \tag{1.35}$$

where $\mathcal{H}_1$ and $\mathcal{H}_2$ are the single-particle Hamiltonians, and $\mathcal{H}_{12}$ denotes the potential energy term due to the exchange interaction obtained above. These can be written as the stationary state wavefunctions, $\phi(\mathbf{r}_i)$ centered at the lattice points, $\mathbf{r}_i$ where the particles are located, via

$$J = \int\int \phi_1^*(\mathbf{r}_1)\phi_2^*(\mathbf{r}_2)\mathcal{H}_{12}(\mathbf{r}_1, \mathbf{r}_2)\phi_2(\mathbf{r}_1)\phi_1(\mathbf{r}_2)d^3\mathbf{r}_1 d^3\mathbf{r}_2. \tag{1.36}$$

The exchange interaction is very strong and can be of the order of a fraction of an eV, which is equivalent to several hundreds of Kelvin. Thus, the exchange interaction is sufficient to align the spins even at room temperature. However, it decays exponentially with distance.

Let us discuss the exchange interaction, $\mathcal{H}_{12}$ in some more details. It is the Coulomb interaction between the particles and, for the simplest case, corresponding to two one-electron atoms [such as two hydrogen (H) atoms],

$$\mathcal{H}_{12}(\mathbf{r}_1, \mathbf{r}_2) = \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} + \frac{e^2}{|\mathbf{R}_1 - \mathbf{R}_2|} - \frac{e^2}{|\mathbf{r}_1 - \mathbf{R}_1|} - \frac{e^2}{|\mathbf{r}_2 - \mathbf{R}_2|} \tag{1.37}$$

where $\mathbf{r}_{1,2}$ refer to the coordinates of the electrons and $\mathbf{R}_{1,2}$ denote the nuclei of the two atoms. Here, the true ground state is the Heitler–London (HL) state, namely,

$$\psi_{HL} = \psi_s = \frac{1}{\sqrt{2}}\left[\psi_1(\mathbf{r}_1)\psi_2(\mathbf{r}_2) + \psi_1(\mathbf{r}_2)\psi_2(\mathbf{r}_1)\right] \tag{1.38}$$

ψ_s is the singlet state associated with χ_{00}. Of course, the HL state applies well to the atoms that are physically separated, however for atoms in a real solid, the magnetic interaction is severely complex, and may not be restricted to a 4×4 Hilbert space (that is a simple two-body term) which we have discussed earlier. Still in cases, where the magnetic atoms (or ions) are well separated, one can extend the two-spin interaction picture for the entire system.

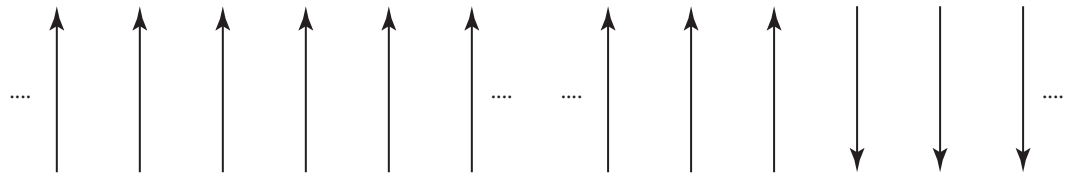


FIG. 1.4
A perfectly ordered (left) and a completely disordered (right) phase in a one-dimensional spin chain.

The prospect of ordering also depends on the dimensionality of the lattice. For example, one-dimensional spin systems really do not order at any finite temperature. The reason for this can be stated through the following illustration of Ising spins.

Assuming a nearest neighbor spin-only Hamiltonian, as in Eq. (1.33), left and right in Fig. 1.4 denotes two different phases, where the left denotes a perfectly ordered phase with an energy $-NJ$ (N being the number of spins), while the right denotes a perfectly disordered phase with an energy $-(N-1)J$ with only one bond is broken. Thus, the relative energy difference between the phases in them, that is, $\frac{\Delta E}{E} \sim \frac{1}{N}$ which vanishes in the limit of large N. Such a vanishingly small energy difference cannot stabilize an ordered state. Thus, dimensions higher than "one" are crucially required for magnetic ordering to exist.

In the following, let us describe a couple of different methods for solving the Heisenberg model using certain approximations. The methods are

i. Mean field theory (MFT),
ii. Holstein–Primakoff (HP) transformation and linear spin wave theory.

1.5 MEAN FIELD THEORY

In the mean field approximation, each spin "*feels*" an average field due to all the other spins of the system. The approximation is valid in any dimension; however, it is more accurate as the dimensionality grows larger (the fluctuations of the mean field state compared to the exact one diminishes with increase in the number of nearest neighbors or the dimensionality). In order to implement the mean field theory (MFT), one decouples $\mathbf{S}_i$ from $\mathcal{H}$ in Eq. (1.33) which makes it a single site Hamiltonian (that is, at a given site i) of the form,

$$\mathcal{H}_{MF}(i) = -\mathbf{S}_i.\left(J\sum_{\delta}\langle\mathbf{S}_{i+\delta}\rangle\right) - g\mu_B\mathbf{B}_{ext}\cdot\sum_i \mathbf{S}_i, \tag{1.39}$$

where, as one can notice that we have included an additional applied magnetic field $\mathbf{B}_{ext}$, and the second term denotes coupling with $\mathbf{B}_{ext}$. Since the bracketed expression in the first term above is just a constant

(being summed over nearest neighbors), the above Hamiltonian in Eq. (1.39) becomes a single spin Hamiltonian in an effective field, $\mathbf{B}_{eff}$, namely,

$$\mathcal{H}_{MF}(i) = -\mu_i \cdot \mathbf{B}_{eff}, \tag{1.40}$$

where $\mathbf{B}_{eff} = \mathbf{B}_{ext} - \frac{J}{g\mu_B}\sum_\delta \mathbf{S}_{i+\delta}$. Here g is the landé g factor, and μ_B denotes the Bohr magneton. In fact $\mathbf{S}_{i+\delta}$ can be replaced by its thermal average $\langle \mathbf{S}_{i+\delta}\rangle$, so that

$$\begin{aligned}\mathbf{B}_{eff} &= \mathbf{B}_{ext} - \frac{J}{g\mu_B}\sum_\delta \langle \mathbf{S}_{i+\delta}\rangle \\ &= \mathbf{B}_{ext} - z\frac{J}{g\mu_B}M,\end{aligned} \tag{1.41}$$

where $M = \langle \mathbf{S}_{i+\delta}\rangle$ is the magnetization and z denotes the coordination number.

Thus, essentially, we have a non-interacting system, and the thermal average can be calculated as if one particular spin is in a bath of an effective magnetic field. In this scenario, the standard approach to calculate the magnetization is by computing the free energy, $F = k_B T \ln \mathcal{Z}$, where $\mathcal{Z}$ is the canonical partition function. The method used is detailed in Sec. 3.1. For convenience, the results are quoted again. The partition function is written as

$$\begin{aligned}\mathcal{Z} &= \sum e^{-\beta\mathcal{H}_{MF}} = \sum e^{-\beta\mathcal{H}_{eff}} = \sum e^{-\beta\gamma S B_{eff}} \\ &= \frac{e^{\beta\gamma B_{eff}(S+\frac{1}{2})} - e^{-\beta\gamma B_{eff}(S+\frac{1}{2})}}{e^{\beta\gamma B_{eff}/2} - e^{-\beta\gamma B_{eff}/2}},\end{aligned} \tag{1.42}$$

where $\gamma = g\mu_B$. Hence, the magnetization is computed using $M = \frac{\partial F}{\partial B_{eff}}$. Thus, one arrives at,

$$M = \gamma S B_S(S\beta\gamma B_{eff}) = \gamma S B_S\left(S\beta\gamma(B - \frac{zJ^2}{\gamma}M)\right) \tag{1.43}$$

$B_S(x)$ being the Brillouin function discussed earlier, and

$$x = \left(S\beta\gamma\left(B - \frac{zJ^2}{\gamma}M\right)\right).$$

The above equation is non-linear, and can be solved for the magnetization, M for given values of B and T (remember $\beta = \frac{1}{k_B T}$). Assuming positive values of J, the transition from a paramagnet to a ferromagnetic state is indicated by the appearance of spontaneous magnetization, (finite value of M) in the zero magnetic field limits ($B \to 0$) below a certain critical temperature, T_c. Thus, in the limit $H \to 0$, $B_S(x)$ can be expanded for small x,

$$M \simeq -\gamma S\left(\frac{S+1}{3S}\right)\frac{zJ}{\gamma k_B T_c}M. \tag{1.44}$$

Solving for the transition temperature, T_c,

$$T_c = \frac{S(S+1)}{3}zJ. \tag{1.45}$$

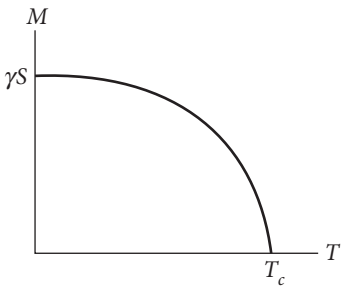


FIG. 1.5
Plot of magnetization as a function of temperature. The magnetization vanishes as $(T - T_c)^{1/2}$.

Thus, the critical temperature scale with the exchange interaction, J, implying that T_c is large for large J. Furthermore, the magnitude of the spin S and the coordination number z (which means dimensionality) play roles in deciding the value of T_c.

In order to see how M varies with T below T_c, Eq. (1.43) has to be solved numerically. The solution is shown schematically in Fig. 1.5. Analytic results exist at the extreme limits, that is, for (i) $T \ll T_c$ and (ii) $T \to T_c$ below. Let us demonstrate how to get these.

i. At $T \ll T_c$, β is large, so the Brillouin function, $B_S(x)$ has to be examined for large x, which yields,

$$\begin{aligned} B_S(x) \simeq &= 1 + \frac{2S+1}{S}\exp\left(-\frac{2S+1}{S}x\right) - \frac{1}{S}\exp\left(-\frac{x}{S}\right) \\ &= 1 + \frac{1}{S}\left((2S+1)\exp(-2x) - 1\right)\exp\left(\frac{x}{S}\right) \\ &= 1 - \frac{1}{S}\exp\left(-\frac{x}{S}\right). \end{aligned} \tag{1.46}$$

Thus,

$$\begin{aligned} M = \gamma S B_S(x) &= \gamma S - \exp\left(-\frac{JzM}{\gamma k_B T}\right) \\ &= \gamma S - \exp\left(-\frac{3}{S+1}\frac{T_c}{\gamma k_B T}\right). \end{aligned} \tag{1.47}$$

In the limit of $T \to 0$ (that is, $T \ll T_c$), the second term (the exponential term) is small. Thus, one gets $M \simeq \gamma S$.

ii. For T close to T_c (from below), we need to expand $B_S(x)$ for small x, where,

$$B_S(x) = \frac{S+1}{3S}x - \frac{(S+1)(2S^2+2S+1)}{90S^3}x^3. \tag{1.48}$$

This yields,

$$\begin{aligned} M &\simeq \frac{S+1}{3}\frac{SM}{k_B T} - \frac{(S+1)(2S^2+2S+1)}{90S^3}\frac{(zJ)^3 S^3 M^{2\,3}}{\gamma k_B T} \\ &= \frac{T_c}{T}M - \frac{3(2S^2+2S+1)}{10\gamma^3 S^2(S+1)^2}\left(\frac{T_c}{T}\right)^3 M^3. \end{aligned} \tag{1.49}$$

Since $M \neq 0$, one can solve for M^2,

$$M^2 \simeq \frac{10S^2(S+1)^2}{3(2S^2+2S+1)\gamma^3}\left(\frac{T}{T_c} - 1\right). \tag{1.50}$$

Thus, $M(T) \sim (T - T_c)^{1/2}$ for T approaching T_c below, that is, from the ordered regime. The exponent 1/2 is the signature of the mean field theory. The temperature dependence of the magnetization is schematically shown in Fig. 1.5. It falls slowly from a value γS at $T = 0$, However, there is a rapid decline in the vicinity of $T = T_c$.

1.6 LINEAR SPIN WAVE THEORY

To continue our discussion on ferromagnets, let us again consider the Heisenberg model given by Eq. (1.33) in a magnetic field B was applied in the z-direction. For positive J, the system minimizes its energy by having all the spins aligned in the z-direction at zero temperature. At small, but finite temperatures, investigating elementary excitations of a spin system is difficult to determine owing to the non-commutativity of the components of the spin operator. However, such an endeavor would still be possible if we can transform the spin operators to bosonic operators via a canonical transformation,[4] known as the Holstein–Primakoff (HP) transformation (Holstein and Primakoff, 1940). The spin operators at a site i are denoted as

$$S_i^+ = S_{ix} + iS_{iy} = \sqrt{2S}\left(1 - \frac{b_i^\dagger b_i}{2S}\right)^{1/2} b_i \tag{1.51}$$

$$S_i^- = S_{ix} - iS_{iy} = \sqrt{2S} b_i^\dagger \left(1 - \frac{b_i^\dagger b_i}{2S}\right)^{1/2}, \tag{1.52}$$

where $S_i^\pm$ are the spin raising and the lowering operators and $b_i^\dagger (b_i)$ are the bosonic creation (annihilation) operators. It is easy to verify that the components of the spin operator obey the commutation relation,

$$\left[S_x, S_y\right] = iS_z, \quad (\hbar = 1), \tag{1.53}$$

where the bosonic operators obey $[b, b^\dagger] = 1$ at each lattice site i. Furthermore, S^2 and S_z commute with the Hamiltonian $\mathcal{H}$ in Eq. (1.33). Some essential mathematical steps include writing down

$$\mathbf{S}^2 = S_x^2 + S_y^2 + S_z^2.$$

Thus, $S_z^2 = \mathbf{S}^2 - S_x^2 - S_y^2$. Using $S^\pm = S_x \pm iS_y$

$$S_z^2 = \mathbf{S}^2 - \frac{1}{2}(S^+S^- + S^-S^+) \quad \forall i \tag{1.54}$$

$\mathbf{S}^2$ and S_z acting on the common eigenfunction of $\mathbf{S}^2$ and S_z yield, $S(S+1)$ and S_z. Using the spin-boson transformation relations [Eqs. (1.51) and Eq. (1.52)], one gets

$$S_z = S - b^\dagger b. \tag{1.55}$$

Assuming translational invariance of the system, we can do a Fourier transform,

$$b_i = \frac{1}{\sqrt{N}} \sum_k e^{-ik.\mathbf{r}_i} b_k. \tag{1.56}$$

Note that the Fourier-transformed operators obey,

$$[b_k, b_{k'}^\dagger] = \delta_{k,k'}.$$

[4] Such transformations always preserve the commutation relations of both the spin and the boson operators.

FIG. 1.6
Plot showing spin wave excitation. Each spin is slightly rotated with respect to its neighbor. A complete cycle of the rotation is shown. These spin wave excitations are called "*magnons*" (analogous to phonons denoting lattice excitations).

A priori, b_k and $b_k^\dagger$, which denotes the annihilation and the creation of a quasiparticle, respectively, are called magnons or spin wave excitations. The spin wave excitations can be denoted by a small rotation of one spin with respect to its preceding neighbor, starting with a perfectly aligned configuration for the first spin. Thus, it is a long wavelength excitation and requires several lattice sites (where the atoms and the ions are localized) for the spins to come back to their original orientation (see Fig. 1.6).

In view of this slow variation (refer to Fig. 1.6), the expansion of the square root in the HP transformation may converge quickly. The reason being $b_i^\dagger b_i$, which denotes a spin deviation at site i has a slow variation. Thus, expanding the square roots,

$$S_i^+ \simeq \sqrt{2S}\left[b_i - \frac{b_i^\dagger b_i b_i}{4S} + \cdots .\right]. \tag{1.57}$$

Fourier transforming the RHS,

$$S_i^+ = \left(\frac{2S}{N}\right)^{1/2}\left[\sum_k e^{-ik.R_i} b_k - \frac{1}{4SN}\sum_{k,k',k''} e^{i(k-k'-k'').R_i} b_k^\dagger b_{k'} b_{k''} + \cdots .\right]. \tag{1.58}$$

Similarly,

$$S_i^- = \left(\frac{2S}{N}\right)^{1/2}\left[\sum_k e^{i\mathbf{k}.R_i} b_k^\dagger - \frac{1}{4SN}\sum_{k,k',k''} e^{i(k+k'-k'').R_i} b_k^\dagger b_{k'}^\dagger b_{k''} + \cdots .\right] \tag{1.59}$$

S^+ and S^- are called magnon operators. Also, without employing any approximation, the z-component of the spin can be written as

$$S_{iz} = S - b_i^\dagger b_i = S - \frac{1}{N}\sum_{k,k'} e^{i(k-k').R_i} b_k^\dagger b_{k'}. \tag{1.60}$$

Summing over all the spins along the z-axis that is along the direction of the magnetic field,

$$S^{\text{tot}} = \sum_i S_{iz} = NS - \sum_k b_k^\dagger b_k \tag{1.61}$$

where $\sum_i e^{i(k-k').R_i} = \delta_{k,k'}$ is used. Neglecting the terms cubic in magnon operators, namely, S_i^+ and S_i^-, the Hamiltonian [Eq. (1.33)] in k-space can be written as

$$\begin{aligned}\mathcal{H} = \frac{JS}{N}\sum_{k,k',\delta_i}[&e^{i(k-k').R_i}e^{ik'.\delta_i}b_k b_{k'}^\dagger + e^{i(k-k').R_i}e^{ik'.\delta_i}b_k^\dagger b_{k'} \\ &- e^{i(k-k').R_i}b_k^\dagger b_{k'} - e^{i(k'-k').R_i}e^{i(k-k').\delta_i}b_k^\dagger b_{k'}] \\ &- \frac{\mu_B B}{N}\sum_{i,k,k'} e^{i(k-k').R_i}b_k^\dagger b_{k'} + \frac{1}{2}JNZS^2 \end{aligned} \tag{1.62}$$

δ connects to the neighbors, and z is the coordination number. Using the definitions of the δ-function one can write

$$\mathcal{H} = JzS\sum_k\left[\gamma_k b_k^\dagger b_k + \gamma_{-k}b_k^\dagger b_k - 2b_k^\dagger b_k\right] + \frac{\mu_B B}{N}\sum_k b_k^\dagger b_k + \frac{1}{2}JNzS^2, \tag{1.63}$$

where $\gamma_k = \frac{1}{z}\sum_\delta e^{ik.\delta}$, and thus $\gamma_k = \gamma_{-k}$.[5] The Hamiltonian can further be simplified to arrive at

$$\mathcal{H} = \sum_k[-2JzS(1-\gamma_k) + \mu_B B]\, b_k^\dagger b_k + \frac{1}{2}JNzS^2, \tag{1.64}$$

where N denotes the total number of spins. Importantly, $\mathcal{H}$ is diagonalized with the first term being quadratic on the basis of magnon operators ($b_k^\dagger b_k$ denotes a number operator for magnons) and the second term is merely a constant. Thus, the energy eigenvalue for the magnon Hamiltonian is

$$\omega_k = -2JSz(1-\gamma_k) + \mu_B B. \tag{1.65}$$

Hence, the magnons (or spin waves) disperse according to this relation. The long wavelength behavior, say, in a square lattice is parabolic, that is, goes as $\sim k^2$ (k being inverse of the wavelength) and owing to the $1-\gamma_k$ factor, it vanishes as $k \to 0$. This vanishing of ω_k is in accordance with the Goldstone theorem (Goldstone, 1961) and the zero energy mode is called the Goldstone mode, which arises due to the spontaneous breaking of the symmetry.

1.6.1 Quantum XY model

The Heisenberg Hamiltonian, in one dimension was exactly solved by H. Bethe in 1931 (Bethe, 1931). In spite of the solution being quite elegant, it still does not enlighten us much about the basic properties, such as long range order. Rather, a simple and more intuitive picture of interacting $s = \frac{1}{2}$ particles emerge from the similarity between the spin and the fermion operators. This similarity was originally exploited by Jordan and Wigner (Jordan and Wigner, 1928) who had converted an interacting problem of $s = \frac{1}{2}$ particles to that of spinless fermions via a canonical transformation, which, for obvious reasons is known as Jordan-Wigner transformation. It is applicable to a simpler variant of the Heisenberg Hamiltonian, where the coupling between the z-component of the spins is switched off. However, quite fortunately, it captures the low energy properties of the antiferromagnetic Heisenberg model.

[5] For example, $\gamma_k = \frac{1}{4}(\cos k_x a + \cos k_y a)$ for a two-dimensional square lattice.

The model is described by the Hamiltonian,

$$\mathcal{H} = J \sum_i (S_{i,x} S_{i+1,x} + S_{i,y} S_{i+1,y}). \tag{1.66}$$

The interaction Hamiltonian is restricted to the x and y components of spins among the nearest neighbors in a one-dimensional chain. The components of the spins obey usual commutation relations,

$$[S_i^\alpha, S_j^\beta] = i\epsilon_{\alpha\beta\gamma}\delta_{ij} S_i^\gamma \tag{1.67}$$

where $\hbar = 1$ and ϵ_{ijk} are the Levi-Civita tensor. The interaction does not include the z-component of the spin, and hence the name XY model. Defining raising and lowering operators,

$$S_i^\pm = S_{i,x} \pm i S_{i,y}. \tag{1.68}$$

In terms of $S^\pm$, the Hamiltonian takes the form,

$$\mathcal{H} = \frac{J}{2} \sum_i (S_i^+ S_{i+1}^- + S_i^- S_{i+1}^+). \tag{1.69}$$

Now, since $\mathbf{S}_i = \frac{1}{2}\sigma_i$ and $S_i^\pm = \sigma_i^\pm$, σ being the Pauli matrices, the Hamiltonian can be written as

$$\mathcal{H} = \frac{J}{2} \sum_i (\sigma_i^+ \sigma_{i+1}^- + h.c.), \tag{1.70}$$

where *h.c.* denotes the Hermitian conjugate. Owing to the 2×2 structure of the Pauli matrices, the vector space is two-dimensional. Now we can derive a fermionic Hamiltonian in terms of c, $c^\dagger$ by defining,

$$c_i^\dagger = \left(\Pi_{j<1}\sigma_j^z\right)\sigma_i^+, \quad c_i = \left(\Pi_{j<1}\sigma_j^z\right)\sigma_i^-. \tag{1.71}$$

Again, these transformations preserve the commutation relations for the fermions and the spin-1/2 particles. From the commutation relation of the Pauli matrices, one can check for the fermionic anticommutation relations for c, $c^\dagger$. The reader is advised to go through a few steps of algebra to derive

$$\sigma_i^z = 1 - 2c_i^\dagger c_i \tag{1.72}$$

or,

$$\sigma_i^z = (-1)^i c_i^\dagger c_i \tag{1.73}$$

such that σ_i^z can take the values ± 1 for $n = 0, 1$. Also, the $\sigma_i^\pm$ are defined by

$$\sigma_i^+ = c_i^\dagger \exp\left[-\pi i \sum_{j=1}^{i-1} c_j^\dagger c_j\right]$$
$$\sigma_i^- = \exp\left[-\pi i \sum_{j=1}^{i-1} c_j^\dagger c_j\right] c_i. \tag{1.74}$$

One can check that,

$$\{c_i, c_i^\dagger\} = \frac{\{\sigma_i^-, \sigma_i^+\}}{\Pi_{j=i}^{i-1}(-1)^j c_j^\dagger c_j \Pi_{j'=1}^{i-1}(-1)^{j'} c_{j'}^\dagger c_{j'}} = 1. \tag{1.75}$$

Also, one can show that, $c_i^\dagger c_i = \sigma_i^- \sigma_i^+$.

A special mention is required for the commutation relations. c operators at the same site obey anti-commutation relations, while at different sites obey commutation relations (similar to bosons), such that the unitary rotations obey neither bosonic nor fermionic relations.

Going back to Eq. (1.70), for $1 \leq i \leq (N-1)$,

$$\begin{aligned}
\sigma_i^{-1}\sigma_{i+1}^+ &= \exp\left[-\pi i \sum_{j=1}^{i-1} c_j^\dagger c_j\right] c_i c_{i+1}^\dagger \exp\left[-\pi i \sum_{j=1}^{i-1} c_j^\dagger c_j\right] \\
&= c_i \exp\left[-\pi i \sum_{j=1}^{i-1} c_j^\dagger c_j\right] \exp\left[-\pi i \sum_{j=1}^{i-1} c_j^\dagger c_j\right] c_{i+1}^\dagger \\
&= c_i \exp\left[-\pi i \sum_{j=1}^{i-1} c_j^\dagger c_j\right] c_{i+1}^\dagger = c_i(1 - 2c_i^\dagger c_i)c_{i+1}^\dagger \\
&= c_i^\dagger c_i.
\end{aligned} \tag{1.76}$$

Similarly, the conjugate is written as

$$\sigma_{i+1}^- \sigma_i^+ = c_i^\dagger c_{i+1}. \tag{1.77}$$

Thus, after expressing the '*boundary particle*' operators $\sigma_N^-\sigma_1^+ + \sigma_1^-\sigma_N^+$ in terms of the fermion operators c_i and $c_i^\dagger$ one gets,

$$\mathcal{H} = \frac{J}{2}\sum_i \left(c_{i+1}^\dagger c_i + c_i^\dagger c_{i+1}\right) - \frac{J}{2}\sum_i \left(c_1^\dagger c_N + c_N^\dagger c_1\right)\left(\exp\left[-\pi i \sum_{j=1}^{i-1} c_j^\dagger c_j + 1\right]\right). \tag{1.78}$$

The first term on the RHS is quadratic in c operators and describes free fermions in a closed chain. The effect of the boundary enters through the second term can be neglected for large N, as it merely denotes a $1/N$ correction to the first term. Thus, the Hamiltonian becomes,

$$\mathcal{H} \simeq \frac{J}{2}\sum_i \left(c_i c_{i+1}^\dagger + c_{i+1} c_i^\dagger\right). \tag{1.79}$$

This finally yields a quadratic term, and hence a non-interacting fermionic Hamiltonian in one dimension which can be solved exactly. The Hamiltonian in Eq. (1.79) commutes with the number operator, $\mathcal{N} = c_i^\dagger c_i$, that is, $[\mathcal{H}, \mathcal{N}] = 0$. Furthermore, the z-component of the spin operator can be written as

$$S_i^z = \frac{1}{2}\left[\sigma_i^\dagger, \sigma_i^-\right] = c_i^\dagger c_i - \frac{1}{2} \tag{1.80}$$

and $S^z = \sum_i S_i^z$. Thus, each spinless fermion created by $c^\dagger$ carries $S^z = 1$.

Furthermore, Fourier transforming the fermionic operators using,

$$c_i = \frac{1}{\sqrt{N}} \sum_k e^{ik.R_i} c_k \tag{1.81}$$

one arrives at a tight binding form,

$$H = \sum_k \epsilon_k c_k^\dagger c_k = J \sum_k \cos ka \; c_k^\dagger c_k, \tag{1.82}$$

where $k \in [-\pi, \pi]$, and a is the lattice constant which, without any loss of generality, can be taken to be unity. Thus, from an interacting spin problem, we arrived at a non-interacting fermionic problem. The spectrum is gapless, that is, there may be gapless excitations, which implies that one extra fermion can be added to the system without any additional cost of energy at the Fermi level. However, the gapless situation will vanish if a nearest neighbor interaction term among the z-components, that is, $J_z S_i^z S_{i+1}^z$ is included.

Finally, the z-component of the spin, S^z yields,

$$S^z = \sum_i c_i^\dagger c_i = \sum_k c_k^\dagger c_k - \frac{N}{2}. \tag{1.83}$$

If we split the above sum into $k < 0$ and $k > 0$ (including $k = 0$)

$$S^z = \sum_{k>0} c_k^\dagger c_k + \sum_{k<0} (1 - c_k^\dagger c_k) - \frac{N}{2} \tag{1.84}$$

which can also be written as

$$S^z = \sum_k sgn(\epsilon_k) c_k^\dagger c_k + \sum_{k<0} 1 - \frac{N}{2}. \tag{1.85}$$

Thus, excitations with $|k| < \frac{\pi}{2}$ carry $S^z = +1$ while those with $|k| > \frac{\pi}{2}$ carry $S^z = -1$. Thus, the z-component of the total spin of the ground state equals zero, and hence, it is non-degenerate. The same result holds for the unrestricted (that includes interaction between the z-component of the spins) antiferromagnetic Heisenberg model. In fact, the ground state energy and the excitation spectrum is identical to the antiferromagnetic case where the ground state is non-degenerate. However, this is very unlike the ferromagnetic Heisenberg model, where the ground state is hugely degenerate and it carries a value for the z-component of the spin, namely, $S_z = \frac{N}{2}$.

Just to put things into perspective, here we have discussed a magnetic Hamiltonian in one dimension which has an exact solution. The solution yields a magnetic metal with gapless excitations. In the event one additionally includes a z component of the spin interaction, the spectrum becomes gapped, and hence corresponds to an insulating scenario.

1.7 ISING MODEL OF FERROMAGNETISM: TRANSFER MATRIX

Consider a spin-only model interacting via the nearest neighbor exchange interaction in the presence of an external magnetic field, B. In order to solve the problem, we shall use a transfer matrix technique. Again, we consider $s = \frac{1}{2}$ particles, which can assume two different orientations, namely, $\uparrow$ and $\downarrow$. The Hamiltonian of such a system is written as

$$\mathcal{H} = g\mu_B \mathbf{B} \sum_i \mathbf{S}_i - \sum_{\langle ij \rangle} J_{ij} \mathbf{S}_i \cdot \mathbf{S}_j, \tag{1.86}$$

where g and μ_B are the landé g factor and Bohr magneton, respectively. Because of the discrete possibilities of the spin orientation, the Hamiltonian can be written in a scalar form as

$$\mathcal{H} = g\mu_B B \sum_i S_i^z - \sum_{\langle ij \rangle} J_{ij} S_i^z S_j^z. \tag{1.87}$$

Writing $m_i = S_i^z$,

$$\mathcal{H} = g\mu_B B \sum_i m_i - \sum_{\langle ij \rangle} J_{ij} m_i m_j. \tag{1.88}$$

A further variable transform, $\alpha_i = 2m_i$ yields

$$\mathcal{H} = \frac{g\mu_B B}{2} \sum_i \alpha_i - \frac{J}{2} \sum_{i=1}^{N} \alpha_i \alpha_{i+1}. \tag{1.89}$$

We can assume a periodic boundary condition denoted by $\alpha_{N+1} = \alpha_1$ (N being the number of spins in the chain). Thus, the last spin is connected to the first one, and the system is in the shape of a closed loop with no free edge. The canonical partition function for the above Hamiltonian is written as

$$\mathcal{Z} = \sum_{\{\alpha_i\}} e^{-\beta E_n\{\alpha_i\}} = \sum_{\{n_i\}} e^{-\beta \mathcal{H}\{\alpha_i\}}, \quad \text{where } \beta = \frac{1}{k_B T}. \tag{1.90}$$

By expanding the partition function, one gets (Chandler *et al.*, 1987),

$$\mathcal{Z} = \sum_{\alpha_1 = \pm 1} \sum_{\alpha_2 = \pm 1} \cdots \sum_{\alpha_N = \pm 1} K(\alpha_1, \alpha_2) K(\alpha_2, \alpha_3) \ldots K(\alpha_N, \alpha_1) \tag{1.91}$$

where

$$K(\alpha_1, \alpha_2) = \exp\left[-\frac{\beta g \mu_B B}{2} (\alpha_1 + \alpha_2) + \frac{\beta J}{2} \alpha_1 \alpha_2 \right]. \tag{1.92}$$

Writing K for $\alpha_i = \pm 1$,

$$K = \begin{bmatrix} e^{(-x+a)} & e^{-a} \\ e^{-a} & e^{(x+a)} \end{bmatrix}, \tag{1.93}$$

where $x = \frac{g\mu_B B}{2k_B T}$ and $a = \frac{J}{2k_B T}$. Thus, the partition function becomes,

$$Z = \mathrm{Tr}(K^N). \tag{1.94}$$

In order to obtain the partition function in the closed form, please note that K being a 2×2 matrix, has two eigenvalues. Let us call them λ_1 and λ_2 which yields

$$Z = \mathrm{Tr}(K^N) = \lambda_1^N + \lambda_2^N = \lambda_1^N \left[1 + \left(\frac{\lambda_2}{\lambda_1}\right)^N\right]. \tag{1.95}$$

We assume one of them to be greater than the other, that is, $\lambda_1 > \lambda_2$. Since the partition function involves terms raised to the power N and with N being large, one can write (neglecting second term inside the bracket),

$$Z = \lambda_1^N, \tag{1.96}$$

where $\lambda_{1,2}$ are given by

$$\lambda_{1,2} = e^a[\cosh x \pm (\sinh^2 x + e^{-4a})^{1/2}]. \tag{1.97}$$

Keeping the relevant one[6] for computing the partition function,

$$\lambda_1 = e^a[\cosh x + (\sinh^2 x + e^{-4a})^{1/2}]. \tag{1.98}$$

The free energy is written as

$$F = -k_B T \ln Z = -N k_B T \ln \lambda_1. \tag{1.99}$$

One can hence compute the magnetization, M using,

$$M = \lim_{B\to 0} -\frac{\partial F}{\partial B} = \lim_{B\to 0} \frac{Ng\mu B}{2} \left\{ \frac{\sinh(g\mu_B B/2k_B T)}{[\sinh^2(\frac{g\mu_B B}{2k_B T}) + e^{-2J/k_B T}]^{1/2}} \right\}, \tag{1.100}$$

which yields,

$$M = \left[\frac{Ng^2\mu_B^2}{4k_B} \frac{e^{J/k_B T}}{T}\right] B. \tag{1.101}$$

The magnetic susceptibility is given by

$$\chi = \lim_{B\to 0} \frac{\partial M}{\partial B} = \frac{e^{J/k_B T}}{T}. \tag{1.102}$$

Finally, one can derive Curie's law,

$$\begin{aligned} \frac{1}{\chi} &= T e^{-J/k_B T} \\ &= T(1 - J/k_B T + \ldots) \\ &= T - J/k_B. \end{aligned} \tag{1.103}$$

Rewriting the above relation in a more familiar form,

$$\chi = \frac{C}{T - \theta}, \tag{1.104}$$

[6] The other one will vanish in the limit of large N.

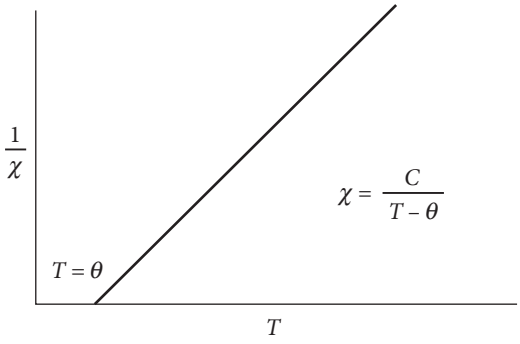


FIG. 1.7
$1/\chi$ is schematically shown as a function of temperature, T. It vanishes at $T = \theta$.

where θ is a characteristic temperature called the Curie temperature ($\theta = J/k_B$) at which ferromagnetic ordering takes place (see Fig. 1.7). Thus, the system is paramagnetic above a temperature $T = \theta$ where the spins are randomly oriented owing thermal effects, while they align below θ. This is the simplest discussion of ferromagnetism in a model Hamiltonian, which arises out of the exchange interaction among the neighboring spins in the presence of a magnetic field. The known ferromagnets, such as Fe and Ni have Curie temperatures 1093 K and 650 K, respectively.

It is worthwhile to point out that the above model has no spontaneous magnetization, which means that as the external magnetic field is switched off, the magnetization vanishes.

1.8 QUANTUM ANTIFERROMAGNET

Antiferromagnetism requires at least two sublattices, say A and B, inter-penetrating with each other to explain their structures and properties. Considering the Heisenberg model,

$$\mathcal{H} = -J \sum_{\langle ij \rangle} \mathbf{S}_i.\mathbf{S}_j \tag{1.105}$$

now with $J < 0$, such the energy is minimized when spins in one sublattice point in the $+z$-direction (up spin), with those in the other sublattice point in the $-z$-direction (down spin). This state with alternate up and the down spins is referred to as the classical Neél state. The lattice called the bipartite lattice. In such a bipartite lattice, one can define a staggered magnetization, $M_S = \sum_i (-1)^i M_i$, where $M_i = \langle S_i \rangle$. Because of the pre-factor $(-1)^i$, it takes a value $+1$ for the sites on the A sublattice, while it takes -1 for those in the B sublattice, or vice versa. Thus, M_S assume a value NS (N being the total number of spins).

One can also define a sublattice magnetization pertaining to each of the two sublattices which takes values $NS/2$ for the Neél state. However, the true ground state is far from classical, which is especially true at low dimensions owing to the presence of quantum fluctuations. These fluctuations lower the saturation value (Neél) of the magnetization. Let us illustrate this in the following by employing a Holstein–Primakoff transformation to an antiferromagnetic Heisenberg model, where one has to separately deal with the sublattices A and B in a bipartite lattice.

The transformations for the A sublattice can now be written as

$$S_i^{A+} = S_{ix}^A + iS_{iy}^A = \sqrt{2S}\left(1 - \frac{a_i^\dagger a_i}{2S}\right)^{1/2} a_i \tag{1.106}$$

and,

$$S_i^{A-} = S_{ix}^A - iS_{iy}^A = \sqrt{2S}a_i^\dagger\left(1 - \frac{a_i^\dagger a_i}{2S}\right)^{1/2} \tag{1.107}$$

where, $a_i^\dagger$ (a_i) denote bosonic creation (annihilation) operators for the A sublattice. Repeating the same for the B sublattice,

$$S_i^{B+} = S_{ix}^B + iS_{iy}^B = \sqrt{2S}\left(1 - \frac{b_i^\dagger b_i}{2S}\right)^{1/2} b_i \tag{1.108}$$

and,

$$S_i^{B-} = S_{ix}^B - iS_{iy}^B = \sqrt{2S}b_i^\dagger\left(1 - \frac{b_i^\dagger b_i}{2S}\right)^{1/2} \tag{1.109}$$

where $b_i^\dagger$ (b_i) again denote bosonic creation (annihilation) operators for the B sublattice. The z-component of the spin operators is written as

$$S_{iz}^A = (S - a_i^\dagger a_i), \quad S_{iz}^B = -(S - b_i^\dagger b_i). \tag{1.110}$$

Now, as earlier done for a ferromagnet, we introduce the Fourier-transformed operators,

$$\alpha_k = \frac{1}{\sqrt{N}}\sum_{i\in A} a_i e^{ik.R_i}, \quad \beta_k = \frac{1}{\sqrt{N}}\sum_{i\in B} b_i e^{ik.R_i}$$
$$\alpha_k^\dagger = \frac{1}{\sqrt{N}}\sum_{i\in A} a_i^\dagger e^{-ikR_i}, \quad \beta_k^\dagger = \frac{1}{\sqrt{N}}\sum_{i\in B} b_i^\dagger e^{-ikR_i}. \tag{1.111}$$

Recognizing that the periodicity of the sublattice is twice that of the crystal lattice, the effective Brillouin zone (over which the momentum index k runs) is half. Also α_k and β_k correspond to excitation of magnons in A and B sublattices. Expanding the spin operators up to linear order for the A sublattice in α_k and β_k,

$$S_i^{A+} \simeq \left(\frac{2S}{N}\right)^{1/2}\left[\sum_k e^{-ikR_i}\alpha_k + \cdots.\right]$$
$$S_i^{A+} \simeq \left(\frac{2S}{N}\right)^{1/2}\left[\sum_k e^{ikR_i}\alpha_k + \cdots.\right]. \tag{1.112}$$

Similarly for the B sublattice,

$$S_i^{B+} \simeq \left(\frac{2S}{N}\right)^{1/2}\left[\sum_k e^{-ikR_i}\beta_k + \cdots.\right]$$

$$S_i^{A+} \simeq \left(\frac{2S}{N}\right)^{1/2} \left[\sum_k e^{ikR_i}\beta_k + \cdots .\right]. \tag{1.113}$$

The z-components yield the exact expressions,

$$S_{iz}^A = S - \frac{1}{N}\sum_{k,k'} e_i^{i(k-k')R}\alpha_k^\dagger \alpha_{k'}$$

$$S_{iz}^B = -\left(S - \frac{1}{N}\sum_{k,k'} e_i^{i(k-k')R}\beta_k^\dagger \beta_{k'}\right). \tag{1.114}$$

Inserting these into the Heisenberg Hamiltonian, the Hamiltonian becomes quadratic and it reads as

$$\mathcal{H} \simeq -NzJS^2 + JSz\sum_k \left[\gamma_k(\alpha_k^\dagger\beta_k^\dagger + \alpha_k\beta_k) + (\alpha_k^\dagger\alpha_k + \beta_k^\dagger\beta_k)\right], \tag{1.115}$$

where,

$$\gamma_k = \frac{1}{z}\sum_{\delta\in \mathrm{nn}} e^{ik\delta} \tag{1.116}$$

where nn denotes the number of nearest neighbors. Although the Hamiltonian is bilinear in $\alpha_{\mathbf{k}}$ and β_k, unlike the ferromagnetic case, it is not readily diagonalizable. Another canonical transformation[7] involves linear combinations of α_k and β_k as

$$\begin{aligned} \eta_k &= u_k\alpha_k - v_k\beta_k^\dagger; \quad \zeta_k = u_k\beta_k - v_k\alpha_k^\dagger \\ \eta_k^\dagger &= u_k\alpha_k^\dagger - v_k\beta_k; \quad \zeta_k^\dagger = u_k\beta_k^\dagger - v_k\alpha_k. \end{aligned} \tag{1.117}$$

To have the bosonic commutation relations intact for η_k and ζ_k, we demand that the coefficients u_k and v_k are related by

$$u_k^2 - v_k^2 = 1. \tag{1.118}$$

This allows us to choose (though not uniquely), $u_k = \cosh\theta_k$ and $v_k = \sinh\theta_k$. For each k, one can put the anomalous terms (the ones which diagonalizing the Hamiltonian) to zero, which yields the following condition,

$$\tanh 2\theta_k = -\gamma_k. \tag{1.119}$$

Thus, the Heisenberg Hamiltonian becomes,

$$\mathcal{H} \simeq -NzJS^2 + NJSz\sum_k E_k\left(\eta_k^\dagger\eta_k + \zeta_k^\dagger\zeta_k + 1\right) \tag{1.120}$$

where the dispersion, E_k is given by

$$E_k = NJSz\sqrt{1-\gamma_k^2}. \tag{1.121}$$

[7] These are called Bogoliubov or Bogoliubov-Valatin transformation and will be used in the chapter on superconductivity

A comparison with the ferromagnetic case (which yields $E_k \sim 1 - \gamma_k$), the antiferromagnetic dispersion is $\sqrt{1-\gamma_k^2}$. Thus, here we get a linear behavior, that is, $E_k \sim k$ as opposed to parabolic ($E_k \sim k^2$) for ferromagnets. Additionally, there are two degenerate modes for each k in the Brillouin zone owing to two sublattice structure of the lattice.

At zero temperature, there are no magnons (or excitations of the spin wave). Thus, $\langle \eta_k^\dagger \eta_k \rangle$ or $\langle \zeta_k^\dagger \zeta_k \rangle$ vanishes and we obtain the ground state energy as

$$E_{GS} = -NJzS(S+1) + \sum_k E_k. \tag{1.122}$$

Introducing a constant κ, the expression for E_{GS} can be re-written as

$$E_{GS} = -NJzS^2\left(1 + \frac{\kappa}{S}\right), \tag{1.123}$$

where κ is given by

$$\kappa = \left(\frac{2}{N}\right)\sum_k \left[1 - (1-\gamma_k^2)^{1/2}\right]. \tag{1.124}$$

Thus, E_{GS} can be obtained for a crystal lattice.

The spin wave excitations lower the value of the sublattice magnetization compared to its saturation value. It is possible to have a quantitative estimate for a given lattice. For either of the sublattices (A sublattice is considered here for concreteness), the average value of the z-component of the spin is written as

$$\langle S_z^A \rangle = S - \frac{1}{N}\sum_k \langle \alpha_k^\dagger \alpha_k \rangle. \tag{1.125}$$

Using the Bogoliubov transformation,

$$\begin{aligned}\langle S_z^A \rangle &= S - \frac{1}{N}\sum_k \langle (u_k \eta_k^\dagger + v_k \zeta_k)(u_k \eta_k + v_k \zeta_k^\dagger)\rangle \\ &= S - \frac{1}{N}\sum_k (u_k^2 \langle \eta_k^\dagger \eta_k \rangle + v_k^2 \langle \zeta_k^\dagger \zeta_k \rangle + v_k^2) \\ &= S - \delta S_A \quad \text{(say)}.\end{aligned} \tag{1.126}$$

Clearly, the second term denotes the departure from the saturation values and hence denotes as fluctuations. The averages $\langle \eta_k^\dagger \eta_k \rangle$ and $\langle \zeta_k^\dagger \zeta_k \rangle$ yield the Bose distribution function at finite temperatures, $f_B(\omega_k)$ $(= \frac{1}{e^{\beta\omega_k}-1})$ such that it yields the number of bosons (spin waves) when summed over all k in the Brillouin zone.

$$n_k = \langle \eta_k^\dagger \eta_k \rangle = \langle \zeta_k^\dagger \zeta_k \rangle = f_B(\omega_k). \tag{1.127}$$

The coefficients u_k and v_k satisfy,

$$u_k^2 + v_k^2 = \cosh 2\theta_k = \frac{1}{\sqrt{1-\gamma_k^2}}. \tag{1.128}$$

Thus, the magnitude of the fluctuations, δS_A can be written as

$$\delta S_A = \frac{1}{N}\sum_k (n_k + \frac{1}{2})\frac{1}{\sqrt{1-\gamma_k^2}}. \tag{1.129}$$

Thus, δS_A can be computed at a given temperature and a crystal lattice.

1.9 ITINERANT ELECTRON MAGNETISM

It is by and large true that identifying the magnetic order present in the system is far easier than understanding the origin of the order. There are competing ways of studying magnetic order, which are exchange interactions between the localized electrons, and those between the itinerant electrons. Even though we feel that conceptualizing magnetic order arising either from a local or an itinerant view should not be viewed as a regimented criterion, and is often perceived as a convenience of description. We have seen so far how model Hamiltonians arising out of the exchange interaction between the localized moments, so it is instructive to discuss the role of itinerant electrons in producing magnetic effects in materials.

One simple way to consider the role of itinerant electrons resulting in magnetic effects is via a magnetic impurity in a free electron model or in the presence of bands. Such a scenario can indeed be realized in dilute magnetic semiconductors. The localized magnetic impurity causes an oscillation in the susceptibility at large distances away from the scatterer (the magnetic impurity). In a slightly different context, if a spin-polarized electron gas is either ferromagnetically or antiferromagnetically coupled to an impurity, a particular kind of exchange interaction ensue, and is known as the RKKY interaction (Ruderman and Kittel, 1954).

An alternative route to magnetic phenomena in itinerant systems arises from the competition between the kinetic and the potential energies. While we describe this scenario later, a priori it can be understood via a population imbalance of one kind of spin with respect to the other, thereby causing a decrease in the potential energy and an increase in kinetic energy. If such an interplay can lower the total energy, such that it becomes favorable for the bands to split and give rise to a net magnetization. The necessary conditions for such a scenario to occur are determined by the density of states at the Fermi level and the energy scale associated with the Coulomb interaction.[8]

In a physical situation, it is impossible to extricate the contributions arising out of the localized electrons from the itinerant ones. For example, magnetism in light actinides (where partially extended

[8] It will be introduced as the Stoner criterion.

5*f* orbitals are involved), is potentially different from the heavy actinides (or the lanthanides) where the 5*f* orbitals are strongly hybridized with the 6*d* and the 7*s* bands, thereby making the itinerant description to be more suited in the former, while a localized description is preferred for the latter. The same is true for the heavy fermion compounds,[9] such as $CeAl_3$, UPt_3 etc, where the localized *f* electrons strongly hybridize with the conduction band electrons, creating a significantly complicated scenario.

Kübler (2000) first attempted to explain magnetic correlations in materials via local density functional approximation (LDA). The method incorporates electronic correlations and indicates the importance of the electronic structure in understanding magnetic properties. Specifically, the itinerant electron picture makes concrete contributions to the exploration of half-metallic ferromagnets, GMR observed in multi-layered systems, etc.

In the previous discussion, we have seen ferromagnetism in the presence of a spin exchange interaction. It is worthwhile to investigate the ordering scenario in detail in an itinerant electron model, such as the Hubbard model. Since this discussion may be new to readers, we include an introduction of the model and emphasize on its properties. We apprise the readers that elementary knowledge of band theory and the second quantization is essential to understand the subsequent discussions.

A brief introduction to an interacting electronic model in terms of creation and annihilation operators is as follows:

$$\mathcal{H} = \sum_{i,i',\sigma,\sigma'} c^{\dagger}_{i\sigma} t^{\sigma,\sigma'}_{ii'} c_{i'\sigma'} + \sum_{i,i',j,j',\sigma,\sigma'} U_{ii'jj'} c^{\dagger}_{i\sigma} c_{j\sigma} c^{\dagger}_{i'\sigma'} c_{j'\sigma'}, \tag{1.130}$$

where i, i', j, j' refer to site indices and σ, σ' denote the spins. The Hamiltonian includes a single-particle term (the kinetic energy), and a two-particle interaction term. There could be the presence of interaction effects involving larger number of particles, however they are mostly weaker (other than being unsolvable), as compared to the two-particle term, thereby making two-particle interactions to be good enough for describing most of the interacting systems. It may be noted that we have considered the most general forms for both the terms in the Hamiltonian, however, either of them, or both may not depend upon the spin indices as shown in Eq. (1.130).

1.10 MAGNETIC SUSCEPTIBILITY: KUBO FORMULA

Here, we shall show the calculation of the magnetic susceptibility using linear response theory. We shall be deriving the Kubo formula, which is essential in a variety of systems. For example, calculation of resistivity for an electron gas in the presence of an external magnetic field (for example, the quantum Hall effect), or the polarizability of a dielectric in the presence of an electric field, etc. We give a thorough derivation of the Kubo formula in Chap. 2.

[9] The effective mass of the fermions is several times larger than the corresponding bare mass.

Consider an applied magnetic field $\mathbf{B}(\mathbf{r}, t)$. It is useful to assume it to be function of both $\mathbf{r}$ and t for reasons that will be clear later. Assume $\mathcal{H}'$ denotes the coupling between the spin and the magnetic field and is given by

$$\mathcal{H}' = -\int \mathbf{B}(\mathbf{r}, t) \cdot \mathbf{S}(\mathbf{r})\, d\mathbf{r} \tag{1.131}$$

where $\mathbf{S}(\mathbf{r}) = \sum_i \delta(\mathbf{r} - \mathbf{r}_i)\mathbf{S}_i$. $\mathbf{S}_i$ denotes the spin vectors at a given site i written in the Heisenberg representation. Also, here we have dropped the Bohr magneton, $\mu_B = -\frac{e\hbar}{2m}$, by taking $\mu_B = 1$. We shall compute the magnetization defined by $\langle \mathbf{S}(\mathbf{r}, t)\rangle$ induced by the magnetic field $\mathbf{B}$ via,

$$\langle \mathbf{S}(\mathbf{r}, t)\rangle = \langle \psi_0(t)|\mathbf{S}(\mathbf{r})|\psi_0(\mathrm{t})\rangle \tag{1.132}$$

where $|\psi_0(t)\rangle$ denotes the ground state at time t. Using the results of linear response theory, at the first order in B in a translationally invariant system,

$$\langle \mathbf{S}_i(\mathbf{r}, t)\rangle_B = \langle \mathbf{S}_i(\mathbf{r}, t)\rangle_{B=0} + \sum_j \int dt' \int d\mathbf{r}'\, \chi_{ij}(\mathbf{r} - \mathbf{r}', t - t')\mathbf{B}_j(\mathbf{r}, t') \tag{1.133}$$

where $\chi_{ij}(\mathbf{r}, t)$ is a retarded two-particle propagator and denotes the susceptibility tensor. The susceptibility tensor is the same as the spin–spin correlation function defined by

$$\chi_{ij}(\mathbf{r} - \mathbf{r}', t - t') = i\theta(t - t')\langle[\mathbf{S}_i(\mathbf{r}, t), \mathbf{S}_j(\mathbf{r}', t')]\rangle \tag{1.134}$$

where,

$$\begin{aligned}\theta(t - t') &= 1 \quad \text{for } t > t' \\ &= 0 \quad \text{for } t < t'.\end{aligned} \tag{1.135}$$

In order to proceed further, we shall Fourier transform the spin operators,

$$\mathbf{S}(\mathbf{r}) = \sum_{\mathrm{q}} e^{i\mathrm{q}\cdot\mathbf{r}}\mathbf{S}(\mathbf{q}), \tag{1.136}$$

where the time variable (t) is withheld for brevity. $\mathbf{S}(\mathbf{q})$ can be written in terms of the fermion operators as

$$\mathbf{S}(\mathbf{q}) = \sum_{\mathrm{k},\alpha,\beta} c^{\alpha\dagger}_{\mathrm{k+q}}\sigma^{\alpha\beta}c^{\beta}_{k}. \tag{1.137}$$

The components of $\mathbf{S}$, or the linear combinations thereof allow us to write the spin raising and lowering operators,

$$S^{\pm} = \frac{1}{2}(S_x \pm iS_y), \tag{1.138}$$

where,

$$S^{+}(\mathbf{r}) = \sum_{\mathrm{q}} e^{i\mathrm{q}\cdot\mathbf{r}} \sum_{\mathrm{k}} a^{\dagger}_{\mathrm{k+q}\uparrow}a_{\mathrm{k}\downarrow}.$$

and,

$$S^{-}(\mathbf{r}) = \sum_{\mathbf{q}} e^{-\mathbf{iq}\cdot\mathbf{r}} \sum_{\mathrm{k}} \mathbf{a}^{\dagger}_{\mathrm{k}+\mathbf{q}\downarrow}\mathbf{a}_{\mathrm{k}\uparrow}$$

Now, we can write down the transverse and the longitudinal susceptibilities χ^{-+} and χ_{zz} using the raising, lowering and z-component of the spin operators,

$$\chi^{-+}(\mathbf{r}-\mathbf{r}', t-t') = i\theta(t-t')\langle[\mathbf{S}(\mathbf{r},t), \mathbf{S}^{\dagger}(\mathbf{r}',t)]\rangle \tag{1.139}$$

and

$$\chi_{zz}(\mathbf{r}-\mathbf{r}', t-t') = i\theta(t-t')\langle[S_z(\mathbf{r},t), S_z(\mathbf{r}',t)]\rangle. \tag{1.140}$$

The equation of motion for χ^{-+} can now be written as (with $\hbar = 1$),[10]

$$i\frac{\partial}{\partial t}\chi^{-+}(k,q,t) = -\delta(t)\langle[c^{\dagger}_{k+q\downarrow}c_{\mathrm{k}\uparrow}, S^{+}(0,0)]\rangle + i\theta(t)\langle[[c^{\dagger}_{k+q\downarrow}c_{k\uparrow}, \mathcal{H}], S^{+}(0,0)]\rangle. \tag{1.141}$$

Here, the initial time t' is set to zero and the derivative of $\theta(t)$ is $\delta(t)$. Substituting the Hubbard Hamiltonian, $\mathcal{H}$ above,

$$\begin{aligned}[c^{\dagger}_{k+q\downarrow}c_{k\uparrow}, \mathcal{H}] = &-(\epsilon_{k+q} - \epsilon_k)c^{\dagger}_{k+q\downarrow}c_{k\uparrow} \\ &+ \frac{U}{N}\sum_{k',q'}(c^{\dagger}_{k+q\downarrow}c_{k-q'\uparrow}c^{\dagger}_{k'-q'\downarrow}c_{k'\uparrow} \\ &- c^{\dagger}_{k'+q'\uparrow}c_{k'\downarrow}c^{\dagger}_{k+q-q'\downarrow}c_{k\uparrow}).\end{aligned} \tag{1.142}$$

Now, the quartic operators in the second term of RHS can be dealt within a Hartree-Fock approximation, and the combinations, such as $\langle c^{\dagger}c\rangle$ can be retained.[11] At finite temperatures, the average values are written as

$$\langle c^{\dagger}_{k\sigma}c_{k'\sigma'}\rangle = \delta_{kk'}\delta_{\sigma\sigma'}f_{k\sigma} \tag{1.143}$$

where $f_{k\sigma}$ denotes the distribution function for fermions with spin σ and momentum k and has a form,

$$f_{k\sigma} = \frac{1}{e^{\beta(\epsilon_{k\sigma}-\mu)}+1} \tag{1.144}$$

where $\epsilon_{k\sigma}$ denotes the band energies. Thus, the second term in the above commutator is written as

$$\frac{U}{N}\sum_{k'}(f_{k\uparrow} - f_{k+q\downarrow})c^{\dagger}_{k+k'+q\downarrow}c_{k+k'\uparrow} + (f_{k'\downarrow} - f_{k'\uparrow})c_{k+q\downarrow}c_{k\uparrow}.$$

Going back to the equation of motion (EOM), the term,

$$\langle[c^{\dagger}_{k+q\downarrow}c_{k\uparrow}, S^{+}(0)]\rangle = \sum_{k,q}\langle[c^{\dagger}_{k+q\downarrow}c_{k\uparrow}, c^{\dagger}_{k+q\uparrow}c_{k\downarrow}]\rangle$$

[10] A time dependent external magnetic field was taken earlier precisely for this reason so that a time derivative can be taken.
[11] A $\langle c^{\dagger}c^{\dagger}\rangle$ or $\langle cc\rangle$ is relevant for studying superconducting correlations.

can be simplified to yield $(f_{k+q\downarrow} - f_{k\uparrow})$. Thus, the EOM for the transverse susceptibility, χ^{-+} can be written as

$$\left[i\frac{d}{dt} + (\widetilde{\epsilon}_{k+q\uparrow} - \widetilde{\epsilon}_{k\downarrow})\right]\chi^{-+}(k,q,t) = -\delta(t)(f_{k+q\downarrow} - f_{k\uparrow}) - (f_{k+q\downarrow} - f_{k\uparrow})\frac{U}{N}\sum_{k'}\chi^{-+}(k',q,t) \tag{1.145}$$

where the renormalized band energies $\widetilde{\epsilon}_{k\sigma} = \epsilon_k - \frac{U}{N}\sum f_{k\sigma}$. Thus, the one particle energies are modified by the interaction term as shown here.

The dynamical susceptibility, $\chi(\omega)$, are calculated by doing a Fourier transform,

$$\chi(\omega) = \int_{-\infty}^{\infty} dt\chi(t)e^{i\omega t}. \tag{1.146}$$

Thus,

$$\chi(k,q,\omega) = \frac{(f_{k\uparrow} - f_{k+q\downarrow})\left(1 + \frac{U}{N}\chi^{-+}(q,\omega)\right)}{\omega + \widetilde{\epsilon}_{k+q\uparrow} - \widetilde{\epsilon}_{k\downarrow}}, \tag{1.147}$$

where $\chi^{-+}(q,\omega) = \sum_k \chi^{-+}(k,q,\omega)$, one can write the expression for susceptibility as

$$\chi^{-+}(q,\omega) = \frac{\chi_0^{-+}(q,\omega)}{1 - U\chi_0^{-+}(q,\omega)} \tag{1.148}$$

where

$$\chi_0^{-+}(q,\omega) = \frac{1}{N}\sum_k \frac{f_{k\uparrow} - f_{k+q\downarrow}}{\omega - (\widetilde{\epsilon}_{k\downarrow} - \widetilde{\epsilon}_{k+q\uparrow}) + i\eta} \tag{1.149}$$

is the free susceptibility tensor. The $+i\eta$ is added in the denominator as is usually done for a propagator (Mahan, 2000), with the positive sign referring to the retarded propagator. The divergence of χ^{-+} denotes instability in the system, and hence, signals a phase transition. This instability is indicated by the divergence of the $\chi^{-+}(q,\omega)$, which happens when,

$$1 - U\chi_0^{-+}(q,\omega) = 0 \tag{1.150}$$

is satisfied. In particular, at $\omega = 0$, that is when the external magnetic field is time independent, the system is lossless. Thus, $\mathrm{Im}\chi^{-+}(\mathrm{q},\omega = 0) = 0$. In such a static situation, any instability in the system is a signature of the ground state instability. Furthermore, the energies of the low-lying excited states begin to merge with those of the ground state. Hence, we can rephrase the instability criterion as

$$U\chi_0^{-+}(q,0) = 1 \tag{1.151}$$

where,

$$\chi_0^{-+}(q,0) = \frac{1}{N}\sum_k \frac{f_k - f_{k+q}}{\epsilon_{k+q} - \epsilon_k}. \tag{1.152}$$

Thus, at any arbitrary value of q, the instability condition is associated with a critical value of the interaction U,

$$U_c = \frac{1}{\chi_0^{-+}(q,0)}. \tag{1.153}$$

In order to relate $\chi_0^{-+}(q,0)$ to something physical, we may note that the terms functions f_{k+q} and ϵ_{k+q} can be expanded in a Taylor series in the large wavelength limit, namely, $q \to 0$,

$$f_{k+q} \approx f_k + q \cdot \frac{\partial \epsilon_k}{\partial k} \frac{\partial f_k}{\partial \epsilon_k} \tag{1.154}$$

$$\epsilon_{k+q} \approx \epsilon_k + q \cdot \frac{\partial \epsilon_k}{\partial k} \tag{1.155}$$

$$\lim_{q\to 0} \chi^{-+}(q,0) = \frac{1}{N} \sum_k \left(-\frac{\partial f_k}{\partial \epsilon_k} \right) = \frac{1}{N} \sum_k \delta(\epsilon - \epsilon_k) \tag{1.156}$$

where we have replaced the derivative of the Fermi function with a δ-function which is valid at zero temperature. Moreover,

$$\frac{1}{N} \sum_k \delta(\epsilon - \epsilon_k) = N(\epsilon)$$

$N(\epsilon)$ is the density of states (DOS). This finally yields us the Stoner criterion that we are familiar with, namely,

$$UN(\epsilon_F) = 1. \tag{1.157}$$

Thus, instability corresponding to $q = 0$ leads to a tendency for the system to acquire a finite magnetization.

However, instability may occur even at finite values of q. To remind ourselves, the spin arrangement in antiferromagnets has an ordering wave vector $q = (\frac{\pi}{a}, \frac{\pi}{a}, \frac{\pi}{a})$ such that $e^{iq\cdot r}$ changes sign as one goes from one lattice site to its neighbor. In this case, the Fermi surface for the half-filled band (one electron per lattice site) coincides with the Brillouin zone. Thus, as opposed to a ferromagnet, infinitesimal interaction strength, U may cause instability in a system with one conduction electron per atom, inducing a transition from a metal to an insulating antiferromagnet.

1.11 JELLIUM MODEL

The electronic description of metals due to Bloch, Bethe and others in the 1930s neglects the electron–electron interaction, and in a vast number of cases, such a simplified description works. Even the distinction between a metal and an insulator can reliably be done by band - filling calculations, that is, without invoking the electron-electron interaction. Nevertheless, the band gaps in semiconductors and insulators are somewhat difficult to calculate quantitatively. However, modifications of the single-particle band structure, in the form Hartree-Fock corrections are computed, which is equivalent to

computing the average energy shift in the single-particle energies in the presence of an average density due to all other electrons have been known. This yields the essence of the mean field theory. Such a description has, by and large, been successful in explaining a host of material properties. Yet, there are reasons for one to deliberate upon, including the electron-electron interaction in studying the physics of materials. The salient ones are as follows:

i. the average inter-electron distance in a typical metal (which goes as $n^{-1/6}$, n being the density) is about 1 nm. For a small interparticle distance, the inclusion of interaction effects is indispensable.
ii. There are physical examples, such as complex materials where significant deviations from the single-particle band theory are noted. Familiar examples are transition metal oxides, cuprate superconductors, etc.

An extension of the free electron gas can be thought of via a jellium model where the electrons interact via the Coulomb potential, and the overall charge neutrality of the system is maintained by a homogeneous positively charged background. We shall mainly concentrate on obtaining the dielectric function via standard many body approaches, such as Hartree and Hartree-Fock approximations, random phase approximation (RPA) . In the following, we consider an electron liquid (because of the involvement of interparticle interactions, we do not call it an electron gas anymore) in 3D at $T = 0$, the so-called jellium model. Jellium is a prototype model for metals, which is a uniform electron gas with a positively charged background. Since the many-electron wavefunction can be solved using computational techniques, it is considered to be a convenient model for testing the characteristics of the density functionals.

1.11.1 The Hamiltonian

The Hamiltonian consists of terms that correspond to the kinetic energy of the electrons (H_{kin}), Coulomb interaction among the electrons (H_{e-e}), the interaction between the electrons and the background positive charge (H_{e-b}) which originates from the electrons interacting with the lattice vibrations, and finally background energy (H_{b-b}) which is basically the Coulomb interaction between the positive (background) charges. An enumeration of different terms in the second quantized notation can be written as (Fetter and Walecka, 1971; Ziman, 1972; and Mahan, 2000)

$$\begin{aligned}
\mathcal{H}_{kin} &= \sum_{\mathbf{k},\sigma} \xi_{\mathbf{k}} c^{\dagger}_{\mathbf{k}\sigma} c_{\mathbf{k}\sigma} \\
\mathcal{H}_{e-e} &= \frac{1}{2\mathcal{V}} \sum_{\mathbf{k},\mathbf{k}',\mathbf{q}} V(\mathbf{q}) c^{\dagger}_{\mathbf{k}\sigma} c^{\dagger}_{\mathbf{k}'+\mathbf{q}\sigma'} c_{\mathbf{k}'\sigma'} c_{\mathbf{k}+\mathbf{q}\sigma} \\
\mathcal{H}_{e-b} &= e^2 \int d^3\mathbf{r} d^3\mathbf{r}' \frac{n(\mathbf{r})}{4\pi\epsilon_0 |\mathbf{r}-\mathbf{r}'|} \\
\mathcal{H}_{b-b} &= \frac{e^2}{2} \int d^3\mathbf{r} d^3\mathbf{r}' \frac{n^2(\mathbf{r})}{4\pi\epsilon_0 |\mathbf{r}-\mathbf{r}'|} \\
\mathcal{H} &= \mathcal{H}_{kin} + \mathcal{H}_{e-e} + \mathcal{H}_{e-b} + \mathcal{H}_{b-b},
\end{aligned} \tag{1.158}$$

where $\mathcal{V}$ is the volume, $c_{\mathbf{k}}$ ($c^{\dagger}_{\mathbf{k}}$) are the single-particle annihilation (creation) operators corresponding to momentum, $\mathbf{k}$, $\xi_{\mathbf{k}} = \epsilon_{\mathbf{k}} - \mu$, μ being the chemical potential, the factor $\frac{1}{2}$ takes into double counting, and $n(\mathbf{r})$ is the density term whose Fourier transform is defined as

$$n_{\mathbf{q}} = \frac{1}{\sqrt{\mathcal{V}}} \int d^3\mathbf{r} e^{-i\mathbf{q}.\mathbf{r}} n(\mathbf{r}) = \frac{1}{\sqrt{\mathcal{V}}} \sum_{\mathbf{k}\sigma} c^{\dagger}_{\mathbf{k}\sigma} c_{\mathbf{k}+\mathbf{q}\sigma}. \tag{1.159}$$

It is somewhat odd to write some of the terms of the Hamiltonian $\mathcal{H}$ in Eq. (1.158) in real space, and others in momentum space. But with the Fourier transformation for the density term shown above, $\mathcal{H}$ can be fully written in the momentum space (by which we also implicitly assume that the electron liquid is homogeneous, and hence translational invariance holds). However, this will not hinder the discussion that is going to follow.

There is a subtle point that deserves a mention. Strictly speaking, a Fourier transform of the Coulomb potential cannot be performed owing to a divergence of the integral.[12] However, a screening term (which is also physically relevant owing to the screening effect from all other charges) of the form, $\frac{e^{-\lambda r}}{\epsilon_0 r}$ kills the divergence and yields $V(q) = \frac{e^2}{\epsilon_0 q^2}$ in the limit of zero screening ($\lambda \to 0$).

1.11.2 Hartree-Fock approximation

It may be noticed that all other terms except the $\mathcal{H}_{e-e}$ are quadratic in c operators and hence can be diagonalized in the single-particle basis, however $\mathcal{H}_{e-e}$ is quartic in c and hence needs an approximation to be at par with other terms for arriving at the solution. In a mean field approximation,[13] the quartic term is decoupled as a sum of all possible quadratic terms, namely,

$$c^{\dagger}_p c^{\dagger}_q c_r c_s \simeq -\langle c^{\dagger}_p c_r\rangle c^{\dagger}_q c_s - \langle c^{\dagger}_q c_s\rangle c^{\dagger}_p c_r + \langle c^{\dagger}_p c_s\rangle c^{\dagger}_q c_r + \langle c^{\dagger}_q c_r\rangle c^{\dagger}_p c_s$$

where the signs are governed by anticommutation relation of the fermions. The above decoupling scheme replaces the interaction term $\mathcal{H}_{e-e}$ by

$$\mathcal{H}^{HF}_{e-e} = \frac{1}{2\mathcal{V}} \sum_{\mathbf{k},\mathbf{k}',\mathbf{q}} [-\langle c^{\dagger}_{\mathbf{k}\sigma} c_{\mathbf{k}\sigma'}\rangle c^{\dagger}_{\mathbf{k}'+\mathbf{q}\sigma'} c_{\mathbf{k}+\mathbf{q}\sigma} - \langle c^{\dagger}_{\mathbf{k}'+\mathbf{q}\sigma'} c_{\mathbf{k}+\mathbf{q}\sigma}\rangle c^{\dagger}_{\mathbf{k}\sigma} c_{\mathbf{k}'\sigma'} + \langle c^{\dagger}_{\mathbf{k}\sigma} c_{\mathbf{k}+\mathbf{q}\sigma}\rangle c^{\dagger}_{\mathbf{k}'+\mathbf{q}\sigma'} c_{\mathbf{k}'\sigma'} + \langle c^{\dagger}_{\mathbf{k}'+\mathbf{q}\sigma'} c_{\mathbf{k}'\sigma'}\rangle c^{\dagger}_{\mathbf{k}\sigma} c_{\mathbf{k}+\mathbf{q}\sigma}]. \tag{1.160}$$

The above Eq. (1.160) can be split into two terms, one for $\mathbf{q} = 0$ and the other for $\mathbf{q} \neq 0$, namely,

$$\mathcal{H}^{HF}_{e-e} = \mathcal{H}^{HF}_{e-e}(\mathbf{q} = 0) + \mathcal{H}^{HF}_{e-e}(\mathbf{q} \neq 0) \tag{1.161}$$

[12] $V(\mathbf{q}) = \frac{e^2}{\epsilon_0} \int_0^{\infty} \frac{1}{r} e^{i\mathbf{q}\cdot\mathbf{r}}$ is a divergent integral.

[13] This type of mean field decoupling was introduced by Weiss, who replaced the magnetic exchange interaction of the form $\mathbf{S}_i \cdot \mathbf{S}_j$ by $\langle \mathbf{S}_i\rangle \mathbf{S}_j + \langle \mathbf{S}_j\rangle \mathbf{S}_i - \langle \mathbf{S}_i\rangle\langle \mathbf{S}_j\rangle$.

where,

$$\mathcal{H}^{HF}_{e-e}(\mathbf{q}=0) = V(\mathbf{q}=0)\sum_{\mathbf{k},\mathbf{k}'\sigma\sigma'}\langle c^\dagger_{\mathbf{k}\sigma}c_{\mathbf{k}\sigma}\rangle c^\dagger_{\mathbf{k}'\sigma'}c_{\mathbf{k}'\sigma'}$$
$$\mathcal{H}^{HF}_{e-e}(\mathbf{q}\neq 0) = \sum_{\mathbf{k},\mathbf{q}\sigma}\left[V(\mathbf{q})n_{\mathbf{k}-\mathbf{q}\sigma}\right]. \tag{1.162}$$

The third and the fourth terms can be written as

$$\mathcal{H}_{e-b} = -V(\mathbf{q}\neq 0)n\mathcal{N}$$
$$\mathcal{H}_{b-b} = \frac{\mathcal{V}}{2}V(\mathbf{q}\neq 0)n^2 \tag{1.163}$$

where $n = \frac{\mathcal{N}}{\mathcal{V}}$ is the electron density. It may be noted that $V(\mathbf{q}=0)$ diverges since $V(\mathbf{q}) \sim 1/q^2$, and because of which, both $\mathcal{H}_{e-b}$ and $\mathcal{H}_{b-b}$ will diverge. However, that is not too much of a concern, as the charge neutrality implies that various infinities would cancel each other to ensure the finiteness of the total energy. Let us combine $\mathcal{H}_{e-b}$ and $\mathcal{H}_{b-b}$, which yields,

$$\mathcal{H}_{eb} = \mathcal{H}_{e-b} + \mathcal{H}_{b-b} = V(\mathbf{q}=0)\left[\frac{\mathcal{V}}{2}n^2 - n\mathcal{N}\right] = \frac{V(\mathbf{q}=0)}{\mathcal{V}}\frac{\mathcal{N}^2}{2} = \frac{e^2}{2\lambda^2\epsilon_0}\frac{\mathcal{N}^2}{\mathcal{V}}. \tag{1.164}$$

Thus, $\mathcal{H}_{eb}$ vanishes in the thermodynamic limit owing the presence of $\mathcal{V}$ in the denominator. There is another interesting cancelation that is shown below.

Let us explicitly calculate the energy due to the background Hamiltonian, $\mathcal{H}_{b-b}$ (it is implicitly assumed that the eigenstates are known so that we can compute the energies), that is,

$$E_{b-b} = \frac{e^2}{2}n^2\int d^3\mathbf{r}d^3\mathbf{r}'\frac{1}{|\mathbf{r}-\mathbf{r}'|}e^{-\lambda|\mathbf{r}-\mathbf{r}'|}. \tag{1.165}$$

Let us assume $\mathbf{r}-\mathbf{r}' = \mathbf{z}$, which upon substituting yields,

$$E_{b-b} = \frac{e^2}{2}n^2\int d^3z d^3\mathbf{r}\frac{1}{|z|}e^{-\lambda|z|} = \frac{e^2}{2}n^2\mathcal{V}\int d^3z\frac{e^{-\lambda z}}{z} = \frac{e^2}{2}n^2\mathcal{V}\frac{4\pi}{\lambda^2}. \tag{1.166}$$

Thus, the energy per particle is,

$$\frac{E_{b-b}}{\mathcal{N}} = \frac{e^2}{2}\frac{4\pi}{\lambda^2}n. \tag{1.167}$$

A similar calculation for the E_{e-b} yields the form,

$$\frac{E_{e-b}}{\mathcal{N}} = -e^2\frac{4\pi}{\lambda^2}n. \tag{1.168}$$

Also the $\mathbf{q}=0$ term of the $\mathcal{H}_{e-e}$ term yields,[14]

$$\frac{E_{e-e}(\mathbf{q}=0)}{\mathcal{N}} = \frac{e^2}{2}\frac{4\pi}{\lambda^2}n. \tag{1.169}$$

[14] This is called as the **Hartree** term or the direct term (because of $\mathbf{q}=0$) in the electron-electron interaction.

Thus, we have,

$$\frac{E_{b-b}}{\mathcal{N}} + \frac{E_{e-b}}{\mathcal{N}} + \frac{E_{e-e}(\mathbf{q}=0)}{\mathcal{N}} = 0. \tag{1.170}$$

Hence, we are left with only the $\mathbf{q} \neq 0$ in the electron-electron interaction term, which is called the *Fock* term or the *exchange* term (as opposed to the direct interaction, called as the Hartree term), apart from the kinetic energy of the electrons.

Thus, we get a Hartree-Fock Hamiltonian, $\mathcal{H}_{HF}$ (since both the direct and the exchange terms are included) which can be written as

$$\mathcal{H}_{HF} = \sum_{\mathbf{k}\sigma} \left(\frac{\hbar^2 k^2}{2m} + \Sigma_{HF}(\mathbf{k}) \right) c^{\dagger}_{\mathbf{k}\sigma} c_{\mathbf{k}\sigma} \tag{1.171}$$

where $\Sigma_{HF}(\mathbf{k})$ the '*self-energy*' at the HF level and is written as

$$\Sigma_{HF}(\mathbf{k}) = -\sum_{\mathbf{q}} n_{\mathbf{k}-\mathbf{q}\sigma} V(\mathbf{q}). \tag{1.172}$$

This is the only term that is relevant to us as it provides a momentum shift, that is, a $\mathbf{k}$-dependent correction to the non-interacting energy.

Here, we have computed the self-energy at the HF level. The self-energy is in general a complex quantity, where the real part of it contributes to the total energy, while the imaginary part denotes the life time of the quasiparticles. Once the self-energy (Σ) is computed, the total energy can be written as

$$E_k = \epsilon_k + Re\Sigma(\mathbf{k}). \tag{1.173}$$

In general $\Sigma(\mathbf{k})$ may include contributions from the exchange, correlation, disorder, electron–phonon scattering, etc., we have only restricted ourselves to the exchange energy.

1.11.3 Hartree-Fock energy

Next, we have to solve the Hartree-Fock (HF) Hamiltonian and obtain its eigensolutions. For this purpose, let us write down the Hamiltonian and the corresponding Schrödinger equation in real space with $(\psi_i(\mathbf{r}), \epsilon_i)$, namely,

$$\frac{\hbar^2}{2m}\nabla^2\psi(\mathbf{r}_i) - \frac{e^2}{2\mathcal{V}} \sum_j \int d^3\mathbf{r}' \frac{1}{|\mathbf{r}-\mathbf{r}'|} \psi_j^*(\mathbf{r}')\psi_i(\mathbf{r})\psi_j(\mathbf{r}) = \epsilon_i\psi_i(\mathbf{r}). \tag{1.174}$$

In the absence of a better estimate for the wavefunction, we assume a plane wave solution of the form, $\psi_i = \frac{1}{\sqrt{\mathcal{V}}} e^{i\mathbf{k}.\mathbf{r}}$, where a box normalization of the plane wave is assumed in the volume $\mathcal{V}$. Plugging this plane wave solution in the above equation (let us only concentrate on the interaction term as the kinetic energy term will trivially yield $\frac{\hbar^2 k^2}{2m}$),

$$\frac{e^2}{2\mathcal{V}} \sum_{\mathbf{k}_j} \int d^3\mathbf{r}' \frac{e^{-i(\mathbf{k}_j-\mathbf{k}_i).\mathbf{r}'}}{|\mathbf{r}-\mathbf{r}'|} \frac{1}{\sqrt{\mathcal{V}}} e^{i\mathbf{k}.\mathbf{r}}.$$

Replace $\mathbf{r} - \mathbf{r}' = \mathbf{z}$, which allows us to write,

$$\frac{e^2}{2\mathcal{V}} \sum_{\mathbf{k}_j} \int d^3\mathbf{z} \frac{e^{-i(\mathbf{k}_j - \mathbf{k}_i).\mathbf{z}}}{|\mathbf{z}|} \frac{1}{\sqrt{\mathcal{V}}} e^{i\mathbf{k}.\mathbf{r}}.$$

This confirms that the plane wave states indeed are the solutions of the HF Hamiltonian as just a coefficient consisting of

$$\frac{e^2}{2\mathcal{V}} \sum_{\mathbf{k}_j} \int d^3\mathbf{z} \frac{e^{-i(\mathbf{k}_j - \mathbf{k}_i).\mathbf{z}}}{|\mathbf{z}|}$$

is multiplied with the plane wave states. This allows us to solve for the energy, ϵ_k as

$$\begin{aligned}
\epsilon_k &= \frac{\hbar^2 k^2}{2m} - \frac{e^2}{\mathcal{V}} \sum_{\mathbf{k}_j} \int d^3\mathbf{z} \frac{e^{-i(\mathbf{k}_j - \mathbf{k}_i).\mathbf{z}}}{|\mathbf{z}|} \\
&= \frac{\hbar^2 k^2}{2m} - e^2 \sum_{\mathbf{k}' < \mathbf{k}_F} \frac{4\pi}{|\mathbf{k} - \mathbf{k}'|^2} \\
&= \epsilon_k^0 + \Sigma_{ex}(\mathbf{k}). \qquad (1.175)
\end{aligned}$$

Thus, the single-particle energies are getting renormalized by the second term, which can now be solved by converting the sum into an integral,

$$\begin{aligned}
\Sigma^{ex}(\mathbf{k}) &= -\int \frac{d^3\mathbf{k}}{(2\pi)^3} \frac{4\pi e^2}{|\mathbf{k} - \mathbf{k}'|^2} = -\frac{e^2}{\pi} \int_0^{k_F} k'^2 dk' \int_{-1}^{+1} \frac{d(\cos\theta)}{k^2 + k'^2 - 2kk'\cos\theta} \\
&= -\frac{e^2}{\pi k} \int_0^{k_F} k' dk' ln \left| \frac{k + k'}{k - k'} \right| = -\frac{e^2 k_F}{\pi} \left[1 + \frac{1 - \alpha^2}{2\alpha} ln \left| \frac{1 + \alpha}{1 - \alpha} \right| \right] = -\frac{e^2 k_F}{\pi} F(\alpha) \qquad (1.176)
\end{aligned}$$

where k_F is the Fermi wave vector and $\alpha = \frac{k}{k_F}$. $F(\alpha)$ contains the wavevector dependence of the self-energy. Here

$$F(\alpha) = 1 + \frac{1 - \alpha^2}{2\alpha} ln \left| \frac{1 + \alpha}{1 - \alpha} \right|.$$

We plot $F(\alpha)$ as a function of α in Fig. 1.8. As k approaches k_F, $F(\alpha)$ decreases indicating an increase in the value of the self-energy, Σ^{ex} for electrons near the Fermi level.

Hence, the ground state energy per particle is obtained after a straightforward calculation as

$$E_g^{ex} = \sum_k n_k \Sigma^{ex}(\mathbf{k}) = \frac{\left(ek_F^2/\pi\right)^2}{2\pi} \int_0^1 F(\alpha) d\alpha = -\frac{3}{4} \frac{e^2 k_F}{\pi}. \qquad (1.177)$$

1.11.4 Magnetic properties

It is known that magnetism does not arise from the interaction between the magnetic moments, but from the exchange part of the Coulomb interaction. In fact, the exchange interaction (Heitler–London type) led us to the spin-only Heisenberg model or the itinerant Hubbard model, which yields a good

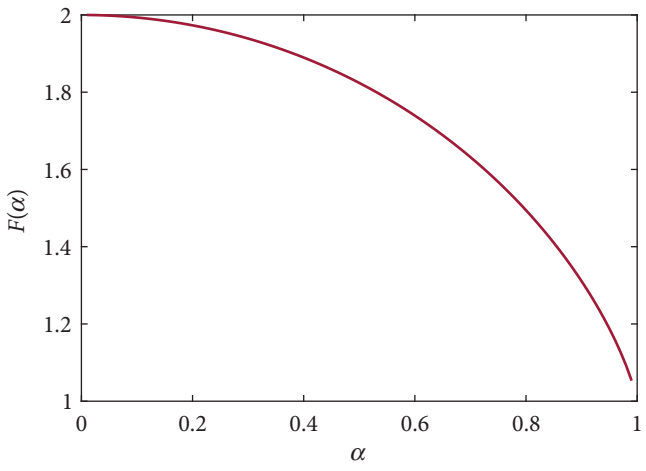


FIG. 1.8
$F(\alpha)$ is plotted as a function of α. As $k \to k_F$, $F(\alpha)$ smoothly decreases.

semi-quantitative basis for describing ferromagnetism in rare earth compounds. However, to describe the itinerant ferromagnetism in the transition metals, it is appropriate to resort to the electron liquid, that is, the carriers are moving in the background of positively charged ions, or the jellium model. In the non-interacting version of it, the kinetic energy of the carriers competes with the ordering energy, which is needed to align (or anti-align) the spins in the presence of an external field. This results in a net effect which is weak and vanishes in the limit of zero external fields. Such a scenario where the magnetic phenomena are absent is referred to as a non-magnetic or a Pauli paramagnetic behavior of the electron liquid.

Now when the electron-electron interaction is included, the spin alignment is mediated by the exchange interaction and can even be strong enough to sustain the order without any external magnetic field. To remind ourselves, the HF energy per electron (without taking into spins) is

$$\left(\frac{E}{\mathcal{N}}\right)_{\rm HF} = \frac{30.1}{(r_s/a_0)^2}\text{eV} - \frac{12.5}{(r_s/a_0)}\text{eV}. \tag{1.178}$$

To discuss magnetic behavior, we need to invoke the role of spins. A straightforward and efficient way to do this is writing down the HF energy in terms of the spin-dependent particle densities. Introducing the total density, $n\ (= \frac{\mathcal{N}}{\mathcal{V}})$ explicitly in the above expression,

$$\left(\frac{E}{\mathcal{N}}\right)_{\rm HF} = \mathcal{N}\left[78.2\left(\frac{\mathcal{N}}{\mathcal{V}}\right)^{2/3}\text{eV} - 20.1\left(\frac{\mathcal{N}}{\mathcal{V}}\right)^{1/3}\text{eV}\right] \tag{1.179}$$

where we used

$$\frac{r_s}{a_0} = \left[\frac{3}{4\pi}\left(\frac{\mathcal{V}}{\mathcal{N}}\right)\right]^{1/3}$$

Introducing the number of $\uparrow$- ($\mathcal{N}_\uparrow$) and $\downarrow$- ($\mathcal{N}_\downarrow$) spins,

$$\begin{aligned}\left(\frac{E}{\mathcal{N}}\right)_{\rm HF}(\mathcal{N}_\uparrow,\mathcal{N}_\downarrow) &= \left(\frac{E}{\mathcal{N}}\right)_{\rm HF}\mathcal{N}_\uparrow + \left(\frac{E}{\mathcal{N}}\right)_{\rm HF}\mathcal{N}_\downarrow \\ &= \left[78.2\text{eV}\left(\frac{\mathcal{N}_\uparrow}{\mathcal{V}}\right)^{2/3} + \left(\frac{\mathcal{N}_\downarrow}{\mathcal{V}}\right)^{2/3}\right] - \left[20.1\text{eV}\left(\frac{\mathcal{N}_\uparrow}{\mathcal{V}}\right)^{1/3} + \left(\frac{\mathcal{N}_\downarrow}{\mathcal{V}}\right)^{2/3}\right]\end{aligned} \tag{1.180}$$

with the condition that $\mathcal{N}_\uparrow + \mathcal{N}_\downarrow = \mathcal{N}$. It is customary to introduce the spin polarization operator, namely,

$$P = \frac{\mathcal{N}_\uparrow - \mathcal{N}_\downarrow}{\mathcal{N}} \tag{1.181}$$

which allows us to write the number of spin-polarized particles as

$$\mathcal{N}_\uparrow = \frac{\mathcal{N}}{2}(1+P); \quad \mathcal{N}_\downarrow = \frac{\mathcal{N}}{2}(1-P). \tag{1.182}$$

This facilitates writing down the HF energy in terms of P as

$$\left(\frac{E}{\mathcal{N}}\right)_{\text{HF}} = \mathcal{N}\left[\frac{1}{2}\{(1+P)^{5/3} + (1-P)^{5/3}\} - \frac{1}{8}\left(\frac{\mathcal{V}}{\mathcal{N}}\right)^{1/3}\{(1+P)^{4/3} + (1-P)^{4/3}\}\right]. \tag{1.183}$$

Let us now explore non-zero polarization ($P \neq 0$), or in other words non-vanishing spontaneous magnetization. Such a scenario will depend upon the sign of $E(\mathcal{N}, P) - E(\mathcal{N}, 0)$ (say $=\Delta E$). If $\Delta E < 0$, then spontaneous magnetization exists, and hence a ferromagnetic state is stabilized. Thus,

$$\frac{\Delta E(P)}{\mathcal{N}} = \frac{1}{2}\left[\{(1+P)^{5/3} + (1-P)^{5/3}\} - \frac{1}{8}\beta\{(1+P)^{4/3} + (1-P)^{4/3}\} - \mathcal{N}\left(1 - \frac{\beta}{4}\right)\right] \tag{1.184}$$

where β is related to the inverse of density, via, $\beta = (\frac{\mathcal{V}}{\mathcal{N}})^{1/3}$. $\frac{\Delta E(P)}{\mathcal{N}}$ plotted as a function of P is shown below (see Fig. 1.9 for a few values of β which shows a change in sign at a critical density given by $\beta_c = \frac{2}{5}(1 + \sqrt[3]{2}) \simeq 0.9$. This value of β_c translates to a density that corresponds to $r_s/a_0 \geq 5.5$. Clearly, this number is related to the density of the electron liquid (or rather the inverse of it) and hence the critical density for ferromagnetic order to occur should be low enough such that the Wigner–Seitz radii (r_s) should at least be $\sim$5.5 a_0. However, as we shall see that this is too stringent a criterion, and is only successfully cleared by Cesium (Cs). Even the best known ferromagnets, such as Fe, Ni and Co, do not meet the criterion.

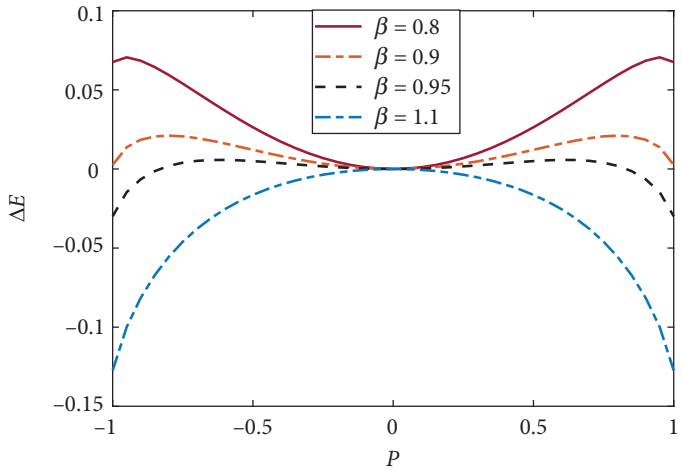


FIG. 1.9
The change in energy per particle is shown as a function of the polarization, P. The magnetic state stabilizes for large β.

Evidently, the premises on which these inferences are based on, that is, the HF approximation, is disconnected from reality. In fact, the maximum density at which ferromagnetism can occur given by $r_s/a_0 \geq 5.5$ is short by one order of

magnitude. It has been found by Zhong *et al.* (2002) that the upper limit of density corresponds to $r_s/a_0 \simeq 50$. The reason for such a large discrepancy lies in the jellium model itself. While the jellium model correctly assesses that completely free electrons (as in metals) can not result in ferromagnetic correlations, however, on the other hand, fails to account for the best known ferromagnets, such as Fe, Ni and Co to have ferromagnetic ordering. A very important fact that the jellium model misses is the combination of the band structure effects along with the exchange interaction.

To make the ongoing discussion more concrete, let us consider a crystal lattice, and more importantly, a specific form of the interaction in this case, the interaction is between the electrons at the same lattice site, and obeys the Pauli exclusion principle. In particular, we wish to talk about the Hubbard model, study its properties, and wish to explore its utility for studying the magnetic properties of solids.

1.12 THE HUBBARD MODEL: AN INTRODUCTION

In a real solid there are atoms or ions which are periodically placed and the electrons are usually free to move through such an array. The atoms or ions have very complex energy levels (or orbitals). The Hubbard model (Hubbard, 1963; Doniach and Sondheimer, 1974) simplifies the description of the constituent atoms with a periodic array of sites with a single energy level. This serves as an approximate description for materials where one energy band is in the vicinity of the Fermi surface, and hence only one orbital is important. With such a postulate, the Hilbert space of the model is restricted to four choices, which are, $|0\rangle, |\uparrow\rangle, |\downarrow\rangle, |\uparrow\downarrow\rangle$. When the electron density is sufficiently large in such a scenario, the electrons can interact pairwise via a screened Coulomb potential, with the largest interaction being between the two electrons residing on the same site, that is, for a '*doublet*' state, $|\uparrow\downarrow\rangle$. Thus, all the other three states are approximated in a manner as if they are experiencing no potential. The Hubbard model enunciates interaction energy, U, if a particular site is doubly occupied (needless to say that by opposite spins to conform to the Pauli exclusion principle). At a site i, this simplified model is represented by an interaction energy of the form $Un_{i\uparrow}n_{j\downarrow}$. Any term such as $Vn_{i\sigma}n_{j\sigma}$, where $i \neq j$ are excluded here and are included in the extended version of the Hubbard model.

Thus, putting these ideas together, the Hubbard Hamiltonian in the grand canonical ensemble is written as

$$\mathcal{H} = -t \sum_{\langle ij\rangle,\sigma} (c^{\dagger}_{i\sigma} c_{j\sigma} + h.c.) - \mu \sum_{i,\sigma} n_{i,\sigma} + U \sum_{i} n_{i\uparrow} n_{i\downarrow}. \tag{1.185}$$

The first term is the kinetic energy, which illustrates the hopping of electrons from a lattice site j to a site i with an energy scale t that could be determined by the overlap of orbitals for the neighboring atoms (c, $c^{\dagger}$ being the fermion annihilation and creation operators). The second term is the chemical potential, which fixes the electron density and the last term is the most 'simplified' The Hubbard (onsite) interaction term that we have discussed earlier. It says that a doubly occupied site, such as $|\uparrow\downarrow\rangle$ will have an energy U. Thus, if U is large (see discussion below), it will cost large energy to form a doubly occupied site, and would correspond to an insulator, known as a Mott insulator.

1.13 SYMMETRIES OF THE HUBBARD MODEL

Exploring symmetries in a model Hamiltonian can be quite helpful in understanding the concepts that govern various physical phenomena. If the onsite interaction, U, is uniform at all lattice sites (the most commonly made assumption), then the U-term is invariant under all symmetry operations of the lattice. Additionally, the spin-rotational invariance and the particle-hole symmetry are important ingredients for understanding magnetism and the electronic properties.

1.13.1 Spin-rotational invariance

The kinetic energy of the model describe the hopping of electrons for both spins from one lattice site to another. Since this term has got nothing to do with spins, it is invariant under the rotation of spins. Now, let us consider the interaction term, $n_{i\uparrow}n_{i\downarrow}$.

$$\begin{aligned} n_{i\uparrow}n_{i\downarrow} &= c^{\dagger}_{i\uparrow}c_{i\uparrow}c^{\dagger}_{i\downarrow}c_{i\downarrow} = c^{\dagger}_{i\uparrow}c_{i\uparrow}(1 - c_{i\downarrow}c^{\dagger}_{i\downarrow}) \\ &= n_{i\uparrow} - c^{\dagger}_{i\uparrow}c_{i\uparrow}c_{i\downarrow}c^{\dagger}_{i\downarrow} = n_{i\uparrow} - c^{\dagger}_{i\uparrow}c_{i\downarrow}c_{i\uparrow}c^{\dagger}_{i\downarrow} \\ &= n_{i\uparrow} - S^{+}_{i}S^{-}_{i}. \end{aligned} \tag{1.186}$$

Here, we have introduced the relation between the spin and the electron operators as

$$S^{\gamma}_{i} = \sum_{\alpha,\beta} c^{\dagger}_{i\alpha}\sigma^{\gamma}_{\alpha\beta}c_{i\beta}. \tag{1.187}$$

Here, for once, we use α, β to denote $\uparrow$ and $\downarrow$ spins. Since σ^{γ} denote $(\sigma_x, \sigma_y, \sigma_z)$, are the components of the Pauli matrix. We can also write, following the above, for the LHS,

$$n_{i\uparrow}n_{i\downarrow} = n_{i\downarrow} - S^{-}_{i}S^{+}_{i}, \tag{1.188}$$

owing to the exclusion principle, $n^{2}_{i\sigma} = n_{i\sigma}$ (which is the property of an idempotent matrix) and squaring $S^{z}_{i} = \frac{1}{2}(n_{i\uparrow} - n_{i\downarrow})$, one gets,

$$(S^{z}_{i})^{2} = \frac{1}{4}(n_{i\uparrow} + n_{i\downarrow} - 2n_{i\uparrow}n_{i\downarrow}). \tag{1.189}$$

Thus, the interaction term takes the form,

$$\mathcal{H}_{int} = \frac{UN}{2} - \frac{2U}{3}\sum_{i} \mathbf{S_i}^{2}. \tag{1.190}$$

It is clearly seen that this term has spin-rotational invariance as $\mathbf{S_i}^{2}$ has the eigenvalues given by

$$\begin{aligned} \mathbf{S_i}^{2} &= \frac{3}{4} \quad \text{for } |\uparrow\rangle, |\downarrow\rangle \\ &= 0 \quad \text{for } |0\rangle, |\uparrow\downarrow\rangle. \end{aligned} \tag{1.191}$$

If one notices the second term in Eq. (1.190) for a moment, it is large for uncompensated configurations, such as $|\uparrow\rangle$ and $|\downarrow\rangle$. Thus, the Hubbard term seeks for such configurations, compared with the other two.

1.13.2 Particle-hole symmetry

We can write down the interaction term in a symmetrized form as

$$\mathcal{H}_{int} = U \sum_i \left(n_{i\uparrow} - \frac{1}{2}\right)\left(n_{i\downarrow} - \frac{1}{2}\right). \tag{1.192}$$

We are going to show particle-hole symmetry of the Hamiltonian, which is important since it provides useful mappings between the repulsive (positive U) and the attractive (negative U) Hubbard Hamiltonians.

It is important to introduce the concept of a bipartite lattice in this context. The entire lattice here splits into two sublattices of the types A and B (which may or may not mean same type of atoms or ions), where A-atoms have neighbors as B-atoms, and vice versa. A square lattice and a honeycomb lattice are examples of bipartite lattices.

Under a particle-hole transformation, the $c_{i\sigma}$ and $c^\dagger_{i\sigma}$ operators transform into a different set of operators via,

$$d^\dagger_{i\sigma} = (-1)^i c_{i\sigma}. \tag{1.193}$$

The $(-1)^i$ takes a value $+1$ in one sublattice and -1 in the other sublattice. The nomenclature of particle-hole transformation is aptly justified since,

$$d^\dagger_{i\sigma} d_{i\sigma} = 1 - c^\dagger_{i\sigma} c_{i\sigma}. \tag{1.194}$$

It can be easily checked that the particle and hole occupations $n = c^\dagger_{i\sigma} c_{i\sigma} = 0, 1$ are interchanged under this transformation. The kinetic energy of course is an invariant under particle-hole transformation, that is,

$$c^\dagger_{i\sigma} c_{j\sigma} \rightarrow (-1)^{i+j} d_{i\sigma} d^\dagger_{j\sigma} = d^\dagger_{j\sigma} d_{i\sigma} \tag{1.195}$$

where we have used the fermionic anticommutation relation, and in a bipartite lattice $(-1)^{i+j} = -1$.

Now let us look at the interaction term, namely, $U(n_{i\uparrow} - \frac{1}{2})(n_{i\downarrow} - \frac{1}{2})$. It is easy to check that,

$$U\left(n_{i\uparrow} - \frac{1}{2}\right)\left(n_{i\downarrow} - \frac{1}{2}\right) = U n_{i\uparrow} n_{i\downarrow} - \frac{U}{2}(n_{i\uparrow} + n_{i\downarrow}) + \frac{U}{4} \tag{1.196}$$

with the last two terms being constants, the first term on the RHS represents the familiar Hubbard interaction. One can trivially show that the form is preserved under the particle-hole transformation, and one gets the same form for the Hubbard Hamiltonian in terms of the d (hole)-operators. The readers are encouraged to complete a few steps of algebra to convince themselves.

1.13.3 Extreme limits of the Hubbard model

It is instructive to explore the two opposite limits of the Hubbard model. They are non-interacting limit, $(U = 0)$ and the opposite of that, which is, $t = 0$, when the system splits into a collection of individual

atoms. The two limits are more familiarly categorized as $U/t \to 0$ and $U/t \to \infty$ respectively. In the non-interacting or the band limit, one gets the tight binding Hamiltonian,

$$\mathcal{H} = \sum_{k,\sigma} (\varepsilon_k - \mu) n_{k\sigma} = \sum_k \xi_k n_{k\sigma} \tag{1.197}$$

with $\xi_k = (\varepsilon_k - \mu)$. Now let us make the case more concrete by taking a two-dimensional square lattice, for which the band energies are given by

$$\varepsilon_k = -2t(\cos k_x a + \cos k_y a) \tag{1.198}$$

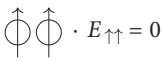
a is the lattice constant. Changing μsmoothly from $-4t$ to $+4t$ ($\pm 4t$ denote the band edges, resulting in a bandwidth of $8t$), n changes from 0 to 2. Of course, in the absence of the interaction, the Hamiltonian in Eq. (1.197) endorses a metallic state, which can further be confirmed from the calculation of the compressibility, $\kappa = (\frac{\partial^2 E}{\partial n^2})^{-1}$ at $T = 0$ which is proportional to the density of states. $\kappa \neq 0$ points toward a metallic behavior.

Furthermore, the band energies given in Eq. (1.198) is shown in the surface plots in Fig. 1.10 where we have taken $t = 1$ and $a = 1$.

On the other hand, in the extreme strong coupling limit or the atomic limit ($t = 0$), the Hubbard Hamiltonian is

$$\mathcal{H} = \sum_i U n_{i\uparrow} n_{i\downarrow}, \tag{1.199}$$

where the chemical potential or the atomic energy has been set to zero. The ground state of this model is tremendously degenerate. For N sites at half filling (number of electrons is also N) the spin degeneracy alone is 2^N. Thus, the degeneracy corresponding to $U \to \infty$ is exponentially large. However, an effective spin Hamiltonian can be obtained, which is also useful for the study of magnetism. At half filling, in the limit $U/t \to \infty$, one can derive the Heisenberg model from the Hubbard model. Here we sketch the derivation pictorially for two sites i and j (see Fig. 1.11). Thus, the energy cost is $E_{\uparrow\downarrow} = -\frac{2t^2}{U}$. The factor 2 in the right arises because of hopping can occur from left to right or right to left. Thus, a state with two spins in a singlet configuration can gain kinetic energy by tunneling to doubly occupied states. Whereas, there will be no hopping corresponding to



owing to the Pauli principle. This leads to an effective Heisenberg Hamiltonian valid for two spins,

$$\mathcal{H}_{\text{Heisenberg}} = J S_1 \cdot S_2 \tag{1.200}$$

where $J = \frac{2t^2}{U}$. In fact, in a d-dimensional hypercubic lattice with z neighbors, $J = \frac{zt^2}{U}$. If we extend the case for two site problem to an array of sites with full spin rotational symmetry (over a unit sphere) for the spin angular momentum, we obtain the familiar Heisenberg model,

$$\mathcal{H} = J \sum_{\langle ij \rangle} \mathbf{S}_i \cdot \mathbf{S}_j. \tag{1.201}$$

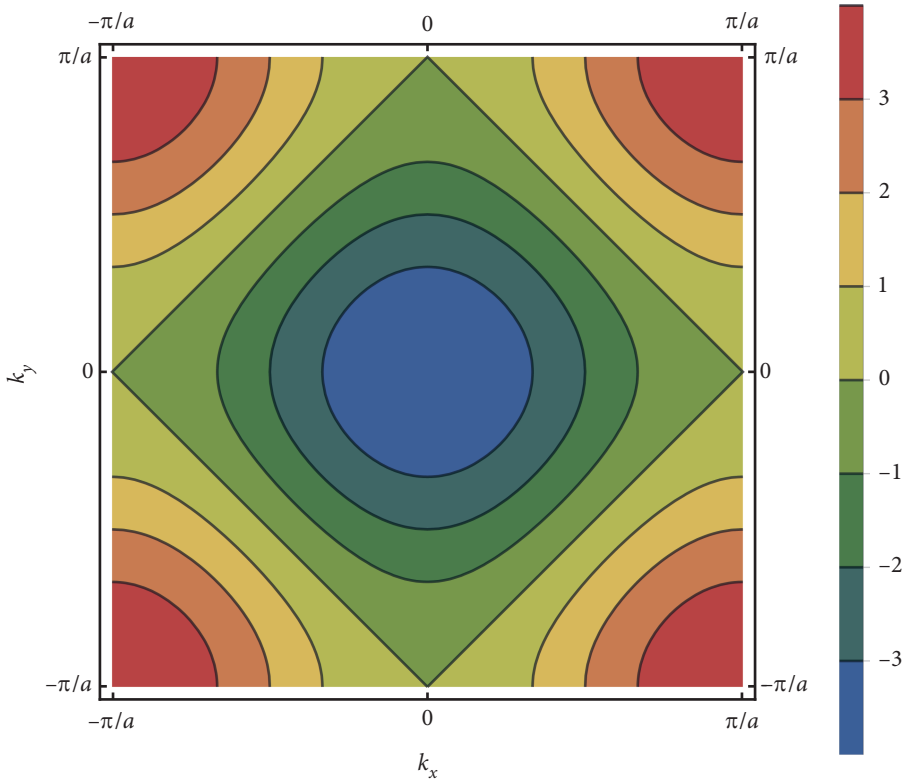


FIG. 1.10
Contour plot of the dispersion surface $\varepsilon_k = -2t(\cos k_x a + \cos k_y a)$. The shapes shown here, for example, blue circle at the center (corresponding to low filling), green rhombus (half filling), etc., demonstrate constant energy surfaces.

We wish to make it clear that we shall study the magnetic phenomena via an electronic model, for example, the fermionic Hubbard model. Except in $d = 1$, where the model can be solved exactly using Bethe ansatz, in higher dimensions, there are no exact solutions. We shall mostly concentrate on $d = 2$, and hence use controlled approximations to solve the Hubbard Hamiltonian. The reader should be careful and clear in mind about the inadequacies of the approximation used, and the extent of their validity.

Consider a generic Hamiltonian of the form,

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}'. \tag{1.202}$$

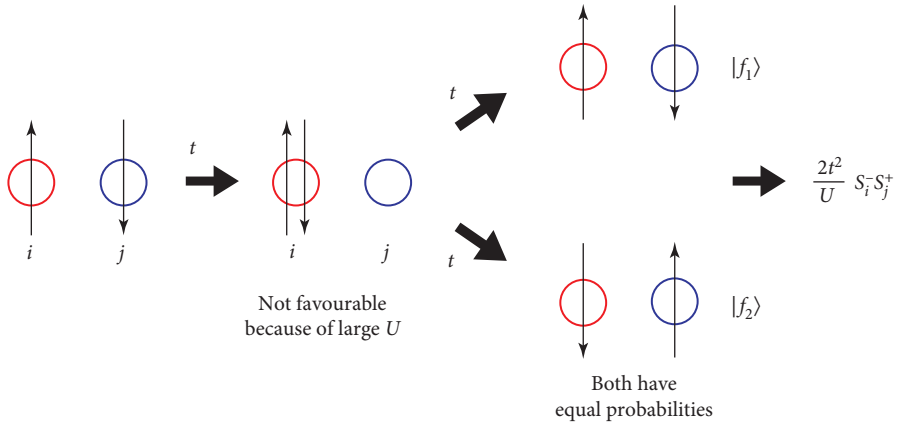


FIG. 1.11
Second order in t (think of perturbation theory with an energy denominator).

Usually, the eigensolutions of $\mathcal{H}_0$ are completely known. Furthermore, it can be assumed that the eigenvalues of $\mathcal{H}_0$ are bunched into different groups. The energy eigenvalues in each group are located in close vicinity to each other, however the levels belonging to different groups are separated far apart. Let us label the groups by $\alpha, \beta \cdots$ etc., while the energy levels in each group are denoted by p, q, etc. For example, $E_{p\alpha}$ denotes the p^{th} energy level in the group α.

Without any further conditions imposed, $\mathcal{H}'$ should have matrix elements involving eigenstates from different groups. Mathematically, stated, this is equivalent to $\langle \alpha, p|\mathcal{H}'|\beta, q\rangle \neq 0$ for $\alpha \neq \beta$. However, if $\mathcal{H}'$ is weak, that is, it can be used as a perturbation term, then we can do a unitary transformation on $\mathcal{H}'$, and hence obtain an effective Hamiltonian, $\mathcal{H}_{eff}$. Thus $\mathcal{H}_{eff}$ does not contain matrix elements between different groups and only the elements involving states within a group. The matrix elements of $\mathcal{H}_{eff}$ upto second order in $\mathcal{H}'$ are written as Cohen-Tonnoudji *et al.* (1977),

$$\langle \alpha, p|\mathcal{H}_{eff}|\beta, q\rangle = E_\alpha \delta_{p,q} + \langle p, \alpha|\mathcal{H}'\alpha, q\rangle + \frac{1}{2}\sum_{\beta,k\neq\alpha} \langle \alpha, p|\mathcal{H}'|\beta, k\rangle\langle \beta, k|\mathcal{H}'|\alpha, q\rangle \left[\frac{1}{E_{\alpha,p} - E_{\beta,k}} + \frac{1}{E_{\alpha,q} - E_{\beta,k}}\right] \tag{1.203}$$

$\mathcal{H}_{eff}$ thus yields matrix elements involving a particular group, say α, and hence is diagonal.

We shall apply the above calculations to the Hubbard Hamiltonian [Eq. (1.185)] and aim to formally arrive at a super-exchange Hamiltonian (as depicted earlier pictorially) in the following. Being one of the simplest cases, let us consider two sites, namely, 1 and 2, and two electrons with spins $\uparrow$ and $\downarrow$. The kinetic energy, t can be taken as the perturbation, that is, $t \ll U$. There are six states corresponding to the extreme strong coupling limit, that is, $t = 0$. They are,

i. both are ↑-spin particles and hence belong to different sites (Pauli's exclusion principle), that is,
$|1\rangle = c^\dagger_{1\uparrow} c^\dagger_{2\uparrow}|0\rangle$, with energy, $E = 0$

ii. both are ↓-spin particles and hence belong to different sites (Pauli's exclusion principle), that is,
$|2\rangle = c^\dagger_{2\downarrow} c^\dagger_{2\downarrow}|0\rangle$, with energy, $E = 0$

iii. ↑-spin particle in site 1, ↓-spin particle in site 2, that is,
$|3\rangle = c^\dagger_{1\uparrow} c^\dagger_{2\downarrow}|0\rangle$, with energy, $E = 0$

iv. ↓-spin particle in site 1, ↑-spin particle in site 2, that is,
$|4\rangle = c^\dagger_{1\downarrow} c^\dagger_{2\uparrow}|0\rangle$, with energy, $E = 0$

v. ↑-spin particle in site 1, ↓-spin particle in site 1, that is,
$|5\rangle = c^\dagger_{1\downarrow} c^\dagger_{1\uparrow}|0\rangle$, with energy, $E = U$

vi. ↑-spin particle in site 2, ↓-spin particle in site 2, that is,
$|6\rangle = c^\dagger_{2\downarrow} c^\dagger_{2\uparrow}|0\rangle$, with energy, $E = U$.

The purpose of writing down the effective Hamiltonian is to eliminate the last two states, that is, (v) and (vi) with large energy (of magnitude U).

We can now switch to a small hopping term, t, so that the Hamiltonian can be written as $\mathcal{H} = \mathcal{H}_0 + \mathcal{H}'$, where,

$$\mathcal{H}_0 = U(n_{1\uparrow} n_{1\downarrow} + n_{2\uparrow} n_{2\downarrow})$$
$$\mathcal{H}' = -t\left(c^\dagger_{1\uparrow} c^\dagger_{2\uparrow} + c^\dagger_{1\downarrow} c^\dagger_{2\downarrow} + c^\dagger_{2\uparrow} c^\dagger_{1\uparrow} + c^\dagger_{2\downarrow} c^\dagger_{1\downarrow}\right). \tag{1.204}$$

Let us evaluate a non-zero matrix element, for example,

$$\langle 6|\mathcal{H}'|3\rangle = -t\langle 0|c_{2\downarrow} c_{2\uparrow}(c^\dagger_{1\uparrow} c_{2\uparrow} + c^\dagger_{1\downarrow} c_{2\downarrow} + c^\dagger_{2\uparrow} c_{1\uparrow} + c^\dagger_{2\downarrow} c_{1\downarrow}) c^\dagger_{2\downarrow} c^\dagger_{1\uparrow})|0\rangle.$$

There are 4 terms here, and they can be shown to have the following values, namely,

a. $-t\langle 0|c_{2\downarrow} c_{2\uparrow} c^\dagger_{1\uparrow} c_{2\uparrow} c^\dagger_{2\downarrow} c^\dagger_{1\uparrow})|0\rangle|0\rangle = 0.$
b. $-t\langle 0|c_{2\downarrow} c_{2\uparrow} c^\dagger_{1\downarrow} c_{2\downarrow} c^\dagger_{2\downarrow} c^\dagger_{1\uparrow})|0\rangle|0\rangle = 0.$
c. $-t\langle 0|c_{2\downarrow} c_{2\uparrow} c^\dagger_{2\uparrow} c_{1\uparrow} c^\dagger_{2\downarrow} c^\dagger_{1\uparrow})|0\rangle|0\rangle = -t.$
d. $-t\langle 0|c_{2\downarrow} c_{2\uparrow} c^\dagger_{2\downarrow} c_{1\downarrow} c^\dagger_{2\downarrow} c^\dagger_{1\uparrow})|0\rangle|0\rangle = 0.$
e. $\langle 5|\mathcal{H}'|3\rangle = -t.$
f. $\langle 6|\mathcal{H}'|4\rangle = -t.$
g. $\langle 5|\mathcal{H}'|4\rangle = t.$

Now, up to the second order in t, using Eq. (1.203), one gets,

$$\langle 3|\mathcal{H}_{eff}|3\rangle = -\frac{2t^2}{U} = \langle 4|\mathcal{H}_{eff}|4\rangle$$

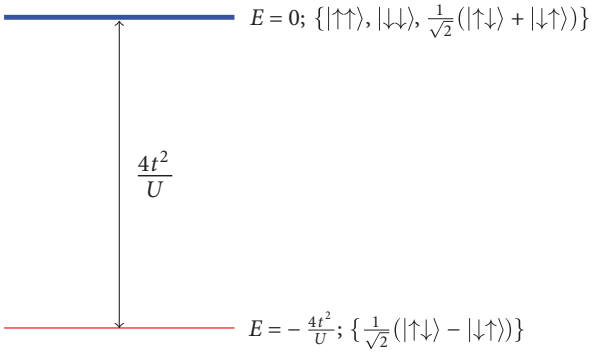


FIG. 1.12
The energy levels and the corresponding states are shown upto order t^2/U.

$$\langle 3|\mathcal{H}_{eff}|4\rangle = \frac{2t^2}{U} = \langle 4|\mathcal{H}_{eff}|3\rangle$$
$$\langle 1|\mathcal{H}_{eff}|1\rangle = \langle 1|\mathcal{H}_{eff}|2\rangle = \langle 2|\mathcal{H}_{eff}|1\rangle = \langle 2|\mathcal{H}_{eff}|2\rangle = 0. \quad (1.205)$$

Thus, $\mathcal{H}_{eff}$ connects the states $|3\rangle$ and $|4\rangle$ only up to order $\frac{t^2}{U}$. Thus on the basis of $|3\rangle$ and $|4\rangle$, One gets the following form for $\mathcal{H}_{eff}$, namely,

$$\mathcal{H}_{eff} = \begin{bmatrix} -\frac{2t^2}{U} & \frac{2t^2}{U} \\ \frac{2t^2}{U} & -\frac{2t^2}{U} \end{bmatrix}. \quad (1.206)$$

The eigenvalues are $E = 0$ and $E = -\frac{4t^2}{U}$. We show the energies and the corresponding states schematically in Fig. 1.12. Thus, we get an effective superexchange Hamiltonian of the form,

$$\mathcal{H}_{eff} = J\mathbf{S}_1 \cdot \mathbf{S}_2, \quad \text{with } J = \frac{4t^2}{U}. \quad (1.207)$$

1.14 FERROMAGNETISM IN THE HUBBARD MODEL: STONER CRITERION

To explore ferromagnetism in the Hubbard model, we write the Hamiltonian once again on a lattice,

$$\mathcal{H} = -t \sum_{\langle ij\rangle,\sigma} c^{\dagger}_{i\sigma} c_{j\sigma} + U \sum_i n_{i\uparrow} n_{i\downarrow}. \quad (1.208)$$

The model can be solved within a mean field approximation, which decouples the quartic term (in terms of the fermionic operator) into quadratic ones. This is known as the Hartree-Fock approximation, which allows,

$$U \sum_i n_{i\uparrow} n_{i\downarrow} \rightarrow U \sum_i \langle n_{i\uparrow}\rangle n_{i\downarrow} + U \sum_i \langle n_{i\downarrow}\rangle n_{i\uparrow}. \quad (1.209)$$

Let us provide a physical feel for the mean field approximation and its validity. Suppose a student is sitting in the class and, he (she) has friends who are sitting right beside him (her) and there are some friends who are sitting very far away, so that no communication (or interaction) is possible while the class is going on. However, the students who are sitting in the vicinity, can interact with that particular student. The situation is similar to a many body system, where the charge carriers interact strongly with other carriers that are in the vicinity, and lesser interaction ensues with the ones that are far off, and even lesser interaction with those which are farther off. However, one can reduce the complexity of the problem by making an approximation that an average field is acting on the charge carrier (that is, one

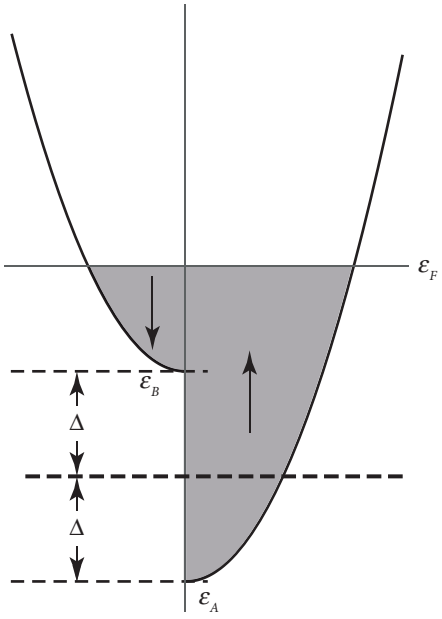


FIG. 1.13
The bands for ↑ and ↓-spins are shown. The Fermi energy (ε_F) is marked by a horizontal line, while the dashed horizontal line denotes a reference, which is at separated by an energy Δ from the bottom of the ↑ and ↓-spin bands. The shaded area denotes filled energy levels.

particular electron) due to the presence of all other carriers (just like an average effect on a particular student due to all other students in the class). This reduces the complicated many body phenomenon to a single particle problem (and hence solvable) at the expense of ignoring fluctuations which arise out of a differential nature of the interaction between charge carriers in the vicinity with those which are further away. This allows us to replace the operators n_i with their expectation values, thereby making the quartic operators at each lattice site to a quadratic ones.

Coming back to the context, in a ferromagnet, in one species of spin, that is, say ↑-spin outnumbers the other one, that is, ↓-spin (see Fig. 1.13), we can assume,

$$\langle n_{i\uparrow}\rangle \gg \langle n_{i\downarrow}\rangle. \tag{1.210}$$

Thus, magnetization, m at a given lattice site defined by $m_i = \langle n_{i\uparrow}\rangle - \langle n_{i\downarrow}\rangle = m$ can be computed via,

$$m = \int_{\varepsilon_A}^{\varepsilon_F} d\varepsilon N_\uparrow(\varepsilon) - \int_{\varepsilon_B}^{\varepsilon_F} d\varepsilon N_\downarrow(\varepsilon). \tag{1.211}$$

where $N_\sigma(\varepsilon)$ denotes the density of states for spin-σ and ϵ_A and ϵ_B denote the bottom of the bands for ↑ and ↓-spins, respectively. We may choose a level symmetrically between the bottom of the ↑ and ↓-spin bands denoted by the dashed horizontal line in the figure. Let us assume that the energy gap between this line and $\varepsilon_{A,B}$ is denoted by Δ one may write the integral as

$$m = \int_{-\Delta}^{\varepsilon_F} d\varepsilon N(\varepsilon + \Delta) - \int_{\Delta}^{\varepsilon_F} d\varepsilon N(\varepsilon - \Delta). \tag{1.212}$$

Here, the ↑ and ↓-spin density of states are denoted by $N(\varepsilon \pm \Delta)$. Assuming Δ to be small, one can perform a Taylor expansion of the density of states,

$$N(\varepsilon + \Delta) = \frac{dN}{d\varepsilon}\Delta \tag{1.213}$$

$$N(\varepsilon - \Delta) = \frac{dN}{d\varepsilon}(-\Delta) \text{ where } \Delta = \frac{Um}{2}. \tag{1.214}$$

Since the Fermi energy ε_F is much larger than Δ, one can set the lower limit to be zero and combine the integrals,

$$m = 2\Delta \int_{\Delta(=0)}^{\varepsilon_F} \frac{dN}{d\varepsilon} d\varepsilon \tag{1.215}$$

$$= 2\Delta N(\varepsilon_F) = mUN(\varepsilon_F). \tag{1.216}$$

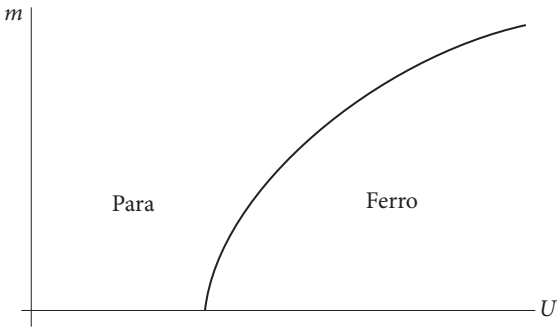


FIG. 1.14
m as a function of Hubbard *U*. The plot emphasizes Stoner criterion and shows the boundary separating the paramagnetic and ferromagnetic phases.

Canceling m from both sides, we arrive at a condition,

$$N(\varepsilon_F) = \frac{1}{U}. \tag{1.217}$$

This is known as the Stoner criterion, which yields a very stringent condition on the occurrence of ferromagnetism. From the figure, it is clear that the interaction term multiplied by the density of states at the Fermi level must be necessarily greater than unity and the boundary between the paramagnetic and ferromagnetic phases is indicated by the curved line (Fig. 1.14).

It is important to note that only Fe, Ni and Co pass the Stoner criterion. As a toy model, if we consider a two-dimensional square lattice whose density of states diverges at the half filling, that is one particle per site ($n = 1$), $U_c \to 0$ which says that the system is unstable against a ferromagnetic ordering even at a negligible interparticle interaction.

1.15 ANTIFERROMAGNETISM IN THE HUBBARD MODEL

Having discussed ferromagnetism in an itinerant electronic model, we shall discuss antiferromagnetism. For concreteness, we consider a two-dimensional square lattice, consisting of two sublattices, namely A and B sublattices. A sublattice contains predominantly $\uparrow$-spin particles (the $\uparrow$-spin density being large, with a small $\downarrow$-spin density). Similarly, the other sublattice, namely, the B sublattice contains predominantly $\downarrow$-spin density (along with a small $\uparrow$-spin density). It may be noted a perfect Neél order hasn't been assumed and a small density of the other type is essential for our purpose. Furthermore, every A sublattice site has 4 neighbors which are that of B sublattice and vice versa. Thus, there an antiferromagnetic ordering with a wave vector $q = (\pi, \pi)$ (Singh and Tesanović, 1990).

Once again, we take the Hubbard Hamiltonian and do an unrestricted Hartree-Fock approximation (with of course ignoring terms, such as $\langle c^\dagger c^\dagger \rangle$ and cc) on the interaction term, which yields,

$$\mathcal{H}_{int}^{HF} = U \sum_i \left[\langle c_{i\uparrow}^\dagger c_{i\uparrow} \rangle c_{i\downarrow}^\dagger c_{i\downarrow} + \langle c_{i\downarrow}^\dagger c_{i\downarrow} \rangle c_{i\uparrow}^\dagger c_{i\uparrow} - c_{i\uparrow}^\dagger c_{i\downarrow} \langle c_{i\downarrow}^\dagger c_{i\uparrow} \rangle - c_{i\downarrow}^\dagger c_{i\uparrow} \langle c_{i\uparrow}^\dagger c_{i\downarrow} \rangle \right] \tag{1.218}$$

where we have neglected a constant term of the form, $U \langle c_{i\sigma}^\dagger c_{i\sigma'} \rangle \langle c_{i\sigma'}^\dagger c_{i\sigma} \rangle$. Rewriting the decoupled interaction term,

$$\mathcal{H}_{int}^{HF} = U \sum_i \left[\langle n_{i\uparrow} \rangle n_{i\downarrow} + \langle n_{i\downarrow} \rangle n_{i\uparrow} \right]. \tag{1.219}$$

This implies that an ↑-spin particle feels a potential $U\langle n_{i\downarrow}\rangle$ and a ↓-spin Hamiltonian feels a potential $U\langle n_{i\uparrow}\rangle$, with the constraint that

$$\langle n_{\uparrow}^{i(A/B)}\rangle + \langle n_{\downarrow}^{i(A/B)}\rangle = 1. \tag{1.220}$$

Also, the sublattice magnetization m is defined by

$$\langle n_{\uparrow}^{i(A/B)}\rangle - \langle n_{\downarrow}^{i(A/B)}\rangle = \pm m \tag{1.221}$$

with m being replaced by $(-1)^i m$ implying a change in sign as one goes from A sublattice to a B sublattice. Thus, in terms of U and m, the potential experienced by ↑-spin and ↓-spin particles are given by

$$\begin{aligned} V_{\uparrow}(i) &= U\langle n_{i\downarrow}\rangle = \frac{U}{2} - \frac{Um}{2}(-1)^i \\ V_{\downarrow}(i) &= U\langle n_{i\uparrow}\rangle = \frac{U}{2} + \frac{Um}{2}(-1)^i. \end{aligned} \tag{1.222}$$

Defining, $\Delta = \frac{Um}{2}(-1)^i$ (distinguish it from the definition of Δ used in ferromagnets), one may write,

$$\begin{aligned} V_{\uparrow}(i \in A) &= \frac{U}{2} - \Delta = V_{\downarrow}(i \in B) \\ V_{\uparrow}(i \in B) &= \frac{U}{2} + \Delta = V_{\downarrow}(i \in A). \end{aligned} \tag{1.223}$$

Thus, the mean field Hamiltonians for each spin say ↑-spin is written as

$$\mathcal{H}_{\uparrow}(i) = -t\sum_{\langle ij\rangle} c_{i\uparrow}^{\dagger} c_{j\uparrow} + \sum_{i\in A,B} V_{\uparrow}(i) n_{i\uparrow} \tag{1.224}$$

where $n_{i\uparrow} = c_{i\uparrow}^{\dagger} c_{i\uparrow}$. Assuming translational invariance, one can Fourier transform the above Hamiltonian,

$$\mathcal{H}_{\uparrow}(k) = \sum_{k} \epsilon_k c_{k\uparrow}^{\dagger A} c_{k\uparrow}^{B} + \sum_{Q}\Big[(-\Delta) c_{k\uparrow}^{\dagger A} c_{k-Q\uparrow}^{A} + \Delta c_{k\uparrow}^{\dagger B} c_{k-Q\uparrow}^{B}\Big]. \tag{1.225}$$

Writing $\mathcal{H}_{\uparrow}(k)$ in the sublattice basis renders a 2×2 form,

$$\mathcal{H}_{\uparrow}(k) = \begin{bmatrix} -\Delta & \varepsilon_k \\ \varepsilon_k & \Delta \end{bmatrix} \begin{bmatrix} c_{k\uparrow}^{A} \\ c_{k-q\uparrow}^{B} \end{bmatrix}. \tag{1.226}$$

Diagonalizing, one gets the eigenvalues as

$$E_{k\uparrow} = \pm\sqrt{\Delta^2 + \varepsilon_k^2}. \tag{1.227}$$

The energy spectrum is plotted for a given value of U in Fig. 1.15. There is clearly a spectral gap that scales with U. Similarly, for the ↓-spin,

$$\mathcal{H}_{\downarrow}(k) = \begin{bmatrix} \Delta & \varepsilon_k \\ \varepsilon_k & -\Delta \end{bmatrix}. \tag{1.228}$$

Again, one gets, $E_{k\downarrow} = \pm\sqrt{\Delta^2 + \varepsilon_k^2} = E_{k\uparrow}$, that is the same as the ↑-spin.

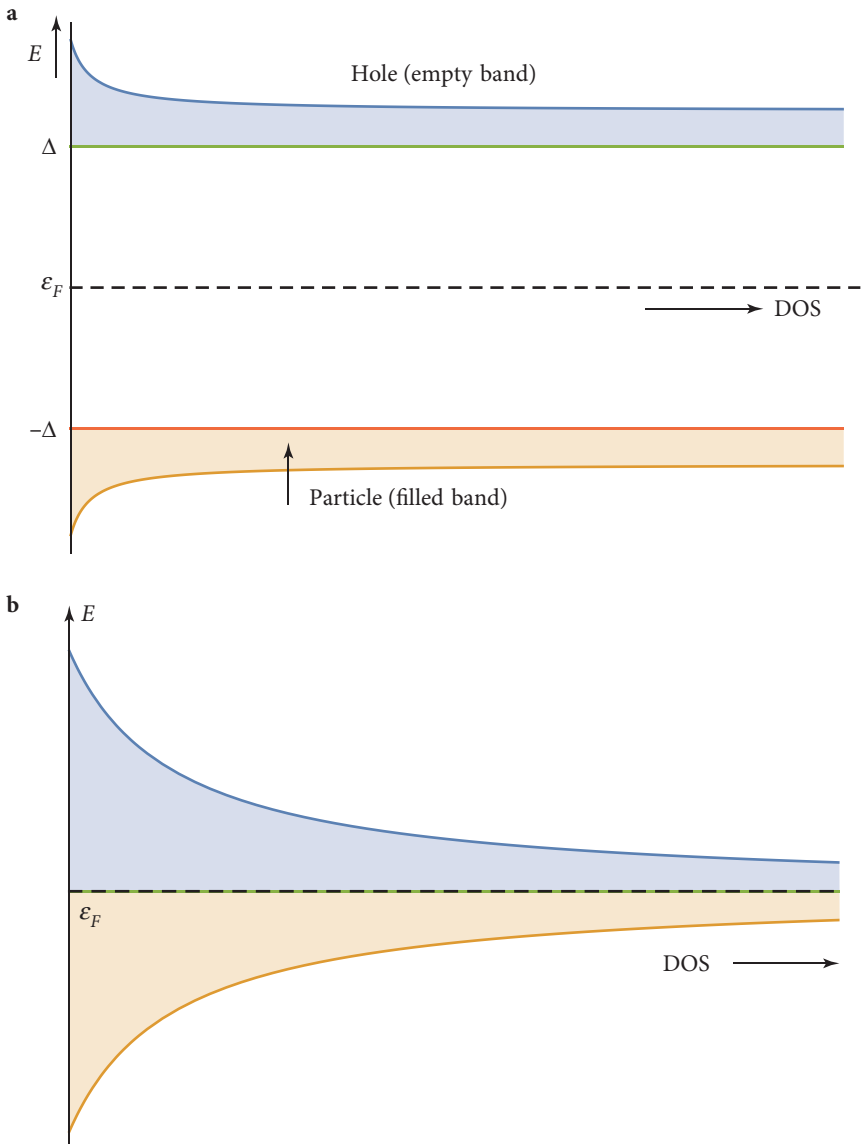


FIG. 1.15
Density of states (DOS) along *x*-axis is shown as a function of energy (*E*) shown along *y*-axis. (a) The lower (light yellow) and the upper (light blue) Hubbard bands are shown. The band gap scales with the Hubbard interaction, *U*. (b) The lower and upper Hubbard bands merge in the limit $U \rightarrow 0$.

The eigenvectors corresponding to the filled band (negative eigenvalue) for the ↑-spin are,

$$\begin{bmatrix}\alpha_{k\uparrow}\\ \beta_{k\uparrow}\end{bmatrix} = \begin{bmatrix}\sqrt{1-\beta_{k\uparrow}^2}\\ \frac{\varepsilon_k}{(E_{k\uparrow}-\Delta)^2+\varepsilon_k^2}\end{bmatrix}. \tag{1.229}$$

The magnetization, m can now be computed using,

$$m = n_\uparrow - n_\downarrow = \frac{1}{N/2}\sum_{k\in \text{ LHB}} (\alpha_{k\uparrow}^2 - \alpha_{k\downarrow}^2) \tag{1.230}$$

where LHB denotes the lower (filled) Hubbard band. Also, $\frac{N}{2}$ arises because we are summing the number of unit cells, which are $N/2$ numbers corresponding to N sites. This yields a self-consistent value of m,

$$m = \frac{2}{N}\sum_k \frac{\Delta}{E_k} = \frac{1}{N}\sum_k \frac{mU}{E_k} \tag{1.231}$$

where, $mU = 2\Delta$ (ignoring the sign that it picks up in going from one lattice to the next). This further can be simplified to,

$$\begin{aligned} 1 &= \frac{1}{N}\sum_k \frac{U}{E_k}\\ \frac{1}{U} &= \frac{1}{N}\sum_{k\in \text{LHB}} \frac{1}{E_k}. \end{aligned} \tag{1.232}$$

The above equation can be considered as a self-consistent equation for calculating magnetization, and requires a numerical solution for Δ $(= mU)$. Instead of resorting to the discussion of the numerical solution, let us consider different limits to explore the behavior of the magnetization.

1.15.1 Strong coupling limit

Here, we consider $U \gg t$ or equivalently $U \to \infty$ (that is greater than any other energy scale of the problem) where it is energetically unfavorable for electrons of opposite spins to occupy the same lattice site. Thus, $\Delta = \frac{mU}{2}$ is very large and $\beta_{k\uparrow}$ in Eq. (1.229) takes the form,

$$\beta_{k\uparrow} = \frac{\varepsilon_k}{(E_{k\uparrow}-\Delta)^2+\varepsilon_k^2} \simeq \frac{\varepsilon_k}{2\Delta} \simeq 0. \tag{1.233}$$

Thus, $\alpha_{k\uparrow} \approx 1$ (since $\alpha_{k\uparrow} = \sqrt{1-\beta_{k\uparrow}^2}$). Hence, the magnetization becomes,

$$m = \frac{2}{N}\sum_k 1 = \frac{2}{N}\times\frac{N}{2} \simeq 1 \tag{1.234}$$

where we used $\sum_k 1 = N$.

The same inference can also be obtained from the self-consistent equation,

$$\frac{1}{U} = \frac{1}{N}\sum_k \frac{1}{2\Delta} \tag{1.235}$$

$2\Delta = U$ implies $m = 1$. Thus, in the strong coupling limit, the sublattice magnetization attains its saturation value, namely, equal to 1. Quantum fluctuations reduce this value. For example, the magnetization obtained from quantum Monte Carlo calculations is 0.6 for a two-dimensional square lattice (Sorella *et al.*, 1989).

Now consider the opposite limit, that is, $U/t \to 0$ or equivalently, $\Delta/t \to 0$, then one gets for the self-consistent equation,

$$\frac{1}{U} = \frac{1}{2N}\sum_k \varepsilon_k. \tag{1.236}$$

Converting the sum into an integral,

$$\frac{1}{U} = \int N(\varepsilon)\frac{1}{\varepsilon}d\epsilon. \tag{1.237}$$

As a particular case in two dimensions (2D), the DOS has a divergence, that is,

$$N(\varepsilon) \sim ln\left(\frac{1}{\epsilon}\right) \text{ (van Hove singularity)}$$

Thus, one gets,

$$\frac{1}{U} \simeq \lim_{\varepsilon\to 0} ln\frac{\left(\frac{1}{\varepsilon}\right)}{\varepsilon} \simeq \left[ln\left(\frac{1}{\varepsilon}\right)\right]^2. \tag{1.238}$$

One can take a tiny cut off δ to satisfy the divergence of the RHS, which renders,

$$\frac{1}{U} \sim \left[ln\left(\frac{1}{\delta}\right)\right]^2 \tag{1.239}$$

yielding,

$t/U \sim [ln(t/\delta)]^2$

Thus, one gets, $\Delta/t \sim e^{-t/U}$, implying that the magnetization takes a form,

$$m \sim \frac{t}{U}e^{-t/U} \sim \frac{1}{U}e^{-1/U}. \tag{1.240}$$

Thus, as $U \to 0$, $m \to 0$, but in a complex fashion, that is,

$$m \simeq \frac{1}{U}e^{-1/U}.$$

A schematic plot of magnetization is shown in Fig. 1.16. Thus, at low interaction strengths, the magnetization vanishes as $\frac{1}{U}e^{-1/U}$, while, for obvious reasons, it saturates at a value 1 for large U. Hence, even without a self-consistent numerical

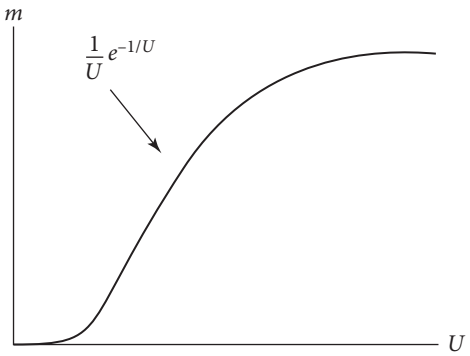


FIG. 1.16
Magnetization as a function of U in strong and weak couples regimes.

solution, we get an idea about the behavior of magnetization as a function of the Hubbard interaction parameter. While it is difficult to guess the exact form at low interaction strengths, it is still possible to reconcile that magnetization vanishes as $t \gg U$.

1.15.2 Summary and outlook

In this chapter, we have given an overview of different magnetic orders in solids. What follows afterward is a brief discussion of magnetism in materials with atoms having filled and partially filled shells. Pauli paramagnetism, Curie's law and Landau diamagnetism are hence discussed. We have underscored the role of electron-electron interactions in explaining ferromagnetism and antiferromagnetism. The mean field theory of ferromagnetism is introduced. As an extension, spin wave theory is applied to ferromagnets using Holstein–Primakoff transformation. A simplified version of the Heisenberg model, that is, without the interaction between the z-component of the spins at neighboring sites, namely, the quantum XY model is solved. Furthermore, antiferromagnetism is studied within a similar formalism. Yet another technique is employed, namely, the transfer matrix approach to study the critical properties of a ferromagnet to a paramagnet phase transition. The magnetic susceptibility is hence studied within a linear response theory. To wind up our discussion, we have introduced an itinerant electronic model to study the properties of materials, namely, the Hubbard model and discuss its properties. The weak and the strong coupling limits of the model are studied. The model is hence solved using Hartree-Fock approximation. The magnetic properties of both ferromagnets and antiferromagnets are studied, where the magnetization is solved self-consistently as a function of interparticle (Hubbard) interaction.

Summarizing, apart from studying diamagnetism and paramagnetism, which can be comprehensively understood without invoking the electron-electron interaction, we have presented interacting models involving spins and electronic degrees of freedom and computed the key magnetic property, that is, magnetization (or equivalently, the magnetic susceptibility) as a function of the interparticle interaction parameter.

1.16 APPENDIX

The electronic configuration of elements can be understood by adding one electron at a time to the available energy levels of the atoms. Each electron can be added adhering to the Pauli exclusion principle, and Hund's rule (discussed later). There are a number of factors that decide the energy of an electron in an orbital. They are, respectively, the mass number, (that is number of protons), average distance of the orbital from the nucleus, and last, but not the least, screening effects. The last one needs a special mention, particularly for many-electron atoms, where the positive charge of the nucleus is partially screened to an electron in the outer shells owing to the presence of intervening electrons. The screening is not uniform for all the orbitals of a given energy level. The orbitals of an atom overlap the region surrounding the nucleus. The more an orbital penetrates onto the negative cloud of the

screening electrons, the more strongly the electron is pulled by the nucleus, and consequently the corresponding energy will be lower. For example, corresponding to the same energy level (same value of the principle quantum number, n), the smaller the value of the orbital quantum number, l, the larger the probability of finding the electron near the nucleus. Thus, an s-electron ($l = 0$) for any n is more penetrating than the p-electron ($l = 1$). Thus, for a given n, the energies are arranged as $E_s < E_p < E_d < E_f \ldots$ (d corresponds to $l = 2$ and f corresponds to $l = 3$).

Even the above classification is not free of loopholes. For instance, the 4s orbital ($n = 4$, $l = 0$) may have a lower energy than a $3d$ orbital ($n = 3$, $l = 2$). Here is where the Aufbau ordering, and Hund's rule come into picture for deriving the correct electronic configuration of atoms. We shall discuss them below.

1.17 RS COUPLING

In a many-electron system, the combination of the orbital and the spin degrees of freedom is quite complicated, owing to the presence of spin–spin, orbit–orbit, and spin–orbit couplings. In the Russel–Saunders's (RS) scheme, it is generally assumed that the strength of the spin–orbit coupling is the weakest, followed by the orbit–orbit coupling and the spin–spin coupling, respectively. This gives rise to two principle coupling schemes, namely, RS coupling (same as LS coupling) and jj coupling. The above hierarchy of the coupling strengths is found to be valid for the first row transition series, where the coupling of the spin and orbit can be ignored. However, for larger elements (mostly with the ones with atomic number greater than thirty), the spin–orbit coupling becomes increasingly important, and the jj coupling scheme should be used.

The coupling schemes are best described by the selection rules that govern the transition probabilities between different atomic energy levels. Not all possible transitions actually occur. These are the constraints imposed by the selection rules. As a particular case, consider the electric dipole transitions. Assuming LS coupling, the selection rules can be stated as in the following.

i. Only one electron makes a transition at a time.
ii. The l-value will only alter by one unit, that is, $\Delta = \pm 1$.
    This ensures that the parity of the wavefunction must change in an electric dipole transition. To remind the readers, the parity of the wavefunction is denoted by the factor $(-1)^l$, which implies that if l is even, the wavefunction has even parity[15] while if l is odd, it has odd parity. Since the electric dipole moment ($\mathbf{p} = e\mathbf{r}$) involves a Hamiltonian of the form $\mathcal{H}' \sim r$, which itself has odd parity, the transition matrix elements $\langle f|\mathcal{H}'|i\rangle$ must have even parity. That dictates that the above matrix elements will be non-zero only when the eigenstates ($|i\rangle$ and $|f\rangle$) have opposite parity.
iii. The quantum numbers for the whole atom must charge as follows. $\Delta s = 0$, $\Delta l = 0, \pm 1$, $\Delta J = 0, \pm 1$ (with the restriction that $J = 0$ to $J = 0$ is prohibited), and finally $\Delta M_J = 0, \pm 1$ (again $M_J = 0$ to $M_J = 0$ is not allowed if $\Delta J = 0$)

[15] $r \to r, \theta \to \pi - \theta, \phi \to \pi + \phi$

1.18 jj COUPLING

In a situation (may be for large atoms) where there is a strong coupling between l_i and s_i is large, one has to apply the jj coupling scheme. Both l_i and s_i discontinue becoming good quantum numbers, while they combine to yield a total angular momentum quantum number, j_i to be a good quantum number. The individual j_i's are loosely coupled. Thus, there is no definite L or S, making room for J $(= L + S)$ to provide a valid description of the system.

In this case, the selection rules are stated as

i. Only one electron makes a transition at a given time.
ii. The value of l must change by $'1'$, that is $\Delta = \pm 1$. More fundamentally, the parity of the initial and final states should be opposite.
iii. $\Delta j = 0, \pm 1$
iv. For the entire atom, $\Delta J = 0, \pm 1$ (with $J = 0$ to $J = 0$ is forbidden), $\Delta M_J = 0, \pm 1$, (however, $M_J = 0$ to $M_J = 0$ is forbidden for $\Delta J = 0$).

1.19 HUND'S RULE

The Aufbau scheme discusses how the electrons fill the energy levels stating with the lowest orbitals. They move to the higher orbitals only when the lower orbitals are filled up. For example, 1*s* orbitals are filled up ahead of the 2*s* orbitals. However, the question remains, in what order do the different 2*p* orbitals get filled up. The answer to this involves Hund's rule.

Hund's rule can be stated as the following:

i. Every orbital is first filled up with a single electron, before any double occupancy occurs.
ii. All the electrons in the singly occupied orbitals have only one type of spin. This maximizes the total spin.

The essence of the first rule is, while assigning electrons to orbitals, an electron first attempts to fill the degenerate orbitals, that is, the ones with the same energy, before pairing with another electron in the half-filled orbital. Atoms, in their ground states, tend to have as many unpaired electrons as possible. In fact, all the electrons repel each other and try to get as far as possible, before they pair up. The electrons tend to minimize repulsion by occupying their own orbital, rather than sharing an orbital with another electron. The singly occupied orbitals are less efficiently screened from the nuclear charge.

According to the second rule, the unpaired electrons in the singly occupied orbitals have the same spin. Assume the electron which is placed first in an orbital has a spin-↑, the spins of all of the other electrons get fixed, which means the unpaired electrons are all spin-↑.

As an example, consider the carbon (C) atom which has 6 electrons and has an electronic configuration $1s^2\ 2s^2\ 2p^2$. The two $2s$ electrons will occupy the same orbital, while the two $2p$ electrons will be placed in different orbital and will be aligned in the same direction (say, all spin-↑) according to Hund's rule.

Similarly, the next atom, namely, Nitrogen ($z = 7$) has the configuration $1s^2\ 2s^2\ 2p^3$. The $1s$ and the $2s$ orbitals get completely filled up, which leaves three (unpaired) electrons. Again, according to the Hund's rule, the remaining electrons will fill all the empty orbitals.

Finally, consider Oxygen, which has $z = 8$ and the corresponding configuration is written as $1s^2\ 2s^2\ 2p^4$. The $1s$ and $2s$ orbitals get paired up electrons, which leaves 4 electrons to be accommodated in the $2p$ orbitals. According to Hund's rule, again all the orbitals will be singly occupied before at least one (which is the case here) of the $2p$ orbitals gets doubly occupied. Thus, $2p$ will have one paired, and the other two unpaired levels comprise the same spin.

REFERENCES

Ashcroft, N. and Mermin, N. D., *Solid State Physics* (Saunders College Publishing, 1976).

Bethe, H., "Zur Theorie der Metalle. Eigenwerte und Eigenfuntionen der linearen Atomkette," Z. Phys. **71**, 205 (1931).

Chandler, D., *Introduction to Modern Statistical Mechanics* (Oxford University Press, 1987).

Cohen-Tonnoudji, C., Diu, B., and Laloë, F., *Quantum Mechanics* (Wiley-VCH, 1977), Vol. II.

Doniach, S. and Sondheimer, E., *Green's Functions for Solid State Physicists* (Addison Wesley, 1974).

Fetter, A. L. and Walecka, J. D., *Quantum Theory of Many Particle Systems* (McGraw Hill, 1971).

Goldstone, J., Nuovo Cimento **19**, 154 (1961).

Holstein, T. and Primakoff, H., Phys. Rev. **58**, 1098 (1940).

Hubbard, J., Proc. R. Soc. London Ser. A, Math. Phys. Sci. **276**, 238 (1963).

Jordan, P. and Wigner, E., Z. Phys. **47**, 631 (1928).

Kübler, J., *The Theory of Itinerant Electron Magnetism* (Oxford University Press, 2000).

Mahan, G. D., *Many Particle Physics*, 3rd ed. (Springer, 2000).

Pathria, R. K. and Beale, P. D., *Statistical Mechanics*, 3rd ed. (Elsevier, 2011).

Peierls, R. E., *Quantum Theory of solids* (Clarendon Press, 1955), pp. 144–149.

Ruderman, M. A. and Kittel, C., Phys. Rev. **96**, 99 (1954); T. Kasuya, Prog. Theor. Phys. **16**, 45 (1956); K. Yoshda, Phys. Rev. **106**, 893 (1957).

Singh, A. and Tesanović, Z., Phys. Rev. B **41**, 11457 (1990).

Sorella, S., Parola, A., Parinello, M., and Tosatti, E., Int. J. Mod. Phys. B **3**, 1875 (1989).

Stauffer, D., *Introduction to Percolation Theory*, 2nd ed. (Taylor & Francis, 1992).

Zhong, F. H., Lin, C., and Ceperley, D. M., Phys. Rev. B **66**, 036703 (2002).

Ziman, J. M., *Principals of the Theory of Solids*, 2nd ed. (Cambridge University Press, 1972).

CHAPTER

2 QUANTUM HALL EFFECT

2.1 INTRODUCTION

The date of discovery of the quantum Hall effect (QHE) is known pretty accurately. It occurred in the night between 4th and 5th February, 1980 at 2:00 AM in the morning at the high magnetic lab in Grenoble, France (see Fig. 2.1). There is ongoing research on the transport properties of silicon field-effect transistors (FET). The main motive was to improve the mobility of these FET devices. They were provided by Dorda and Pepper, which allowed direct measurement of the resistivity tensor. The system is a highly degenerate two-dimensional electron gas contained in the inversion layer of a metal-oxide semiconductor field-effect transistor (MOSFET) operated at low temperatures and strong magnetic fields. The original notes appear in Fig. 2.1, where it is clearly stated that the Hall resistivity involves universal constants and hence signals toward the involvement of a very fundamental phenomenon.

In the classical version of the phenomenon discovered by E. Hall in 1879, just over a hundred years before the discovery of its quantum analog, one may consider a sample with a planar geometry so as to restrict the carriers to move in a two-dimensional (2D) plane. Next, turn on a bias voltage so that current flows in one of the longitudinal directions and a strong magnetic field perpendicular to the plane of the gas (see Fig. 2.2). Because of the Lorentz force, the carriers drift toward a direction transverse to the direction of the current flowing in the sample. At equilibrium, a voltage that develops in the transverse direction, which is known as the Hall voltage. The Hall resistivity, R defined as the Hall voltage divided by the longitudinal current, is found to depend linearly on the magnetic field, B and inversely on the carrier density, n through $R = \frac{B}{nq}$ (q is the charge). A related and possibly more familiar quantity is the Hall coefficient, denoted by $R_H = R/B$ which, via its sign, yields information on the type of the majority carriers, that is, whether they are electrons or holes.

At very low temperature, or at very high values of the magnetic field (or at both), the resistivity of the sample assumes quantized values of the form, $\rho_{xy} = \frac{h}{ne^2}$. Initially n was found to be an integer with extraordinary precession (one part in $\sim 10^8$). This is shown in Fig. 2.3. The quantization of the Hall resistivity yields the name "*quantum*" (or quantized) Hall effect, which we refer to as QHE throughout this chapter.

Klaus von Klitzing and his co-workers Klitzing *et al.* (1980); and Landwehr (1986) while measuring the electrical transport properties of planar systems formed at the interface of two different semiconducting samples in the strong magnetic field facility at Grenoble, France, noted that the Hall resistivity

https://doi.org/10.1063/9780735422537_002

Modern Perspectives in the Study of Electronic Systems, Saurabh Basu and Sourav Chattopadhyay

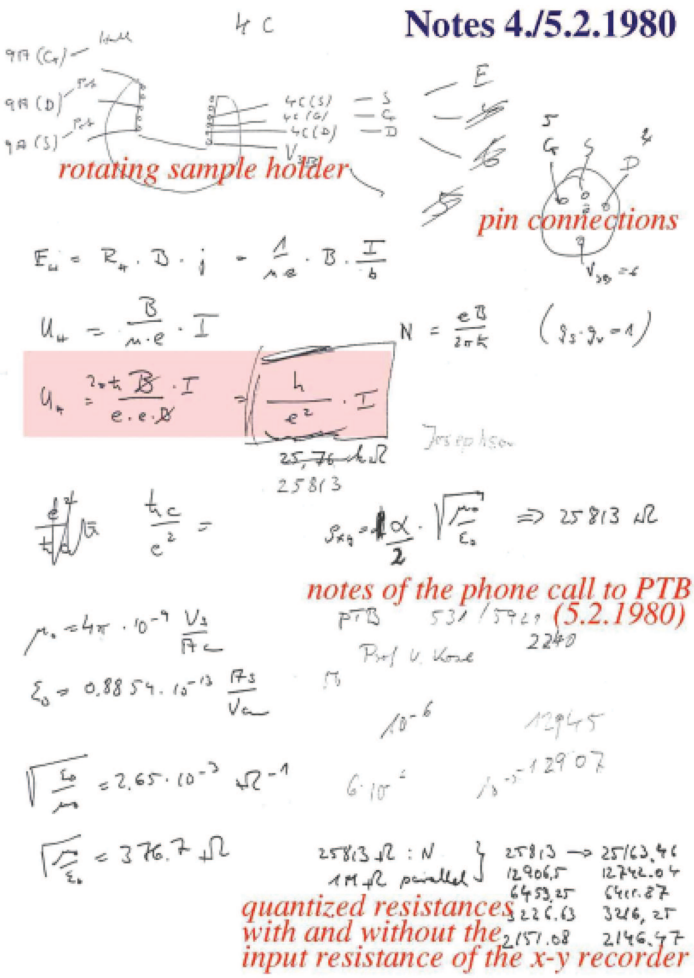


FIG. 2.1
Copy of the original notes by Professor K.v. Klitzing on the discovery of the quantum Hall effect. It documents that the Hall resistance ($\frac{U_H}{I}$) involves a fundamental constant h/e^2.

is quantized in units of h/e^2 as a function of the external magnetic field. The flatness of the plateaus occurring at integer or fractional values of h/e^2 has an unprecedented precession and is independent of the geometry of the sample (as long as it is two dimensional), the density of the charge carriers and its purity. The accuracy of the quantization aids in fixing the unit of resistance, namely, $h/e^2 = 25.813$ KΩ, also known as the Klitzing constant. Thus, among other significant properties of QHE that we shall be discussing in due course of time, an experiment performed at a macroscopic scale can be used for

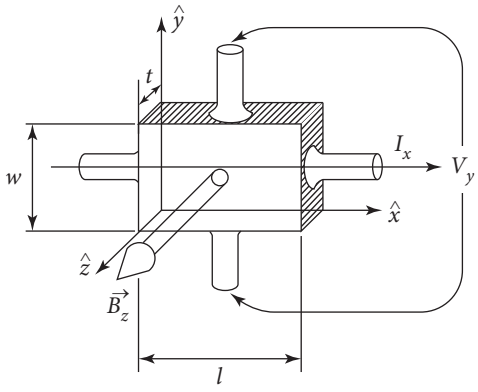


FIG. 2.2
Typical Hall experiment set up showing direction of the current, I_x and the magnetic field B_z. V_y denotes the Hall voltage.

metrology or yields the values for the fundamental constants used in quantum physics is truly amazing and hence calls for an intense scrutiny. The effect occurs when the density of the carriers, n are such that they are encoded in the integers (that come as proportionality constants to the Hall resistivity in terms of h/e^2) as if the charges locked their separation at some particular values. The phenomenon remains resilient to changing the carrier density by a small amount, however changing it by a large amount does destroy the effect.

The Hall resistivity (in red) in Fig. 2.3 becomes constant for certain ranges of the external magnetic fields, which are called plateaus. Further the longitudinal resistivity (in green) in the same plot vanishes everywhere. Although it shows peaks wherever there is a jump in the Hall resistivity from one plateau to another. Later on it was found that ν is not only restricted to integer values, but also takes values which are rational fractions, such as $\nu = \frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{3}{7}, \frac{4}{9}, \frac{5}{9}$ etc. There

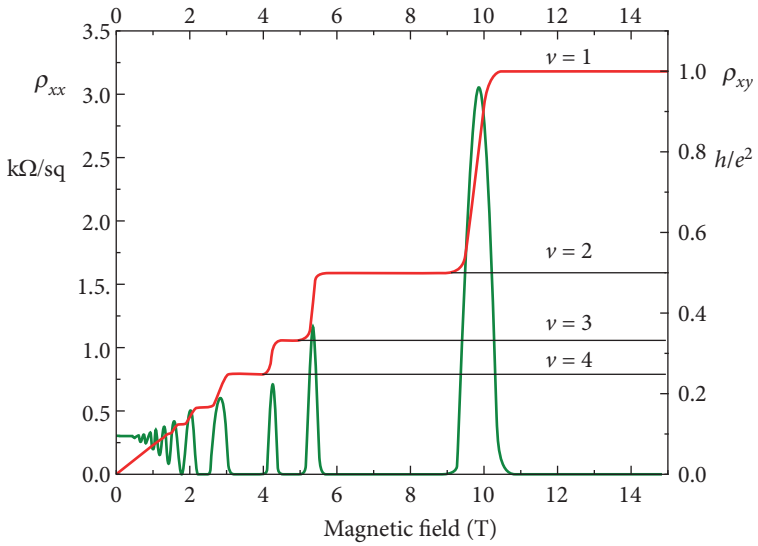


FIG. 2.3
Schematic plot of integer quantum Hall effect (IQHE) as a function of the applied magnetic field. The plot in red denotes the Hall resistivity and the one in green shows the longitudinal resistivity (or the magnetoresistivity).

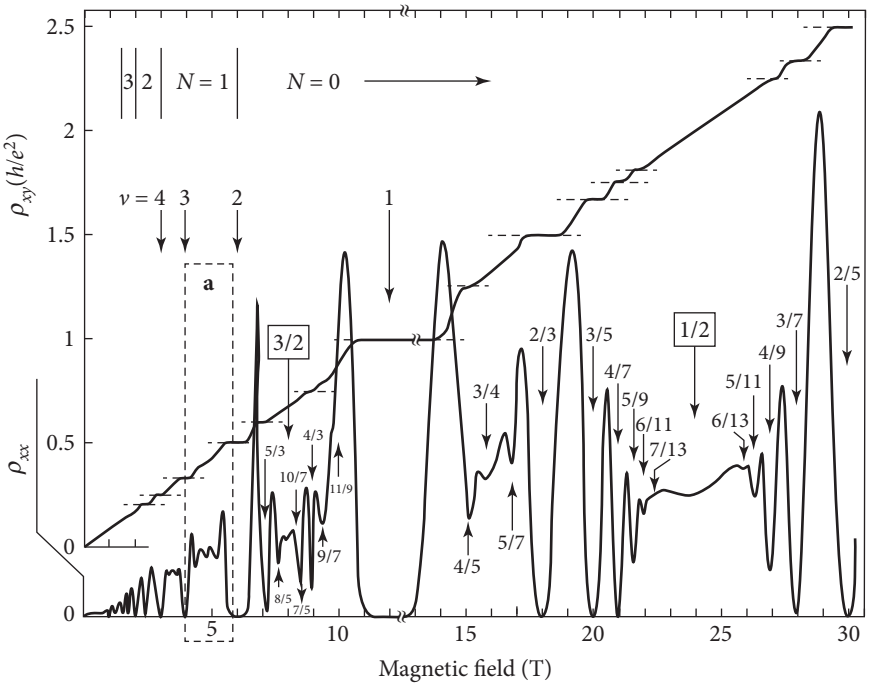


FIG. 2.4
The plot shows fractional quantum Hall effect (FQHE). The plateaus are shown at fractional values in units of h/e^2. Taken from Tsui *et al.* (1982).

are about 100 fractions (including the improper ones) that have been noted in experiments so far. The corresponding plot appears in Fig. 2.4.

2.1.1 General Perspectives

The charge carriers being confined in 2D wells have a longer history. Since 1966, it is known that the electrons accumulated at the surface of a silicon single crystal induced by a positive gate voltage form a 2D electron gas (2DEG). The energy of the electrons corresponding to a motion perpendicular to the surface is quantized (box quantization) and on top of it, the free motion of the electrons in 2D becomes quantized when a strong magnetic field is applied perpendicular to the plane (Landau quantization). Thus, QHE has both the quantization phenomena built into it.

An important recent development in the study of semiconductors is the achievement of structures in which the electrons are restricted to move essentially in 2D. This immediately says that the carriers

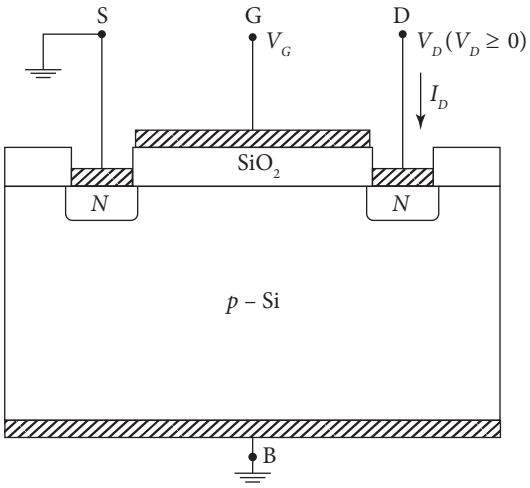


FIG. 2.5
MOSFET structure showing base (B), gate (G), source (S) and drain (D). The substrate is a *p*-type Si. SiO_2 denotes the insulating oxide layer.

are prohibited from moving along the direction transverse to the plane. Hence, the motion is quantized. Such 2D behavior of the carriers can be found in metal-oxide-semiconductor (MOS) structures, quantum wells and superlattices. An excellent prototype is the metal-insulator-semiconductor (MIS) layered structure of which the insulator is usually an oxide, such as Al_2O_3 (thereby making it a MOS structure). In Fig. 2.5 we show a typical MOS device where the substrate is a doped *p*-type silicon (Si) which is grounded and is called a base (shown by *B* in Fig. 2.5). On the top, there is a metallic layer (shown by the hatched regime) followed by an insulating layer formed by SiO_2. The metallic layer is called the gate, denoted by *G*, which is biased by a voltage V_G. The source (grounded) and the drains are denoted by *S* and *D* respectively, which in Fig. 2.5 consist of *n*-type materials. The gate voltage causes the carriers beneath the gate electrode to drift between the source and the drain. The layer of charge carriers below the oxide layer forms the 2DEG, which is central to our discussion. The energy dispersion in this case reads

$$E_n(k_x, k_y) = \frac{\hbar^2 k_x^2}{2m_{xx}^*} + \frac{\hbar^2 k_y^2}{2m_{yy}^*}, \tag{2.1}$$

where m_{xx}^* and m_{yy}^* are the components of the effective mass tensor defined by the inverse of the curvature of the band structure, namely,

$$m_{\alpha\beta}^* = \hbar^2 \left(\frac{\partial^2 E(k)}{\partial k_\alpha \partial k_\beta} \right)^{-1}. \tag{2.2}$$

The physical properties of all systems are governed by their density of states (DOS) which plays a crucial role in deciding the dependence of the temperature, density of carriers, etc. In 2D systems with a parabolic dispersion, as elaborated above in Eq. (2.1), the DOS is a constant and assumes the form,

$$g(E) = g_{2D}(E) = \frac{m^*}{\pi \hbar^2}. \tag{2.3}$$

The energy independent DOS is very special to 2D and is in sharp contrast to three dimensions (3D) where it goes as $E^{1/2}$, and in one dimension (1D) where it goes as $E^{-1/2}$. In a general sense, and not restricted to the discussion on the Hall effect, the DOS enters while calculating the average quantities, such as the average energy or the average number of particles. For example, the average of a physical

observable, O, of a fermionic system, is computed using

$$\langle O \rangle = \int_0^{\mu} O f(E) g(E) dE, \tag{2.4}$$

where $f(E)$ is the Fermi distribution function given by

$$f(E) = \frac{1}{e^{\beta(E-\mu)} + 1}$$

with $\beta = \frac{1}{k_B T}$ and μ denotes the chemical potential. In general, this integral is quite challenging to compute analytically because of the Fermi distribution function ($f(E)$) present in the integrand.

Meanwhile, there is a wonderful simplification where $f(E)$ assumes a value unity at all temperatures for which the experiments are performed. Only at temperatures close to the Fermi temperature, T_F defined via, $\epsilon_F = k_B T_F$ (ϵ_F being the Fermi energy), $f(E)$ starts deviating from unity and its exact form needs to be incorporated in the integral. However, T_F is usually of the order of tens of thousands of Kelvin for typical metals (such as Cu, Al, etc.) which is too high for them to appear in experimental situations. Moreover, the DOS only depends on energy and is independent of the temperature to a very good approximation. Thus, computation of Eq. (2.4) becomes trivial as the integrand becomes independent of temperature or weakly dependent on temperature (Kittel, 2004; and Ashcroft and Mermin, 1976).

In the following, we mention that there is something interesting about the transport properties of two-dimensional systems. In the linear response regime, Ohm's law is valid and says that

$$V_\alpha = R_{\alpha\beta} I_\beta$$

where $R_{\alpha\beta}$ denotes the resistivity tensor, and α, β denote spatial variables, x, y etc. One can equivalently invert this equation to write, $I_\alpha = G_{\alpha\beta} V_\beta$, where $G_{\alpha\beta}$ represents the conductivity tensor with $G = R^{-1}$. Equivalent relations in terms of the components of the electric field (E) and current density (j) read as

$$E_\alpha = \rho_{\alpha\beta} j_\beta \quad \text{and} \quad j_\alpha = \sigma_{\alpha\beta} E_\beta \tag{2.5}$$

where ρ and σ denote the resistivity and the conductivity tensors, respectively.

An interesting (and useful too) artifact of 2D physics is an accidental similarity that exists in decoding some of the key features of the transport properties. For example, the resistivity, ρ (or the conductivity, σ) is a quantity that is independent of the system geometry, and hence is useful for a theoretical analysis. Whereas in experiments, one measures the resistance of a sample, R (or the conductance, G). For a sample in the shape of a hypercube of sides, L, the resistance and the resistivity are related by

$$R = \rho L^{2-d}, \tag{2.6}$$

where d denotes the dimensionality. Only for $d = 2$ the resistance is a scale invariant quantity. This puts the experimentalists and the theorists on the same page, as the geometry of the sample does not explicitly enter the analysis of its transport properties.

As we dig more into the details of the transport properties of 2DEG in the presence of a magnetic field, further useful information emerges. The off-diagonal elements of both the conductivity and the

resistivity tensors are antisymmetric with regard to the direction of the applied field, **B**. Consider a planar sample with the dimensions $L_x \times L_y$. The conductivity tensor is of the form,

$$\sigma = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} \\ \sigma_{yx} & \sigma_{yy} \end{pmatrix}. \tag{2.7}$$

Let us try to understand the nature of the tensor, σ in the presence of a magnetic field. A conductor in an external magnetic field obeys,

$$j_\alpha = \sigma_{\alpha\beta} E_\beta. \tag{2.8}$$

Onsager's reciprocity principle does not hold in the presence of a magnetic field, **B**, which implies (Onsager, 1931).

$$\sigma_{\alpha\beta}(\mathbf{B}) \neq \sigma_{\beta\alpha}(\mathbf{B}). \tag{2.9}$$

Instead, one has $\sigma_{\alpha\beta}(\mathbf{B}) = \sigma_{\beta\alpha}(-\mathbf{B})$ to make sure that the time reversal holds only if **B** changes sign. Let us write the conductivity tensor as the sum of a symmetric and antisymmetric tensors (note that this is always possible for a rank 2 tensor). Thus,

$$\sigma_{\alpha\beta} = S_{\alpha\beta} + A_{\alpha\beta} \tag{2.10}$$

where **S** and **A** are symmetric and the antisymmetric tensors which obey the following relations,

$$\begin{aligned} S_{\alpha\beta}(\mathbf{B}) &= S_{\beta\alpha}(-\mathbf{B}) = S_{\alpha\beta}(-\mathbf{B}) \\ A_{\alpha\beta}(\mathbf{B}) &= A_{\beta\alpha}(-\mathbf{B}) = -A_{\alpha\beta}(-\mathbf{B}) \end{aligned} \tag{2.11}$$

such that the components of $S_{\alpha\beta}$ are even functions of **B**, while those of $A_{\alpha\beta}$ are odd functions of **B**. Putting in Eq. (2.8),

$$j_\alpha = S_{\alpha\beta} E_\beta + A_{\alpha\beta} E_\beta. \tag{2.12}$$

But owing to the antisymmetry,

$$A_{\alpha\beta} = \epsilon_{\gamma\alpha\beta} A_\gamma = -\epsilon_{\alpha\gamma\beta} A_\gamma. \tag{2.13}$$

Putting it in Eq. (2.12),

$$\begin{aligned} j_\alpha &= S_{\alpha\beta} E_\beta - \epsilon_{\alpha\beta\gamma} A_\beta E_\gamma \\ &= S_{\alpha\beta} E_\beta - (\mathbf{A} \times \mathbf{E})_\alpha \\ &= S_{\alpha\beta} E_\beta + (\mathbf{E} \times \mathbf{A})_\alpha. \end{aligned} \tag{2.14}$$

Assuming that we can expand $\sigma(\mathbf{B})$ in powers of **B**,[1] such that the antisymmetric part contains odd powers of **B**, then we can write,

$$A_\alpha = \eta_{\alpha\beta} B_\beta \tag{2.15}$$

and $S_{\alpha\beta}(\mathbf{B})$ consists of even powers of **B**,

$$S_\alpha = (\sigma_0)_{\alpha\beta} + \zeta_{\alpha\beta\gamma\delta} B_\gamma B_\delta. \tag{2.16}$$

[1] It should be valid for weak magnetic fields, and is not exactly true for quantum Hall effect, but nevertheless it serves our purpose.

The first term is the zero field conductivity tensor. Thus, putting things together up to terms linear in **B**,

$$j_\alpha = S_{\alpha\beta}E_\beta + (\mathbf{E}\times\mathbf{A})_\alpha. \tag{2.17}$$

The second term denotes the Hall effect, which is linear in **B**. This implies that the Hall current is perpendicular to the electric field, **E** and is proportional to **E** and **B**. Thus, an antisymmetric tensor is relevant to the study of the Hall effect, which is why the conductivity and the resistivity tensors are antisymmetric. This is an important result which deviates from the corresponding scenario that arises in the absence of an external magnetic field.

2.1.2 Translationally Invariant system: Classical limit of QHE

It is quite an irony that the extreme universal signature of the transport properties of a 2DEG characterized by the flatness of the plateaus, that not only occurs, but survives even in the presence of disorder, impurity and imperfection. In the absence of the magnetic field, Anderson localization would have governed the transport signatures of non-interacting electrons which says that in any dimension less than three, all eigenstates of a system are exponentially localized even for an infinitesimal disorder strength. Only in three dimensions, there is a critical disorder at which a metal-insulation transition occurs. However, the scenario is strongly altered by the presence of the magnetic field, which yields, as we shall shortly see, a series of unique phase transitions from a perfect conductor to a perfect insulator. No other system demonstrates re-occurrence of the same phases over and over again as the magnetic field is gradually ramped up.

To begin with, we shall consider the case which is free from disorder, or equivalently, a translationally (Lorentz) invariant system that possesses no preferred frame of reference. Thus, we can think of a reference frame that is moving with a velocity $-\mathbf{v}$ with respect to the lab frame, where the current density is given by $\mathbf{j} = -ne\mathbf{v}$ (n: areal electron density, e: electronic charge). In this frame, the electric and the magnetic fields are given by[2]

$$\mathbf{E} = -\mathbf{v}\times\mathbf{B}; \quad \text{and} \quad \mathbf{B} = B\hat{z}. \tag{2.18}$$

The above transformation ensures that an electric field must exist to balance the Lorentz force $-e\mathbf{v}\times\mathbf{B}$ in order to conduct without deflection. For the electric field, this yields,

$$\mathbf{E} = \frac{1}{ne}\mathbf{J}\times\mathbf{B}. \tag{2.19}$$

This is equivalent to the tensor equation,

$$E^\mu = \rho_{\mu\nu}j^\nu \tag{2.20}$$

[2] Remember that $\mathbf{E} = 0$ in the lab frame, though **B** remains unchanged.

with the resistivity tensor given by

$$\rho_{\mu\nu} = \frac{B}{ne}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{2.21}$$

By inverting the tensor equation, one can obtain,

$$j^{\mu} = \sigma_{\mu\nu}E^{\nu} \tag{2.22}$$

where the conductivity tensor, $\sigma_{\mu\nu}$ is defined via,

$$\sigma_{\mu\nu} = \frac{ne}{B}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{2.23}$$

There is an interesting paradox which states that, $\sigma_{xx} = \rho_{xx} = 0$ (see above) which is of course contradictory. However, we reserve this rather interesting topic for a discussion immediately afterwards. Here we wish to point out that we get the results for a classical Hall effect, that is, $\sigma_{xy} = \frac{ne}{B}$ (or $\rho_{xy} = \frac{B}{ne}$). It is important to realize that the result is an artefact of Lorentz invariance, where the characteristics of the sample or the 2DEG enter only through the carrier density, n for a translationally invariant system. Thus, in the absence of defect, disorder and impurity, the Hall effect is concerned with the carrier density of the sample and nothing else. Hall resistivity depends linearly on the magnetic field. The quantum Hall effect, which is much more versatile (than merely depending on the density) where disorder, that jeopardizes the translational invariance, plays an indispensable part in causing plateaus in the conductivity (or the resistivity) to occur and survive. We show a schematic plot in Fig. 2.6 to emphasize the difference between the classical and the quantum Hall effects. Both the Hall and the longitudinal resistivities are significantly different in these two cases.

Writing the equation of motion (EOM) for a charge particle of mass m, moving with a velocity, $\mathbf{v}$ $(=\frac{\mathbf{p}}{m})$ in the presence of a longitudinal electric field, $\mathbf{E}$ and a perpendicular magnetic field, $\mathbf{B}$,

$$\frac{d\mathbf{p}}{dt} = -e\mathbf{E} - e\frac{\mathbf{p}}{m} \times \mathbf{B} - \frac{\mathbf{p}}{\tau}. \tag{2.24}$$

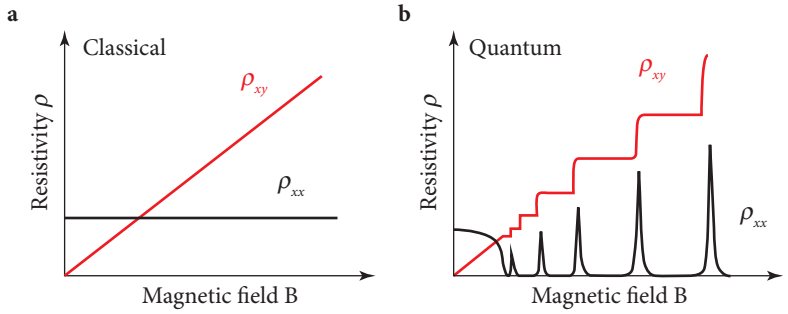


FIG. 2.6
Schematic plot showing Hall and magnetoresistivities for both classical (a) and quantum (b) Hall effects.

The last term is the resistive force arising from electron-impurity scattering with τ being the relaxation time. Using the current density, $\mathbf{j} = -ne\mathbf{p}/m$ and the cyclotron frequency, $\omega_B = eB/m$, at the steady state ($\frac{d\mathbf{p}}{dt} = 0$),

$$\frac{ne^2\tau}{m}\mathbf{E} + \omega_B\tau\mathbf{j} \times \hat{\mathbf{z}} + \mathbf{j} = 0. \tag{2.25}$$

Assuming the motion of carriers along the x-direction, that is, $\mathbf{j} = j\hat{\mathbf{x}}$, casting it in the form, $\mathbf{j} = \sigma\mathbf{E}$, σ assumes the form,

$$\sigma = \frac{ne^2\tau/m}{1+\omega_B\tau^2}\begin{pmatrix} 1 & -\omega_B\tau \\ \omega_c\tau & 1 \end{pmatrix}. \tag{2.26}$$

This yields the Drude conductivity, which can be written as

$$\sigma_{xx} = \frac{\sigma_0}{1+\omega_B\tau^2} \tag{2.27}$$

where $\sigma_0 = ne^2\tau/m$. In the absence of any scattering by the impurities, the relaxation time, τ is infinitely large, which yields $\sigma_{xx} \to 0$. This induces Ohm's law to assume the form (now writing in terms of the resistivity),

$$\mathbf{E} = \begin{pmatrix} E_x \\ E_y \end{pmatrix} = \begin{pmatrix} 0 & \rho_{xy} \\ -\rho_{xy} & 0 \end{pmatrix}\begin{pmatrix} j_x \\ j_y \end{pmatrix} = \begin{pmatrix} \rho_{xy}j_y \\ -\rho_{xy}j_x \end{pmatrix}. \tag{2.28}$$

Hence, the electric field $\mathbf{E}$ is perpendicular to the current density, $\mathbf{j}$, which says that $\mathbf{j}\cdot\mathbf{E} = 0$. The physical significance of $\mathbf{j}\cdot\mathbf{E}$ is the work done that accelerates the charges, which being zero in this case implies that a steady current flow in the sample without requiring any work, and hence causes no dissipation. Thus, $\sigma_{xx} = 0$ implies that no current flows in the longitudinal direction, which is actually a signature of a perfectly insulating state. Since the components of the resistivity and the conductivity tensors are related by

$$\sigma_{xx} = \frac{\rho_{xx}}{\rho_{xx}^2+\rho_{xy}^2}; \quad \sigma_{xy} = \frac{-\rho_{xy}}{\rho_{xx}^2+\rho_{xy}^2}. \tag{2.29}$$

Let us examine the possible scenarios:

i. If $\rho_{xy} = 0$, one gets $\sigma_{xx} = \frac{1}{\rho_{xx}}$ and $\sigma_{xy} = 0$ which is a familiar scenario.
ii. If $\rho_{xy} \neq 0$, σ_{xx} and σ_{xy} both exist.
iii. Now consider $\rho_{xx} = 0$, $\sigma_{xx} = 0$, if $\rho_{xy} \neq 0$. This is truly interesting, since if $\rho_{xx} = 0$, it implies a perfect conductor, and at the same time $\sigma_{xx} = 0$ implies a perfect insulator. This is surprising, but truly occurs in the presence of an external magnetic field. This is reflected in the plots presented in Figs. 2.3 and 2.4. We shall return for a more thorough discussion later.

2.1.3 Charge particles in a magnetic field: Landau levels

Let us examine the fate of the electrons confined in a 2D plane in the presence of a magnetic field. Consider non-interacting spinless electrons in an external field, $\mathbf{B}$ and write down the Schrödinger

equation to solve for the eigenvalues and eigenfunctions. The canonical momentum can now be written as $\mathbf{p} \rightarrow \mathbf{p} - q\mathbf{A} = \mathbf{p} + e\mathbf{A}$ (Goldstein *et al.*, 2002) where $q = -e$ is the electronic charge and $\mathbf{A}$ is the vector potential corresponding to the field, $\mathbf{B}$. $\mathbf{B}$ and $\mathbf{A}$ are related via, $\nabla \times \mathbf{A} = \mathbf{B}$. The time-independent Schrödinger equation becomes,

$$\frac{1}{2m}(\mathbf{p} + e\mathbf{A})^2\psi(\mathbf{r}) = E\psi(\mathbf{r}). \tag{2.30}$$

In order to solve this, we need to fix a gauge or a choice of the vector potential. Note that the choice of the gauge will not alter the solution of the equation, which in other words can be stated as the Schrödinger equation being gauge invariant. However, a particular choice is essential for us to go ahead.

Corresponding to a magnetic field, $\mathbf{B}$ in the z-direction, such that $\mathbf{B} = B\hat{z}$ (just as the case of a 2DEG subjected to a perpendicular magnetic field), the vector potential can be chosen as

$$A_x = -By, \quad A_y = A_z = 0. \tag{2.31}$$

This is known as the Landau gauge. It also allows us to assume $A_y = Bx$ and $A_x = A_z = 0$.[3] In the Landau gauge, the Schrödinger equation becomes,

$$\left[\frac{1}{2m}(p_x - eBy)^2 + \frac{p_y^2}{2m} + \frac{p_z^2}{2m}\right]\psi(\mathbf{r}) = \epsilon\psi(\mathbf{r}). \tag{2.32}$$

Clearly, in the z-direction, the particle behaves like a free particle with energy, $\epsilon_z = \frac{p_z^2}{2m}$ with the eigenfunction same as that of particle in a box in the z-direction. Thus, in the $x - y$ plane, the above equation becomes,

$$\left[\frac{1}{2m}(p_x - eBy)^2 + \frac{p_y^2}{2m}\right]g(x,y) = \epsilon g(x,y)$$
$$\mathcal{H}(x,y)g(x,y) = \epsilon g(x,y) \tag{2.33}$$

where for the 2D case, $\psi(\mathbf{r})$ becomes $g(x, y)$ and $E = \epsilon + \epsilon_z$. It is easy to see that p_x commutes with $\mathcal{H}(x, y)$, that is,

$$[\mathcal{H}(x,y), p_x] = 0. \tag{2.34}$$

Hence, p_x is a constant of motion. Thus, for a p_x given by $p_x = \frac{2\pi\hbar}{L_x}n_x$, one can write Eq. (2.33) as

$$\left[\frac{p_y^2}{2m} + \frac{1}{2}m\left(\frac{eB}{m}\right)^2(y - y_0)^2\right]f(y) = \epsilon f(y) \tag{2.35}$$

where $y_0 = \frac{p_x}{eB} = k_x l_B^2 = k l_B^2$ (say), and $f(y)$ is only a function of y. y_0 has the dimension of length, and hence l_B is denoted as the magnetic length, which is an important quantity in all subsequent discussions. Moreover, $f(y)$ is the eigenfunction corresponding to the one-dimensional Hamiltonian written above.

[3] A combination of the two yields a "symmetric" gauge which we shall introduce and employ later.

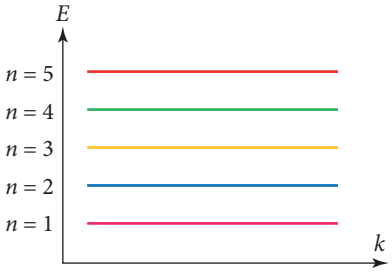


FIG. 2.7
Schematic plot showing the Landau levels. Each of these levels has a large degeneracy.

Interestingly, the left-hand side of Eq. (2.35) denotes the Hamiltonian for a simple harmonic oscillator (SHO) which oscillates in the y-direction about a mean position y_0 with a frequency, $\omega_c = \frac{eB}{m}$. ω_c is known as the cyclotron frequency. Taking results from the SHO problem in quantum mechanics, the energy eigenvalues can be found as

$$\epsilon_n = \left(n + \frac{1}{2}\right)\hbar\omega_c = \left(n + \frac{1}{2}\right)\hbar\frac{eB}{m}. \tag{2.36}$$

The energy levels in Eq. (2.36) are referred to as the Landau levels. The levels are schematically shown in Fig. 2.7. They are equidistant for a given value of the magnetic field, however their separation increases with the increase in B. Each of these levels has a large degeneracy (see below).

Further, the eigenfunction, $g(x, y)$ in Eq. (2.33) corresponding to an oscillatory motion (as an SHO) in the y-direction and a free motion (like a particle in the absence of any potential) in x-direction assumes the form,

$$g(x, y) = \frac{1}{L_x} e^{ik_x x} A_n e^{\frac{-eB(y-y_0)^2}{\hbar}} H_n\left(\frac{eB(y - y_0)}{\hbar}\right) \tag{2.37}$$

where $H_n(\xi)$ with $\xi = \frac{eB(y-y_0)}{\hbar}$ denote the Hermite polynomials that are familiar in the context of SHO and A_n denote the normalization constants. Thus, the trajectory of the particle is similar to that of a simple harmonic oscillator centered about a certain value of y (instead of the origin), namely, y_0 and freely propagates along the x-direction. y_0 is controlled by the strength of the magnetic field B, and is inversely proportional to it.

2.1.4 Degeneracy of the Landau levels

The Landau levels given by Eq. (2.36) are hugely degenerate. Since p_x is a constant of motion, the energy is independent of p_x. Thus, all possible values of the quantum numbers corresponding to the motion in the x-direction, namely, n_x that are defined by $k_x = \frac{2\pi}{L_x} n_x$, ($L_x$ denotes the length of the sample in the x-direction and $n_x = 0, 1, 2 \cdots$.) will make the levels degenerate. The degree of degeneracy is limited only by the length of the sample in the y-direction, namely, L_y. For the magnetic length, $y_0 = kl_B^2 = k(\frac{\hbar}{eB})$ for which the simple harmonic motion occurs should not exceed the length L_y of the sample, the maximum degeneracy of the Landau levels can be found by substituting $y_0 = L_y = (\frac{\hbar}{eBL_x})n_x$, and hence using the maximum possible value of n_x, that is,

$$(n_x)_{\max} = g = \frac{eBL_xL_y}{h} = \frac{eB\mathcal{A}}{h} \tag{2.38}$$

where $\mathcal{A} = L_xL_y$ is the area of the sample in the $x - y$ plane. This yields the degeneracy, g to be identified as the flux, ϕ ($=B\mathcal{A}$) threading the planar sample, via, $g = \Phi/\Phi_0$ where $\Phi_0 = h/e$.[4]

[4] The value of $\phi_0 \simeq 4.13 \times 10^{-15}$ Wb-m^2.

A few comments are in order:

i. The degeneracy, g, is independent of the effective mass of the carriers, and hence independent of the material.
ii. The degeneracy is proportional to the area of the sample and the value of the magnetic field. Thus, the degeneracy can be controlled by the applied magnetic field.

To remind ourselves, we have solved for the properties of a single electron confined in a plane in the presence of a perpendicular magnetic field. These energies are called the Landau levels, and these levels are highly degenerate. The scenario is a prototype of what happens in a Hall effect experiment where the external magnetic field is varied and the resistivities (both Hall and the longitudinal) are measured. The Hall resistivity shows plateaus in multiples of h/ne^2 whenever the filling fraction, ν of the Landau levels (defined below) is close to an integer n. Consider n_0 is the density of charge carriers of the sample, then ν is defined using, $\nu = \frac{n_0}{g/\mathcal{A}} = \frac{n_0 h}{eB}$. ν denotes the filling fraction, which is also defined as

$$\nu = \frac{\text{number of electrons}}{\text{flux quantum}}$$

As and when ν assumes a value near an integer, n (or a rational fraction as in the case of fractional quantum Hall effect) as the magnetic field is tuned, one observes a plateau in the Hall resistivity, ρ_{xy}. Something else happens at the same time that is equally interesting. The longitudinal resistivity, ρ_{xx} drops to zero whenever ρ_{xy} acquires a plateau. The vanishing of ρ_{xx} makes the system dissipationless. However, the diagonal conductivity (σ_{xx}) also vanishes, which makes the system insulating. This paradox is explained in detail elsewhere.

2.1.5 Conductivity of the Landau levels: Role of the edge modes

Here, we discuss some of the key properties of the Landau levels. Let us now calculate the current carried by the Landau levels. The expression for the current can be found using $\langle \mathbf{J} \rangle = e\langle \mathbf{v} \rangle$ where the expectation value must be computed within the Landau states.

$$\langle \mathbf{J} \rangle = -e\langle \psi_k | \mathbf{v} | \psi_k \rangle = -\frac{e}{m}\langle \psi_k | \mathbf{p} + e\mathbf{A} | \psi_k \rangle. \tag{2.39}$$

The longitudinal current in the x-direction carried by the Landau levels is obtained via,

$$\langle J_x \rangle = -\frac{e}{ml_B\sqrt{\pi}} \int dy\, e^{-\frac{1}{l_B^2}(y-y_0)^2} (\hbar k - eBy) = 0. \tag{2.40}$$

The integrand has an even function (the first one) and an odd function (the second one). Thus, the integral vanishes. We can get the average velocity, $\langle v \rangle = \frac{1}{\hbar}\frac{\partial \epsilon_k}{\partial k} = 0$ as ϵ does not depend upon k. Thus, the Landau wavefunctions carry no current by themselves. They only carry current in the presence of an electric field in the x-direction as shown below.

2.1.6 Spin and the electric field

So far we have been talking about spinless fermions. It is in general a worthwhile exercise to include the spin of the electrons and explore if there is any significant development to the quantization phenomena

discussed above. The spin degrees of freedom placed in an external magnetic field introduce a Zeeman energy scale owing to a coupling between the spin and the magnetic field. The Zeeman term is written as $\Delta_Z = g\mu_B B$, where $\mu_B = e\hbar/2m$ is the Bohr magneton and $g = 2$. Now the splitting between the Landau levels originating from the orbital effect ($\mathbf{p} \rightarrow \mathbf{p} - e\mathbf{A}$) is $\Delta = \hbar\omega_B = \frac{e\hbar B}{m} = \Delta_B(say)$. Δ_B is the so-called cyclotron energy. But for electrons, this precisely coincides with the Zeeman splitting $\Delta = g\mu_B B$ between the $\uparrow$ and the $\downarrow$-spins. Thus, it looks that the spin-$\uparrow$ particles in a Landau level with index, n, have exactly the same energy as the spin-$\downarrow$ particles in the next higher Landau level index, that is, $n + 1$. However, in real materials this does not occur. For example, in GaAs, the Zeeman energy is typically about 70 times smaller than the cyclotron energy.

Now consider an external electric field, E, applied in the x-direction. This creates an electric potential of the form, $\phi = -Ex$. Thus, the resulting Hamiltonian, including external electric and magnetic fields, can be written as

$$\mathcal{H} = \frac{1}{2m}\left[p_x^2 + (p_y + eBx)^2\right] + eEx. \tag{2.41}$$

It is important to note that we have considered a different choice of the gauge here, namely, $A_y = Bx$. Instead of a rigorous derivation, we can complete square in the expression for energy of the particles,

$$E_{n,k} = \hbar\omega_B\left(n + \frac{1}{2}\right) - eE\left(kl_B^2 + \frac{eE}{m\omega_B^2}\right) + \frac{1}{2}m\frac{E^2}{B^2}. \tag{2.42}$$

The resultant spectrum is plotted in Fig. 2.8. The levels are dispersive because of the k-dependent middle term in Eq. (2.42). This is interesting because the degeneracy of the Landau level has now been lifted. The energy in each level now depends linearly on k as shown, which was earlier independent of k. The eigensolution is simply that of a harmonic oscillator shifted from the origin and displaced along the x-axis by an amount mE/eB^2 and is written as

$$\psi(x, y) = \psi_{n,k}\left(x + \frac{mE}{eB^2}, y\right). \tag{2.43}$$

Among the other properties, the group velocity is given by

$$v_y = \frac{1}{\hbar}\frac{\partial E_{n,k}}{\partial k} = \frac{e}{\hbar}El_B^2. \tag{2.44}$$

Putting $l_B = \sqrt{\frac{\hbar}{eB}}$ (as said earlier, l_B is an important length scale of the problem, which we shall see throughout the discussion),

$$v_y = \left(\frac{eE}{\hbar}\right)\cdot\left(\frac{\hbar}{eB}\right) = \frac{E}{B}. \tag{2.45}$$

Thus, the energy has three terms in the Eq. (2.42),

i. the first one is that of a harmonic oscillator,
ii. the second one is the potential energy, of a wave packet localized at $x = (-kl_B^2 - \frac{mE}{e\omega_B^2})$,
iii. and finally, the last one denotes the kinetic energy of the particle, namely, $\frac{1}{2}mv_y^2$.

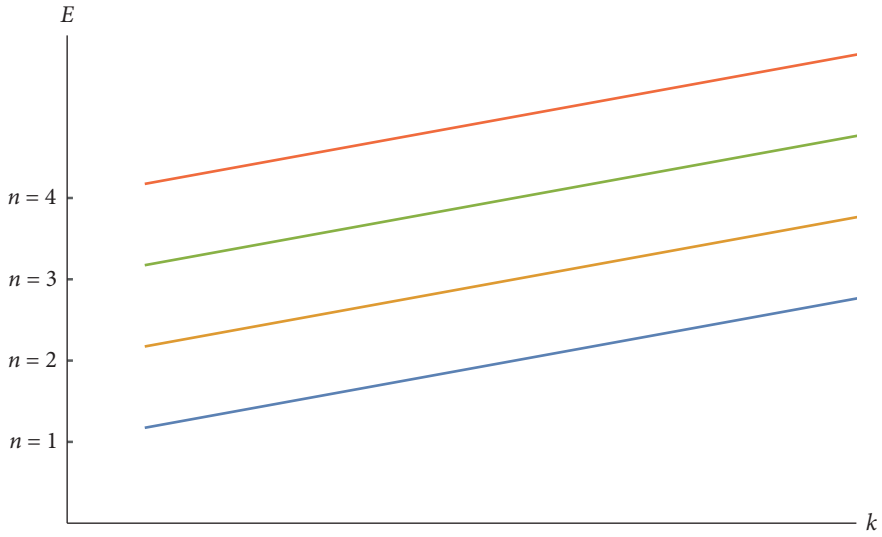


FIG. 2.8
Schematic plot showing the Landau levels ($n = 1, 2, 3, 4$) in the presence of an electric field. The levels are tilted because of the electric field.

2.1.7 Laughlin's argument: Corbino ring

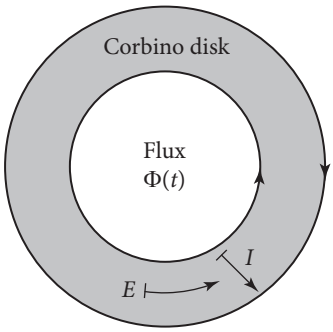


FIG. 2.9
Schematic plot showing Corbino ring.

Laughlin intuitively considered the phenomenon of QHE as a quantum pump. Consider a ring (see Fig. 2.9) where the vacant region admits a magnetic field and hence a flux ϕ. For this argument to be valid, the geometry of the ring is important. Here, in addition to the background magnetic field **B** that threads the sample, we can thread an additional flux Φ through the center of the ring. This Φ can affect the quantum state of the electrons. In addition, the temperature is low such that the thermal effects can be neglected.

Let us first see what this flux Φ has got to do with the Hall conductivity. Suppose we slowly increase Φ from 0 to $\Phi_0 (= \frac{h}{e})$, that is within a time $t_0 \gg \frac{1}{\omega_B}$. This induces an emf around the annular region $\varepsilon = \frac{\partial \Phi}{\partial t} = \frac{-\Phi_0}{t_0}$. The purpose of this emf is to transport "n" electrons from the inner circumference to the outer circumference. This would result in a current in the radial direction, $I_r = -ne/t_0$. Thus, the Hall resistivity is

$$\rho_{xy} = \frac{\varepsilon}{I_r} = -\frac{\Phi_0}{t_0} \cdot \frac{t_0}{(-ne)} = \frac{h}{e^2} \cdot \frac{1}{n}. \tag{2.46}$$

The same arguments hold equally for the IQHE and FQHE, in the former n is an integer, while n is a fraction for the latter. In FQHE, the interpretation is as follows: as we increase the flux from Φ to Φ_0, a charge of magnitude e/m is transported from the inner circumference to the outer one when the flux is increased by Φ_0 units. The resultant Hall conductivity (or equivalently the resistivity) becomes,

$$\sigma_{xy} = \frac{e^2}{h} \cdot \frac{1}{m}. \tag{2.47}$$

Thus, a whole electron is transferred only when the flux is increased by $m\Phi_0$ units.

2.1.8 Edge modes and the conductivity of the single Landau level

When a particle is restricted to move only in one direction, the motion is said to be chiral where backscattering is prohibited. Thus, the particles propagate in one direction at one edge of the sample and move in the other direction at the other end of the sample. Let us understand how the edge modes appear.

An edge can be modeled by a potential $V(x)$ in the y-direction which rises steeply, as shown in Fig. 2.10. Let us continue working on the Landau gauge, such that the Hamiltonian is given by

$$\mathcal{H} = \frac{1}{2m}[p_x^2 + (p_y + eBx)^2] + V(x). \tag{2.48}$$

In the absence of the potential $V(x)$, the (lowest) wavefunction is a Gaussian of width $l_B \left(= \sqrt{\frac{\hbar}{eB}}\right)$. If we assume that $V(x)$ is smooth over a distance l_B, and hence assume that the center of each Gaussian to be localized at $x = x_0$, we can Taylor expand the potential $V(x)$ about x_0 in the following fashion,

$$V(x) = V(x_0) + \frac{\partial V}{\partial x}(x - x_0) + \cdots. \tag{2.49}$$

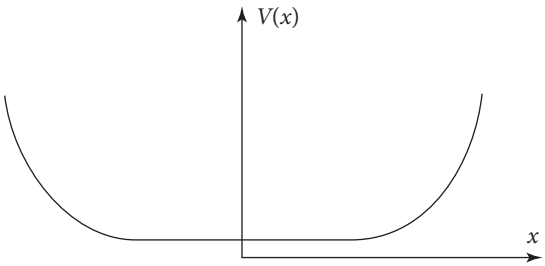


FIG. 2.10
Schematic plot showing the potential seen by the charge due to edge of the quantum Hall sample.

Dropping terms after the second one and assuming the constant term (first term) to be zero, the second term looks like the potential due to the electric field. So the particle acquires a drift velocity in the y-direction,

$$v_y = -\frac{1}{eB}\frac{\partial V}{\partial x}. \tag{2.50}$$

Each wavefunction labeled by momentum, k is located at different x-positions, namely, $x = -kl_B^2$, and thus has a different drift velocity. Additionally, the slopes of the two edges drawn above has different slopes, so $\frac{\partial V}{\partial x}$ has different signs at the edges. Thus, v_y at the left edge has a different sign with respect to the right edge. Further, because of the drift in y-direction, there will be a current I_y which is known as the Hall current and is calculated as

$$\begin{aligned} I_y &= -e\int \frac{dk}{2\pi} v_y(k) \\ &= \frac{e}{2\pi l_B^2}\int dx \frac{1}{eB}\frac{dV}{dx}; \text{ using } l_B^2 = \frac{\hbar}{eB} \\ &= \frac{e^2}{2\pi\hbar} V_H \end{aligned} \tag{2.51}$$

V_H is the Hall voltage. Now, $\sigma_{xy} = \frac{I_y}{V_H} = \frac{e^2}{2\pi\hbar} = \frac{e^2}{h}$, which is indeed the expected conductivity for a single Landau level.

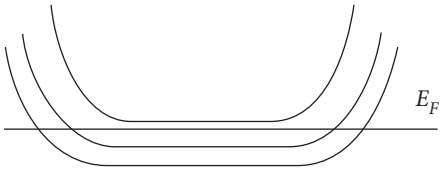


FIG. 2.11
Schematic plot showing the appearance of edge modes.

The above schematic diagram (Fig. 2.11) shows that the current is entirely carried by the edge states, since the bulk Landau level is absolutely flat (having no k dependence) and hence does not carry any current. The argument is also elegant, as it does not depend upon the form of the potential $V(x)$.

Everything that we have discussed so far holds for a single Landau level, however the argument is equally valid for a large number of Landau levels, as long as the Fermi energy lies in between the filled and the unfilled Landau levels. Also, the chiral edge modes are robust to any impurity or disorder as there is no phase space available for scattering. If a left moving electron is to scatter over onto a right moving electron, it has to cross the entire sample edge, which is not allowed as the probability of scattering would be infinitesimally small owing to the macroscopic physical dimension of the sample. Thus, the Hall plateaus are robust to disorder, defects and impurities.

A supremely important issue stands out: how do the plateaus exist in the first place? To see this, let us fix the electron density, n. Then, we shall only have filled Landau levels when the magnetic field is exactly $B = \frac{n}{\nu}\Phi_0$ $\left(\text{with } \Phi_0 = \frac{h}{e}\right)$ for some integer ν. But what happens when $B \neq \frac{n}{\nu}\Phi_0$, that is, when the Landau levels are partially filled? Also on top of that, there is also a (small) electric field. In the partially filled last Landau level, the longitudinal conductivity will be non-zero, while the Hall conductivity

will not be quantized. So how do the plateaus appear and the longitudinal conductivities vanish? The disorder comes to the rescue. This gives a finite width to the Landau levels. So, even if the $B \neq \frac{n}{\nu}\Phi_0$, there is a finite regime over which the Hall conductivity, σ_{xy} remains constant and hence the plateaus form so that the Hall resistivity, ρ_{xy} freezes at that value, until the magnetic field is increased further.

2.1.9 Incompressibility and the QH states

The key feature of the QH states is that the states are incompressible. The compressibility, κ is defined by

$$\kappa = -\frac{1}{A}\frac{\partial A}{\partial P}\Big|_N \tag{2.52}$$

where A, P and N are the area, pressure and the number of particles, respectively. The system is said to be incompressible when $\kappa = 0$, that is, when the area is insensitive to the pressure applied. The pressure is defined as the change of energy as a result of change of area, that is,

$$P = -\frac{\partial E}{\partial A}.$$

Thus, the inverse of the compressibility is defined as

$$\kappa^{-1} = -A\frac{\partial P}{\partial A}\Big|_N = A\frac{\partial^2 E}{\partial A^2}\Big|_N.$$

Since the energy is an extensive quantity, that is, it depends on the number of particles and hence can be written as

$$E = N\epsilon(n)$$

where ϵ is the energy per particle, and n is the particle density, such that the total number of particles is defined by $N = An$. Hence,

$$\begin{aligned}\kappa^{-1} &= \frac{1}{n}\frac{d}{d(\frac{1}{n})}\frac{d\epsilon(n)}{d(\frac{1}{n})} \\ &= n^2\left(2\frac{d\epsilon(n)}{dn} + n\frac{d^2\epsilon(n)}{dn^2}\right) \\ &= n^2\frac{d^2(n\epsilon)}{dn^2}.\end{aligned} \tag{2.53}$$

Also, the chemical potential is given by

$$\mu = \frac{\partial E}{\partial N}\Big|_V = \frac{d(n\epsilon)}{dn}. \tag{2.54}$$

Thus, comparing Eqs. (2.53) and (2.54),

$$\kappa^{-1} = n^2\frac{d\mu}{dn}. \tag{2.55}$$

The system is incompressible ($\kappa = 0$) when the chemical potential μ increases discontinuously as a function of density, that is, $\frac{\partial n}{\partial \mu} = 0$.

2.2 DERIVATION OF THE HALL RESISTANCE

Before we go on to derive the conductivity (or the resistivity) using the Kubo formula, let us present a simpler derivation. Consider a length "l" of the sample. The electric current carried by each charge in this length is given by $-ev/l$, where v denotes the group velocity, and e is the magnitude of the electronic charge. Now the total number of electrons between the momentum range p and $p + dp$ can be found by multiplying the current carried per unit charge ($=-ev/l$) and $(l/h)dp$ where h is the Planck's constant. Thus, the elemental current is given by

$$dI = -(ev/l) \times \left(\frac{l}{h}\right) dp = -\frac{ev}{h} dp.$$

Thus, the arbitrary length, l gets canceled. Replacing the velocity by $v = \frac{dE}{dp}$, the elemental current becomes

$$dI = -\frac{e}{h}\frac{dE}{dp} dp. \tag{2.56}$$

Hence, the total current can be obtained by integrating Eq. (2.56) from p_1 to p_2 denote arbitrary magnitudes of the momenta corresponding to the top (TE) and the bottom (BE) edges[5], respectively of a typical Hall sample (see Fig. 2.2). Thus, the total current is given by

$$I = \int_{p_1}^{p_2} \left(-\frac{e}{h}\right)\left(\frac{dE}{dp}\right) dp = -\frac{e}{h}[V(BE) - V(TE)] \tag{2.57}$$

where V denotes the potential energies at the two edges. Thus, the current is driven by the potential energy difference between the BE and the TE, which of course arises out of the potential difference that exists between them. Thus,

$$I = -\frac{e}{h}[(-eV_2) - (-eV_1)] \tag{2.58}$$

where V_2 and V_1 denote the voltages at the BE and TE, respectively. This yields a familiar form for the current, that is,

$$I = \frac{e^2}{h}(V_2 - V_1). \tag{2.59}$$

This is the Hall current, which is independent of the dimensions of the sample. The resistance, or more precisely the Hall resistance, is given by

$$R_H = \frac{V_2 - V_1}{I} = \frac{h}{e^2}$$

which is precisely the Hall resistivity for the filling fraction, $\nu = 1$. For an arbitrary filling fraction ν, the number of electrons will be ν times what it is for $\nu = 1$, where the Hall resistance can be written as

$$R_{\mathrm{H}} = \frac{h}{\nu e^2} \tag{2.60}$$

[5] The top and the bottom edges are perpendicular to the direction of the applied electric field.

2.3 KUBO FORMULA AND THE HALL CONDUCTIVITY

The important question at this stage is what protects the plateaus in the Hall conductivity (or the resistivity)? Why are they so flat and robust at the integer values (or at certain rational fractions)? Remember the system does not have either translational invariance (because of the presence of disorder) or time reversal invariance (because of the presence of a magnetic field). Thus, two known symmetries are lost and still the plateaus persist.[6] We shall now show that the Hall conductivity assumes $\sigma_{xy} = \frac{\nu e^2}{h}$ (or equivalently $\rho_{xy} = \frac{h}{\nu e^2}$) under these conditions.

To compute the conductivity, we resort to the Kubo formula, which arises from a more generalized concept, known as the "*linear response theory*." We shall derive the Kubo formula under a few conditions for the sake of simplicity. They are,

i. Before the fields are applied, at $t = -\infty$, the system is in a non-interacting many-particle state which obeys $\mathcal{H}_0|\psi_m\rangle = E_m|\psi_m\rangle$, where $\{E_m, \psi_m\}$ are the eigensolutions of $\mathcal{H}_0$.
ii. Even if we actually apply a constant electric field, **E**, it is helpful to consider an alternating field with frequency, ω of the form, $\mathbf{E}(t) = \mathbf{E}e^{-\mathbf{i}\omega t}$ and at the end of the calculation, take the zero frequency limit, that is, $\omega \to 0$.
iii. Consider a gauge in which the transverse components of the vector potential are zero, that is, $A_t = 0$. In other words, $\mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t}$ with no $\nabla\phi$ term, ϕ being the scalar potential. Equivalently, one can assume that $\phi =$ constant.

Now let us write the full Hamiltonian as

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}' \tag{2.61}$$

where $\mathcal{H}_0$ is the non-interacting Hamiltonian whose exact solutions are known (as described earlier) and $\mathcal{H}'$ is the interaction term due to the coupling of the electrons to the external field. Thus, $\mathcal{H}'$ involves the current due to the motion of the electrons coupling with the vector potential arising due to the presence of a magnetic field. Thus,

$$\mathcal{H}' = -\mathbf{J} \cdot \mathbf{A} \tag{2.62}$$

where **J** and **A** denote the electric current density and the vector potential, respectively. **J** is related to the mechanical momentum $\mathbf{p} + e\mathbf{A}$. By using $\mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t}$, one can write

$$\mathbf{A} = \frac{\mathbf{E}}{i\omega} e^{-i\omega t}. \tag{2.63}$$

The aim is to compute the expectation value of the current density and find out how it depends on the applied electric field, such that the proportionality constant yields the conductivity. In particular, we are interested in computing the Hall conductivity, σ_{xy}.

[6] In fact, naively it seems is ironical that a broken time reversal symmetry is solely responsible for quantization of the Hall plateaus.

Here we shall consider the interaction picture where the time evolution of an arbitrary operator $\hat{O}$ is written as

$$\hat{O}(t) = e^{i\mathcal{H}_0 t/\hbar}\hat{O}(0)e^{-i\mathcal{H}_0 t/\hbar}. \tag{2.64}$$

Here $\hat{O}$ can be any operator, such as $\mathbf{J}$ or $\mathcal{H}'$. Further, the eigenstates in the interaction picture evolve with time according to,

$$|\psi(t)\rangle = U(t, t_0)\,|\psi(t_0)\rangle, \tag{2.65}$$

where t_0 refers to an earlier time when the interaction is switched on and t denotes a later time. The time evolution operator, $U(t, t_0)$ is a unitary operator having a form,

$$U(t, t_0) = T\exp\left(-\frac{i}{\hbar}\int_{t_0}^{t}\mathcal{H}'(t')dt'\right) \tag{2.66}$$

T denotes time ordering in the above equation (Mahan, 2000). If the interval $[t : t_0]$ is split into several time steps, T keeps the earliest time to the right. Now let us consider that as $t_0 \to -\infty$, that is, before the perturbation is switched on, the system is in the ground state $|\psi_0(t)\rangle$. Hence, the time evolution operator can be written as

$$U(t, t_0 = -\infty) = T\exp\left(-\frac{i}{\hbar}\int_{-\infty}^{t}\mathcal{H}'(t')dt'\right) = U(t)\ \text{ say.} \tag{2.67}$$

The ground state expectation value of the current operator is given by

$$\begin{aligned}\langle \boldsymbol{J}(t)\rangle &= \langle\psi_0(t)|\,\boldsymbol{J}(t)\,|\psi_0(t)\rangle \\ &= \langle\psi_0|U^{-1}(t)\boldsymbol{J}(t)U(t)\,|\psi_0\rangle \\ &= \langle\psi_0|\left[Te^{\frac{i}{\hbar}\int_{-\infty}^{t}\mathcal{H}'(t')dt'}\boldsymbol{J}(t)Te^{-\frac{i}{\hbar}\int_{-\infty}^{t}H'(t')dt'}\right]|\psi_0\rangle.\end{aligned} \tag{2.68}$$

An expansion of the exponentials (assuming the interaction term to be weak) and retaining terms up to first order in $\mathcal{H}'$ yields,

$$\langle\mathbf{J}(t)\rangle \approx \langle\psi_0|\left[\mathbf{J}(t) + \frac{i}{\hbar}\int_{-\infty}^{t}dt'[\mathcal{H}'(t'), \mathbf{J}(t)]\right]|\psi_0\rangle. \tag{2.69}$$

The second term inside the bracket of RHS involves a commutator of $\mathcal{H}'$ and $\mathbf{J}$. It is to be kept in mind that the commutator does not vanish at two arbitrary times t and t'. Using $\mathbf{A}(t) = \frac{\mathbf{E}}{i\omega}e^{-i\omega t}$ and $\mathcal{H}'(t) = -\mathbf{J}\cdot\mathbf{A}$, the $\langle\mathbf{J}(t)\rangle$ takes the form,

$$\langle\mathbf{J}(t)\rangle \approx \langle\psi_0|\left[\mathbf{J}(t) + \frac{i}{\hbar}\int_{-\infty}^{t}dt'[-\mathbf{J}(t')\cdot\frac{\mathbf{E}}{i\omega}e^{-i\omega t'}, \mathbf{J}(t)]\right]|\psi_0\rangle. \tag{2.70}$$

The first term inside the bracket in RHS is the current due to the absence of an external electric field, which can safely be ignored in this study. So only the second term survives. Thus, writing for components α, β ($\alpha, \beta \in x, y, z$)

$$\langle J_\alpha(t)\rangle = \frac{1}{\hbar\omega}\int_{-\infty}^{t}dt'\langle\psi_0|\,[J_\beta(t'), J_\alpha(t)]\,|\psi_0\rangle E_\beta e^{-i\omega t'} \tag{2.71}$$

where $\boldsymbol{J}\cdot\mathbf{E}$ is written as $J_\beta E_\beta$.

Since the system is invariant under time translation, the above correlation depends on $t - t'$ and not on t and t' individually. Introducing a new variable $t - t' = \tilde{t}$,

$$\langle J_\alpha(t)\rangle = \frac{1}{\hbar\omega}\left(\int_0^\infty d\tilde{t} e^{i\omega\tilde{t}}\langle\psi_0|\,[J_\beta(0), J_\alpha(\tilde{t})]\,|\psi_0\rangle\right)E_\beta e^{-i\omega t} \tag{2.72}$$

where the term inside the bracket in RHS can be written as $\sigma_{\alpha\beta}E_\beta e^{-i\omega t}$, $\sigma_{\alpha\beta}$ being the components of the conductivity tensor. Note that the time t dependence is outside the integral and appears as $e^{-i\omega t}$. Thus, in the linear response regime, if an electric field of frequency, ω is applied, the current responds to the external field by oscillating with the same frequency as the external field.

The Hall conductivity is the off-diagonal component and can be computed using,

$$\sigma_{xy}(\omega) = \frac{1}{\hbar\omega}\int_0^\infty dt e^{i\omega t}\langle\psi_0|\,[J_y(0),\, J_x(t)]\,|\psi_0\rangle \tag{2.73}$$

$J_x(t)$ can be written as

$$J_x(t) = e^{i\mathcal{H}_0 t/\hbar}J_x(0)e^{-i\mathcal{H}_0 t/\hbar} \tag{2.74}$$

in the above expression, and using the complete relation of the states, namely, $|\psi_n\rangle\langle\psi_n| = \mathbf{1}$, one obtains

$$\begin{aligned}\sigma_{xy}(\omega) = \frac{1}{\hbar\omega}\int_0^\infty dt\, e^{i\omega t}\sum_n &\langle\psi_0|\,J_y\,|\psi_n\rangle\langle\psi_n|\,J_x\,|\psi_0\rangle e^{i(E_n-E_0)t/\hbar} \\ &- \langle\psi_0|\,J_x\,|\psi_n\rangle\langle\psi_n|\,J_y\,|\psi_0\rangle e^{i(E_0-E_n)t/\hbar}.\end{aligned} \tag{2.75}$$

Now, we shall perform the integral over time, t. Let us write $\frac{E_n}{\hbar} = \omega_n$ and introduce $\omega \to \omega + i\epsilon$ where ϵ is the infinitesimal quantity considered for the convergence of the integral. Then, the Hall conductivity is obtained as

$$\sigma_{xy}(\omega) = -\frac{i}{\omega}\sum_{E_n\neq E_0}\left[\frac{\langle\psi_0|\,J_y\,|\psi_n\rangle\langle\psi_n|\,J_x\,|\psi_0\rangle}{\hbar\omega + E_n - E_0} - \frac{\langle\psi_0|\,J_x\,|\psi_n\rangle\langle\psi_n|\,J_y\,|\psi_0\rangle}{\hbar\omega + E_0 - E_n}\right]. \tag{2.76}$$

Finally, we shall take the limit $\omega \to 0$ to account for the constant field. For that, let us expand the denominator as follows:

$$\begin{aligned}\frac{1}{\hbar\omega + E_n - E_0} &\approx \frac{1}{E_n - E_0} - \frac{\hbar\omega}{(E_n - E_0)^2} + O(\omega^2) \\ \frac{1}{\hbar\omega + E_0 - E_n} &\approx \frac{1}{E_0 - E_n} - \frac{\hbar\omega}{(E_0 - E_n)^2}.\end{aligned} \tag{2.77}$$

The first term looks divergent and such divergence is responsible for the peak in the longitudinal conductivity, which, by now we are familiar with. Moreover, $\sigma_{xy}(\omega)$ should not contain a term which is independent of ω. This is the DC conductivity ($\omega = 0$) which is absent in a translationally invariant system. Thus, finally one arrives at,

$$\sigma_{xy}(\omega) = i\hbar\sum_{n\neq 0}\frac{\langle\psi_0|\,J_y\,|\psi_n\rangle\langle\psi_n|\,J_x\,|\psi_0\rangle - \langle\psi_0|\,J_x\,|\psi_n\rangle\langle\psi_n|\,J_y\,|\psi_0\rangle}{(E_n - E_0)^2}. \tag{2.78}$$

This is the Kubo formula for the Hall conductivity.

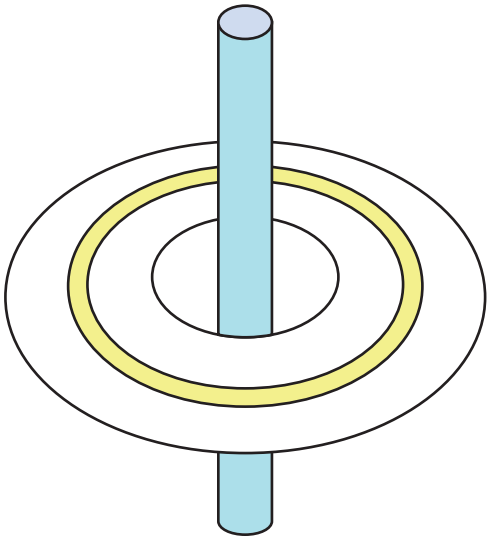


FIG. 2.12
The schematic plot shows a disc threaded by a flux perpendicular to the plane and another one in the plane of the disc.

To proceed further, let us assume that a specific case of perturbing a system below (Tong, 2016). Consider a quantum Hall sample in the form of a torus (or a donut). Let us thread two fluxes Φ_x and Φ_y (instead of one) as shown in Fig. 2.12. Owing to this, the gauge potentials can be written as

$$A_x = \frac{\Phi_x}{L_x},$$
$$A_y = \frac{\Phi_y}{L_y} + Bx. \tag{2.79}$$

It is clear that the states of the quantum system are sensitive to non-integer values of Φ_i/Φ_0 $(i = x, y)$ where, $\Phi_0 = \frac{h}{e}$. Specifically, if we increase either Φ_x or Φ_y from 0 to Φ_0 then the spectrum of the quantum system must remain invariant. Hence, writing the perturbation Hamiltonian in terms of the fluxes,

$$\mathcal{H}' = -\sum_{i\in 1,2} \frac{J_i \Phi_i}{L_i}. \tag{2.80}$$

To the first order corresponding to this perturbation term, the modified ground state becomes,

$$|\psi_0'\rangle = |\psi_0\rangle + \sum_{\substack{\psi_n\neq\psi_0 \\ E_n\neq E_0}} \frac{\langle\psi_n| H' |\psi_0\rangle}{E_n - E_0} |\psi_n\rangle. \tag{2.81}$$

Thus, writing,

$$\Delta|\psi\rangle = |\psi_0'\rangle - |\psi_0\rangle = \sum_{\substack{\psi_n\neq\psi_0 \\ E_n\neq E_0}} \frac{\langle\psi_n| H' |\psi_0\rangle}{E_n - E_0} |\psi_n\rangle. \tag{2.82}$$

Considering infinitesimal changes in Φ_i we can write,

$$\left|\frac{\partial \Delta\psi}{\partial \Phi_i}\right\rangle = -\frac{1}{L_i}\sum_n \frac{\langle\psi_n| J_i |\psi_0\rangle}{E_n - E_0} |\psi_n\rangle. \tag{2.83}$$

Terms like those in the RHS of the above equation appeared in the Hall conductivity. Let us write the total Hall conductivity including the area factor L_xL_y of the sample, which can be written as

$$\sigma_{xy} = i\hbar L_x L_y \sum_{\substack{\psi_n\neq\psi_0 \\ E_n\neq E_0}} \frac{\langle\psi_0| J_y |\psi_n\rangle\langle\psi_n| J_x |\psi_0\rangle - \langle\psi_0| J_x |\psi_n\rangle\langle\psi_n| J_y |\psi_0\rangle}{(E_n - E_0)^2}$$
$$= i\hbar\left[\left\langle\frac{\partial\psi_0}{\partial\Phi_y}\middle|\middle|\frac{\partial\psi_0}{\partial\Phi_x}\right\rangle - \left\langle\frac{\partial\psi_0}{\partial\Phi_x}\middle|\middle|\frac{\partial\psi_0}{\partial\Phi_y}\right\rangle\right] \tag{2.84}$$

$$= i\hbar\left[\frac{\partial}{\partial\Phi_y}\left\langle\psi_0\left|\frac{\partial\psi_0}{\partial\Phi_x}\right\rangle - \frac{\partial}{\partial\Phi_x}\left\langle\psi_0\left|\frac{\partial\psi_0}{\partial\Phi_y}\right\rangle\right.\right] \tag{2.85}$$

$\left\langle\psi_0\left|\frac{\partial\psi_0}{\partial\Phi_x\partial\Phi_y}\right\rangle\right.$ will be canceled from both the terms.

2.3.1 Hall conductivity and the Chern number

Remember that the spectrum of the Hamiltonian depends upon Φ_i mod Φ_0.[7] If there is a remainder, then the division does not yield an integer, and if there is none, then $\Phi_i/\Phi_0 =$ integer. Φ_i being parameters of $\mathcal{H}'$, Φ_i's are periodic functions. To emphasize the periodicity, we shall introduce angular variables, θ_i such that,

$$\theta_i = \frac{2\pi\Phi_i}{\Phi_0} \qquad \text{where } \theta_i \in [0, 2\pi]. \tag{2.86}$$

As θ_i increases from 0 to 2π, Φ_i increases from $0 \to \Phi_0$. Now rewrite $\frac{\partial}{\partial\Phi_i}$ as $\frac{\partial}{\partial\theta_i}$ and introduce a quantity called Berry connection, which is defined on the surface of the torus as

$$\mathcal{A}_i(\Phi) = -i\langle\psi_0|\frac{\partial}{\partial\theta_i}|\psi_0\rangle. \tag{2.87}$$

Further, we define a quantity called the Berry curvature (analogous to the magnetic field),

$$\mathcal{F}_{xy} = \frac{\partial\mathcal{A}_x}{\partial\theta_y} - \frac{\partial\mathcal{A}_y}{\partial\theta_x} = (\nabla_\theta \times \mathcal{A}) = -i\left[\frac{\partial}{\partial\theta_y}\left\langle\psi_0\left|\frac{\partial\psi_0}{\partial\theta_x}\right\rangle - \frac{\partial}{\partial\theta_x}\left\langle\psi_0\left|\frac{\partial\psi_0}{\partial\theta_y}\right\rangle\right.\right]. \tag{2.88}$$

Note that the last term in Eq. (2.88) is the Hall conductivity, which is formally written in terms of the Berry curvature as

$$\sigma_{xy} = -\frac{e^2}{h}\mathcal{F}_{xy}. \tag{2.89}$$

We are still left with the task of understanding the quantization of σ_{xy} which is central to the discussion on quantum Hall effect. We can now integrate over the surface of the torus to get the total conductivity,

$$\sigma_{xy} = -\frac{e^2}{h}\int_{\text{torus}}\frac{d^2\theta}{(2\pi)^2}\mathcal{F}_{xy}. \tag{2.90}$$

The quantity $C = \frac{1}{2\pi}\int d^2\theta\mathcal{F}_{xy}$ is called the first Chern number. Thus, if we average over the fluxes the conductivity assumes a form,

$$\sigma_{xy} = -\frac{e^2}{h}C \tag{2.91}$$

C is necessarily an integer. It is also referred to as the TKNN originates after Thouless *et al.* (1982).

Here, we provide an argument that the Chern number, which is the integral over the Berry curvature, is indeed an integer. For simplicity, let us assume a translationally invariant system in which the eigenstates can be represented by the Bloch functions. That is, $|\psi\rangle_0$ appearing above can be written as $u_k e^{i\phi_k}$

[7] Φ_i mod Φ_0 refers to the remainder when Φ_i is divided by Φ_0.

where u_k captures the periodicity of the lattice. Since the Berry connection requires a derivative to be taken, namely, $\langle\psi_0|\,\frac{\partial}{\partial\theta}\,|\psi_0\rangle$, which would be equivalent to,

$$\frac{\partial}{\partial\phi_k}(u_k e^{i\phi_k}) \approx \nabla_k\phi_k. \tag{2.92}$$

Now, when one takes an integral over the Brillouin zone (which is equivalent to the surface of the torus in real space), then $(\nabla_k\phi_k)\cdot dk$ over a closed surface is zero.

$$\oint \nabla_k\phi_k dk = 0. \tag{2.93}$$

This says that the measurable quantity, $e^{i\phi}$ obeys,[8]

$$e^{i\phi(0)} = e^{i\phi(2\pi)}. \tag{2.94}$$

Thus,

$$\begin{aligned}|\phi(0) - \phi(2\pi)| &= 2\pi \times (\text{some integer})\\ &= 2\pi C.\end{aligned} \tag{2.95}$$

Thus, the integral over the curvature,

$$\oint \mathcal{F}\frac{d^2k}{2\pi} = \oint \nabla_k\phi_k dk = \frac{2\pi C}{2\pi} = C. \tag{2.96}$$

Thus, Chern number is an integer and hence we get the Hall conductivity to be quantized in units of e^2/h. The above calculations are of course applicable to IQHE.

2.4 QUANTUM HALL EFFECT IN GRAPHENE

Having studied the quantized Hall effect in a 2DEG in detail, we focus on another system which is of topical interest, namely, graphene. Apart from reviewing the basic electronic properties of graphene, we discuss the properties of the Landau levels. The unequal spacing between the successive Landau levels is a feature that shows up in many experiments. We discuss some of them.

Graphene is made of a single layer of carbon (C) atoms arranged in a honeycomb lattice structure. Each C atom has six electrons with an electronic configuration $1s^2 2s^2 2p^2$. Four electrons in the 2s and the 2p orbitals create a hybridized sp^2 bonding orbital. The orbital connects to the three nearest neighbors in the plane of the crystal lattice, while the fourth electron occupies a π orbital that projects out above and below the plane. This π orbital has a significant overlap with those from the neighboring C atoms. This is responsible for rendering excellent mobility of graphene, while the σ bonds that connect a C atom to its three neighbors yield the stability of the crystal structure. The discovery of such one-atom thick planar material (graphene is a perfect example of a two-dimensional (2D) material

[8] It should be remembered that ϕ is not a measurable quantity, while $e^{i\phi}$ is a measurable quantity.

realized so far) earned a Nobel prize to A. Geim and K. Novoselov, both from the University of Manchester in the UK in 2010. Before we embark on the Hall effect in graphene, let us review its electronic properties.

2.4.1 Basic electronic properties of graphene

Owing to the large mobility of π electrons in graphene, a nearest neighbor tight-binding Hamiltonian of the following form is most suitable, namely,

$$\mathcal{H} = -t \sum_{\langle ij \rangle, \sigma} (a^{\dagger}_{i\sigma} b_{j\sigma} + \text{h.c.}) \tag{2.97}$$

where $a^{\dagger}(b)$ denote creation (annihilation) operators for electrons at the A(B) sublattice sites. σ denotes the spin of the electrons, however that will be suppressed in the next step onwards owing to the active role played by the spin either in the band structure or in QHE. We shall make the spin degrees of freedom apparent only when it is needed. Here $t \simeq 2.7$ eV which is considerably large and allows us to ignore the electron-electron interaction. The vectors connecting the nearest neighbors, $\boldsymbol{\delta}_i$, direct, $\mathbf{a}_i$ and the reciprocal $\mathbf{b}_i$ lattice vectors, are written as

$$\begin{aligned}
\boldsymbol{\delta}_1 &= \frac{a}{2}(\sqrt{3}\hat{x} + \hat{y}); \quad \boldsymbol{\delta}_2 = \frac{a}{2}(-\sqrt{3}\hat{x} + \hat{y}); \quad \boldsymbol{\delta}_3 = -a\hat{x} \\
\mathbf{a}_1 &= \frac{a}{2}(\sqrt{3}\hat{x} + 3\hat{y}); \quad \mathbf{a}_2 = \frac{a}{2}(-\sqrt{3}\hat{x} + 3\hat{y}) \\
\mathbf{b}_1 &= \frac{2\pi}{3a}(\sqrt{3}\hat{k}_x + \hat{k}_y); \quad \mathbf{b}_2 = \frac{2\pi}{3a}(-\sqrt{3}\hat{k}_x + 3\hat{k}_y),
\end{aligned} \tag{2.98}$$

with the lattice constant, $a = 1.42$Å (see Fig. 2.13).

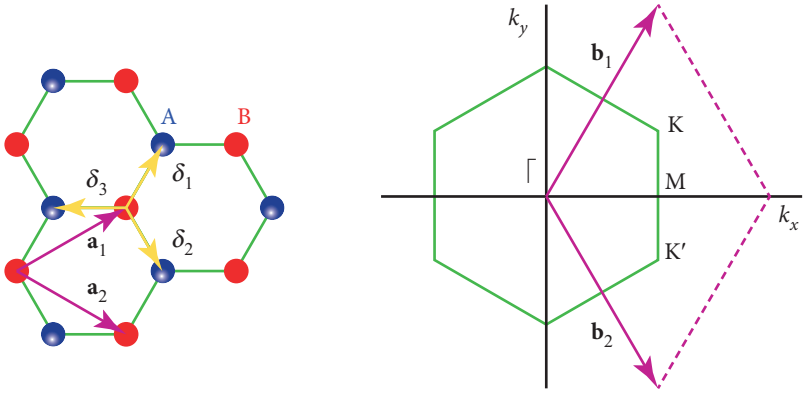


FIG. 2.13
Plot showing $\boldsymbol{\delta}_i$, direct, $\mathbf{a}_i$ and the reciprocal $\mathbf{b}_i$ lattice vectors for graphene.

Using the nearest neighbor vectors, $\boldsymbol{\delta}_i$, we explicitly write the tight-binding Hamiltonian as

$$\mathcal{H} = -t \sum_{\mathbf{R},\boldsymbol{\delta}} \left[b^\dagger(\mathbf{R}+\boldsymbol{\delta}_i) a(\mathbf{R}) + a^\dagger(\mathbf{R}) b(\mathbf{R}+\boldsymbol{\delta}_i) \right]. \tag{2.99}$$

The lattice vector $\mathbf{R}$ at an arbitrary site is given by

$$\mathbf{R} = n\mathbf{a}_1 + m\mathbf{a}_2 \quad n, m \in N. \tag{2.100}$$

The Fourier transforms for these operators are performed using,

$$a_{\mathbf{k}} = \frac{1}{\sqrt{N}} \sum_{\mathbf{R}} e^{-i\mathbf{k}\cdot\mathbf{R}} a(\mathbf{R}). \tag{2.101}$$

This yields the Hamiltonian in the momentum space as follows,

$$\mathcal{H} = -\frac{t}{N} \sum_{\mathbf{k},\mathbf{q}} \sum_{\mathbf{R}} \sum_{i=1}^{3} e^{i(\mathbf{k}-\mathbf{q})\cdot\mathbf{R}} \left[e^{i(\mathbf{q}-\mathbf{k})\cdot\mathbf{R}} e^{i\mathbf{q}\cdot\boldsymbol{\delta}_i} b_{\mathbf{q}}^\dagger a_{\mathbf{k}} + e^{i(\mathbf{k}-\mathbf{q})\cdot\mathbf{R}} e^{-i\mathbf{q}\cdot\boldsymbol{\delta}_i} a_{\mathbf{q}}^\dagger b_{\mathbf{k}} \right]. \tag{2.102}$$

Using the following definition of the Kronecker delta,

$$\delta_{\mathbf{k},\mathbf{q}} = \frac{1}{N} \sum_{\mathbf{R}} e^{i(\mathbf{k}-\mathbf{q})\cdot\mathbf{R}} \tag{2.103}$$

one gets,

$$\begin{aligned}
\mathcal{H} &= -t \sum_{\mathbf{k}} \sum_{i=1}^{3} \left[e^{-i\mathbf{k}\cdot\boldsymbol{\delta}_i} b_{\mathbf{k}}^\dagger a_{\mathbf{k}} + e^{i\mathbf{k}\cdot\boldsymbol{\delta}_i} a_{\mathbf{k}}^\dagger b_{\mathbf{k}} \right] \\
&= -t \sum_{\mathbf{k}} \sum_{i=1}^{3} \begin{pmatrix} a_{\mathbf{k}}^\dagger & b_{\mathbf{k}}^\dagger \end{pmatrix} \begin{pmatrix} 0 & e^{-i\mathbf{k}\cdot\boldsymbol{\delta}_i} \\ e^{i\mathbf{k}\cdot\boldsymbol{\delta}_i} & 0 \end{pmatrix} \begin{pmatrix} a_{\mathbf{k}} \\ b_{\mathbf{k}} \end{pmatrix} \\
&= -t \sum_{\mathbf{k}} \sum_{i=1}^{3} \begin{pmatrix} a_{\mathbf{k}}^\dagger & b_{\mathbf{k}}^\dagger \end{pmatrix} h(\mathbf{k}) \begin{pmatrix} a_{\mathbf{k}} \\ b_{\mathbf{k}} \end{pmatrix}
\end{aligned} \tag{2.104}$$

where $h(\mathbf{k})$ is the Hamiltonian matrix defined by

$$h(\mathbf{k}) = -t \begin{pmatrix} 0 & \left(e^{i\mathbf{k}\cdot\boldsymbol{\delta}_1} + e^{i\mathbf{k}\cdot\boldsymbol{\delta}_2} + e^{i\mathbf{k}\cdot\boldsymbol{\delta}_3}\right) \\ \left(e^{-i\mathbf{k}\cdot\boldsymbol{\delta}_1} + e^{-i\mathbf{k}\cdot\boldsymbol{\delta}_2} + e^{-i\mathbf{k}\cdot\boldsymbol{\delta}_3}\right) & 0 \end{pmatrix}. \tag{2.105}$$

Since the difference between two nearest neighbor lattice vectors, $\boldsymbol{\delta}_i$ and $\boldsymbol{\delta}_j$ must yield a lattice vector, $\mathbf{R}$, we can do a transformation,

$a_{\mathbf{k}} \rightarrow e^{i\mathbf{k}\cdot\boldsymbol{\delta}_3} a_{\mathbf{k}}$, and $a_{\mathbf{k}}^\dagger \rightarrow e^{-i\mathbf{k}\cdot\boldsymbol{\delta}_3} a_{\mathbf{k}}^\dagger$.

The above transformation yields a new Hamiltonian matrix,

$$\tilde{h}(\mathbf{k}) = -t \begin{pmatrix} 0 & \left(e^{i\mathbf{k}\cdot(\boldsymbol{\delta}_1-\boldsymbol{\delta}_3)} + e^{i\mathbf{k}\cdot(\boldsymbol{\delta}_2-\boldsymbol{\delta}_3)} + 1\right) \\ \left(e^{-i\mathbf{k}\cdot(\boldsymbol{\delta}_1-\boldsymbol{\delta}_3)} + e^{-i\mathbf{k}\cdot(\boldsymbol{\delta}_2-\boldsymbol{\delta}_3)} + 1\right) & 0 \end{pmatrix}. \tag{2.106}$$

Using the definitions of $\boldsymbol{\delta}_i$, one gets,

$$\tilde{h}(\mathbf{k}) = -t \begin{pmatrix} 0 & -\left(e^{i\mathbf{k}\cdot\mathbf{a}_1} + e^{i\mathbf{k}\cdot\mathbf{a}_2} + 1\right) \\ -\left(e^{-i\mathbf{k}\cdot\mathbf{a}_1} + e^{-i\mathbf{k}\cdot\mathbf{a}_2} + 1\right) & 0 \end{pmatrix}. \tag{2.107}$$

One can check that $\tilde{h}(\mathbf{k})$ obeys $\tilde{h}(\mathbf{k}) = \tilde{h}(\mathbf{k}+\mathbf{G})$, where $\mathbf{G}$ is the reciprocal lattice vector, defined as $\mathbf{G} = p\mathbf{b}_1 + q\mathbf{b}_2$, with p and q being integers. Thus,

$$\tilde{h}(\mathbf{k}) = -t \begin{pmatrix} 0 & f(\mathbf{k}) \\ f^*(\mathbf{k}) & 0 \end{pmatrix} \tag{2.108}$$

where

$$f(\mathbf{k}) = -t\left(e^{-ik_x a} + 2e^{ik_x a/2}\cos\left(\frac{k_y\sqrt{3}a}{2}\right)\right).$$

The tight-binding energy is obtained by diagonalizing $\tilde{h}(\mathbf{k})$ which yields,

$$\epsilon_{\mathbf{k}} = \pm t\sqrt{|f(k)|^2} = \pm t\sqrt{3 + 2\cos(\sqrt{3}ak_y) + 4\cos(\sqrt{3}ak_y/2)\cos(3ak_x/2)}. \tag{2.109}$$

The two bands described by the "+" and the "−" signs in the above dispersion touch at six points in the Brillouin zone. Since graphene has one accessible electron per C atom, one can assume a half-filled system where the lower band is completely filled. Further, we wish to discuss the low-lying excitations just above the ground state of the system.

This necessitates exploring the low-energy theory of graphene. To achieve that we have to identify the band touching points which can be obtained from the condition, $f(\mathbf{k}) = 0$. Separately, putting the real and the imaginary parts equal to zero yield,

$$\begin{aligned} &\cos(k_x a) + 2\cos(k_x a/2)\cos(\sqrt{3}k_y a/2) = 0 \\ &-\sin(k_x a) + 2\sin(k_x a/2)\cos(\sqrt{3}k_y a/2) = 0. \end{aligned} \tag{2.110}$$

Eq. (2.110) can be manipulated as follows,

$$\sin(k_x a/2)\left[-\cos(k_x a/2) + \cos(k_y a\sqrt{3}/2)\right] = 0. \tag{2.111}$$

Thus, one is left with two options, namely,

either (i) $\sin(k_x a/2) = 0$; which means $\cos(k_x a/2) = \pm 1$;

or, (ii) $\cos(k_x a/2) = \cos(\sqrt{3}k_y a/2)$.

Option (i) gives us,

$1 + 2\cos(k_y\sqrt{3}a/2) = 0$

which yields the points $(0, \pm\frac{4\pi}{3\sqrt{3}a})$ (plus or minus the reciprocal lattice vector, $\mathbf{G}$). Whereas, option (ii) can be written as

$\cos(k_y a\sqrt{3}) + 2\cos^2(k_y a\sqrt{3}/2) = 0.$

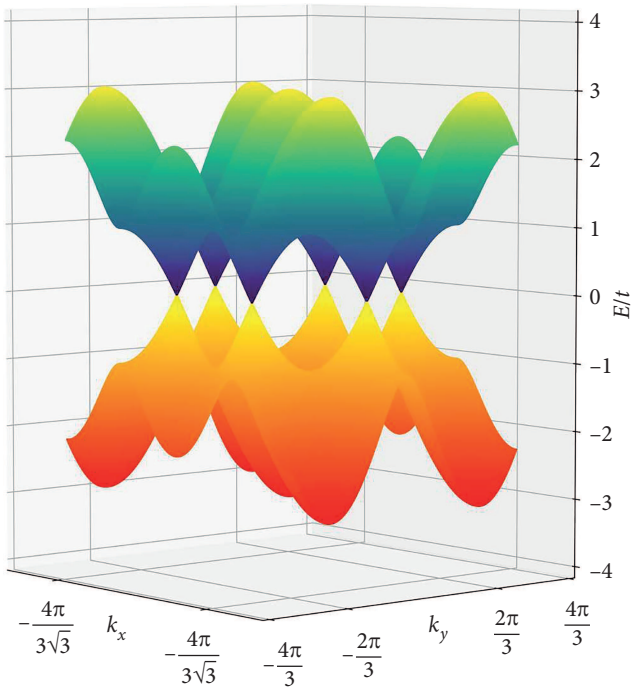


FIG. 2.14
Plot showing the two tight-binding bands of graphene in the first BZ. The two bands touch at six points. In the vicinity of those points, the bands are linearly dispersing bearing signatures of (pseudo-)relativistic physics.

Thus, we get four more points, which are, $\pm\frac{2\pi}{3a}(1, \frac{1}{\sqrt{3}})$, and $\pm\frac{2\pi}{3a}(1, -\frac{1}{\sqrt{3}})$ (again plus or minus the reciprocal lattice vector, **G**). The electronic dispersion of graphene is shown in Fig. 2.14. The six points where the two bands touch are clearly visible.

A closer inspection yields all the six points are not independent. For example, the set of vectors, namely, $\left(0, \pm\frac{4\pi}{3\sqrt{3}a}\right)$, $\frac{2\pi}{3a}\left(1, -\frac{1}{\sqrt{3}}\right)$ and $\frac{2\pi}{3a}\left(-1, -\frac{1}{\sqrt{3}}\right)$ can be connected to each other via the combination of the reciprocal lattice vectors, $\mathbf{b}_1$ and $\mathbf{b}_2$. For example,

$$\left(0, \frac{4\pi}{3\sqrt{3}a}\right) + \mathbf{b}_2 = \frac{2\pi}{3a}\left(1, -\frac{1}{\sqrt{3}}\right)$$
$$\left(0, \frac{4\pi}{3\sqrt{3}a}\right) - \mathbf{b}_1 = \frac{2\pi}{3a}\left(-1, -\frac{1}{\sqrt{3}}\right). \tag{2.112}$$

The same is true for the other vectors,

$$\left(0, \frac{4\pi}{3\sqrt{3}a}\right), \quad \frac{2\pi}{3a}\left(-1, \frac{1}{\sqrt{3}}\right), \frac{2\pi}{3a}\left(1, \frac{1}{\sqrt{3}}\right).$$

Thus, only two of them were found to be independent. Traditionally, they are called **K** and **K′** and can be written as

$$\mathbf{K} = \frac{2\pi}{3a}\left(1, \frac{1}{\sqrt{3}}\right), \text{ and } \mathbf{K'} = \frac{2\pi}{3a}\left(1, -\frac{1}{\sqrt{3}}\right).$$

Any other independent pair is also a valid choice for **K** and **K′**.

It may be kept in mind that the two bands touch at these points and the gap between the conduction and the valence bands closes. Thus, there are two branches of low-energy excitations, namely one of them with momentum close to **K** and the other close to **K′**. Since $f(\mathbf{k})$ becomes zero at $\mathbf{k} = \mathbf{K}$. By defining $\mathbf{q} = \mathbf{k} - \mathbf{K}$, one can expand $f(\mathbf{k})$ near **K** in Taylor series about $\mathbf{q} = 0$,

$$f'(\mathbf{q}) = \frac{\partial f(\mathbf{k})}{\partial k_x}\Big|_{(k_x - K_x)}(k_x - K_x) + \frac{\partial f(\mathbf{k})}{\partial k_y}\Big|_{(k_y - K_y)}(k_y - K_y) = \frac{3at}{2}(q_x + iq_y).$$

Thus, the energy spectrum assumes the form,

$$\epsilon_{\mathbf{K}}(\mathbf{q}) = \hbar v_F(q_x + iq_y) \tag{2.113}$$

where v_F is the Fermi velocity defined by $v_F = \frac{3at}{2\hbar} \simeq 10^6 ms^{-1}$. Similarly, if we expand around **K′**, one gets,

$$\epsilon_{\mathbf{K'}}(\mathbf{q}) = \hbar v_F(q_x - iq_y). \tag{2.114}$$

Thus, in a general notation, we can write,

$$\epsilon_{\mathbf{K},\mathbf{K'}} = \hbar v_F \mathbf{q} \cdot \boldsymbol{\sigma} \tag{2.115}$$

where **q** is a planar vector (q_x, q_y) and $\boldsymbol{\sigma}$ is the Pauli matrix vector (σ_x, σ_y). The electrons close to the **K** and **K′** points are called massless Dirac fermions, as they obey the Dirac equation without the "mass" term[9]. It may be noted that,

$$\epsilon_{\mathbf{K'}}(\mathbf{q}) = \epsilon^*_{\mathbf{K}}(\mathbf{q}). \tag{2.116}$$

This implies that (as will be seen later) that the "helicity" of the electrons is opposite at **K′** with respect to **K**.

To sum up our preliminary discussion on graphene, we note that the low-energy properties are governed by the dispersion,

$$\epsilon(q) = \pm v_F|q| \tag{2.117}$$

which implies that the eigenvalues are only functions of the magnitude of the wave vector **q**, and do not depend upon its direction in the 2D plane. Also, the Hamiltonian on a formal note denotes that of a massless $s = 1/2$ particle, such as a neutrino, however the velocity of the particles is reduced by a factor of 300 compared to the speed of light. Further, the handedness (or the helicity) feature of neutrinos is inbuilt, where the electrons behave similarly to the "left handed" neutrino at the Dirac point **K** and as a "right handed" neutrino at **K′** or vice versa.

[9] The Dirac equation is written in conventional notations as $\mathcal{H} = c\boldsymbol{\alpha} \cdot \mathbf{p} + \beta mc^2$ where α and β are Hermitian operators which do not operate on the space and time variables. In the case of graphene, the second term is absent.

2.4.2 Experimental confirmation of the Dirac spectrum

When a beam of monochromatic photons with an energy larger than the work function of a particular material interacts with the constituent charges (electrons) by incidenting on the surface of the sample, the electrons absorb the photons and thus possess sufficient energy to escape from the sample. By measuring the energy and the momentum of the photoelectrons and using energy-momentum conservation laws, one can derive the properties of the electrons prior to their being incident on the surface and relate them with those getting scattered. The angle resolved photoemission spectroscopy (ARPES) can be a direct probe to resolve the momentum-dependent band structure and the topology of the Fermi surface. In ARPES, a photon is employed to eject an electron from the surface of the graphene layer. The intensity of the ARPES is proportional to the transition probability from an initial Bloch state with crystal momentum, $\mathbf{k}$ and energy, E to a final state, $\mathbf{k}'$. The method conclusively establishes the existence of Dirac fermions seen via linearly dispersing bands in the vicinity of the Dirac points. The experimental setup, ARPES data and the hexagonal Brillouin zone (which we have discussed before) are shown in Fig. 2.15.

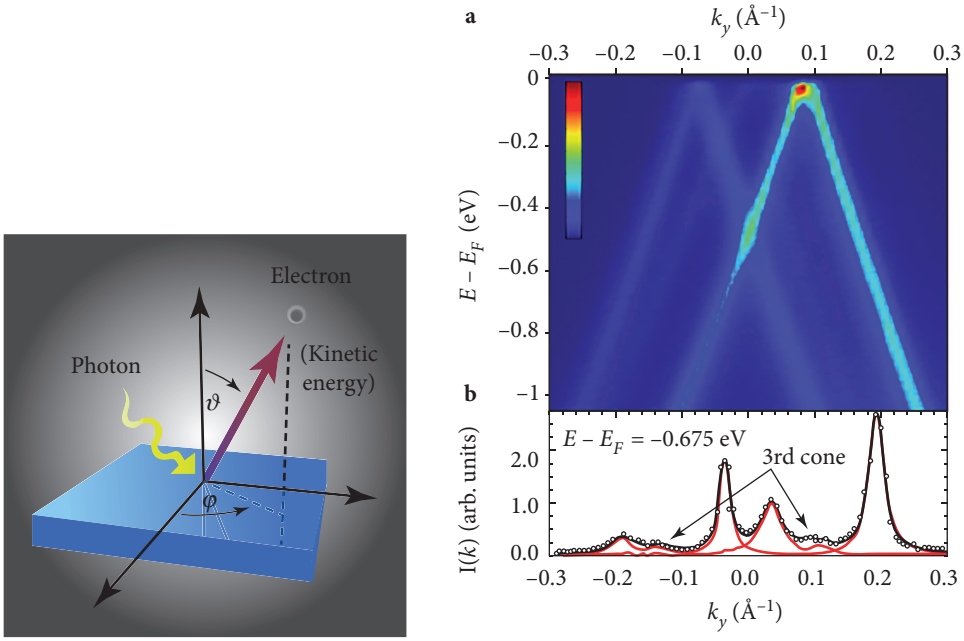


FIG. 2.15
Plot showing the experimental setup for ARPES (left), the two linearly dispersing bands (middle) and the Dirac points are shown on the hexagonal BZ of graphene (right). Taken from Sprinkle *et al.* (2009).

2.4.3 Hofstadter Butterfly

The fate of an electron gas or graphene described by a tight-binding model subjected to an external magnetic field is to show quantized plateaus of the Hall resistance. Consequently, the band energies of the electrons transform into discrete Landau levels. The presence of a periodic crystal potential adds further exciting features to the spectrum. The Hamiltonian of such a system is given by

$$\mathcal{H} = \frac{(\mathbf{p} - e\mathbf{A})^2}{2m} + V(\mathbf{r}) \tag{2.118}$$

where $V(\mathbf{r}) = V(\mathbf{r} + \mathbf{a})$ is the periodic potential with lattice periodicity $\mathbf{a}$. The electrons are described by the Bloch states leading to the formation of bands. In the presence of the vector potential, $\mathbf{A}$, each Bloch band gets further divided into sub-bands, and the resultant energy spectra as a function of the flux give rise to a fractal structure known as the Hofstadter butterfly. These rather complex energy spectra arise owing to a delicate interplay between the two length scales, namely, a and l_B that are associated with two different quantization phenomena. In fact, the Hofstadter butterfly arises when the ratio of these two lengths is a rational fraction. An even more interesting scenario emerges when the ratio is not a rational fraction, however, we shall not discuss it here.

The fractal nature of the spectrum was observed by Hofstadter (1976) which he obtained by solving Harper's equation (Azbel, 1964), and demonstrated that for commensurate values of the magnetic flux such as $\Phi/\Phi_0 = p/q$, where the single-particle Bloch bands split into q sub-bands, which themselves are p-fold degenerate (p and q being co-prime integers). Each of these p sub-bands further split, yielding a continued fraction as a function of the magnetic flux. The distance between the levels, sub-levels, etc., and the width of each of the "superstructures" oscillates with a variation of the magnetic field flux with a period that is universal and is independent of the form of the quasiparticle dispersion relation. Consequently, one observes a quasi-continuous distribution of incommensurate quantum states that form a self-similar recursive structure, like that of a butterfly. Schlösser *et al.* (1996) have realized the Hofstadter spectrum for the first time in semiconductor superlattice structures. It was later observed for a number of systems, such as cold atomic systems in optical lattices. In continuation of our present discussion, we shall discuss the Hofstadter butterfly in graphene.

In order to demonstrate the Hofstadter butterfly (Hofstadter, 1976) we have taken a semi-infinite ribbon of graphene. The ribbon geometry is such that it has zigzag edges, as in Fig. 2.16. In tight-binding approximation, the external magnetic field enters through the hopping integral, which is replaced by the Peierls substitution, namely,

$$\exp\left(\frac{ie}{\hbar}\int_i^j \mathbf{A}\cdot d\mathbf{r}\right) t_{ij} = \exp\left(i(2\pi/\Phi_0)\int_i^j \mathbf{A}\cdot d\mathbf{r}\right) t_{ij} \tag{2.119}$$

where, t_{ij} is the hopping integral between the sites i and j with no field present. The flux is denoted in terms of flux quantum $\Phi_0 = h/e$. To include the magnetic field, we have taken $\mathbf{B} = B\hat{z}$ such that

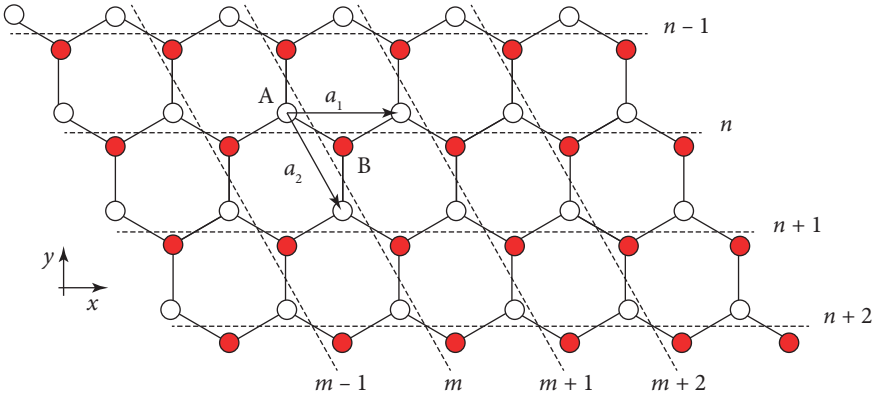


FIG. 2.16
Geometry of a zigzag graphene nanoribbon is shown. The white and the red circles represent A and B sublattices respectively. (m, n) denote the unit cell index in x and y) directions.

the vector potential ($\mathbf{A}$) in the Landau gauge takes the form $\mathbf{A} = Bx\hat{y}$. With this modification, the tight-binding Hamiltonian for graphene introduced in Fig. 2.16 can be written as

$$\mathcal{H} = -\sum_{m,n}\left[te^{i\pi\frac{\Phi}{\Phi_0}n}a^{\dagger}_{m,n}b_{m,n} + te^{-i\pi\frac{\Phi}{\Phi_0}n}a^{\dagger}_{m,n}b_{m-1,n} + t_1 a^{\dagger}_{m,n}b_{m,n-1} + h.c.\right]. \tag{2.120}$$

Here, $a^{\dagger}_{m,n}$, $b_{m,n}$ represent electron creation and annihilation operators of sublattice A and B, respectively, at site index (m, n). Since the ribbon is infinite in x direction, one can use the Fourier decomposition of the operators for the index m giving the following equation (Castro Neto *et al.*, 2009).

$$\mathcal{H} = -\sum_{k,n}\left[te^{i\pi\frac{\Phi}{\Phi_0}n}a^{\dagger}_{k,n}b_{k,n} + te^{-i\pi\frac{\Phi}{\Phi_0}n}e^{ika}a^{\dagger}_{k,n}b_{k,n} + t_1 a^{\dagger}_{k,n}b_{k,n-1} + h.c.\right] \tag{2.121}$$

Assuming the eigenfunction to be $|\psi(k)\rangle = \sum_n[\alpha_{k,n}|a,k,n\rangle + \beta_{k,n}|b,k,n\rangle]$, the eigenvalue equation of Hamiltonian (2.121) gives the following two Harper (Harper, 1955) equations.

$$E_k\alpha_{k,n} = -\left[e^{ika/2}2t\cos\left(\pi\frac{\Phi}{\Phi_0}n - \frac{ka}{2}\right)\beta_{k,n} + t_1\beta_{k,n-1}\right] \tag{2.122}$$

$$E_k\beta_{k,n} = -\left[e^{-ika/2}2t\cos\left(\pi\frac{\Phi}{\Phi_0}n - \frac{ka}{2}\right)\alpha_{k,n} + t_1\alpha_{k,n+1}\right] \tag{2.123}$$

In order to get the spectra for the Haldane and semi-Dirac Haldane model, equations (2.122) and (2.123) together with the following two characteristic equations of NNN hopping are used.

$$E_k\alpha_{k,n} = -t_2\left[e^{ika/2}\cos\left(\frac{ka}{2}+\phi\right)\alpha_{k,n+1} + \cos(ka-\phi)\alpha_{k,n} + e^{-ika/2}\cos\left(\frac{ka}{2}+\phi\right)\alpha_{k,n-1}\right] \tag{2.124}$$

$$E_k\beta_{k,n} = -t_2\left[e^{ika/2}\cos\left(\frac{ka}{2}-\phi\right)\beta_{k,n+1} + \cos(ka+\phi)\beta_{k,n} + e^{-ika/2}\cos\left(\frac{ka}{2}+\phi\right)\beta_{k,n-1}\right] \tag{2.125}$$

Using these equations, we have numerically calculated the Hofstadter butterfly spectrum for the semi-infinite nanoribbons with $q = 200$. The fractal spectra are depicted in Fig. 2.17 for graphene (Rammal, 1985). Energies are taken in units of t. The fractal spectra result from the competition of magnetic field and lattice effects. For each effective flux $f = \phi/\phi_0 = p/q$ there are $2q$ number of sub-bands. The spectrum is periodic in f with periodicity $3\sqrt{3}Ba^2/2\phi_0$.

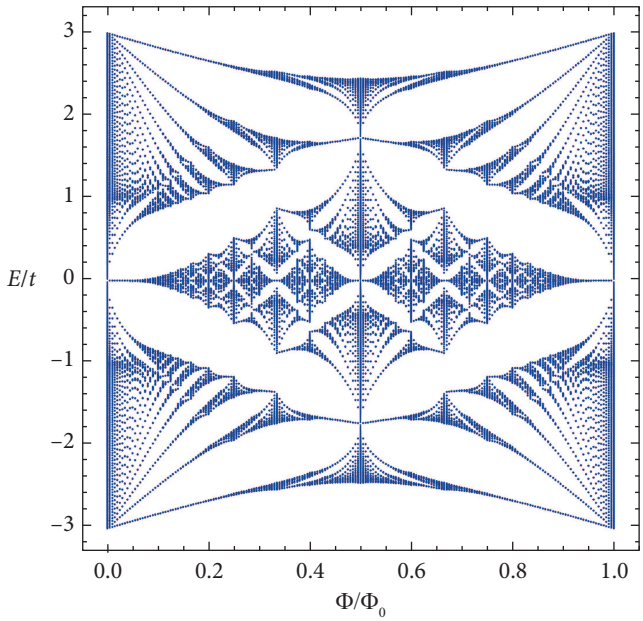


FIG. 2.17
The Hofstadter butterfly is shown for graphene. The fractal structure as a function of the external flux (scaled by the flux quantum), that is, Φ/Φ_0 can be seen.

2.4.4 Landau levels in Graphene

In order to proceed, we can write the Hamiltonian in a unified way that includes the description of both the Dirac points (valleys) **K** and **K′**. So for each valley, we have one two-dimensional spinor Hamiltonian. Thus, augmenting the Hilbert space, we can write the eigenfunctions as

$$\psi = (\psi_{\mathbf{K}'}, \psi_{\mathbf{K}})^T$$

and the Hamiltonian is given by

$$\mathcal{H} = \hbar v_F \begin{pmatrix} -\sigma^* \cdot \mathbf{k} & 0 \\ 0 & \sigma \cdot \mathbf{k} \end{pmatrix} = v_F \begin{pmatrix} -\sigma^* \cdot \mathbf{p} & 0 \\ 0 & \sigma \cdot \mathbf{p} \end{pmatrix}. \tag{2.126}$$

Now we shall discuss the motion of the massless relativistic electrons in a magnetic field. In a standard 2D electron gas, the Landau quantization produces equidistant levels [see Eq. (2.36)], which is an artifact of the non-relativistic parabolic dispersion of the free carriers. We need to ascertain how the quantization formula is modified for the case of graphene.

As earlier we do the Peierls substitution, $\mathbf{p} \to \mathbf{p} + \mathbf{eA}$. Thus, the Hamiltonian becomes,

$$\mathcal{H}_{\mathbf{K},\mathbf{K}'} = v_F \begin{pmatrix} 0 & -(p_x + ip_y) & 0 & 0 \\ -(p_x - ip_y) & 0 & 0 & 0 \\ 0 & 0 & 0 & (p_x - ip_y) \\ 0 & 0 & (p_x + ip_y) & 0 \end{pmatrix}.$$

The wavefunction has now 4-components, namely,

$$\psi = \begin{pmatrix} \phi_{\mathrm{A}}^{\mathbf{K}'} \\ \phi_{\mathrm{B}}^{\mathbf{K}'} \\ \phi_{\mathrm{A}}^{\mathbf{K}} \\ \phi_{\mathrm{B}}^{\mathbf{K}} \end{pmatrix} \tag{2.127}$$

where $\phi_{\mathrm{A,B}}^{\mathbf{K}}$ are the wavefunctions for an electron at momentum values corresponding to the valley **K** at the two sublattice sites A and B. Similar notations carry on for the other valley **K′**.

For a perpendicular magnetic field, $\mathbf{B} = B\hat{z}$, one can choose a Landau gauge, $\mathbf{A} = (-By, 0, 0)$. Since with this choice, the Hamiltonian is independent of the spatial variable, x. So $[\mathcal{H}, p_x] = 0$ and hence p_x continues to be a good quantum number.

Further the Hamiltonian in Eq. (2.127) is valley decoupled, that is, there are no matrix elements that connect the two valleys, namely, **K** and **K′**. Thus, it allows us to look at the solutions at each valley separately. For the **K** point, we have a coupled equation for the wavefunctions, ϕ^{A} and ϕ^{B}

$$\epsilon \phi_{\mathrm{A}}^{\mathbf{K}} = v_F(p_x - ip_y)\phi_{\mathrm{B}}^{\mathbf{K}} \tag{2.128}$$

$$\epsilon \phi_{\mathrm{B}}^{\mathbf{K}} = v_F(p_x + ip_y)\phi_{\mathrm{A}}^{\mathbf{K}}. \tag{2.129}$$

One can insert Eq. (2.128) in Eqs. (2.129) and (2.129) in Eq. (2.128) to obtain,

$$\epsilon^2 \phi_{\mathrm{A}}^{\mathbf{K}} = \hbar^2 v_F^2 (p_x - ip_y)(p_x + ip_y)\phi_{\mathrm{A}}^{\mathbf{K}} \tag{2.130}$$

$$\epsilon^2 \phi_{\mathrm{B}}^{\mathbf{K}} = \hbar^2 v_F^2 (p_x + ip_y)(p_x - ip_y)\phi_{\mathrm{B}}^{\mathbf{K}}. \tag{2.131}$$

Inserting the Landau gauge such that $p_x \to p_x + eBy$,

$$\frac{\epsilon^2}{\hbar^2 v_F^2}\phi_{\mathrm{B}}^{\mathbf{K}} = (p_x + eBy + ip_y)(p_x + eBy - ip_y)\phi_{\mathrm{B}}^{\mathbf{K}}$$
$$= \left[(p_x + eBy)^2 - i[(p_x + eBy), p_y] + p_y^2\right]\phi_{\mathrm{B}}^{\mathbf{K}}. \tag{2.132}$$

Since $[p_x, p_y] = 0$ and $[y, p_y] = i\hbar$, one gets,

$$\frac{\epsilon^2}{\hbar^2 v_F^2}\phi_{\mathrm{B}}^{\mathbf{K}} = [(p_x + eBy)^2 + e\hbar B + p_y^2]\phi_{\mathrm{B}}^{\mathbf{K}}. \tag{2.133}$$

Thus, we arrive at,

$$\left(\frac{\epsilon^2}{\hbar^2 v_F^2} - e\hbar B\right)\phi_{\mathrm{B}}^{\mathbf{K}} = (\tilde{p}_x^2 + \tilde{p}_y^2)\phi_{\mathrm{B}}^{\mathbf{K}} \tag{2.134}$$

where $\tilde{p}_x^2 = (p_x + eBy)^2$ and $\tilde{p}_y^2 = p_y^2$. Dividing both sides by $2m$,

$$\frac{1}{2m}\left(\frac{\epsilon^2}{\hbar^2 v_F^2} - e\hbar B\right)\phi_{\mathrm{B}}^{\mathbf{K}} = \left(\frac{\tilde{p}_x^2 + \tilde{p}_y^2}{2m}\right)\phi_{\mathrm{B}}^{\mathbf{K}} = \left(\frac{1}{2}\tilde{k}(y - y_0)^2 + \frac{\tilde{p}_y^2}{2m}\right)\phi_{\mathrm{B}}^{\mathbf{K}} \tag{2.135}$$

with $\tilde{k} = \frac{e^2B^2}{m}$, $y_0 = \frac{p_x}{eB}$. Thus, the RHS is identified as the Hamiltonian for a particle executing an SHM in two dimensions about a coordinate point $(0, y_0)$. Thus, it is obvious that the energy spectrum is given by $\epsilon_n = (n + \frac{1}{2})\hbar\omega_B$ with $\omega_B = \frac{eB}{m}$. Hence,

$$\frac{1}{m}\frac{\epsilon^2}{\hbar^2 v_F^2} = 2\left(n + \frac{1}{2}\right)\hbar\omega_B + \hbar\omega_B = \frac{2n\hbar\omega_B}{m} \qquad \text{where } n = 0, 1, 2, \ldots. \tag{2.136}$$

Eq. (2.136) allows positive and negative roots for ϵ. So we can obtain the energy spectrum as

$$\epsilon = sgn(n)\sqrt{n}\, v_F\, \frac{(2\hbar eB)^{1/2}}{m}. \tag{2.137}$$

Let us define another quantity $\tilde{\omega} = v_F\,(2\hbar eB)^{1/2}$, so as to formally write the energy expression as that of a harmonic oscillator. We rewrite the above expression as

$$\epsilon = \hbar\tilde{\omega}\mathrm{sgn}(n)\sqrt{|n|}. \tag{2.138}$$

Thus, as opposed to the familiar harmonic oscillators when n can take positive integer values (including zero), however, here in graphene, all integers, that is, both positive and negative numbers are allowed. The positive integers denote particles (or electrons) in the conduction band and the negative ones denote holes in the valence band. Furthermore, unlike the 2DEG, here, the Landau levels are not equidistant (see Fig. 2.18). The largest separation occurs between the lowest Landau level ($n = 0$) and the first one ($n = \pm 1$). This large gap essentially facilitates observation of QHE in graphene at high temperatures, which is even true for room temperature.[10]

[10] The title of the paper by Geim and Novoselov is "*Room temperature quantum Hall effect in Graphene.*" See Novoselov *et al.* (2007)

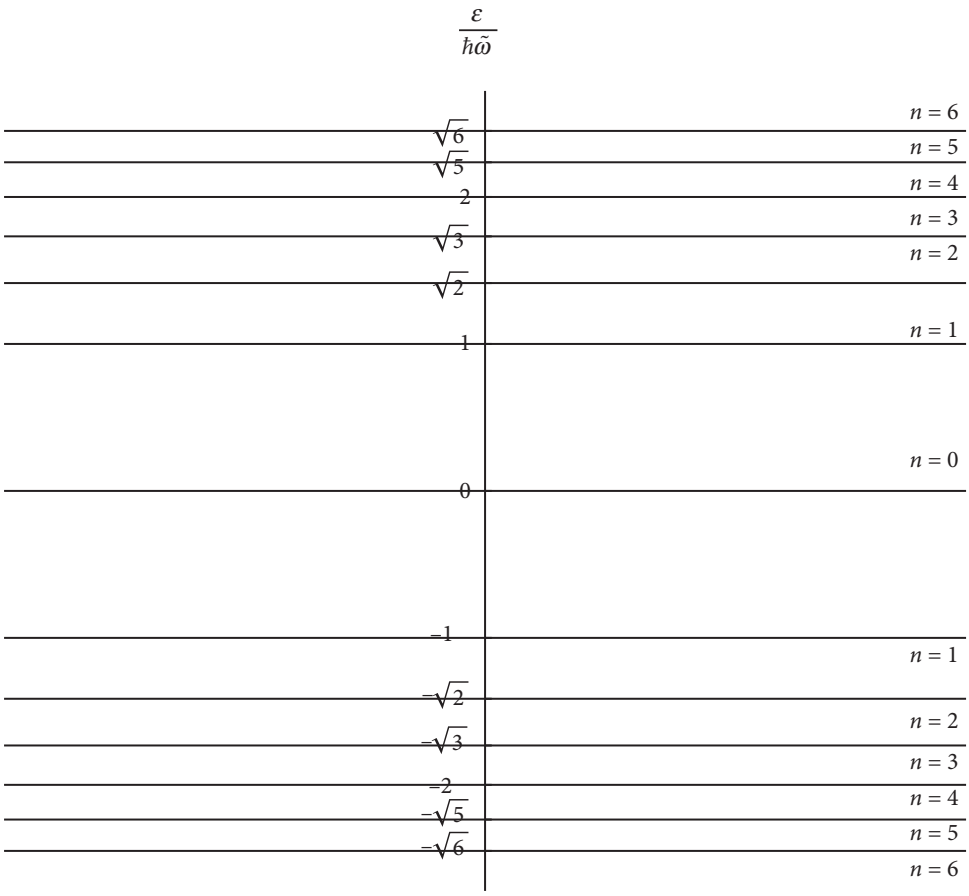


FIG. 2.18
Plot showing the Landau levels in graphene for different indices, n.

So far we have been discussing spinless particles. Including the spin, there is be an additional two-fold degeneracy of the Landau levels owing to Zeeman spitting. A hierarchy of energy scales needs to be ascertained here. Let us compare the energy gap between the two lowest Landau levels and the Zeeman splitting corresponding to a typical magnetic field, B, for example, $B = 10$ T.

$$\Delta E_{\rm LL} = \frac{\hbar\omega_B/2}{v_F\sqrt{e\hbar B}} \text{ (for the successive Landau levels)} \tag{2.139}$$

$$\Delta E_z = \sqrt{e\hbar B} \text{ (for the Zeeman term).} \tag{2.140}$$

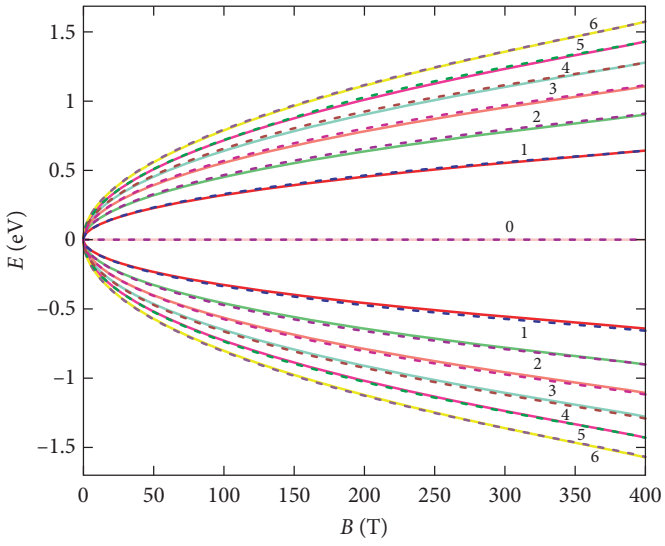


FIG. 2.19
Plot showing the Landau levels for different indices, n as a function of the magnetic field, B. The Landau levels vary as $\sqrt{n}$.

For typical $v_F \simeq c/300$ (c: speed of light),

$$\frac{\Delta E_z}{\Delta E_{\rm LL}} \simeq 10^2. \tag{2.141}$$

Thus, the Zeeman energy scale is much larger than the Landau-level splitting, which makes it imperative to include spin degeneracy. Thus, including the Zeeman term, the energy can be written as,

$$\frac{\epsilon^2}{v_F^2} = 2\hbar eB(n+1) \quad n = 0, 1, \ldots \tag{2.142}$$

where the additional term in the RHS (denoted by $2\hbar eB$) accounts for the spin. The spacing between the consecutive Landau levels are shown in Fig. 2.19. The energy level $\epsilon = 0$ is not present in the spectrum, even for $n = 0$. This lowest Landau level is somewhat special in the following sense. The $n = 0$ level receives a contribution from only one sublattice at each of the Dirac points. For example, "A" sublattice contributes to the wave function at the Dirac point $\mathbf{K}$, and the "B" sublattice contribute to $\mathbf{K}'$. However, the $n \neq 0$ Landau levels have non-zero amplitudes at both the A and the B sublattice.

Finally, the wavefunctions corresponding to an arbitrary Landau-level index at the two Dirac points, corresponding to the gauge we have chosen are,

$$\psi^{\mathbf{K}}_{n,k} = \frac{c_n}{\sqrt{L}} e^{-ikx} \begin{pmatrix} 0 \\ 0 \\ sgn(n)(-i)\phi_{|n|-1,k} \\ \phi_{|n|,k} \end{pmatrix} \tag{2.143}$$

and

$$\psi^{\mathbf{K}'}_{n,k} = \frac{c_n}{\sqrt{L}} e^{-ikx} \begin{pmatrix} \phi_{|n|,k} \\ sgn(n)(-i)\phi_{|n|-1,k} \\ 0 \\ 0 \end{pmatrix} \tag{2.144}$$

with

$$\begin{aligned} c_n(x) &= 1 \quad \text{for } n = 0 \\ &= \frac{1}{\sqrt{2}} \quad \text{for } n \neq 0. \end{aligned} \tag{2.145}$$

Further,

$$sgn(n) = 0 \quad \text{for } n = 0$$
$$= \frac{n}{|n|} \quad \text{for } n \neq 0 \tag{2.146}$$

with

$$\phi_{n,k} = exp\left[-\frac{1}{2}\frac{(y - kl_B^2)^2}{l_B^2}\right] H_n\left[\frac{(y - kl_B^2)}{l_B}\right] \tag{2.147}$$

l_B is the magnetic length ($= \sqrt{\frac{\hbar}{eB}}$) as defined before, and $H_n(x)$ are the Hermite polynomials. $\phi_{n,k}$ denotes the eigenfunctions of an electron in the presence of a magnetic field. n refers to the Landau-level index. The Landau levels for different values of the flux, ϕ are shown in Fig. 2.20 as a function

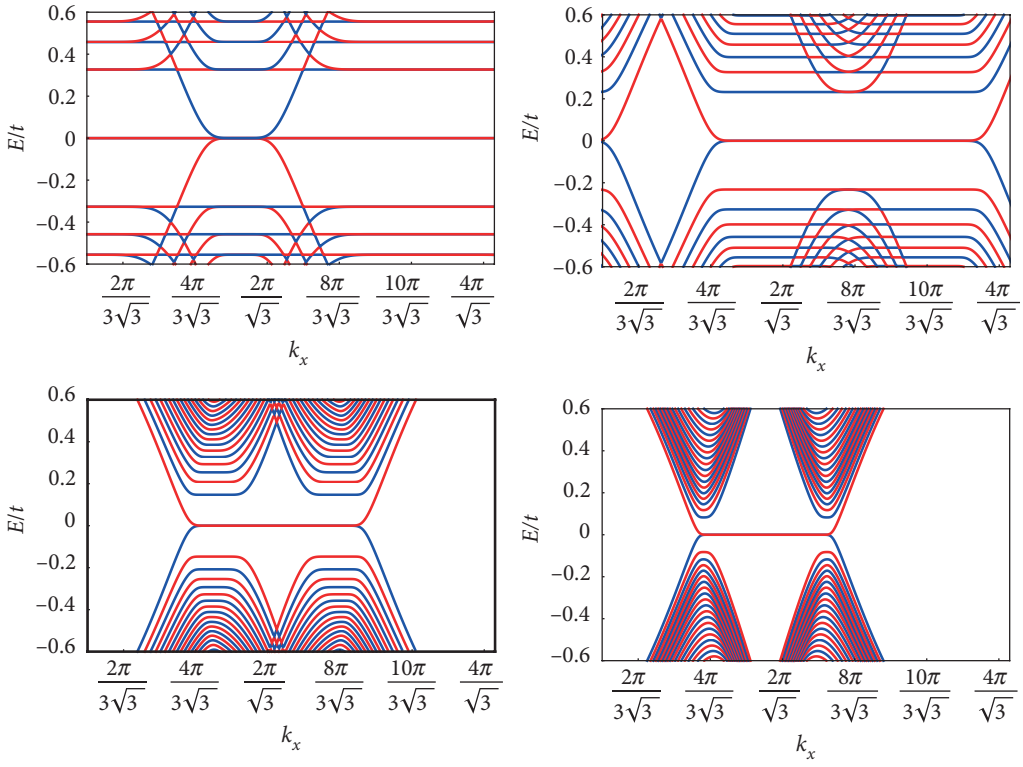


FIG. 2.20

Plot showing the Landau levels in graphene for different values of flux, ϕ. The values of the fluxes are $\phi = \frac{\Phi_0}{100}$ (for left upper panel), $\phi = \frac{\Phi_0}{200}$ (for right upper panel), $\phi = \frac{\Phi_0}{500}$ (for left lower panel) and $\phi = \frac{\Phi_0}{1600}$ (for right lower panel). Here $\Phi_0 = \frac{h}{e}$ is the flux quantum.

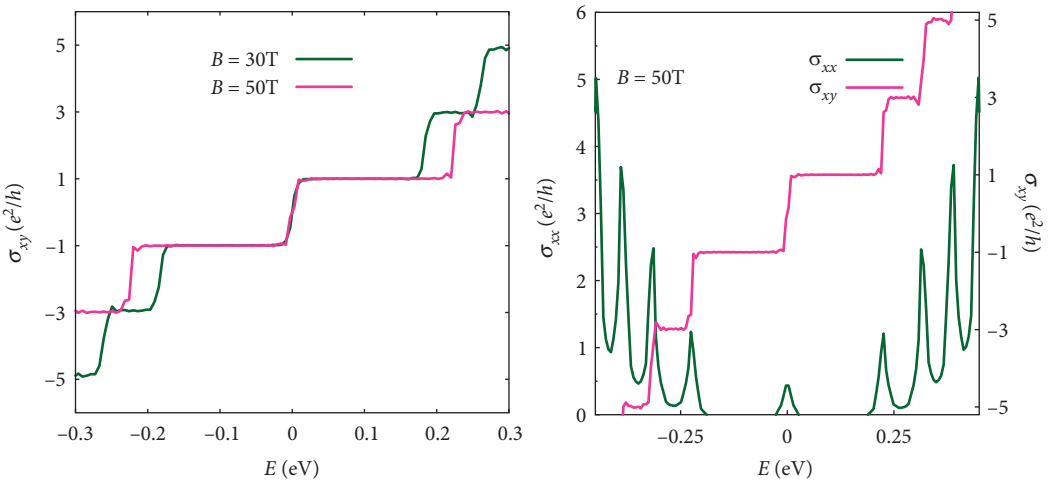


FIG. 2.21
Plot showing the variation of Hall conductivity as a function of the bias voltage for a particular value of the magnetic field, such as $B = 30$T and 50T. The plateaus in the Hall conductivity are clearly visible. The longitudinal conductivity (σ_{xx}) is shown for a specific value, namely, $B = 50$T.

of one of the wavevectors, k_x. As the magnitude of the flux is decreased, the width of the flat band appearing at the Fermi energy (at $E/t = 0$) decreases. Furthermore, the flat bands become dissipative in the bulk, corresponding to larger values of the Landau-level index, n and lower values of the flux ϕ. For a weak magnetic field, such that, $\Phi/\Phi_0 = 1/1600$, the energy bands of the bulk regain their Dirac-like structure similar to the single-particle energy levels, while the zero mode flat band continues to exist.

We have included numeric computation of the Hall conductivity as a function of the Fermi energy at comparable values (to those in experiments) of the magnetic field, namely, $B = 30$T and 50T using the Kubo formula in Fig. 2.21. The quantization of the Hall conductivity is clearly visible. Further, the longitudinal conductivity, σ_{xx} is shown for one of them, namely, 50T. σ_{xx} shows vanishingly small values corresponding to the plateaus of the Hall conductivities, while it shows spikes when the Hall conductivity jumps from one plateau to another.

2.4.5 Experimental observation of the Landau levels in graphene

There are primarily two experimental techniques for observing the existence of Landau levels. They are (i) the infra-red (IR) spectroscopy (Deacon *et al.*, 2007) and (ii) scanning tunneling microscopy

(STM) (Li *et al.*, 2009) experiments. In the following we include a brief discussion on each of them and their utilities in observing the non-equidistant Landau levels in graphene.

In IR spectroscopy, the optical transitions from one Landau level to another are studied by measuring the cyclotron frequencies. The Landau levels are proportional to $\sqrt{n}$ (n: Landau level index), all the frequencies of the optical transitions are distinct as the energy spacing between each pair of Landau levels is different than the other pairs. These optical transitions are of two types which correspond to transitions between the electron or the hole states in the conduction or the valence bands (intra-band transition), or the transitions between the electron and the hole states pertaining to the valence and conduction bands (inter-band transition) respectively. The photoconductive response and the resistive voltage show the existence of differently spaced Landau levels in graphene at particular values of the magnetic field, longitudinal current and the IR frequency. The photoconductive intensity as a function of the carrier density, n (not to be confused with the Landau level index) shows distinct peaks which are proportional to the energy absorbed from the incident IR radiation.

In the STM experiment, the specific energy levels can be identified by varying the bias voltage between the tip and the surface of the sample, and the tunneling current generated is proportional to the local density of states. In graphene, the Landau levels are directly observed via the peaks in the tunneling spectrum. From the positions of the peaks as a function of the sample bias (shown in Fig. 2.22), the energies of the Landau levels can be extracted.

Let us add a bit more details on the experimental results presented in Fig. 2.22. The differential conductance, $\frac{dI}{dV}$ is plotted as a function of the bias voltage (in mV) in (a). Within a proportionality constant, it yields the local DOS, which is seen to vanish at zero bias, while it remains finite for graphite. The latter clearly implies the opening of a gap at the Dirac points as a single-layer graphene is coupled to the substrate for the latter. Figure 2.22(b) shows the surface map of the low-energy tunneling conductance for two regions, one of which corresponds to single-layer graphene (darker region marked by "A"), while the other (lighter region marked by "B") denotes that of graphite. In (c) the tunneling spectra is shown as a function of the bias voltage for different values of the external field. These peaks correspond to the positions of the Landau levels. Finally, the massless nature of the Dirac fermions, and the Landau-level spacing varying as $\sqrt{n}$ are shown via plotting the energy, E as a function of $\sqrt{|n|B}$ which allows the collapse of the data onto a single straight line shown in Fig. 2.22(d). The corresponding slope yields a fairly good estimate of the velocity of the Dirac fermions ($\sim 0.8 \times 10^6$ ms^{-1}). Such a collapse is obviously absent for graphite.

Now we shall wrap up the discussion of the quantum Hall effect in graphene and give an introduction to the fractional quantum Hall effect (FQHE) in our subsequent discussion.

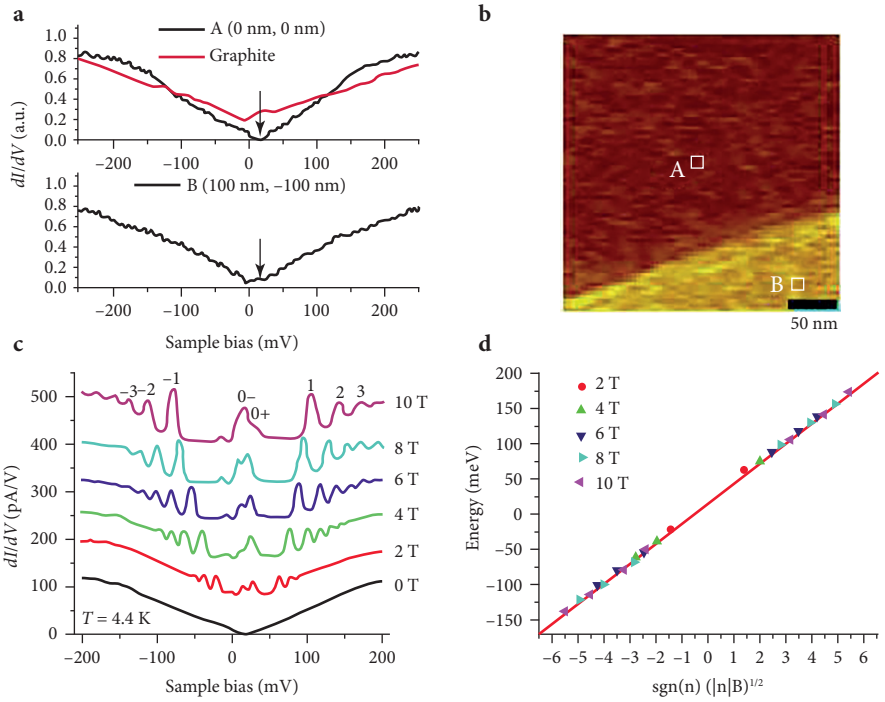


FIG. 2.22

Plot of the scanning tunneling spectra ($\frac{dI}{dV}$) for graphene. (a) shows spectra at zero field taken in the regions "A" and "B" marked by squares in (b). A single-layer graphene is shown by the black curve, which vanishes at zero voltage. The spectra for graphite is shown for comparison. (b) shows the map of $\frac{dI}{dV}$ at energy as marked by arrows in (a). $\frac{dI}{dV}$ vanishes in the dark region, but is finite in the bright region. (c) shows the field dependence of the tunneling spectra in the region "A." The peaks are labeled with LL index n. In (d) the energies of the Landau levels showing a square-root dependence on the level index, n (that is, $\sqrt{n}$) and for a few distinct field values. The symbols correspond to the peaks in (c). The figure is taken from Li *et al.* (2009).

2.5 FRACTIONAL QUANTUM HALL EFFECT

The fractional quantum Hall effect (FQHE) was discovered by Tsui, Stormer and Gossard in 1982 at Bell labs.[11] They observed that at very high magnetic fields, a 2DEG shows fractional quantization of the Hall conductance. In particular they got a quantized Hall plateau of magnitude $\rho_{xy} = \frac{3h}{e^2}$ which is accompanied by the vanishing of the longitudinal conductivity, ρ_{xx} at low temperature ($T < 5K$) in GaAs and AlGaAs samples. As opposed to the IQHE, where an integer number of Landau levels

[11] ΔE is called as the activation energy in the original paper by Novoselov *et al.* (2007).

are occupied, here in FQHE the Landau levels are partially occupied. If one makes the magnetic field large enough, the lowest Landau level will be partially filled. We can expect that the system will form some kind of a lattice, for example, a Wigner crystal or a charge density wave. Thus, it naively seems reasonable that the system would like to minimize its potential energy, since there is no kinetic energy left in the system corresponding to the lowest Landau level, and only trivial zero point energy is present in the system. Thus, the ions should tend to stay away from each other and thus would form something similar to a crystal lattice. However, surprisingly that does not happen, and instead the system becomes an incompressible quantum liquid, which has gaps in the energy spectrum at filling $1/m$ (m: odd, or a rational fraction of the form n/m). So it is inevitable that the system minimizes its energy by having gaps at fractional values of filling. The reason is that, owing to the presence of a large number of electrons (macroscopically degenerate in any of the Landau levels), a many-body interaction is induced which in fact makes the excitations above this incompressible ground state fractional. So in essence, the Hall current carries a fractional charge. In fact, the excitations are called abelian anions, which are neither fermions, nor bosons, if they are taken twice around a complete circle, they will pick up a phase which is either 0 or π. The phenomenology is put forward by Jain (2007) in terms of composite fermions, which we shall describe later.

One important thing that becomes apparent from the preceding discussion is that FQHE is impossible to explain by invoking the interaction between the electrons that eventually split the degeneracy of these enormously degenerate Landau levels, leading to the opening of a gap. The gap is in principle similar to the cyclotron gap, $\hbar\omega_c$ seen earlier in IQHE. However, this introduces another energy scale, leading to a hierarchy of energy scales, for example, the Coulomb energy scale, namely, $e^2/\epsilon a_0$ (the length scale in the denominator is taken as the Bohr radius) and the cyclotron energy. However, the kinetic degrees of freedom of the electrons are frozen and hence get eliminated from the problem. As said earlier, while the expectation is that of a trivial-insulating state, but at fractional fillings (less than 1), something more interesting happens.

Now consider the effect of the Coulomb interaction between the electrons,

$$V = \frac{e^2}{|\mathbf{r_i} - \mathbf{r_j}|} \tag{2.148}$$

This interaction should lift the degeneracy of the ground states. A degenerate perturbation theory seems may be an answer to this. However, degenerate perturbation theory problems are only solvable if the degree of degeneracy is such that it is analytically or numerically tractable. But here we are stuck with extraordinarily large matrices to diagonalize. Even numerically, with the best computational resources available nowadays, one cannot diagonalize for more than a few particles.

In the absence of these interactions, we have sharply defined Landau levels. In the presence of interactions, the degeneracy goes away and the Landau-level broadens, resulting in a spectrum of states width $\sim E_{Coulomb}$. Now the experimental data would be nicely explained, if a (tiny) gap exists for a filling fraction $\nu = \frac{1}{3}$ ($\nu = \frac{1}{3}$ is just an example), and the gap can exist at any of the rational fractions for which FQHE is seen. Of course, more prominent the plateau in the FQHE plot shown in Fig. 2.4, larger is the

gap. So we include some disorders that introduce localized states within the gap, which then give rise to plateaus in ρ_{xy} (and $\rho_{xy} = 0$). The whole description requires a hierarchy of energy scales, namely,

$$\hbar\omega_B \gg E_{Coulomb} \gg V_{disorder} \tag{2.149}$$

2.5.1 Electrons in the symmetric gauge

The vector potential $\mathbf{A}$ in the mixed gauge can be written as (again yields $\mathbf{B} = B\hat{\mathbf{z}}$),

$$\mathbf{A} = -\frac{1}{2}(\mathbf{r} \times \mathbf{B}) = \frac{1}{2}(-By\hat{\mathbf{x}} + Bx\hat{\mathbf{y}}) \tag{2.150}$$

The choice of the symmetric gauge breaks the translation symmetry in both x and y-directions. However, it preserves rotational symmetry about the origin. This of course means that angular momentum is a good quantum number. This is the most convenient gauge to study the fractional quantum Hall effect (FQHE). Furthermore, we can write down the non-canonical momentum as

$$\pi = \mathbf{p} + e\mathbf{A} = m\dot{\mathbf{r}}. \tag{2.151}$$

This is gauge invariant but non-canonical. One can use this to form the raising and lowering operators,

$$\begin{aligned} a &= \frac{1}{\sqrt{\frac{eBh}{\pi}}}(\pi_x - i\pi_y) = \sqrt{\frac{\pi}{heB}}(\pi_x - i\pi_y) \\ a^\dagger &= \sqrt{\frac{\pi}{heB}}(\pi_x + i\pi_y) \end{aligned} \tag{2.152}$$

where $2e\hbar B = 2e\frac{h}{2\pi}B = \frac{eBh}{\pi}$. Now,

$$\begin{aligned} [a, a^\dagger] &= aa^\dagger - a^\dagger a = \left(\frac{\pi}{heB}\right)(\pi_x - i\pi_y)(\pi_x + i\pi_y) - (\pi_x + i\pi_y)(\pi_x - i\pi_y) \\ &= \left(\frac{\pi}{eBh}\right)[\pi_x^2 + i\pi_x\pi_y - i\pi_y\pi_x + \pi_y^2 - \pi_x^2 + i\pi_x\pi_y - i\pi_y\pi_x - \pi_y^2] \\ &= \left(\frac{2\pi i}{eBh}\right)[\pi_x\pi_y - \pi_y\pi_x] \\ &= \left(\frac{i}{eB\hbar}\right)[\pi_x, \pi_y] \end{aligned} \tag{2.153}$$

$\boldsymbol{\pi}$ is not the canonical momenta in the sense, $[x_i, \pi_j] \neq \delta_{ij}$ and $[\pi_i, \pi_j] = \delta_{ij}$. However, they are gauge invariant. The numerical value of $\boldsymbol{\pi}$ does not depend upon the choice of gauge. This can be proved from the commutation relations,

$$\{p_i + eA_i, p_j + eA_j\} = -e\left(\frac{\partial A_j}{\partial x^i} - \frac{\partial A_i}{\partial x^j}\right) = -e\varepsilon_{ijk}B_k. \tag{2.154}$$

Thus, $[a, a^\dagger] = (\frac{i}{eB\hbar})(-ie\hbar B) = 1$

To explore whether the Landau levels yield the expected degeneracy, we can introduce another momentum variable, namely,

$$\tilde{\pi} = \pi - e\mathbf{A}$$

$\tilde{\pi}$ is not gauge invariant, and depends on the gauge potential chosen. The vector potential, $\mathbf{A}$ enjoys the gauge freedom, that is,

$$\mathbf{A}' = \mathbf{A} - \nabla\chi$$

where χ is an arbitrary scalar. This yields

$$\tilde{\pi}' = \mathbf{p}' - e(\mathbf{A}' - \nabla\chi) = \mathbf{p}' - e\mathbf{A}' + e\nabla\chi.$$

The commutation relation for $\tilde{\pi}'$ are,

$$[\tilde{\pi}'_x, \tilde{\pi}'_y] = ie\hbar B \tag{2.155}$$

which differ only by a sign with respect to the π momenta. This is also an advantage of this new momenta in that they obey,

$$[\tilde{\pi}_i, \tilde{\pi}_j] = 0.$$

Finally, the Hamiltonian is written in terms of the a and $a^\dagger$ operators as

$$\mathcal{H} = \hbar\omega_B\left(a^\dagger a + \frac{1}{2}\right) = \hbar\omega_B\left(n + \frac{1}{2}\right) \tag{2.156}$$

where $n = 0, 1, \ldots$ and denote the indices of the Landau levels.

Now, as a matter of exercise, in order to explore the degeneracy of the Landau levels, we introduce a second pair of raising and lowering operators, namely,

$$b = \frac{1}{\sqrt{2e\hbar B}}(\tilde{\pi}_x + i\tilde{\pi}_y) \tag{2.157}$$

$$b^\dagger = \frac{1}{\sqrt{2e\hbar B}}(\tilde{\pi}_x - i\tilde{\pi}_y). \tag{2.158}$$

They also obey $[b, b^\dagger] = 1$. These b, $b^\dagger$ will yield the degeneracy of the Landau levels, as shown in the following. Thus, a general state in the Hilbert space $|n, m\rangle$ is defined by

$$|n, m\rangle = \frac{(a^\dagger)^n (b^\dagger)^m}{\sqrt{n!m!}}\,|0, 0\rangle \tag{2.159}$$

where $a\,|0, 0\rangle = b\,|0, 0\rangle = 0$ and $\mathcal{H} = \frac{1}{2m}\pi \cdot \pi = \frac{1}{2m}(\mathbf{p} + e\mathbf{A})^2$.

Let's now construct the wavefunction in the symmetric gauge. We are going to focus on the lowest Landau level, $n = 0$ as it is of primary interest for discussing FQHE. The trick is to convert the definition of a into a differential equation,

$$a = \frac{1}{\sqrt{2e\hbar B}}(\pi_x - i\pi_y) = \frac{1}{\sqrt{2e\hbar B}}[p_x - ip_y + e(A_x - iA_y)] \tag{2.160}$$

$$= \frac{1}{\sqrt{2e\hbar B}}\left[-i\hbar\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right) + \frac{eB}{2}(-y - ix)\right] \tag{2.161}$$

using $z = x - iy$ and $\tilde{z} = x + iy$. Remember this is not usually how we define z and z^* (or $\tilde{z}$), however, we shall stick to this definition. Also, define,

$$\partial = \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right) \tag{2.162}$$

and,

$$\tilde{\partial} = \frac{1}{2}\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right) \tag{2.163}$$

which obey $\partial z = \tilde{\partial}\tilde{z} = 1$ and $\partial\tilde{z} = \tilde{\partial}z = 0$.

$$\partial z = \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right)(x - iy) = \frac{1}{2}(1 + 1) = 1.$$

So a and $a^\dagger$ in terms of the coordinates z can be written as

$$a = -i\sqrt{2}\left(l_B\tilde{\partial} + \frac{z}{4l_B}\right); \quad a^\dagger = -i\sqrt{2}\left(l_B\tilde{\partial} - \frac{\tilde{z}}{4l_B}\right). \tag{2.164}$$

Now, the lowest Landau level is found by the one which is annihilated by the operator a.

$$a\,|0, m\rangle = 0$$

$$-i\sqrt{2}\left(l_B\tilde{\partial} + \frac{z}{4l_B}\right)|0, m\rangle = 0$$

$|0, m\rangle$ is called $\psi_{\mathrm{LLL}}(z, \tilde{z})$, where LLL stands for lowest Landau level.

$\psi_{\mathrm{LLL},m=0} \simeq e^{-|z|^2/4l_B^2}$

The ground state is known to be a Gaussian for a linear harmonic oscillator.[12]

One can construct the higher Landau level wavefunctions by employing $b^\dagger$ successively to the $m = 0$ state. This yields,

$$\psi_{\mathrm{LLL},m} \simeq \left(\frac{z}{l_B}\right)^m e^{-|z|^2/4l_B^2}. \tag{2.165}$$

It is straightforward to ascertain that $\psi_{\mathrm{LLL},m}$ are eigenfunctions of J_z, defined by

$$J_z = \hbar(z\partial - \tilde{z}\tilde{\partial}) \tag{2.166}$$

and obey, $J_z\psi_{\mathrm{LLL},m} = m\hbar\psi_{\mathrm{LLL},m}$.

Let us explore the degeneracy associated with the Landau levels $\psi_{\mathrm{LLL},m}$, which is obtained by noting that the wavefunction with angular momentum, m is peaked on a circular ring of radius, $r = l_B\sqrt{2m}$. The number of states in an area, $A = \pi R^2$ is $\mathcal{N} = \pi R^2/\pi r^2 \simeq eBA/2\pi\hbar$ which is a result we have seen earlier.

[12] $(y + \frac{\partial}{\partial y})u(y) = 0$, or $\frac{\partial u}{u} = -ydy$. The solution is $u = e^{-y^2/2}$.

A quick recap on the ongoing discussion reveals that for a constant magnetic field B, the vector potential can be obtained as

$$\mathbf{A} = \frac{1}{2}(\mathbf{r} \times \mathbf{B}) = \frac{B}{2}(-y\hat{\mathbf{x}}, x\hat{\mathbf{y}}, 0). \tag{2.167}$$

Writing down the free particle Hamiltonian $\mathcal{H} = \frac{1}{2m}(\mathbf{p} + e\mathbf{A})^2$ in the above gauge,

$$\mathcal{H} = \frac{1}{2l_B^2}\left[\left(-i\frac{\partial}{\partial x} - \frac{y}{2}\right)^2 + \left(-i\frac{\partial}{\partial y} + \frac{x}{2}\right)^2\right]. \tag{2.168}$$

It is convenient to introduce the complex variables z and z^* via,

$$z = x - iy = re^{-i\theta}, \quad z^* = x + iy = re^{i\theta}$$

The derivatives can be written as[13]

$$\frac{\partial}{\partial x} = \frac{\partial}{\partial z} + \frac{\partial}{\partial z^*}, \quad \frac{\partial}{\partial y} = -i\left(\frac{\partial}{\partial z} - \frac{\partial}{\partial z^*}\right)$$

In terms of z and z^*, the Hamiltonian reads as

$$\mathcal{H} = \frac{1}{2l_B^2}\left[\frac{1}{4}|z|^2 - 4\frac{\partial^2}{\partial z \partial z^*} - z\frac{\partial}{\partial z} + z^*\frac{\partial}{\partial z^*}\right]. \tag{2.169}$$

The Hamiltonian has few similarities with that of the harmonic oscillator, especially because of the mixed second derivative and the two first derivatives. In order to solve the Hamiltonian, a set of ladder operators can be introduced, namely,

$$\begin{aligned} b &= \frac{1}{\sqrt{2}}\left(\frac{z^*}{2} + 2\frac{\partial}{\partial z}\right), \quad b^\dagger = \frac{1}{\sqrt{2}}\left(\frac{z}{2} - 2\frac{\partial}{\partial z^*}\right) \\ a &= \frac{1}{\sqrt{2}}\left(\frac{z}{2} + 2\frac{\partial}{\partial z^*}\right), \quad a^\dagger = \frac{1}{\sqrt{2}}\left(\frac{z^*}{2} - 2\frac{\partial}{\partial z}\right). \end{aligned} \tag{2.170}$$

These operators obey

$$[a, a^\dagger] = [b, b^\dagger] = 1 \tag{2.171}$$

with all other commutators vanish. This facilitates writing the Hamiltonian in a familiar form,

$$\mathcal{H} = a^\dagger a + \frac{1}{2}. \tag{2.172}$$

The eigenvalue of $a^\dagger a$ denotes the Landau level index n.

Let us now concentrate on the b-operators. They play a role to play in writing down the z-component of the angular momentum, namely, J_z, where

$$J_z = -i\hbar\frac{\partial}{\partial \phi} = -\hbar\left(z\frac{\partial}{\partial z} - z^*\frac{\partial}{\partial z^*}\right) = a^\dagger a - b^\dagger b \tag{2.173}$$

[13] Somewhat non-trivial definitions of z and z^* are adopted to make sure of the analyticity of the wavefunction in z.

The eigenvalue of J_z is $-m\hbar$ where m takes values from $-n$ to $+n$, n being the Landau level (LL) index. $b^\dagger(b)$ increases (decreases) the value of m by one unit, while keeping the LL index unchanged. However, $a^\dagger(a)$ increases (decreases) n and decreases (increases) m by one unit.

Clearly, the harmonic oscillator problem here has two indices, namely, n and m, which is apparent through,

$$\mathcal{H}|n,m\rangle = E_n\,|n,m\rangle \tag{2.174}$$

where

$$|n,m\rangle = \frac{(b^\dagger)^{n+m}}{\sqrt{(n+m)!}}\frac{(a^\dagger)^n}{\sqrt{n!}}\,|0,0\rangle$$

where $|0,0\rangle$ denotes a Gaussian,

$$|0,0\rangle = \frac{1}{\sqrt{2\pi}}e^{-|z|^2/4l_B^2}$$

and obeys,

$$a\,|0,0\rangle = 0 = b\,|0,0\rangle.$$

One can generate the family of the lowest Landau levels by successively operating $b^\dagger$ on $|0,0\rangle$, that is,

$$(b^\dagger)^m\,|0,0\rangle = |0,m\rangle = \frac{z^m e^{-|z|^2/4l_B^2}}{\sqrt{2\pi\,2^m m!}}$$

$\sqrt{2\pi\,2^m m!}$ appears the normalization constant and z^m is a polynomial in z, where each z appears due to acting $b^\dagger$ on $|0,0\rangle$ each time.

It can be checked that $|0,m\rangle$ is an analytic function of z, since

$$\frac{\partial}{\partial z^*}\,|0,m\rangle = 0,$$

which is evident from the definition of $\frac{\partial}{\partial z^*}(=\frac{1}{2}(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}))$ as the Cauchy-Riemann condition is satisfied.

Now getting a Landau level with arbitrary indices, namely, $|n,m\rangle$, we need to simultaneously act $a^\dagger$ n times and $b^\dagger$ $(n+m)$ times. This is necessary because $a^\dagger$ reduces the m index by one unit each time. Thus,

$$\begin{aligned}|n,m\rangle &= \frac{1}{2\pi\,2^{m+2n}n!(n+m)!}\left(a^\dagger\right)^n\left(b^\dagger\right)^{n+m}e^{-|z|^2/4l_B^2}\\ &= \frac{1}{2\pi\,2^{m+2n}n!(n+m)!}\left(z^* - 2\frac{\partial}{\partial z^*}\right)^n\left(z - 2\frac{\partial}{\partial z}\right)^{m+n}e^{-|z|^2/4l_B^2}.\end{aligned} \tag{2.175}$$

It has a somewhat complicated form, but at least denotes an expression which can be evaluated for a given value of n and m.

Let us more closely discuss the physics of these fractional quantum Hall states. The first approach to the fractional quantum Hall states was due to Laughlin, who described the filling fractions given by $\nu = 1/m$, (m is an odd integer). Since the resultant matrix was too difficult to diagonalize, so Laughlin wrote down the answer in the following sense. He wrote down the wavefunction by intuition that preserves the physical properties and the symmetries of the system. To understand the Laughlin wavefunction, let us consider only two particles at the lowest Landau level. Consider a potential, $V = V(|\mathbf{r_1} - \mathbf{r_2}|)$. For such a potential, the wavefunction is an eigenstate of the angular momentum (recall the H-atom problem). In order for the angular momentum basis to be used, we need a symmetric gauge, namely,

$$\mathbf{A} = -\frac{1}{2}(\mathbf{r} \times \mathbf{B}) = -\frac{By}{2}\hat{\mathbf{x}} + \frac{Bx}{2}\hat{\mathbf{y}}. \tag{2.176}$$

The choice of the gauge breaks the translation symmetry in both the x and the y direction, however it preserves rotational symmetry about the origin. This means that angular momentum is a good quantum number and hence justifies using the angular momentum basis.

2.5.2 The lowest Landau level (LLL)

The (unnormalized) single-particle wave function in the lowest Landau level takes the form,

$$\psi_m = z^m e^{-|z|^2/4l_B^2} \tag{2.177}$$

with $z = x - iy$. These states are located on a ring of radius $r = l_B\sqrt{2m}$. The exponent m labels the angular momentum. The largest value of m for which the state falls inside the ring is given by $m = R^2/2l_B^2$, R being the corresponding value of the radius of the ring. m now denotes the total number of eigenstates in the lowest Landau level that falls inside the ring. That is, $0 \leq m \leq N_\Phi$, where $N_\Phi = \frac{A}{\Phi_0}$ is the number of flux quantum. Hence, the degeneracy per unit area is given by $\frac{1}{2\pi l_B^2} = \frac{eB}{h} = \frac{B}{h/e}$. As long as we neglect mixing between the successive Landau levels, a condition that should be valid for $V \ll \hbar\omega_B$ then a two particle eigenstate takes its form,

$$\psi = (z_1 + z_2)^M (z_1 - z_2)^m e^{-(|z_1|^2 + |z_2|^2)/4l_B^2} \tag{2.178}$$

where M is the center of mass angular momentum, m is the relative angular momentum. M and m are non-negative integers (see discussion below), which make the pre-factor a polynomial in $z_1 + z_2$ and $z_1 - z_2$. Hence, a state formed out of the linear combinations of ψ_m in Eq. (2.178), and hence lies in the lowest Landau level. Of course, m must be odd to yield an antisymmetric wavefunction.

Now we can write down the wavefunction without explicitly solving the Schrödinger equation for any general potential of the form $V(\mathbf{r})$ (in principle, it should be an unsolvable problem). We remind ourselves that we are in the lowest Landau level that allows us to perform the above form of the wavefunction. For N-particles, the many-body state can be written as

$$\psi(z_1, z_2, \ldots, z_N) = f(z_1, z_2, \ldots, z_N) e^{-\sum_{i=1}^{N} |z_i|^2/4l_B^2} \tag{2.179}$$

where $f(z)$ is a polynomial in z_i and contains the maximum power of any z_i occurring in N_Φ. It also takes care of the statistics, that is, for interchanging $z_i \leftrightarrow z_j$, $f(z)$ picks up a negative sign.

With all the points discussed above, Laughlin made a suggestion for the ground state wavefunction (Laughlin, 1983) as

$$\psi(z_1, z_2, z_3 \cdots z_N) = \prod_{i<j}(z_i - z_j)^m e^{-\sum_{i=1}^{N}|z_i|^2/4l_B^2}. \tag{2.180}$$

So the symmetry function is fixed as $(z_i - z_j)^m$. The above form makes sense as it keeps the electrons apart and thus reduces the Coulomb interaction (or the potential energy, as discussed earlier). This is the celebrated Laughlin wavefunction of FQHE. This can be proved easily by writing,

$$J_z|\psi_m\rangle = i\hbar\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right)|\psi_m\rangle = m\hbar|\psi_m\rangle \tag{2.181}$$

If m is an odd integer, for example, 1, 3, 5 etc. then f is an antisymmetric function.

The pre-factor vanishes of order m if two electrons happen to come close. Meanwhile, the exponential factor vanishes as i and j get far away from the origin. The angular momentum of these states is $m\hbar$. For the lowest Landau level, there is only one state for a given value of m, and it should only be a positive value of m. The exclusion of the negative values is an artefact of the chirality created by the presence of the magnetic field.

It is worth comparing this with the scenario corresponding to the Landau gauge, which had discrete levels as the basis states. However, physical properties, such as the density of states (per unit area) should be independent of the choice of the gauge. It can be tested by taking the mod square of Eq. (2.177). The maximum value of this probability density occurs at $R = \sqrt{2ml_B^2}$. Thus, the area of a circle given by $\pi R^2 (= 2m\pi l_B^2)$ contains m flux quanta, which yields the familiar result of one state per Landau level per flux quantum piercing the quantum Hall sample.

Let us show that this wavefunction has the desired filling fraction, which is an important aspect of the Laughlin state. The exponent m is not arbitrary and is related to the filling fraction. For this purpose, let us focus on the wavefunction for a single-particle coordinate, say, z_1. The terms that depend upon z_1 in the pre-factor can be written as

$$\prod_{i<j}(z_i - z_j)^m \sim \prod_{i=2}^{N}(z_1 - z_i)^m \tag{2.182}$$

$$= (z_1 - z_2)^m (z_1 - z_3)^m \cdots (z_1 - z_N)^m \tag{2.183}$$

Thus, there are $m(N-1)$ powers of z_1. Hence, the maximum angular momentum of the first particle is $m(N-1)$. So the maximum extent of the wavefunction is given by the radius, $R \approx \sqrt{2mN}l_B$, where $(N-1)$ is approximated by N for large N. Correspondingly, the area over which it spans is given by $\pi R^2 \approx 2\pi mNl_B^2$.

Now, recall that the number of states in the filled landau level is $AB/\Phi_0 \approx A/2\pi l_B^2$ where B/Φ_0 is the inverse of the area. Thus putting $A = \pi R^2 = 2\pi mNl_B^2$, $l_B = \sqrt{\frac{\hbar}{eB}}$, the total number of states is

given by

$$\frac{2\pi mNl_B^2}{2\pi l_B^2} = mN \tag{2.184}$$

Since the total number of states is mN, the filling fraction is $\frac{1}{m}$ (m: odd) as we have discussed earlier. An exact diagonalization of the Hamiltonian matrix for a small number of particles shows that the Laughlin wavefunction is extremely accurate.

Let us revisit the above discussion in a slightly different language. The total angular momentum is the sum of the angular momentum of individual electrons. So the total angular momentum carried by the Laughlin state is as follows.

A typical term in the Laughlin wave function is

$$z_1^0 z_2^m z_3^{2m} \cdots z_N^{(N-1)m} e^{-\sum_{i=1}^N |z_i^2|}$$

where N denotes the number of electrons. Since the above state has angular momentum $m\hbar$ for the individual electrons, so the total angular momentum is

$$|\mathbf{J}_{tot}| = m\hbar \sum_{n=0}^{N-1} n = \frac{(N-1)Nm}{2}\hbar$$

All other contributions will yield the same angular momentum.

On the other hand, the maximum angular momentum, namely, $n_{max}\hbar$ that an electron can have is given by the maximum power of the variable z in the wavefunction. Here,

$$n_{max} = (N-1)m.$$

Consider the form of the lowest Landau level.

$$\psi_m = z^m e^{-|z|^2/4l_B^2}$$

The probability of finding the electron at a given z is given by $|\psi(z)|^2$ (or $|\psi(r)|^2$). This quantity has a sharp peak at $r = r_m$.

$$r_m = \sqrt{2m}l_B$$

Thus, the area over which the Landau levels are located is given by

$$\begin{aligned} A &= \pi r_{max}^2 = 2\pi n_{max} l_B^2 \\ &= 2\pi m(N-1)l_B^2 \end{aligned} \tag{2.185}$$

Thus, the Laughlin state is realized at the filling fraction

$$\nu = \frac{N\Delta A}{A} = \frac{N}{m(N-1)} = \frac{1}{m}$$

for large N.

Another important feature of the Laughlin wavefunction is that the argument of the Gaussian can be written as

$$\exp\left(-\sum_i z_i^2\right) = \exp\left(-\sum_{i\neq j} |z_i - z_j|^2 + \sum_i z_i^2\right).$$

By writing $\tilde{z}^2 = \sum_i z_i^2$, we note that $\tilde{z}$ is the coordinate of the center of mass. Thus, apart from this trivial factor in the exponent, the wavefunction depends on the relative (complex) coordinates, namely, $|z_i - z_j|$, which implies that the wavefunction is uncorrelated and similar to the wavefunction seen for the integer version of the QHE.

This brings us to the most potential issue: whether the much talked about Laughlin wavefunction yields the Hall conductivity quantized as $\nu\frac{e^2}{h}$ with $\nu = \frac{1}{m}$, m is an odd integer? Furthermore, what does a fractional coefficient of $\frac{e^2}{h}$ exactly mean in terms of fractionalizing the unit of charge?

2.5.3 Fractional charge and the Hall conductivity

The existence of the fractional electron charge is surprising, as electrons are indivisible objects. Also, and probably more importantly, how does the Laughlin wavefunction produce the Hall conductivity with plateaus at rational fractions? Consider the Corbino disc argument applicable to the IQHE (see Fig. 2.12). In order to introduce a connection between the Corbino disc model and the Laughlin wavefunction written above, we consider enhancing the magnetic field B in a controlled manner so as to introduce one flux quantum, Φ_0 in any region of the disc. The questions are that: how does that modify the Laughlin wave function? Recall that the pre-factor of the Gaussian in the Laughlin wavefunction is given by z^m (see Eq. (2.177)), where the index m counts the number of flux quantum. Thus, with the above modification, "m" increases "$m+1$," which makes the maximum power of z_i increase from $m_{max} = Nm$ to $Nm+1$. The situation to the Laughlin wavefunction is incorporated by introducing,

$$\psi' = \left[\Pi_{i=1}^N z_i\right]\psi_0(z_1, z_2 \cdots z_N) \tag{2.186}$$

where $\psi_0(z_1, z_2 \cdots z_N)$ is the usual Laughlin state. The above wavefunction ψ' dismisses the origin to be a special point about which the original Laughlin state was centered (or equivalently, the single-particle density had a peak at the origin). Since the origin, (z_0 is the origin) does not play an important role, we are allowed to write

$$\psi'(z_0) = \left[\Pi_{i=1}^N (z_i - z_0)\right]\psi_0(z_1, z_2 \cdots z_N), \tag{2.187}$$

assuming z_0 is not within a distance l_B from the edge of the disc, so that the single particle density remains uniform up to a distance l_B relative to the edge of the disc. In effect, the probability of finding an electron at the origin is missing and its density within an area l_B^2 about the origin is reduced. In the basic language of solid state physics, a missing electron is equivalent to the appearance of a "hole" which is what precisely happens here.

Now the magnetic field B is increased such that m flux quanta are added in the process and m holes are created. This yields the Laughlin state to assume the form,

$$\psi_m = \Pi_{i=1}^{N}(z_i - z_0)^m \psi_0(z_1, z_2 \cdots z_N) = \Pi_{i=1}^{N+1}(z_i - z_0)^m \exp\left(-\sum_{i=1}^{N+1} \frac{z_i^2}{4l_B^2}\right) \tag{2.188}$$

The above form precisely coincides with the case of $N+1$ electrons and consequently corresponds to $m(N+1)$ flux quanta. Thus, the addition of an extra electron compensates for the m added holes. Hence, the charge of the hole is

$$e_h = -\frac{e}{m}. \tag{2.189}$$

This accounts for the fractional charge. But at the same time, this means that the probability of finding an electron near the origin is reduced by $1/m$. This probability helps us to reconcile the very idea of the Corbino disc, that is transfer the "fractional" charge from the center to the edge of the disc.

We rerun the same argument as done for the IQHE. For this argument to be valid, the geometry of the ring is important. Here, in addition to the background magnetic field **B** that threads the sample, we can thread an additional flux Φ through the center of the ring. This Φ can affect the quantum state of the electrons.

Let us first see what this flux Φ has got to do with the Hall conductivity. Suppose we slowly increase Φ from 0 to $\Phi_0(=\frac{h}{e})$ *i.e.*, within a time $t_0 \gg \frac{1}{\omega_B}$. This induces an emf around the annular region $\varepsilon = -\frac{\partial\Phi}{\partial t} = \frac{-\Phi_0}{t_0}$. The purpose of this emf is to transport "n" electrons from the inner circumference to the outer circumference. This would result in a current in the radial direction, $I_r = -ne/t_0$. Thus, the Hall resistivity is

$$\rho_{xy} = \frac{\varepsilon}{I_r} = -\frac{\Phi_0}{t_0} \cdot \frac{t_0}{(-ne)} = \frac{h}{e^2} \cdot \frac{1}{n} \tag{2.190}$$

The same arguments hold equally for the IQHE and FQHE, in the former n is an integer, while n is a fraction for the latter. In FQHE, the interpretation is as follows: as we increase the flux from Φ to Φ_0, a charge of e/m is transported from the inner circumference to the outer one when the flux is increased by Φ_0 units. The resultant Hall conductivity (or equivalently the resistivity) becomes,

$$\sigma_{xy} = \frac{e^2}{h} \cdot \frac{1}{m}. \tag{2.191}$$

Thus, a whole electron is transferred only when the flux is increased by $m\Phi_0$ units.

Shot noise measures of the fractional plateaus indeed confirm the existence of fractional charge. In the experimental setup (Glattli *et al.*, 2000), the two energy modes at opposite edges of the Hall sample are coupled by a quantum point contact, which facilitates a flow of current between the two edge channels. The random transfer of charges yields fluctuations (noise) in the current. In the weak coupling regime, the noise intensity is proportional to the backscattered current and the (fractional) charge. Thus, the intensity of the shot noise experiments detect fractional electronic charges.

2.5.4 Fractional Hall fluid and the plasma

The variational wavefunction corresponding to the lowest Landau level can be written as

$$\psi_m(z_1 \cdots z_N) = \Pi_{j<k}^N (z_j - z_k)^m e^{-\sum_j |z_j|^2/4l_B^2} \tag{2.192}$$

The wavefunction is applicable to a filling fraction $\nu = \frac{1}{m}$, (m is an odd integer) such as $\nu = \frac{1}{3}$ (for $m = 3$). However, if only $1/m$ fractions were found in experiments, life would have been simpler. Unfortunately, many other fractions were found, including the improper ones.

Let us enumerate a few properties of the wavefunctions.

i. It is antisymmetric with respect to the swapping of coordinates of the fermions ($z_j \leftrightarrow z_k$),

   $\psi_m(z_1 \ldots . z_j, z_k \ldots . z_N) = -\psi_m(z_1 \ldots . z_k, z_j \ldots . z_N)$

   In order for this to be valid, the index m must be odd.

ii. A surprising fact is revealed when the unnormalized probability density of the Laughlin states is computed. That is,

   $|\psi_m(z)|^2 = |\Pi_{j<k}^N (z_j - z_k)^m e^{-\sum_j |z_j|^2/4l_B^2}|^2$.

   The two terms (the Jastrow factor and the Gaussian) in the RHS behave differently. That is, prefactor (of the Gaussian) or the Jastrow factor tries to keep the fermions away, and it grows larger as they move further away. However, the exponential term shrinks as the fermions spread out. Under this competing scenario, can it ensure a uniform density?

There is an answer to the rather complicated problem, and again it is due to Laughlin, via an analogy with classical plasma, albeit the Hall fluid is at a very low temperature. The norm of the wavefunction is written as

$$|\psi_m(z)|^2 = e^{-\beta V_{\text{plasma}}}, \tag{2.193}$$

where

$$V_{\text{plasma}} = 2m^2 \sum_{j<k} ln|z_j - z_k| + \frac{m}{2l_B^2} \sum_j |z_j|^2. \tag{2.194}$$

β is identified as $1/m$. So Eq. (2.193) yields the density of the plasma. It should be remembered that it is only an analogy where a classical plasma constitutes of particles with charge m in a uniform (neutral) background. The existence of a plasma-like state at a very low temperature should be contradictory to each other, and thus suggests of a liquid phase of fermions. It would have been a crystalline state (for example, a Wigner crystal) at large values of the charge "m." It is indeed a new state of matter, and is often denoted as the Laughlin state.

In order to understand the potential V_{plasma}, it can be noted that the electric field, $\mathbf{E(r)}$ and the potential, $\phi(\mathbf{r})$ due to a point charge q are given by

$$\mathbf{E(r)} = \frac{q\mathbf{r}}{r^2}, \text{ and } \phi(\mathbf{r}) = -qln(r) \tag{2.195}$$

These yield the Laplace's equation in 2D as

$$\nabla \cdot \mathbf{E} = -\nabla^2 \phi(\mathbf{r}) = 2\pi q \delta^2(\mathbf{r}), \tag{2.196}$$

where $\delta^2(\mathbf{r})$ is the two-dimensional Dirac delta function. The first term in Eq. (2.194) is explained by the logarithmic dependence of the potential, which yields

$$V^{(1)}_{\text{plasma}} = m^2 \sum_{j<k} \left(-ln\,|z_j - z_k|\right). \tag{2.197}$$

We are missing a factor of "2" here, but that can be absorbed in the definition of β (β can be redefined as $2/m$). The second term can be understood by noting that,

$$\nabla^2 \frac{|z|^2}{4} = \frac{1}{l_B^2}, \text{ or, } \rho = -\frac{1}{2\pi l_B^2}, \tag{2.198}$$

where the charge density ρ satisfies the Poisson equation with a potential,

$$V^{(2)\text{plasma}} = \frac{|z|^2}{4}. \tag{2.199}$$

This term denotes the energy of "m" charges interacting with the negative charge density. It is obvious that $2\pi l_B^2$ is the area that contains one quantum of flux ($\Phi_0 = h/e$), which makes the background charge density to be B/Φ_0, which also denotes the density of flux in the unit of the flux quantum.

While the Laughlin wavefunction demonstrates a lot of merit, one pertinent question remains. How good is the wavefunction in practical cases? To enumerate its success, the overlap between the Laughlin state and the exact state (obtained via exact diagonalization) for 3 particles, and a few representative (odd) values of m and corresponding to a few different forms of the potential $V(\mathbf{r})$ (for example, $V(\mathbf{r}) = 1/r, -ln\,r, e^{-r^2/2}$ etc) are obtained. The overlap between the two results is very close to 100%, and thus ensures that the proposed wavefunction is indeed a good one (Laughlin, 1983).

2.5.5 Fractional statistics

The occurrence of quantization of the Hall plateaus at fractional filling in the expression bears a testimony for the quantization of the fractional charge. Ironically, the fractionally charged quasiparticles being localized at the quantum Hall plateaus, do not conduct and the contribution to the Hall current is due to the background (neutral) state, which is incompressible and does not contain any quasiparticles. Nevertheless, the fractional charge requires particle statistics just as the integral (including zero) charge particles. It should be kept in mind that the particle statistics show up in the form of collective phenomena, such as the formation of a Bose Einstein condensate, or a degenerate Fermi gas.

In this context, we may recall that we have been exposed to the concept of identical particles, which are either Bosons or Fermions and obey the properties of symmetry or antisymmetry, respectively, under the exchange of particles pairwise. That is, for the two particles,

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \pm\psi(\mathbf{r}_2, \mathbf{r}_1), \tag{2.200}$$

where the + sign refers to Bosons and the − sign is applicable to Fermions. However, it is also true that probabilities are the same, namely,

$$|\psi(\mathbf{r}_1, \mathbf{r}_2)|^2 = |\psi(\mathbf{r}_2, \mathbf{r}_1)|^2 . \tag{2.201}$$

Such that the wave functions, upon exchange of particles, at most pick up a phase,

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = e^{i\pi\alpha}\psi(\mathbf{r}_2, \mathbf{r}_1). \tag{2.202}$$

Repeating the exchange brings back the same state, which implies $e^{2\pi i\alpha} = 1$ where $\alpha = 0$ for Bosons and $\alpha = 1$ for Fermions. There is a subtle point here, which is not explicitly stated. In 3 dimensions, a rotation by 2π brings us back to the original state which should be equivalent to changing a pair of particles twice. Such exchanges (or equivalently a rotation by 2π) are continuously, also known as world lines, connected and do not cross each other's path. However, in two dimensions, there is a big difference where such paths cross and wind around each other. For example, there is a distinction between a clockwise and an anti-clockwise exchange of particles, though in either case, their paths get tangled and form braids. A braid describes a pattern formed by the interaction of two or more strands of wire (or hair). Thus, a clockwise braid is distinct from an anti-clockwise braid and are said to belong to different topological sectors. A tangled path implies an arbitrary phase involved in the exchange of particles because the paths cross, there arises an ambiguity in the phase and hence α can assume any value in the interval $[0 : 1]$. More concretely, say, for an anti-clockwise exchange,

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = e^{i\pi\alpha}\psi(\mathbf{r}_2, \mathbf{r}_1). \tag{2.203}$$

A clockwise exchange results in

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = e^{-i\pi\alpha}\psi(\mathbf{r}_2, \mathbf{r}_1). \tag{2.204}$$

These particles are known as anyons owing to their allegiance to "any" statistics that interpolates between bosons and fermions. "Any" statistics here refers to fractional statistics and apply to the quasi-particles (or more aptly called as the quasiholes). The charge of the quasiparticles or the anyons is fractional.

Without going into elaborate calculations, we refer to published works by D. Arovas, J. Schrieffer and F. Wilczek (Phys. Rev. Lett. **53**, 772 (1984)) and D. Tong in "The Quantum Hall Effect" (TIFR Infosys lectures), the Berry phase around a closed loop for those quasiparticles can be obtained as

$$\Phi_B = \frac{e\Phi}{mh} = \frac{\Phi}{m\Phi_0} \quad (\text{m} : \text{odd}) \tag{2.205}$$

where Φ is the flux enclosed by a quasiparticle around a closed contour and $\Phi_0 =$ the flux quantum (h/e). It also has the interpretation that

$$\Phi_B = \frac{e^*\Phi}{h}, \tag{2.206}$$

where e^* refers to a fractional charge.

2.6 COMPOSITE FERMIONS

To understand the concept of composite fermions, let us review the concept of vortices. In a reasonably different sense than the supercomputers (where magnetic flux lines puncturing a superconductor creates a vortex in a type-II material where gradient in the phase of a condensate wave function produces circulating current that is detectable in experiments), vortices in FQHE can be understood as follows. A complex number, $z\ (= re^{i\theta})$ has a vortex at the origin, implying that a complete circle around the origin creates a phase change of 2π. Similarly, the FQHE wave function contains $(z - z_0)^{2p}$ which implies particle "1" sees $2p$ vortices to be carried by particle "2" and vice versa. Thus, every particle can be imagined to carry an even number of vortices. The particles carrying the vortices are certainly not real. There are no real magnetic flux quanta attached to them. However, the intuitive picture leads us to the concept of composite Fermions (CF) (Jain, 2007) and their relevance to the fractional statistics (Leinaas *et al.*, 1977; and Wilczek, 1982).

A composite fermion is defined as the bound state of a fermion and an even number of vortices or equivalently flux quanta (remember the equivalence of flux quantum and vortices come from the fundamental fact they are topologically similar in the sense they both result in phase change of 2π upon circling about a closed path around it). The picture portrayed above is only for convenience of the visualization of a physical scenario, however in reality, there are no bound states of fermions and vortices (see Fig. 2.23).

Thus a two-dimensional systems in presence of a magnetic field yields a scenario where as if the electrons capture a significant fraction of the external field, thereby the "original" electrons transform into the CF. Several experiments, such as thermopower (Zeitler *et al.*, 1993; and Ying *et al.*, 1994), Shubnikov-de Haas (SdH) oscillations and their cyclotron orbits (Leadley *et al.*, 1994), the observation of a Fermi sea (Halperin *et al.*, 1993). Importantly, the formalism of CF was capable of explaining the physics associated with the quantum Hall plateau observed at fractional filling of the Landau levels.

The role of interaction in the context of FQHE has been illustrated earlier and interestingly, it is the only energy scale left in the problem (the kinetic energy of the electrons become an irrelevant constant).

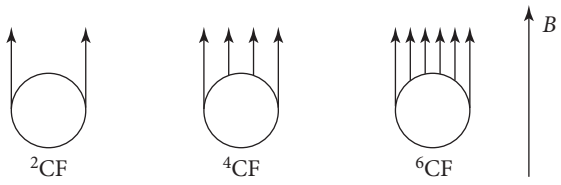


FIG. 2.23
Electrons capture 2, 4 and 6 flux quanta and the composite particles are known as composite Fermions (CF).

These strongly interacting electrons in the presence of a magnetic field, B transforms into such weakly interacting composite particles (CF) in a much weaker field, B^* where B^* is reduced from B (being absorbed by the electrons) by $2p\phi_0\rho$, namely,[14]

$$B^* = B - 2p\phi_0\rho \qquad \begin{cases} 2p & : \text{an integer} \\ \phi_0 & : \text{flux quantum} \\ \rho & : \text{density} \end{cases} \tag{2.207}$$

The situation is analogous to a filling fraction ν^* for the CF that corresponds to a fraction ν of the original electrons, and they are related to each other by

$$1 = \frac{\nu^*}{\nu} - 2p\nu^*$$

$$1 = \nu^*\left(\frac{1}{\nu} - 2p\right) \tag{2.208}$$

$$\nu = \nu^*(1 - 2p\nu). \tag{2.209}$$

Equivalently, one can write,

$$\frac{1}{\nu^*} = \pm\left(\frac{1}{\nu} - 2p\right) \tag{2.210}$$

Not only is the fundamental property of the fermions obeyed, for example the exchange of fermions involves a negative sign, the Aharonov–Bohm phase associated with the cyclic variation of the wave function around a closed path is unity.[15]

However, there is an important development that has occurred here. The highly degenerate many particle ground state for fractional filling ($\nu < 1$) in the absence of interaction transforms into the ground state of the CF with drastically reduced degeneracy corresponding to a filling $\nu^* > 1$. For integral values of ν^*, the situation yields a non-degenerate ground state. The loss of degeneracy makes the interaction among the CF to be vanishingly small. Thus, the FQHE of the original electrons transforms into IQHE of the CF and should yield a great simplification to a rather complex problem. The wavefunction for the CF can be obtained from the same variational state written down by Laughlin as in the following,

$$\psi_\nu = \prod_{j<k}(z_j - z_k)^{2p}\phi_{\nu^*} \tag{2.211}$$

where ϕ_{ν^*} denotes the wavefunction for the non-interacting electrons a filling ν^* and $(z_j - z_k)^{2p}$ is the similar Jastrow factor that was present in the Laughlin's proposal, which ensures keeping the CF apart, and no two of them come close to one another.[16]

[14] To arrive at the expression, consider the AB phase of a particle executing a circular motion of area A (disregarding the motion of all other particles) in an effective field B^* which is a result of $2\pi AB^*/\phi_0 \equiv 2\pi AB/\phi_0 - 2\pi(2p)\rho A$, where the last term denotes the flux due to 2p vortices.

[15] a phase, $e^{2\pi\phi/\phi_0}$ with $\phi = 2p\phi_0$ yields, $e^{2\pi(2p)} = 1$.

[16] The CF function vanishes as $\nu^{2(2p+1)}$ instead of ν^2 that is familiar for electrons obeying Pauli principle.

The success of the composite fermion picture can be tested by plotting the magneto (longitudinal) resistivities as a function of the inverse filling fraction $\frac{1}{\nu^*}$ which is proportional to B^*. A comparison between the IQHE and FQHE resistivities demonstrates the dips in the resistivity (see Figs. 2.3 and 2.4) correspond to integer filling fractions, while the ones in the lower panel correspond to the fractional fillings, but are related by

$$\nu = \frac{\nu^*}{2\nu^* + 1} \tag{2.212}$$

which can be obtained by putting $p = 1$ in Eq. (2.210). Thus, the IQHE of CF occurs at filling fractions $\nu^* = n$ where n is an integer. The filling fractions, despite being very different, look quite similar. This similarity re-emphasizes the dynamics of the interacting electrons that resembles that of the non-interacting CF very closely at a reduced magnetic field B^*.

Many of the observed fractional values for the quantum Hall plateaus can be obtained by putting $p = 1, 2, 3$ etc in Eq. (2.210) which can be viewed potentially as in Fig. 2.23.

Other fractions can be obtained by putting a negative sign in the following equation:

$$\frac{1}{\nu^*} = \pm\left(\frac{1}{\nu} - 2p\right), \tag{2.213}$$

where, negative sign denotes a B^* that is directed antiparallel to B. A few observed fractions and the corresponding (even) number of flux quanta (vortices) are tabulated in Table 2.1.

Table 2.1
A few representative p values and the corresponding ν values are shown.

p	$\nu = \frac{\nu^*}{2p\nu^*+1}$
1	$\frac{1}{3}, \frac{2}{5}, \frac{3}{7} \cdots$
−1	$\frac{2}{3}, \frac{3}{5}, \frac{5}{7} \cdots$
2	$\frac{1}{5}, \frac{2}{9} \cdots$
−2	$\frac{2}{7}, \frac{3}{11} \cdots$

Direct numerical solution of the Schrödinger equation for 10-15 particles in the presence of a pairwise Coulomb potential of the form,

$$\mathcal{H} = \frac{1}{4\pi\epsilon_0} \sum_{j \neq k} \frac{e^2}{|\mathbf{r}_j - \mathbf{r}_k|} \tag{2.214}$$

has been solved exactly.[17] Corresponding to ν^* equal to an integer (say, n) one expects a gap in the spectrum. At such values of filling, the Coulomb interaction lowers the enormous degeneracy of the lowest Landau level by a great extent, so as to produce a non-degenerate ground state. This state notionally denotes $\nu^*(= n)$ filled with Landau levels of CF. For the electrons, the filling fraction is of course ν which denotes a fraction (with odd denominator). Further, the wavefunctions of this CF has an excellent overlap with the wavefunctions stated in Eq. (2.179).[18]

In order to recover Laughlin's wavefunction, consider $\nu^* = 1$ which denotes $\nu = \frac{1}{2p+1}$. The wavefunction corresponding to the filling ν is

[17] Since the Coulomb potential is the only energy scale left in the problem, we have represented it by the Hamiltonian, $\mathcal{H}$.
[18] For an excellent discussion on the subject see Jain (2007).

$$\begin{aligned}\psi_{\nu=\frac{1}{2p+1}} &= \prod_{j<k}(z_j - z_k)^{2p}\Phi_{\nu^*=1} \\ &= \prod_{j<k}(z_j - z_k)^{2p}\left[\prod_{j<k}(z_j - z_k)\; e^{-\frac{1}{4l_B^2}\sum_l |z_l|^2}\right] \\ &= \prod_{j<k}(z_j - z_k)^{2p+1}\; e^{-\frac{1}{4l_B^2}\sum_l |z_l|^2} \end{aligned} \tag{2.215}$$

This is exactly Laughlin's wavefunction if one identifies $m = 2p + 1$.

Thus far, we have identified that the challenges of FQHE are too many, and most importantly it has to do with the absence of any "*small*" parameter of the problem. The carriers are frozen, and hence there is no kinetic energy, leaving the interparticle interaction to be the only energy scale of the problem. This impedes known methods of solution to be applicable. Furthermore, the enormous degeneracy of the Landau levels (without invoking the Coulomb interaction) aggravates this problem. The third, and quite a crucial one, is the unavailability of a so-called "*normal state*," that is, there is no known state which becomes unstable in favor of a fractional quantum Hall state by turning on a weak interparticle interaction. Hence, there is no small expansion parameter, humongously large number of ground states that are degenerate in the absence of interaction, and the non-existence of a normal fluid state put together narrate about an enormously complicated problem.

However, this enormously complicated problem is intuitively solved by writing a wavefunction (all credits to Laughlin), which we have denoted by $|\psi(\nu)\rangle$. Using this wavefunction, the energy eigenvalues for the problem can be written as

$$\langle \mathcal{H} \rangle = E_\nu = \langle \psi_\nu | \sum_{j<k} \frac{1}{|r_j - r_k|} |\psi_\nu\rangle + V_{e-b} + V_{b-b}, \tag{2.216}$$

where the first term on the RHS of Eq. (2.216) is the Coulomb energy, and the last two denote the interaction energy of the electron-background and between the background entities, such as ions etc. Importantly for us, the LLL index ν is given by

$$\nu = \frac{\nu^*}{2p\nu^* \pm 1}, \tag{2.217}$$

where the index ν^* is defined for the functions of the non-interacting electrons, namely, ϕ_{ν^*}. Thus, a strongly interacting problem in the presence of an external magnetic field B is reduced to a non-interacting one in an effective field given by

$$B^* = B - 2p\nu\Phi_0. \tag{2.218}$$

Thus, B^* can be zero if B is exactly canceled by the second term in Eq. (2.218), thereby leading to a non-degenerate state (since the degeneracy is proportional to B), or even be negative, where the vortices carried by composite particles point opposite to the applied field. Further, ψ_ν contains a Jastrow factor, namely, $\Pi_{j<k}\,(z_j - z_k)^{2p}$ which projects the wavefunction onto the LLL.

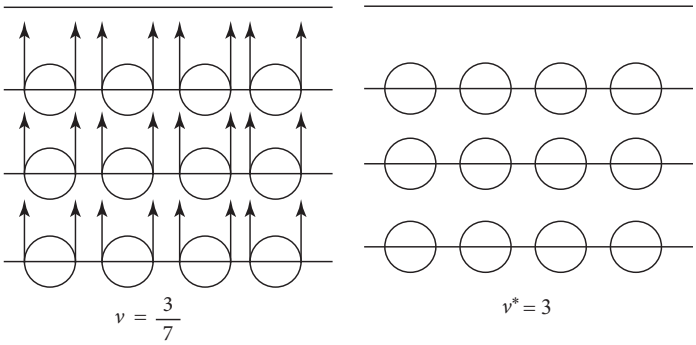


FIG. 2.24
The schematic plots show the integer quantum Hall effect for $\nu^* = 3$ (right) and the fractional quantum Hall effect for $\nu = \frac{3}{7}$ (left).

Let us consider the following example. A correspondence between the integer quantum Hall ground state with $\nu^* = 3$, that is, three filled levels and the corresponding fermion picture with $\nu = \frac{\nu^*}{2p\nu^*+1}$ is shown schematically in Fig. 2.24.

The above discussion yields an intuitive picture of composite fermions which are "*bound states*" of fermions and $2p$ "vortices" (p is an integer). By absorbing the "vortices" or the flux quanta, the electrons minimize the interparticle interaction energies. As these composite structures (particles + vortices) propagate in a quantum Hall fluid, they create irreducible phases, which (partially or fully) cancel the phase due to the external magnetic field.

The composite fermions truly are new entities, and not seen before. Since a vortex has a quantum mechanical origin, the composite fermions are quantum objects (owing to the quantum mechanical phases associated with the vortices) as well. However, in a fluid they behave as free particles. Cooper pairs, which are bound states of two electrons formed in the presence of lattice excitations, that is, phonons, bring up a close analogy, but are distinct in many ways as well. In fact, the topological quantization of the vorticity is directly linked with the quantization of the Hall plateaus.

We prefer to stop here on the topic of composite fermions and suggest more specialized articles and books on the subject. Particularly, the book by Jain (2007) has given an extensive account of composite fermions and, in general, on the theory of the fractional quantum Hall effect in a fairly lucid manner.

2.7 HIERARCHY APPROACH TO FQHE

We have so far discussed the Laughlin wavefunction and a composite fermion scenario to understand the physics of the fractional quantum Hall fluid. There are significant merits in both approaches. The Laughlin wavefunction is similar to the Jastrow function earlier employed to the superfluids, such as ^{4}He. A number of fractions that are experimentally observed could be explained by the Laughlin states, while a very large number of them remained elusive. On the other hand, the composite fermion ideas were motivated by considering FQHE to be analogous to IQHE, which eliminates the effects of strong interaction among the fermions by absorbing $2p$ flux quanta. Several experiments corroborate the composite fermion picture.

The non-Laughlin fractions still needed an explanation. Besides, we have a ground state wavefunction in some form, and thus, a natural expectation is to obtain information on the excited states. This forms the foundation of the hierarchy scenario, which assumes the Laughlin wavefunction as the starting point. The construction of a hierarchical wavefunction was put forward by Haldane (1983) and Halperin (1984) to provide justification to the non-Laughlin fractions, that is those ones that are not of the form $1/m$. The idea is to create "*Laughlin-like daughter states*" from a given "*parent state*." Additional states with new fractions are iteratively generated from the Laughlin fractions. Thus, all the odd-denominator fractions are reproduced. The fractions at any level of the hierarchical scheme are represented as continuous fractions.

A physical picture relevant to the hierarchy description may be enunciated as follows. At the centers of the quantum Hall plateau, which corresponds to an incompressible fluid with uniform density. As the magnetic field is ramped up, localized quasiparticles or quasiholes develop corresponding to an excess or deficit of densities that are initially pinned to the impurity centers. However, as the external field becomes larger, the excitations split from the impurities and become mobile, leading to a loss of the quantized Hall response. These quasiparticles (or the quasiholes) being charged may be considered to condense into Laughlin-like states on top of the original quantum Hall fluid. The scenario can be iterated which will lead to arbitrary filling fractions (basically the non-Laughlin fractions) with odd denominators.

To understand how the continued fractions arise, it is useful to review the excited states of the quantum Hall fluid, that is the quasiparticles and the quasiholes. The quasiparticles denote excitations, each of which carries a charge $-\frac{e}{m}$, while the quasiholes denote those with charge $\frac{e}{m}$. These excitations behave as individual particles, that are non-interacting, each with a fractional charge, however the total charge adds up to an integer, as it should be.

Halperin (1984) modified the Laughlin wavefunction and made an ansatz for including the quasiparticle and the quasihole excitations as

$$\psi = P(z_k)Q(z_k)e^{-\sum_j |z_j|^2/4l_B^2} \tag{2.219}$$

where $Q(z_k)$ accounts for the quasiparticle and the quasihole excitations and is given by

$$Q(z_k) = \Pi_{j<k}^{N}(z_j - z_k)^{\pm 1/m} \tag{2.220}$$

and $P(z_k)$ is the ubiquitous Jastrow factor given by

$$P(z_k) = \Pi_{j<k}^{N}(z_j - z_k)^{2p} \tag{2.221}$$

Thus,

$$\psi = \Pi_{j<k}^{N}(z_j - z_k)^{2p\pm 1/m} e^{-\sum_j |z_j|^2/4l_B^2} \tag{2.222}$$

with m still being an odd integer.

Put together $P(z_k)$ and $Q(z_k)$ give rise to an interchange that is more distinct than that of the anti-symmetric exchange of fermions. It is further evident that the maximum angular momentum of the state is

$$|J_{max}| = N\left(2p \pm \frac{1}{m}\right).$$

The filling fraction can be obtained by noting that the area A to be given by

$$A = N\left(2p \pm \frac{1}{m}\right)(2\pi m l_B^2).$$

This yields the number of states within an area A, which is given by

$$\mathcal{N} = \frac{BA}{\Phi_0} = \frac{\Phi}{\Phi_0} = \left(2p \pm \frac{1}{m}\right) m^2 N.$$

Hence, the filling fraction is

$$\nu = \frac{1}{2pm^2 \pm m}. \tag{2.223}$$

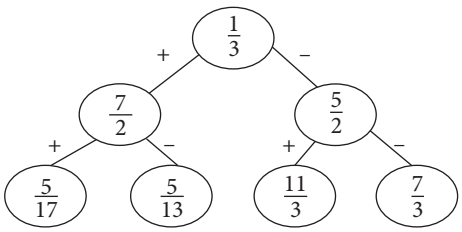


FIG. 2.25
Schematic plot showing the hierarchy scheme for $m = 3$ and $p_j = 1$. At the third level, one obtains non-Laughlin fractions, such as 5/17, 5/13, 11/3, and 7/3.

The above expression allows all fractions and not only the Laughlin $(1/m)$ types. At the third level of the hierarchy, the filling fractions are denoted by

$$\nu = \frac{1}{m \pm \frac{1}{2p_1 \pm \frac{1}{2p_2 \pm \cdots}}} \tag{2.224}$$

For example, for $p_j = 1$ and $m = 3$, at the third level one gets,

$$\nu = \frac{1}{3 \pm \frac{1}{2 \pm \frac{1}{2 \pm \cdots}}} \tag{2.225}$$

For the $+$-branch one gets, 5/17 and 5/13, and the $-$-branch, the fractions are 11/3 and 7/3. Schematically, the hierarchy at this level is expressed via Fig. 2.25.

2.8 SUMMARY AND OUTLOOK

We begin with a historical overview of the quantum Hall effect. The experiment and the physical systems are described with an emphasis on the two-dimensional nature of the 'dirty' electronic system in the presence of a strong perpendicular magnetic field at low temperature. The Hall resistivity as a function of the field shows quantized plateaus in the unit of h/e^2 with an accuracy of one part in more than a billion. Very surprisingly, the longitudinal resistivity synergetically vanishes at the positions of the plateaus for the Hall resistivity. This indicates the emergence of a phase with an inherent ambiguousness of being a perfect conductor and a perfect insulator at the same time. However, such an ambiguity can only be reconciled for an electron gas confined in a plane in the presence of a magnetic field.

Quite intriguingly, the presence of a perpendicular magnetic field introduces "*another*" quantization, which replaces the band structure (energy as a function of the wavevector) of the electronic system. This quantization was shown by solving the Schrödinger equation in the presence of a Landau gauge. The resultant energy levels of this problem are the infinitely degenerate Landau levels, which slightly broaden due to the presence of impurity and disorder, but remain distinct and cause quantization of the Hall conductivity as the magnetic field is ramped up gradually. Further, this quantization is visioned as a quantum pump by Laughlin, where an electron gas in a planar disk geometry subjected to a magnetic field shows a transfer of one unit of charge from the inner to the outer edge of the disk as the magnetic flux changes by one quantum ($= \frac{h}{e}$). Further, we have studied the same problem in the circular gauge, which is relevant to the study of the fractional quantum Hall effect.

The quantum Hall state is the first realization of a topological insulator, where the transmission of charges occurs through the edges of the electronic system, while the bulk remains insulating. The quantization of the Hall conductivity[19] is shown to be a topological invariant called the Chern number, which can only assume integer values. We further derived the Kubo formula to compute the Hall conductivity.

Hence, to compare and contrast between the effects of an external magnetic field in systems other than a two-dimensional electron gas, we have chosen graphene as another candidate, which is often the hobby horse to demonstrate the occurrence of the quantum Hall effect and other topological phenomena. The choice of graphene stems from the fact that the energy scale of the problem allows the quantum Hall effect to be realizable at temperatures as high as the room temperature, or even larger than that. The Landau levels of graphene are computed and are shown to be distinct than those for the two-dimensional electron gas. Lastly, the experimental demonstration and observation of the quantum Hall effect in graphene are discussed.

Finally, we have discussed the fractional quantum Hall effect that occurs in cleaner systems where the interparticle interaction becomes dominant and constitutes the only energy scale of the problem.

[19] It is more appropriate to talk about conductivity, rather than the resistivity.

To deal with the overwhelming complexity, Laughlin wrote down a variational wavefunction for the ground state, which, other than a ubiquitous Gaussian term, contains a Jastrow factor that keeps two fermions away from each other, thereby enforcing Pauli's exclusion principle. The wavefunction corresponds to a filling fraction of $1/m$, m being an odd integer and denotes the eigenvalue of the z-component of the angular momentum for that state. The Laughlin wavefunction is validated against the exact diagonalization of a system of a few electrons, and is found to have a near perfect overlap between the two. One shortcoming of the Laughlin formalism that continued to bother theorists working in the field is the experimental realization of several other values of the fraction (other than $1/m$) at which plateaus are observed. A hierarchy approach is proposed which yields several non-Laughlin fractions encoded via a '*parent-daughter*' relationship starting with a ("*parent*") Laughlin fraction. However, such a scenario suffers from all '*daughter*' to appear with equal weightage, which is not an experimental reality. Another intuitive solution was proposed by Jain that enunciates an effective medium comprising each fermion trapping an even number of flux quanta. Such a scenario greatly reduces the effective field that the system is subjected to, and transforms the fractionally quantized Hall system of interacting fermions to an effectively non-interacting system showing integer quantum Hall effect.

REFERENCES

Aharonov, Y. and Bohm, D., Phys. Rev. **115**, 485 (1959).

Ando, Y., Arakawa, K., Furuya, S. K., and Nakashima, H., *Mesoscopic Physics and Electronics* (Springer, New York, 1998).

Ashcroft, N. W. and Mermin, N. D., *Solid State Physics* (Harcourt College Publishers, 1976).

Azbel, M. Ya., JETP **19**, 634 (1964), available at http://jetp.ras.ru/cgi-bin/e/index/e/19/3/p634?a=list.

Bergmann, G., Phys. Rep. **107**, 1 (1984).

Castro Neto, A. H., Guinea, F., Peres, N. M. R., Novoselov, K. S., and Geim, A. K., Rev. Mod. Phys. **81**, 209 (2009).

Deacon, R. S. *et al.*, Phys. Rev. B **76**, 081406 (2007).

Dutta, S., *Electronic Transport Properties in Mesoscopic Systems* (Cambridge University Press, 1995).

Garcia, R. G., Appl. Phys. Lett. **60**, 1960 (1992).

Glattli, D. *et al.*, Physica E **6**, 22 (2000).

Goldstein, H., Poole, C. P., and Safko, J., *Classical Mechanics*, 3rd ed. (Addison Wesley, 2002).

Haldane, F. D. M., Phys. Rev. Lett. **51**, 605 (1983).

Halperin, B. I., Phys. Rev. Lett. **52**, 1583 (1984).

Halperin, B. I., Phys. Rev. Lett. **52**, 2390 (1984).

Halperin, B. I., Lee, P. A., and Read, N., Phys. Rev. B **47**, 7312 (1993).

Harper, P. G., Proc. Phys. Soc. A **68**, 874 (1955).

Hess, K., "Boltzmann transport equation," in *The Physics of Submicron Semiconducting Devices* edited by H. L. Grubin (Springer, 1988).

Hofstadter, D. R., Phys. Rev. B **14**, 2239 (1976).
Imry, Y., *Electronic Transport in Mesoscopic Systems* (Oxford University Press, New York, 1997).
Jain, J. K., *Composite Fermions* (Cambridge University press, 2007).
Kittel, C., *Introduction to Solid State Physics*, 8th ed. (Wiley, 2004).
Klitzing, K. v., Dorda, G., and Pepper, M., Phys. Rev. Lett. **45**, 494 (1980).
Landauer, R., IBM J. Res. Dev. **1**, 223 (1957).
Landwehr, G., Metrologia **22**, 118 (1986).
Laughlin, R. B., Phys. Rev. Lett. **50**, 1395 (1983).
Leadley, D. R., Nicholas, R. J., Foxon, C. T., and Harris, J. J., Phys. Rev. Lett. **72**, 1906 (1994).
Lee, P. A., and Ramakrishnan, T. V., Rev. Mod. Phys. **57**, 287 (1985).
Leinaas, J. M. and Myrheim, J., Nuovo Cimento B **37**, 1 (1977), available at https://www.ifi.unicamp.br/∼cabrera/teaching/referencia.pdf.
Li, G., Luican, A., and Andrei, E. Y., Phys. Rev. Lett. **102**, 176804 (2009).
Mahan, G. D., *Many Particle Physics*, 3rd ed. (Springer, 2000).
Mello, N. K. P. R., *Quantum Transport in Mesoscopic Systems: Complexity and Statistical Fluctuations* (Oxford University Press, New York, 2004).
Novoselov, K. S. *et al.*, Science **315**, 1379 (2007).
Onsager, L., Phys. Rev. **37**, 405 (1931).
Rammal, R., J. Phys. **46**, 1345 (1985).
Schlösser, T., Ensslin, K., Kotthaus, J. P., and Holland, M., Semicond. Sci. Technol. **11**, 1582 (1996).
Shankar, R., *Principles of Quantum Mechanics*, 2nd ed. (Kluwer Academic/Plenum Publishers, 1994).
Sprinkle, M. *et al.*, Phys. Rev. Lett. **103**, 226803 (2009).
Thouless, D. J., Kohmoto, M., Nightingale, M. P., and den Nijs, M., Phys. Rev. Lett. **49**, 405 (1982).
Tong, D., TIFR, Infosys Lectures on The Quantum Hall Effect, arXiv:1606.06687 (2016).
Tsui, D. C., Stormer, H. L., and Gossard, A. C., Phys. Rev. Lett. **48**, 1559 (1982).
van Wees, B. J., van Houten, H., Beenakker, C. W. J., Williamson, J. G., Kouwen-hoven, L. P., van der Marel, D., and Foxon, C. T., Phys. Rev. Lett. **60**, 848 (1988).
Wharam, D. A. *et al.*, J. Phys. C **21**, L209 (1988).
Wilczek, F., Phys. Rev. Lett. **49**, 957 (1982).
Ying, X., Bayot, V., Santos, M. B., and Shaygon, M., Phys. Rev. B **50**, 4969 (1994).
Zeitler, U. *et al.*, Phys. Rev. B **47**, 16008(R) (1993).

CHAPTER

3 SYMMETRY AND TOPOLOGY

3.1 INTRODUCTION

"*Point set topology is a disease from which the human race will soon recover*" - H. Poincaré (1908)

The Poincaré conjecture was the first conjecture made on topology, which asserts that a 3D manifold is equivalent to a sphere in 3D subject to the fulfillment of a certain algebraic condition of the form $f(x, y, z) = 0$, where x, y, and z are complex numbers. G. Perelman (arguably) solved this conjecture in 2006 Perelman (2006). However, on practical aspects, just the reverse of what Poincaré had predicted, happened. Topology and its relevance to condensed matter physics has emerged in a big way in the recent times. The 2016 *Nobel prize* awarded to D. J. Thouless, J. M. Kosterlitz, F. D. M. Haldane and C. L. Kane and E. Mele getting the *Breakthrough Prize* for contributions to fundamental physics in 2019 bear testimony to that.

Topology and geometry are related, but they have a profound difference. Geometry can differentiate between a square from a circle, or between a triangle and a rhombus, however, topology cannot distinguish between them. All it can say is that, individually all these shapes are connected by continuous lines, and hence are identical. However, topology indeed refers to the study of geometric shapes, where the focus is on how properties of objects change under continuous deformation, such as stretching and bending, however tearing or puncturing is not allowed. The objective is to determine whether such a continuous deformation can lead to a change from one geometric shape to another. The connection to the problem of deformation of geometrical shapes in condensed matter physics may be established if the Hamiltonian for a particular system can be continuously transformed via tuning of one (or more) of the parameter(s) that the Hamiltonian, depends on. Should there be no change in the number of energy modes below the Fermi energy during the process of transformation, then the two systems (that is, before and after the transformation) belong to the same topology class. In the process "*something*" remains invariant. If that *something* does not remain invariant, then there occurs a topological phase transition. This phase transition can occur from one topological phase to another, or from a topological phase to a trivial phase.

In the following, we present the geometric aspects of topology and relate the integral of the geometric properties to closed surfaces to the topological invariants. It turns out that the "*geometric property*" and the "*closed surface*" have smooth connections to physical observables. As we shall see soon that in 1982 Thouless *et al.* (1982) linked the topological invariant to the quantized Hall conductivity.

https://doi.org/10.1063/9780735422537_003 3-1

To test many of the concepts that we are going to discuss in this chapter, we choose two prototype systems, one each in one (1D), and two dimensions (2D). In 1D, we consider a tight binding model, with dimerized hopping, and in 2D, we consider graphene, which has been a hobby horse even for several years before its experimental discovery. The theme is to discuss the interplay of symmetry and the topological properties. Particularly, in 2D, an important highlight in this direction is put forward by Haldane (1988) who had proposed a non-trivial topological phase by breaking one of the fundamental symmetries, namely, the time reversal symmetry. Finally, after the experimental discovery of graphene, yet another distinct topological state of matter was discovered by Kane and Mele (2005) which has culminated into an emerging field of spintronics.

In Chap. 3 we saw that the Hall conductivity (or the resistivity) is quantized in the unit of e^2/h (or h/e^2) within a splendid precision, so much so, that the quantity h/e^2 can define the standard of resistance ($= 25.5\ \text{k}\Omega$). Clearly the quantization is independent of the details of the Hamiltonian, for example, the nature of the sample, the strength of the magnetic field, and disorder present in the system. It is realized later that the universality of the phenomenon arises due to "topological" protection of the energy modes that exist at the edges of a quantum Hall sample, which possess completely different character as compared to the ones that exist in the bulk of the sample. Thus, an understanding emerges, that says that a physical observable (which is either the resistivity or the conductivity) can be represented mathematically by a topological invariant. This invariant does not change even when the Hamiltonian changes (for example, when the strength of the magnetic field is varied), until and unless a phase transition occurs, which will show up via an abrupt change in the value of the topological invariant. There is an elegant explanation of the physics involved with such a universal phenomenon, which brings us to the subject of topology.

Topology in its usual sense deals with the geometry of the objects; in the same spirit, here we shall study the geometrical properties of the Hilbert space for the system under consideration. The idea is best demonstrated for a quantum Hall system, which undergoes a series of transitions from a conducting to an insulating state as a function of the external magnetic field. In the process, the topological invariant, for example, the Chern number in this case (we shall discuss this later) jumps from one integral value to another. Thus, the system repeatedly undergoes through a series of topological phase transitions. In the following we describe this topological phase transition in more general terms.

Consider two Hamiltonians, $\mathcal{H}_1$ and $\mathcal{H}_2$, both of which are functions of a tunable parameter, say, β. If the corresponding energy spectra $\epsilon_1(\beta)$ and $\epsilon_2(\beta)$ are such that the number of energy levels below the zero energy (zero energy is usually the Fermi energy) always remain the same for all values of β, then the Hamiltonians can be continuously transformed (or deformed as we see the analogy of a cup and donut later), and there is no phase transition. Now, consider either $\mathcal{H}_1$ or $\mathcal{H}_2$. If, for either of them, the spectrum is such that the number of energy levels varies as a function of β, that is, if any (or more) levels cross the zero energy, then the "invariant" changes (from one integer value to another), and one encounters that the system is going from one topological phase to another. A quantum Hall system shows a similar transition, where the Hall conductivity changes from ne^2/h to $(n+1)e^2/h$, where n is strictly an integer.

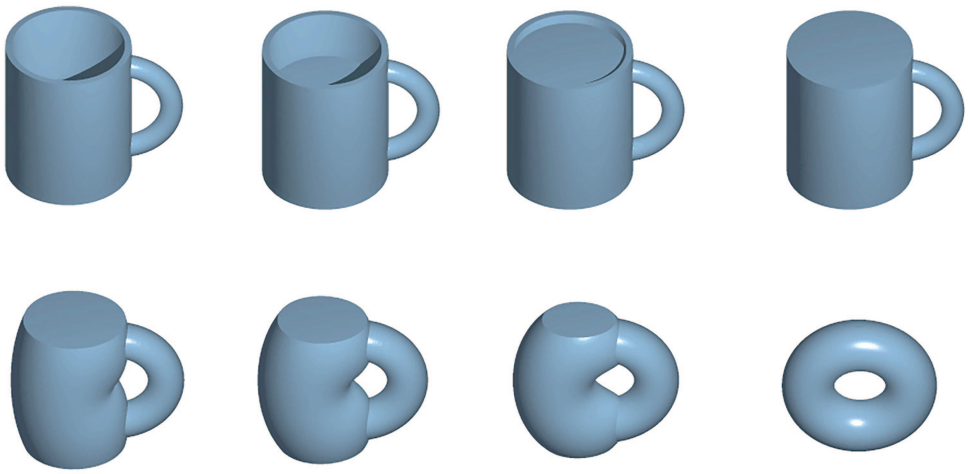


FIG. 3.1
A mug can be transformed smoothly into a donut. The handle of the mug remains invariant and emerges as "hole" of the donut. Thus, the mug and the donut belong to the same universality class.

Thus, the study of topology deals with objects (or Hamiltonians) that can be continuously transformed (or deformed) from one to another without puncturing or tearing the object (or without even closing the energy gap for the quantum system). For geometrical objects, being able to continuously transform depends on the number of "holes" or "genus" that are preserved during the course of the transformation. For example, a soccer ball can be deformed smoothly into a wine glass since both of them have no holes (zero genus), while a mug (as shown in Fig. 3.1) can be transformed smoothly into a donut with one hole (genus equal to 1). The first case with zero hole is called topologically trivial, and the second with a finite number of holes (one in this case) is termed as topologically non-trivial.

3.1.1 Gauss-Bonnet theorem

Gauss-Bonnet theorem in differential geometry that is about the evaluation of the surface integral of a Gaussian surface. Here we state the theorem without proof. In the most general form, for a closed polyhedral surface, the theorem can be stated as

$$\int_{\partial R} k_g(s)\mathrm{d}s + \iint_R K\mathrm{d}A = 2\pi\,\chi(R) \tag{3.1}$$

where R denotes a regular region with the boundary ∂R of R, K is the Gaussian curvature, s denotes the arc length of the curves, C_i and the integral is over C_i. Further $\chi(R)$ is called the Euler-Poincare characteristic. The first term on the left is the integral of the Gaussian curvature over the surface, the second

one is the integral of the geodesic curvature of the boundary of the surface. Thus, the Gauss-Bonnet theorem simply states that the total curvature of R plus the total geodesic curvature of ∂R is a constant.

As an example, we consider the simplest case, that is a sphere of radius R. The Gaussian curvature is $1/r^2$ [1] and the corresponding area is,

$$\iint_R K\mathrm{d}A = K \times \text{Area} = \frac{1}{r^2} \times 4\pi r^2 = 4\pi. \tag{3.2}$$

Again,

$$\iint_R K\mathrm{d}A = 2\pi\chi(R). \tag{3.3}$$

Thus,

$$\chi(R) = 2$$

Thus, the Euler-Poincare characteristic of a sphere is 2 although the genus is equal to zero (see Fig. 3.2). Suppose we wish to extend this argument to other closed, however not necessarily convex

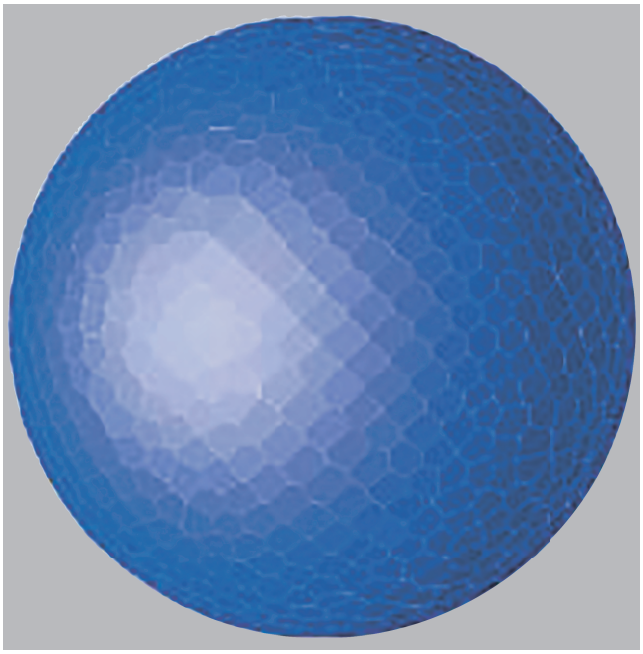


FIG. 3.2
A sphere with no hole (or genus). It represents a trivial phase.

[1] For a geometry with two different radii of curvature, such as a convex lens, the Gaussian curvature is $\frac{1}{r_1 r_2}$.

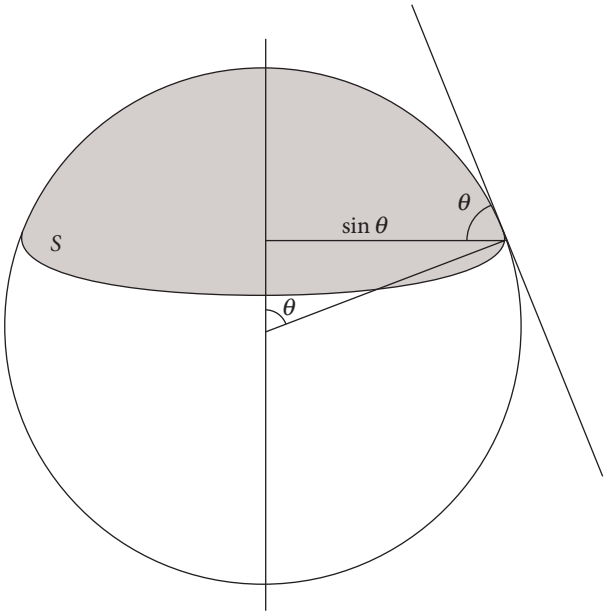


FIG. 3.3
A sphere with a polar cap. The figure will aid in calculating the area S of the polar cap.

surfaces in a three-dimensional space. For that, consider the polar cap of unit radius (see Fig. 3.3). The area is given by

$$S = \int_0^\theta 2\pi \sin\theta \mathrm{d}\theta = 2\pi(1 - \cos\theta). \tag{3.4}$$

Thus,

$$\int_R K\mathrm{d}A = 1 \times (\text{Area of } S) = 2\pi(1 - \cos\theta).$$

The geodesic curvature K is $1/\tan\theta$. Thus,

$$\int_s K_g \mathrm{d}s = K_g \times \text{length}(S) = \frac{1}{\tan\theta}. \tag{3.5}$$

Hence,

$$\int_s K_g \mathrm{d}s + \int_R K\mathrm{d}A = 2\pi(1 - \cos\theta) + 2\pi\cos\theta$$
$$= 2\pi = 2\pi\chi(R)$$

thereby yielding,

$$\chi(R) = 1.$$

In fact, an alternate form for the Gauss-Bonnet theorem is more useful for our purpose, which states that, for a closed convex surface, the integral over the Gaussian curvature can be expressed in terms of the number of holes or the genus of the surface. Thus, a simplified (and more relatable for us) version reads

$$\iint K\mathrm{d}A = 2\pi(2 - 2g) \tag{3.6}$$

Since a sphere has no holes ($g = 0$), the integral of the curvature yields

$$\iint K\mathrm{d}A = 4\pi \tag{3.7}$$

a result that we have seen earlier. Let us look at a case where the genus is non-zero ($g \neq 0$), such as a torus which is topologically equivalent to a mug, as we have seen earlier.

For a torus, the Euler-Poincare characteristic has a value of zero. This implies that, irrespective of how we bend or deform it, the integrated curvature vanishes. Refer to Fig. 3.4 (left panel) where there is a positive curvature on the outer surface, and negative curvature on the inner surface, thereby resulting in zero total curvature. This is consistent with the Gauss-Bonnet theorem which states that the integral of the Gaussian curvature is $2\pi(2 - 2g)$. Since $g = 1$ here, the integral is zero and so is the

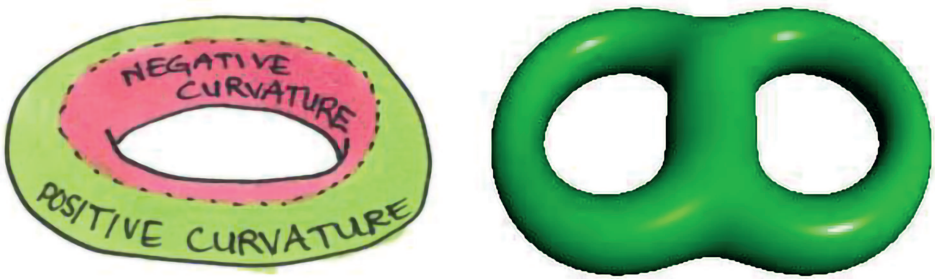


FIG. 3.4
(Left) A donut with genus (or a hole) equal to 1. (Right) A two-hole object has genus equal to 2.

Euler-Poincare characteristic $\chi(R)(= 2 - 2g)$ is zero as well. Similarly, a two-holed donut (see right panel of Fig. 3.4) will have $\chi(R) = -2$, and hence negative integrated Gaussian curvature.

Based on the preceding discussion, a sketchy idea emerges on the relationship between topology, and properties of quantum systems. However, it remains unclear how these ideas can relate to the properties of materials. At this moment, let us talk about crystalline solids for which the electron wavefunction is given by Bloch's theorem, namely,

$$\psi(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} u_{\mathbf{k}}(\mathbf{r}) \tag{3.8}$$

where the periodicity of the crystal potential, that is $V(\mathbf{r}) = V(\mathbf{r} + \mathbf{R})$ is captured by the amplitude function $u_{\mathbf{k}}(\mathbf{r})$, such that

$$u_{\mathbf{k}}(\mathbf{r} + \mathbf{R}) = u_{\mathbf{k}}(\mathbf{r}), \tag{3.9}$$

where $\mathbf{k}$ denotes the crystal momentum, and is distinct than the usual momentum $(= -i\hbar\nabla)$ (Kittel, 1986). The crystal momentum is restricted within the first Brillouin zone (BZ), where the latter is a region in the k-space with periodic boundaries. As the crystal momentum is varied, we map out the energy bands, and one obtains the band structure. The BZ plays the role of the surface over which the integral of the Gaussian curvature is taken, which we have discussed earlier.

Now that brings us to the question: *what is the analog of the Gaussian curvature* for a crystalline solid? To understand this, consider the (non-degenerate) ground state of a Hamiltonian, which depends upon a number of parameters that are time dependent. The adiabatic theorem states that if the Hamiltonian is now changed slowly[2] with respect to the parameters, the system remains in its time-dependent ground state. However, there is something more to it. As the ground state is evolved in time, in addition to the trivial dynamical phase, there may emerge an irreducible geometric phase that comes into play, namely, the Berry phase put forward by Berry (1984). In the following, we discuss the origin of the

[2] The time scale for change is larger than the inverse level spacing (level spacing implies difference between subsequent energy levels) of the system.

Berry phase, and the Berry curvature, which is analogous to the Gaussian curvature. The integral of the Berry curvature over the BZ is shown to yield a constant (or more appropriately an invariant) known as the Chern number, which is analogous to the RHS of Eq. (3.6) or the Euler-Poincare characteristic.

3.1.2 Berry phase

Consider that a particle is in the ground state of a box of length L. Suppose the box slowly expands, such that $L(t)$ is a slow function of time. The adiabatic principle says that, if the expansion is slow, then the particle always remains in the ground state at any time t. This is true for any state of the system. More generally, consider a Hamiltonian $\mathcal{H}(\lambda(t))$ where λ is a parameter which changes slowly. Now the adiabatic principle says that if the particle starts out in the nth eigenstate of $\mathcal{H}(\lambda(0))$, it will land nth instantaneous eigenstate of $\mathcal{H}(\lambda(t))$ at a time t.

The question is, what is the solution of the Schrödinger equation in this approximation? A reasonable guess is

$$|\psi(t)\rangle = \exp\left(-\frac{i}{\hbar}\int \varepsilon_n(t')\mathrm{d}t'\right)|\phi_n(t)\rangle, \tag{3.10}$$

where

$$\mathcal{H}(t)\,|\phi_n(t)\rangle = \varepsilon_n(t)\,|\phi_n(t)\rangle. \tag{3.11}$$

If $\mathcal{H}$ does not vary with time, then the phase is clearly correct. However, it is not so in the case $\mathcal{H}$ depends on time.

To see what is missing in the above ansatz, let us modify it slightly.

$$|\psi(t)\rangle = C(t)\exp\left[-\frac{i}{\hbar}\int_0^t \varepsilon_n(t')\mathrm{d}t'\right]|\phi_n(t)\rangle. \tag{3.12}$$

In Eq. (3.10), $C(t) = 1$. Applying the Schrödinger equation $(i\hbar\frac{\partial}{\partial t} - \mathcal{H})$ to Eq. (3.12) and simplifying, for the time dependence of C, one gets

$$\dot{C}(t) = -C(t)\langle\phi_n(t)|\frac{\mathrm{d}}{\mathrm{d}t}|\phi_n(t)\rangle. \tag{3.13}$$

This yields a solution of the form,

$$C(t) = C(0)\exp\underbrace{\left[-\int_0^t \langle\phi_n(t')|\frac{d}{dt'}|\phi_n(t')\rangle dt'\right]}_{e^{i\gamma}} = C(0)e^{i\gamma}, \tag{3.14}$$

where

$$\gamma = i\int_0^t \langle\phi_n(t')|\frac{d}{dt'}|\phi_n(t')\rangle dt'. \tag{3.15}$$

This extra phase is called the Berry phase or the geometric phase. It is also called the Berry-Panchratnam phase, and is quite a familiar quantity in the field of optics.

In general, phases do not give rise to measurable consequences since the eigenstates are defined only up to a phase factor. Even here, it may be thought that we can define new Berry states to absorb the phase, namely,

$$|\phi_n'(t)\rangle = e^{i\chi(t)}\,|\phi_n(t)\rangle \tag{3.16}$$

then,

$$i\langle\phi_n'(t)|\frac{d}{dt}|\phi_n'(t)\rangle = i\langle\phi_n(t)|\frac{d}{dt}|\phi_n(t)\rangle - \frac{d\chi}{dt}. \tag{3.17}$$

Now suppose the parameter $\chi(t)$ changes the Hamiltonian in such a manner that after a complex cycle,

$$\mathcal{H}(0) = \mathcal{H}(t = T)$$

The end result is,

$$i\oint\langle\phi_n'(t)|\frac{d}{dt}|\phi_n'(t)\rangle = i\oint\langle\phi_n(t)|\frac{d}{dt}|\phi_n(t)\rangle - (\chi(T) - \chi(0))\,. \tag{3.18}$$

The last term on the RHS is an irreducible phase that does not cancel under the redefinition of the Berry states. Remember, χ arises from γ, which we denote as the Berry phase. Single valuedness of χ demands

$$\chi(T) - \chi(0) = 2\pi n, \quad \text{where } n \text{ is an integer}$$

The surface integral of the Berry curvature is called the Chern number. This is analogous to Gauss Bonnet theorem which connects the surface integral of radii of curvature. In the Gauss-Bonnet theorem, just like an object with a genus "1" can not be smoothly transformed into another with genus "zero" or "2" (unless, of course, something drastic, that is, tearing or puncturing is done to the object), a system with a non-zero Chern number cannot be transformed into that with zero Chern number.

In a quantum Hall system, the Hall conductivity is given by

$$\sigma_{xy} = Ce^2/h = ne^2/h, \quad \text{where } C = \text{ Chern number.}$$

It can be argued that the Chern number is always an integer. Further, the Berry curvature, $\mathcal{F}$ is defined as the curl of the Berry connection, namely,

$$\mathcal{F} = \nabla \times \mathcal{A} \tag{3.19}$$

$\mathcal{F}$ is analogous to the magnetic field. The Chern number is defined as the surface integral of the Berry curvature over a surface enclosed.

3.2 SYMMETRIES AND TOPOLOGY

To elucidate more on the topological invariance in materials, we discuss a few discrete symmetries of the Hamiltonian, and how they interplay with the topological properties. In this context, we wish to discuss three symmetries, namely, the inversion symmetry (also known as sublattice symmetry), and

the time reversal symmetry. A third symmetry that we shall talk about is the contribution of the above two, and is known as the particle-hole symmetry. We shall not worry about the third one here, since, it is not relevant for the present discussion, and it has been discussed in the context of the Hubbard model (see Chap. 2).

3.2.1 Inversion symmetry

Let us consider an eigenstate in the position based on $|\psi(\mathbf{r})\rangle$, so that

$$\mathbf{r}\,|\psi(\mathbf{r})\rangle = |\mathbf{r}|\,|\psi(\mathbf{r})\rangle. \tag{3.20}$$

Now we define the inversion symmetry or the parity operator $\mathcal{P}$ such that

$$\mathcal{P}\,|\psi(\mathbf{r})\rangle = |\psi(-\mathbf{r})\rangle. \tag{3.21}$$

Now,

$$\mathbf{r}\mathcal{P}\,|\psi(\mathbf{r})\rangle = \mathbf{r}\,|\psi(-\mathbf{r})\rangle = -|\mathbf{r}|\,|\psi(-\mathbf{r})\rangle. \tag{3.22}$$

If we act $\mathcal{P}^\dagger$ on both sides (remembering $\mathcal{P}^\dagger = \mathcal{P}^{-1} = \mathcal{P}$)

$$\mathcal{P}^\dagger\,\mathbf{r}\,\mathcal{P}\,|\psi(\mathbf{r})\rangle = -|\mathbf{r}|\mathcal{P}^\dagger\,|\psi(-\mathbf{r})\rangle = -|\mathbf{r}|\mathcal{P}\,|\psi(-\mathbf{r})\rangle = -|\mathbf{r}|\,|\psi(\mathbf{r})\rangle = -\mathbf{r}\,|\psi(\mathbf{r})\rangle. \tag{3.23}$$

Thus,

$$\mathcal{P}^\dagger\,\mathbf{r}\,\mathcal{P} = -\mathbf{r}. \tag{3.24}$$

This yields

$$\mathcal{P}^\dagger\,\mathbf{r} = -\mathbf{r}\mathcal{P} \tag{3.25}$$

$$\text{or,}\quad \{\mathcal{P}, \mathbf{r}\} = 0. \tag{3.26}$$

Hence, the parity operator anticommutes with the position operator.

Let us now explore the analogous scenario for the momentum operator. For this purpose, it is convenient to introduce the transformation operator $T(\mathbf{a})$[3] that translates a state $|\psi(\mathbf{r})\rangle$ to $|\psi(\mathbf{r}+\mathbf{a})\rangle$, where $\mathbf{a}$ denotes a fixed length, for example, $\mathbf{a}$ can be the lattice constant. That is,

$$\begin{aligned} &T(\mathbf{a})\,|\psi(\mathbf{r})\rangle = |\psi(\mathbf{r}+\mathbf{a})\rangle \\ &\mathcal{P}^\dagger T(\mathbf{a})\mathcal{P}\,|\psi(\mathbf{r})\rangle = T(-\mathbf{a})\,|\psi(\mathbf{r})\rangle, \end{aligned} \tag{3.27}$$

which yields

$$\mathcal{P}^\dagger\,T(\mathbf{a})\,\mathcal{P} = T(-\mathbf{a}). \tag{3.28}$$

This demands that the translation operator is of the form,

$$T(\mathbf{a}) = e^{i\,\mathbf{k}\cdot\mathbf{a}}. \tag{3.29}$$

Expanding for infinitesimal translations,

$$\mathcal{P}^\dagger\,\mathbf{p}\,\mathcal{P} = -\mathbf{p}. \tag{3.30}$$

[3] Distinguish between $T(a)$ for the translation operator, and $\mathcal{T}$ for the time reversal operator.

Thus, similar to the position operator, the momentum operator also anticommutes with the parity operator.

Since, both $\mathbf{r}$ and $\mathbf{p}$ anticommute, the angular momentum, $\mathbf{L}$ $(= \mathbf{r} \times \mathbf{p})$ commutes with $\mathcal{P}$. In a 3D orthogonal coordinate system, one can invert it about any of the axes. For example, in a Cartesian coordinate system,

i. an inversion about the z-axis is denoted as $\sigma_h(xy)$.
ii. about the y-axis it is $\sigma_v(xz)$ and
iii. $\sigma_v(yz)$ denotes the inversion about x-axis.

Here σ denotes an inversion operation, and has got nothing to do with the Pauli matrices. Under these operations, the position and the angular momentum variable transform as

i. $\sigma_h(xy)$: $x \to x, y \to y, z \to -z$
   $L_x \to -L_x, L_y \to -L_y, L_z \to L_z$
ii. $\sigma_v(xz)$: $x \to x, y \to -y, z \to z$
   $L_x \to -L_x, L_y \to L_y, L_z \to -L_z$
iii. $\sigma_v(yz)$: $x \to -x, y \to y, z \to z$
   $L_x \to L_x, L_y \to -L_y, L_z \to -L_z$

3.2.2 Time reversal symmetry

Now we shall discuss time reversal symmetry. It is obvious that under time reversal, the time variable, t changes to $-t$. This makes the position $(\mathbf{r}(t))$ and the momentum variable $(\mathbf{p}(t))$ transform under time reversal as $\mathbf{r}(-t)$ and $-\mathbf{p}(t)$ respectively. The angular momentum $\mathbf{L}(t)(= \mathbf{r} \times \mathbf{p})$ thus also becomes, $-\mathbf{L}(-t)$ under time reversal. Similar outcomes are expected when $\mathbf{r}(t)$, $\mathbf{p}(t)$ and $\mathbf{L}(t)$ are quantum mechanical operators. Additional inputs to the ongoing discussion can be received from the behavior of the electric field, $\mathbf{E}(\mathbf{r}, t)$ and the magnetic field $\mathbf{B}(\mathbf{r}, t)$ vectors under time reversal. $\mathbf{E}(\mathbf{r}, t)$ does not change sign under time reversal (refer to the Maxwell's equations, $\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}$ where charge density $\rho(\mathbf{r})$ does not change sign), however $\mathbf{B}(\mathbf{r}, t)$ changes sign (owing to, $\nabla \times \mathbf{B} = \mu_0 \mathbf{J}$, where $\mathbf{J}(\mathbf{r}, t)$ is the current density and it changes sign under time reversal).

Now, consider a quantum state $\psi(t)$ that obeys Schrödinger equation,

$$i\hbar \frac{\partial \psi(\mathbf{r}, t)}{\partial t} = \mathcal{H}\psi(\mathbf{r}, t). \tag{3.31}$$

In the following, we suppress the $\mathbf{r}$ dependence of ψ, and simply write $\psi(t)$ which upon the application of the time reversal operator yields $\psi'(-t)$. Mathematically,

$$\mathcal{T}\,|\psi(t)\rangle = |\psi'(-t)\rangle. \tag{3.32}$$

In order to find $\psi'(-t)$, let us look at the solution of Eq. (3.31),

$$|\psi(t)\rangle = e^{-i\mathcal{H}t/\hbar}\,|\psi(0)\rangle. \tag{3.33}$$

For $t = 0$, apply the time reversal operator, that is, $\mathcal{T}\,|\psi(0)\rangle$. Now, let it evolve forward in time, which means we get a state,

$e^{-i\mathcal{H}t/\hbar}\mathcal{T}\,|\psi(0)\rangle$.

For the Hamiltonian to be invariant under time reversal, this state should be the same as $\mathcal{T}\psi(-t)$ which is equivalent to,

$\mathcal{T}e^{i\mathcal{H}t/\hbar}\,|\psi(0)\rangle$.

Thus,

$\mathcal{T}e^{i\mathcal{H}t/\hbar}\,|\psi(0)\rangle = e^{-i\mathcal{H}t/\hbar}\mathcal{T}\psi(0)$.

For small time δt, we can expand the exponential and write,

$$\mathcal{T}i\mathcal{H} = -i\mathcal{H}\mathcal{T}. \tag{3.34}$$

A natural intuition (albeit wrong as shown later) is to cancel the "i" from both sides of Eq. (3.34). This yields

$$\mathcal{T}\mathcal{H} = -\mathcal{H}\mathcal{T} \tag{3.35}$$

which implies,

$$\mathcal{T}\mathcal{H} + \mathcal{H}\mathcal{T} = 0 \qquad \text{or,} \quad \{\mathcal{T}, \mathcal{H}\} = 0. \tag{3.36}$$

But that cannot be correct, since we have assumed that the time reversal operation to be a valid symmetry operation. This means that canceling the "i" from both sides in Eq. (3.34) was not a legitimate step.

Reconciliation is possible if we understand that time reversal, unlike most other operations in quantum mechanics, is an antiunitary operation. To remind ourselves that a unitary operation U satisfies $UU^\dagger = \mathbb{1}$ or a unitary operator acting on a state $\alpha\,|\psi\rangle$ yields

$$U(\alpha\,|\psi\rangle) = \alpha U\,|\psi\rangle. \tag{3.37}$$

This is also a property of a linear operator. However, for an anti-linear operator, A one gets

$$A(\alpha\,|\psi\rangle) = \alpha^* A\,|\psi\rangle, \tag{3.38}$$

which means that the anti-linear operator involves a complex conjugation. This resolves the dilemma caused by the naive cancelation of "i" in Eq. (3.34). Thus, the factor "i" on the LHS of Eq. (3.34) is complex conjugated when encounters $\mathcal{T}$ on the way pulling it. This yields an extra minus sign which cancels with the one on the RHS yielding

$$[\mathcal{T}, \mathcal{H}] = 0. \tag{3.39}$$

This is familiar with the notion that $\mathcal{H}$ is a time reversal invariant (we have to deliberately break the time reversal invariance of the Hamiltonian, either via an external magnetic field or some other means) and hence, the Hamiltonian should commute with the time reversal operator.

Any antiunitary operator can be written as a product of a unitary operator, U multiplied by a complex conjugation operator, $\mathcal{K}$ such that,

$$\mathcal{T} = U\mathcal{K}.$$

A special case in this regard deserves a mention, that is, the case for a $S = \frac{1}{2}$ particle. Spin being an angular momentum, it is odd under time reversal, that is,

$$\mathcal{T}\sigma\mathcal{T}^{-1} = -\sigma. \tag{3.40}$$

This implies, $\mathcal{K}\sigma_x\mathcal{K}^{-1} = \sigma_x$, $\mathcal{K}\sigma_y\mathcal{K}^{-1} = -\sigma_y$ and $\mathcal{K}\sigma_z\mathcal{K}^{-1} = \sigma_z$ which is reasonable as σ_y contains imaginary entries. Thus, for the unitary operator, $U\sigma_x U^{-1} = -\sigma_x$, $U\sigma_y U^{-1} = \sigma_y$ and $U\sigma_z U^{-1} = -\sigma_z$. Thus, the unitary operator, U commutes with σ_y, but anticommutes with σ_x and σ_z.

Finally, the form of the time reversal operator for a Hamiltonian corresponding to a $S = \frac{1}{2}$ system is given by (without proof, readers are encouraged to try using $\mathcal{T} = Ke^{-\frac{\pi}{2}\sigma_y/\hbar}$),

$$\mathcal{T} = -i\sigma_y\mathcal{K}. \tag{3.41}$$

Further for spinless particles (or integer spin), $\mathcal{T}^2 = 1$, while for $S = \frac{1}{2}$ particles, $\mathcal{T}^2 = -1$.

3.3 SSH MODEL

3.3.1 Introduction

To make our concepts clear on the topological phase, and whether a model involves a topological phase transition, we apply it to the simplest model available in the literature. The Su-Schrieffer-Heeger (SSH) model denotes a paradigmatic one-dimensional (1D) model which hosts a topological phase. It also possesses a physical realization of polyacetylene, which is a long chain organic polymer (polymerization of acetylene) with the formula $[C_2H_2]_n$ (shown in Fig. 3.5). The C-C bond lengths are measured by NMR spectroscopy technique and are found to be 1.36 Å and 1.44 Å for the double, and the single bonds, respectively. The chain consists of a number of methyne (= CH−) groups covalently bonded to yield a 1D structure with each C-atom having a π electron. This renders the connectivity to the polymer chain.

Possibly intrigued by this bond length asymmetry, one can write down a tight binding Hamiltonian of such a system with two different hopping parameters for spinless fermions hopping along the single, and the double bonds. These *staggered* hopping amplitudes are represented by t_1 and t_2. Let us consider that the chain consists of N unit cells with two sites (that is, two C atoms) per unit cell, and denote these two sites as A and B. The hopping between A and B sites in a cell is denoted by t_1, while that from B to A across the cell is denoted by t_2. Because of the presence of a single π electron at each of the C atoms, the interparticle interaction effects are completely neglected. We shall show that the staggered hopping or the dimerization has got serious consequences on the topological properties of even such a simple model.

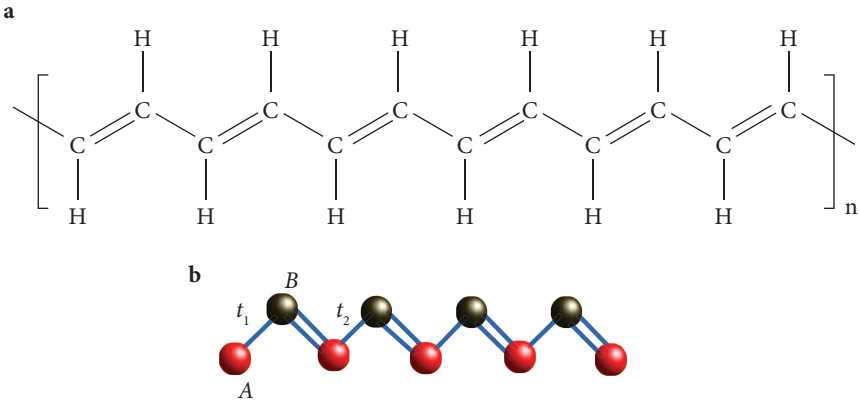


FIG. 3.5
(a) A polyacetylene chain with formula $(C_2H_2)_n$ is shown. (b) The SSH model as a paradigmatic model for Polyacetylene with dimerized hopping.

3.4 THE SSH HAMILTONIAN

The above considerations yield the following Hamiltonian,

$$\mathcal{H} = -t_1 \sum_{n=1}^{N} (c^\dagger_{n,\mathrm{A}} c_{n,\mathrm{B}} + \mathrm{h.c.}) - t_2 \sum_{n=1}^{N-1} (c^\dagger_{n,\mathrm{B}} c_{n+1,\mathrm{A}} + \mathrm{h.c.}). \tag{3.42}$$

For simplicity and concreteness t_1 and t_2 are assumed to be real and non-negative and $c^\dagger_{n,\alpha}(c_{n,\alpha})$ denotes electron creation (annihilation) operator at site n belonging to the α sublattice ($\alpha \in$ A, B).

It is clear that N denotes the total number of cells, which implies $M = 2N$ where M represents the total number of sites. Thus, for an open chain with M atoms, we have $t_M = 0$. On the site basis, the Hamiltonian can be explicitly written as

$$\mathcal{H} = (c_1^\dagger, c_2^\dagger, \ldots, c_M^\dagger) \begin{pmatrix} 0 & t_1 & 0 & . & . & 0 \\ t_1^* & 0 & t_2 & . & . & . \\ 0 & t_2^* & 0 & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 0 & . & . & . & . & 0 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ . \\ . \\ c_M \end{pmatrix}. \tag{3.43}$$

If M is an even number, then $t_{M-1} = t_1$, otherwise, $t_{M-1} = t_2$.

We shall show in the following that a staggered hopping is responsible for opening of a gap in the dispersion, and subject to the fulfillment of a particular condition, the nature of the gap can be topological.

To see that, let us study the band structure. We can Fourier transform the electron operators using,

$$c_\alpha(k) = \sum_n e^{ikn} c_{n\alpha} \quad (\alpha \in \mathrm{A}, \mathrm{B}). \tag{3.44}$$

This yields a tight binding Hamiltonian in the sublattice basis, namely, $(c_{k\mathrm{A}}, c_{k\mathrm{B}})$ as

$$\mathcal{H} = \sum_k c^\dagger_{k\alpha} h_{\alpha\beta}(k) c_{k\beta} \tag{3.45}$$

where,

$$h_{\alpha\beta}(\mathbf{k}) = \begin{pmatrix} 0 & t_1 + t_2 e^{-ik} \\ t_1 + t_2 e^{ik} & 0 \end{pmatrix} = \begin{pmatrix} 0 & f(k) \\ f^*(k) & 0 \end{pmatrix}. \tag{3.46}$$

The 2×2 structure of the matrix $h_{\alpha\beta}(k)$ allows us to write,

$$h_{\alpha\beta}(k) = \mathbf{d}(k) \cdot \sigma \tag{3.47}$$

where $\mathbf{d}(k)$ is a vector given by

$$\mathbf{d}(k) = (d_x(k), d_y(k), d_z(k)) = (t_1 + t_2 \cos k, t_2 \sin k, 0) \tag{3.48}$$

and $\sigma = (\sigma_x \sigma_y, \sigma_z)$ denote the Pauli matrices. The energy dispersion is given by

$$E(k) = \pm|\mathbf{d}(k)| = \pm\sqrt{(t_1 + t_2 \cos k)^2 + t_2^2 \sin^2 k}. \tag{3.49}$$

A slight manipulation of the terms inside the square root yields,

$$E(k) = \pm\sqrt{(t_1 - t_2)^2 + 4t_1 t_2 \cos^2 \frac{k}{2}} \tag{3.50}$$

where k is contained in the BZ, that is, $-\pi \leq k \leq +\pi$. The corresponding normalized eigenvectors are given by

$$|\psi_\pm\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} \pm e^{-i\phi(k)} \\ 1 \end{pmatrix} \tag{3.51}$$

where,

$$\phi(k) = \tan^{-1}\left(\frac{t_2 \sin k}{t_1 + t_2 \cos k}\right)$$

We shall explore a few representative cases to make our ongoing discussion clear, namely,

i. $t_2 = 0$: Extreme dimerized limit (see upper panel of Fig. 3.6).
ii. $t_1 > t_2$: Intra-cell hopping is larger than the inter-cell hopping.
iii. $t_1 = t_2$: Intra-cell hopping is the same as the inter-cell hopping (see lower panel of Fig. 3.6).
iv. $t_1 < t_2$: Intra-cell hopping is smaller than the inter-cell hopping.
v. $t_1 = 0$: Extreme dimerized limit (however, different than (i)).

We plot the band structure and the components of the **d**-vector as a function of the crystal momentum k (see Figs. 3.9–3.13). The purpose here is to define a bulk winding number, which is the topological

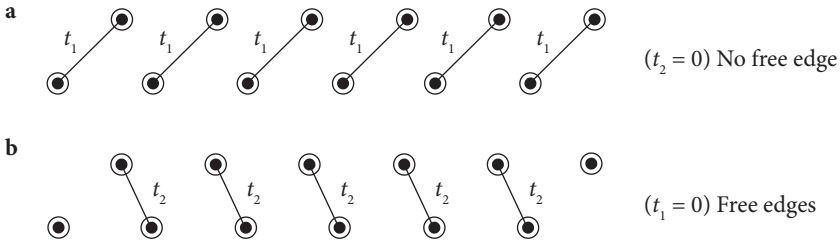


FIG. 3.6
The extreme dimerized limit for the trivial (upper panel), and the topological (lower panel) phases of the SSH model are shown.

invariant here. The plot d_x vs d_y for k in the BZ defines a surface (except for the critical case, $t_1 = t_2$). Whether the surface encloses the origin is decided on its topological properties. Furthermore, the unit vector, $\hat{\mathbf{d}}$ defines the direction of the d vector via, $\hat{\mathbf{d}} = \mathbf{d}/|\mathbf{d}|$. At half filling, the lower band is filled. The two bands are gapped by an amount $2\delta t$, where $\delta t = |t_1 - t_2|$ at $k = \pm\pi$. This is also an insulating phase. However, as we shall see that this phase is distinct from the case (ii).

3.4.1 Topological properties

The SSH chain hosts both the bulk and edge states. The distinction between the bulk and the edge can be understood from the real space analysis. The plot for the energy spectrum as a function of the ratio t_2/t_1 (see Fig. 3.7) shows that the zero modes start to appear just beyond the critical point. Prior to that, for $0 \leq t_2/t_1 \leq 1$, the system behaves like a trivial insulator with a bulk band gap. The gap closes at $t_1 = t_2$, and eventually for $t_2 > t_1$, the bulk gap opens again, however a pair of zero modes appear in the spectrum. These zero modes yield a topological characteristic of the phase. They originate from the two solitary C atoms that reside at the two edges of the chain. The fact that these zero modes indeed arise out of the edges is shown in Fig. 3.8 by plotting the probability densities, $|\psi_i|^2$ at all sites of the chain. The amplitudes at the left and the right edges are shown by red and blue colors, respectively and they vanish everywhere in the bulk of the chain.

Further to emphasize the robustness of the edge modes, we show the inverse participation ratio (IPR) defined by

$$\text{IPR} = \sum_{i=1}^{L} |\psi_i|^4 \tag{3.52}$$

IPR = 0 or 1 denotes the extended or the localized phases. However, these extreme values (namely, 0 and 1) can only be obtained in the thermodynamic limit ($L \to \infty$). Here we denote the edge modes coming from the two edges of the system by red in Fig. 3.7. In this figure, red denotes a finite value of the IPR. Evidently, the zero modes are seen to be localized.

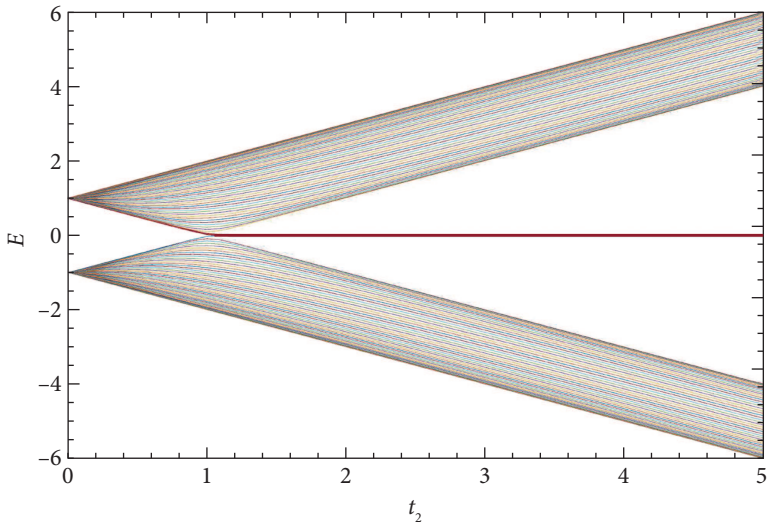


FIG. 3.7
The energy is plotted as a function of t_2. A zero mode exists for $t_2 \geq 1$ (in unit of t_1). A pair of zero modes is shown in red color, which implies a finite value of IPR (see text).

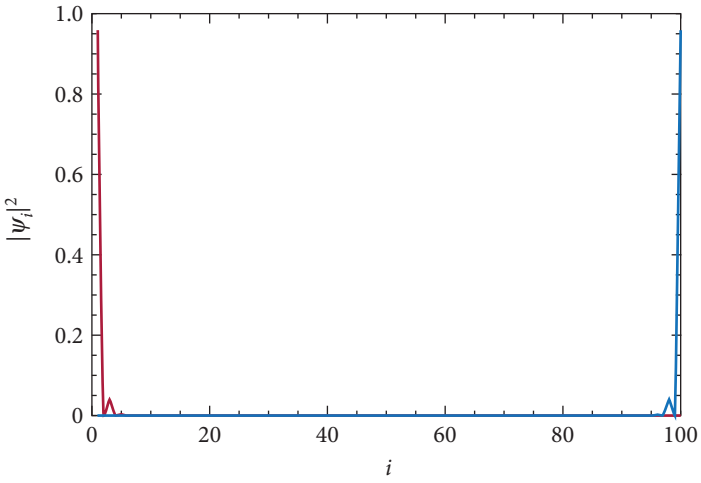


FIG. 3.8
The probability amplitude is plotted as a function of the sites of the chain. Here we have taken the length, $L = 100$.

Let us return to the behavior of the vector $\mathbf{d}(k)$. The components $((t_1 + t_2 \cos k), t_2 \sin k, 0)$ in the BZ defined by $-\pi \leq k \leq +\pi$ of $\mathbf{d}(k)$ denote the eigenstates with the energy spectrum given by

$E(k) = |\mathbf{d}(k)|$

Corresponding to one of the cases, namely $t_2 > t_1$, the vector $\mathbf{d}(k)$ winds about the origin, while for the other, $t_2 < t_1$, it does not. The origin of the $d_x - d_y$ plane is $\mathbf{d}(k) = 0$ and denotes the gapless (critical) condition. Based on the above information, it is possible to define a winding number, ν which would tell us whether the trajectory of $\mathbf{d}(k)$ winds the origin, as k is varied over the BZ. Thus, the winding number is capable of distinguishing the two seemingly equivalent (gapped) scenarios.

Mathematically, the winding number, ν can be written down using the unit $\hat{\mathbf{d}}$ vector defined via,

$$\hat{\mathbf{d}} = \frac{\mathbf{d}(k)}{|\mathbf{d}(k)|}. \tag{3.53}$$

One can now define ν using,

$$\nu = \frac{1}{2\pi} \int_{-\pi}^{+\pi} \left(\hat{\mathbf{d}} \times \frac{d}{dk} \hat{\mathbf{d}} \right)_z dk. \tag{3.54}$$

Let us justify how the above expression on the RHS denotes the winding number. Writing it more explicitly,

$$\nu = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\mathbf{d}(k)}{|\mathbf{d}(k)|} \times \frac{d}{dk} \frac{\mathbf{d}(k)}{|\mathbf{d}(k)|} dk = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\mathbf{d}(k)}{|\mathbf{d}(k)|} \times \left(\frac{d}{dk} \frac{\mathbf{d}(k)}{|\mathbf{d}(k)|} - \frac{\mathbf{d}(k) \frac{d}{dk} |\mathbf{d}(k)|}{|\mathbf{d}(k)|^2} \right) dk$$

since $\mathbf{d}(k) \times \mathbf{d}(k) = 0$

$$\nu = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\mathbf{d} \times \frac{d}{dk} \mathbf{d}(k)}{|\mathbf{d}(k)|^2} dk = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\mathbf{d}(k) \times \delta \mathbf{d}(k)}{|\mathbf{d}(\mathbf{k})|^2} dk$$

$$\frac{d}{dk} \mathbf{d}(k) = \frac{d}{dk} (\mathbf{d}_0 + k \delta \mathbf{d})$$

where in the last line we have used a Taylor expansion of $\frac{d}{dk} \mathbf{d}(k)$. From the definition of the cross product, $\mathbf{d}(k) \times \delta \mathbf{d}(k)$ is the angle (in radian) between $\mathbf{d}(k)$ and $\mathbf{d} + \delta \mathbf{d}$. Thus, integrating this over the BZ yields 2π, which when divided by 2π gives 1.

Another useful form for the winding number is given by

$$\nu = \frac{1}{2\pi i} \int_{-\pi}^{+\pi} dk \frac{d}{dk} \log f(k) \tag{3.55}$$

where, $f(k) = t_1 + t_2 e^{-ik}$. Thus,

$\log f(k) = \log(|f|) e^{i \arg(f)}$.

Consequently, the winding number becomes,

$$\nu = \frac{1}{2\pi i} \int_{-\pi}^{+\pi} dk \frac{d}{dk} \log(f(k)) = \frac{1}{2\pi} \arg(f)_{-\pi}^{+\pi} = 1 \text{ or } 0 \tag{3.56}$$

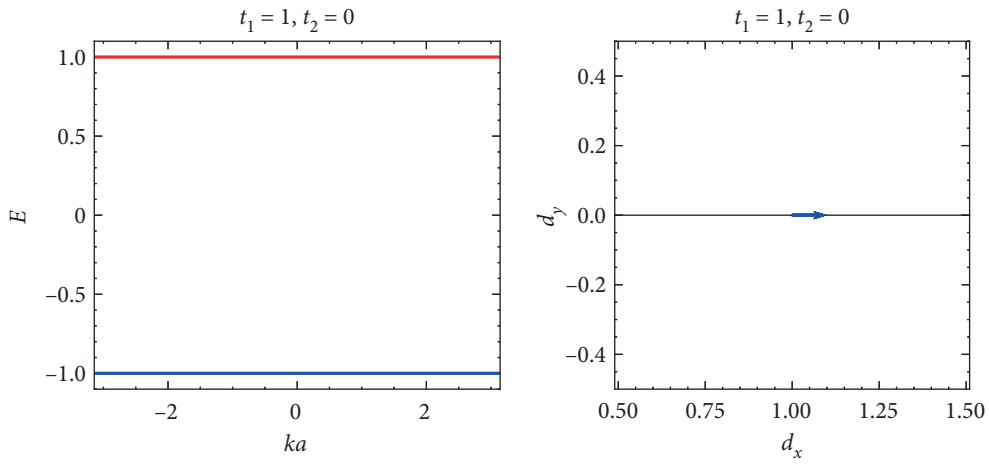


FIG. 3.9
The band structure and *d*-vectors are plotted corresponding to $t_1 = 1, t_2 = 0$.

depending on whether $\arg(f)$ falls in the region of integration, that is, it encloses the origin. The winding number is 1 for the topological phase, and 0 for the trivial phase.

We correlate the band structure with the corresponding winding number calculated above and show various cases as mentioned above. In the extreme dimerized limit, namely, $t_2 = 0$ and $t_1 = 1$, we get two flat bands at $E = \pm 1$, and the **d**-vector is simply shown by an arrow in Fig. 3.9. For $t_1 > t_2$, the spectrum is gapped, and corresponds to a trivial insulator because of the absence of winding of the **d**-vector as shown in Fig. 3.10. Furthermore, the undimerized tight binding chain ($t_1 = t_2$) is shown in Fig. 3.11 where the gap closes, and the tip of the **d**-vector which executes a circle (shown by red colour) in the $d_x - d_y$ plane, just touches the origin at the left, but does not wind it. The fourth case, namely, $t_1 < t_2$ again shows a spectral gap, but it is topological in nature as shown by the winding of the **d**-vector in Fig. 3.12. Finally, the other dimerized case, that is, $t_1 = 0$ and $t_2 = 1$ (see Fig. 3.13) again shows two flat bands at $E = \pm 1$, however the trajectory of the **d**-vector shown by the red circle in the right panel of Fig. 3.13 winds the origin, and hence denotes a topological scenario.

One can also define a Zak phase, Φ_Z (another geometric phase, similar to the Berry phase) defined via,[4]

$$\Phi_Z = i \oint \langle \psi | \nabla_k | \psi \rangle dk. \tag{3.57}$$

[4] Usually geometric phases that characterize the topological properties of the band structure play a crucial role in the band theory of solids. See Zak (1988).

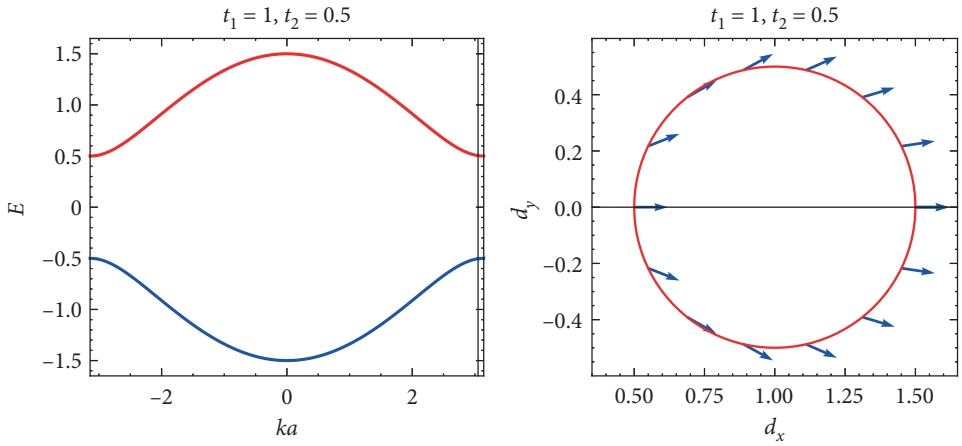


FIG. 3.10
The band structure and *d*-vectors are plotted corresponding to $t_1 = 1, t_2 = 0.5$.

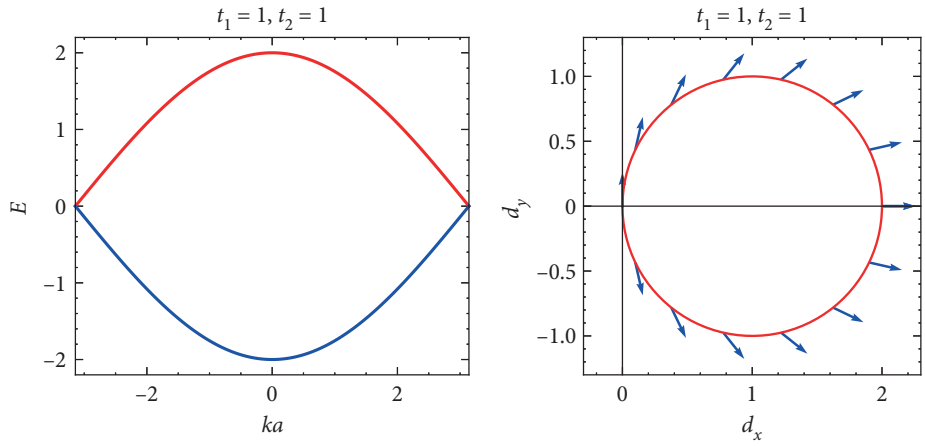


FIG. 3.11
The band structure and *d*-vectors are plotted corresponding to $t_1 = 1, t_2 = 1$, that is a simple tight binding chain.

Using

$$|\psi_{\pm}\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} \pm e^{-i\phi(k)} \\ 1 \end{pmatrix} \tag{3.58}$$

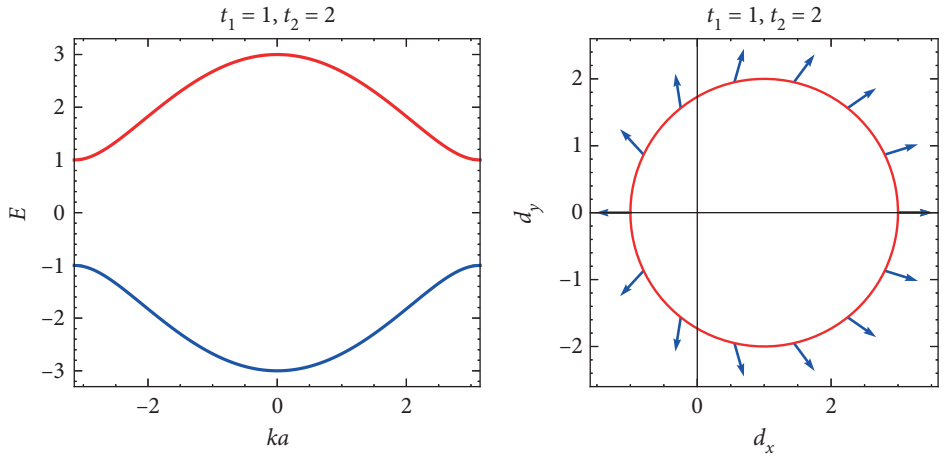


FIG. 3.12
The band structure and *d*-vectors are plotted corresponding to $t_1 = 1, t_2 = 2$.

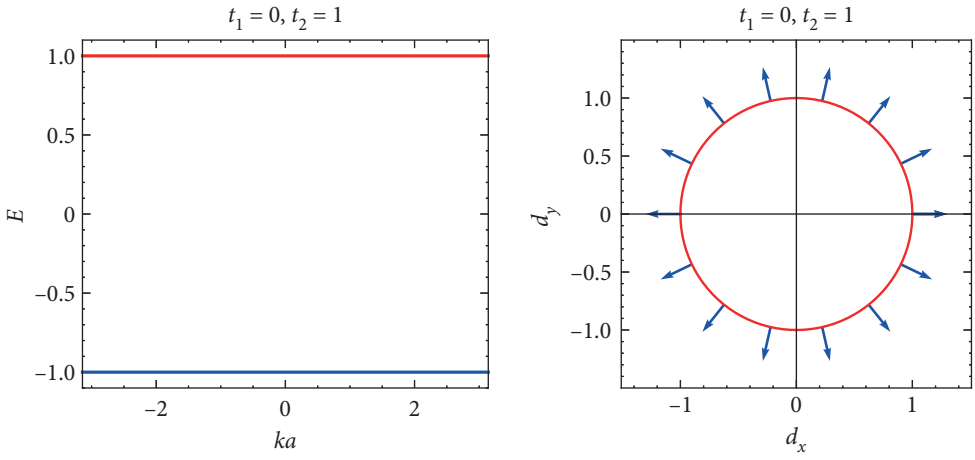


FIG. 3.13
The band structure and *d*-vectors are plotted corresponding to $t_1 = 0, t_2 = 1$.

$$\phi_Z = \frac{1}{2}\oint \frac{d}{dk}\phi(k)dk = \pm\pi \text{ or } 0, \tag{3.59}$$

which are the values respectively for $t_2 > t_1$ and $t_2 < t_1$.

Please note that we have obtained this result without plugging in the explicit form of $\phi(k)$ since the result should be independent of the form of $\phi(k)$. However, if we consider the explicit form of $\phi(k)$, namely,

$$\phi(k) = \tan^{-1}\left(\frac{t_2 \sin k}{t_1 + t_2 \cos k}\right)$$

it throws some subtlety that we need to take care of. If we are in the trivial phase, then the inverse tangent function is present in the first and fourth quadrants because of $t_1 > t_2$. Here, the function does not acquire any extra factor, because of which $\phi_Z = 0$. For the topological phase ($t_2 > t_1$), that is, when the inverse tangent function is in the second quadrant, it picks up a phase $\pi - \tan^{-1} x$, while in the third quadrant, the corresponding value is $\pi + \tan^{-1} x$. This can be seen from the sharp change of $\phi(k)$ twice in BZ, as seen from Fig. 3.14. This means that the inverse tangent function acquires an extra phase of 2π. This yields $\phi_Z = -\pi$. Frankly, the negative sign does not mean anything specific. It arises because we have chosen a positive sign for the wavefunction. Here, the winding number, and the Zak phase are related to $\nu = -\phi_Z/\pi$. Think about $\mathbf{d}(k)$ and ϕ_Z, both of which are obtained from the bulk of the material, yet give information about the edges of the system. This is traditionally referred to as the *bulk-boundary correspondence* (we shall discuss more about it later). It may also be noted that the behavior of $\phi(k)$ is smooth over the BZ, corresponding to the trivial case (see Fig. 3.15).

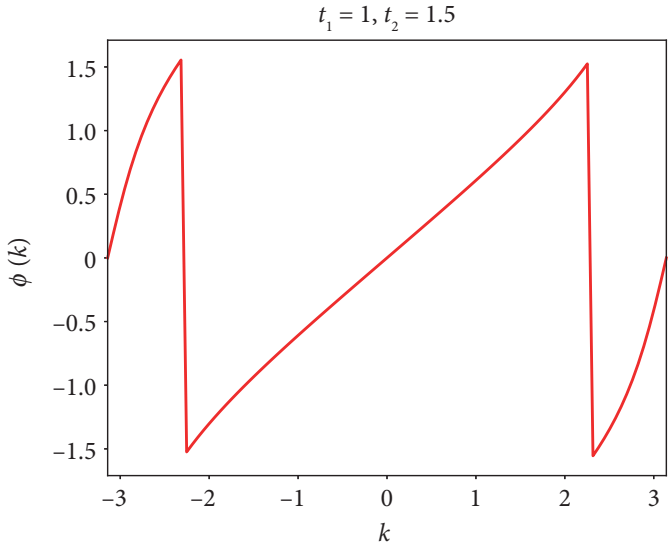


FIG. 3.14
ϕ is plotted as a function of k for the topological phase. There are abrupt jumps in the behavior of ϕ.

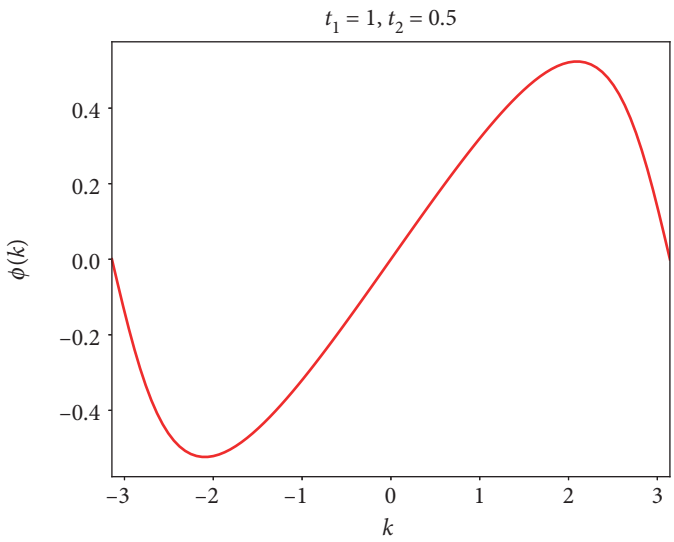


FIG. 3.15
ϕ is plotted as a function of k for the trivial phase. It smoothly varies from $-\pi$ to $+\pi$.

From the preceding discussion it is clear that the two apparently similar insulating phases are topologically different. We have shown that the winding numbers are different (finite in the topological phase and zero in the trivial one) and so are the Zak phases. However, what does it physically entail having different values of the winding number or the Zak phase? Suppose we smoothly deform the Hamiltonian corresponding to the SSH chain via tuning the hopping parameters, t_1 and t_2, we shall get two insulating phases for $t_2 > t_1$, and $t_2 < t_1$ all the while keeping the symmetries preserved and the band gap around $E = 0$ is finite. In tuning from one limit to the other, we need to cross the origin in the $d_x - d_y$ plane (note that $d_z \equiv 0$). This implies that at the intermediate stage, one would obtain the eigensolutions corresponding to the trajectory of $\mathbf{d}(k)$ in the $d_x - d_y$ plane through the origin. Thus, a smooth transition from one insulating phase to another is impossible without closing the gap, or satisfying the metallic condition, $t_1 = t_2$. Hence, quite apparently, there is a topological phase transition occurring here.

It is clear that the above discussion will be invalid if there is a z-component of the $\mathbf{d}(k)$ (or a term proportional to σ_z) is present in the system. A simple way to incorporate such a term is through the inclusion of an onsite potential. The onsite potential will destroy the zero modes, thereby making it meaningless to talk about the topological properties of the system. However, a disorder in the off-diagonal (hopping) term would retain the zero modes, and hence, the system should have a transition from a topological to a trivial phase. The reason for such a distinction between the diagonal and the off-diagonal terms arises because of certain fundamental symmetries that the system possesses. We shall discuss them below.

3.4.2 Chiral symmetry

In standard quantum mechanics, the symmetry of a Hamiltonian, $\mathcal{H}$ is represented by

$$U\mathcal{H}U^\dagger = \mathcal{H}, \text{ or } U\mathcal{H} = \mathcal{H}U, \text{ or } [\mathcal{H}, U] = 0, \tag{3.60}$$

where U denotes a unitary operator. This implies that U and $\mathcal{H}$ have the same eigenstates and hence can be diagonalized simultaneously. However, in general for topological insulators, such usual unitary

symmetries do not have interesting consequences. The reason is that it is mostly possible to make the Hamiltonian block diagonal, thereby reducing the problem to be confined to a single block. In the case of massless Dirac problems, one usually runs out of unitary symmetries, and is left with an irreducible block Hamiltonian which cannot be diagonalized. Thus, for the SSH model, there is a different symmetry, that is called the chiral symmetry, which is operative here.[5] Here $\mathcal{H}$ obeys,

$$\Gamma\mathcal{H}\Gamma^\dagger = -\mathcal{H}, \text{ or } \Gamma\mathcal{H} = \mathcal{H}\Gamma, \text{ or } \{\mathcal{H}, \Gamma\} = 0. \tag{3.61}$$

Here Γ is a unitary operator corresponding to the chiral symmetry. Instead of commuting, it anticommutes with the Hamiltonian. Furthermore, Γ is unitary and Hermitian, implying,

$$\Gamma = \Gamma^\dagger, \quad \text{or} \quad \Gamma^\dagger\Gamma = \Gamma^2 = \mathbf{1} \tag{3.62}$$

where $\mathbf{1}$ denotes the identity matrix. The above requirement raises the possibility that $\Gamma = e^{i\phi}$ where ϕ is an arbitrary phase. However, this possibility can be eliminated by redefining $\Gamma \to \Gamma e^{-i\phi/2}$. The second requirement is that Γ be a local operator. Thus, the matrix elements of Γ survive only within each unit cell, and between the cells they vanish. Hence, for the SSH model, the chiral symmetry is equivalent to the sublattice symmetry, which can be expressed through the projectors P_A and P_B corresponding to the A and B sublattices, namely,

$$P_\mathrm{A} = \frac{1}{2}(\mathbf{1} + \Gamma); \quad P_\mathrm{B} = \frac{1}{2}(\mathbf{1} - \Gamma). \tag{3.63}$$

It can be checked that

$$P_\mathrm{A} + P_\mathrm{B} = \mathbf{1}, \quad \text{and} \quad P_\mathrm{A} \cdot P_\mathrm{B} = 0.$$

It is also possible to show that

$$P_\mathrm{A}\mathcal{H}P_\mathrm{A} = P_\mathrm{B}\mathcal{H}P_\mathrm{B} = 0. \tag{3.64}$$

The consequence of the chiral symmetry results in a symmetric energy spectrum. That is, corresponding to an energy, E, there is a chiral partner with energy, $-E$. This fact can be seen from the following:

$$\begin{aligned} &\mathcal{H}\,|\psi\rangle = E\,|\psi\rangle \\ &\mathcal{H}\Gamma\,|\psi\rangle = -\Gamma\mathcal{H}\,|\psi\rangle = -\Gamma E\,|\psi\rangle = -E\Gamma\,|\psi\rangle. \end{aligned} \tag{3.65}$$

Of course, the above argument is true for $E \neq 0$. Since the SSH model hosts zero modes, one of them is a partner of the other. Besides, for all $E \neq 0$, $|\psi\rangle$ and $\Gamma|\psi\rangle$ correspond to distinct and orthogonal eigenstates, which suggests that every non-zero eigenstate of $\mathcal{H}$ derives an equal contribution from both the sublattices, that is,

$$\langle\psi|\Gamma|\psi\rangle = \langle\psi|P_\mathrm{A}|\psi\rangle - \langle\psi|P_\mathrm{B}|\psi\rangle = 0. \tag{3.66}$$

Whereas, for $E = 0$, $\mathcal{H}\,|\psi\rangle = 0$. Thus,

$$\mathcal{H}P_{\mathrm{A/B}}\,|\psi\rangle = \mathcal{H}\,[|\psi\rangle + \Gamma\,|\psi\rangle] = 0. \tag{3.67}$$

[5] In condensed matter physics, bipartite systems with nearest neighbour hopping, that is when hopping connects sites with opposite sublattices, obey chiral symmetry.

Thus, the zero energy eigenstates are eigenstates of Γ, and hence are chiral symmetric partners of themselves. It is also due to the robustness of the chiral symmetry, the zero modes are robust.

One can also define operators for the sublattice symmetry, namely,

$$\Sigma_z = P_{\mathrm{A}} - P_{\mathrm{B}}. \tag{3.68}$$

It can be shown that $\Sigma_z \mathcal{H} \Sigma_z = -\mathcal{H}$ which is similar to the symmetry relation stated earlier, $\Gamma \mathcal{H} \Gamma = -\mathcal{H}$. Thus, the chiral symmetry of the SSH model is a re-statement of the sublattice symmetry of the Hamiltonian.

In simple language, the chiral symmetry operator, Γ is actually the z-component of the Pauli matrix, σ_z which yields,

$$\sigma_z \mathcal{H} \sigma_z = -\mathcal{H}. \tag{3.69}$$

A direct multiplication of the three matrices on the LHS can be performed for the proof. We have presented above a crisp description of the SSH model, which even being simple enough, possesses both trivial and topological phases where the latter shows up via the presence of robust[6] zero energy edge modes, along with a finite value of the winding number. Further, the band structure shows that a phase transition occurs from a trivial to a topological phase (or vice versa) through a gap closing point, where the staggered hopping amplitudes are equal ($t_1 = t_2$).

3.5 KITAEV MODEL

3.5.1 Introduction

In a one-dimensional dimerized chain comprising two atoms per unit cell within a tight binding approximation, which is known as the Su-Schrieffer-Heeger (SSH) model, there are localized zero energy modes. These modes are robust to adiabatic deformation of the Hamiltonian. Thus, inclusion of disorder (or defect) that respects the chiral symmetry does not harm these zero modes. However, it is difficult to satisfy such a chiral symmetry in real physical systems.

In an analogous model, however, with additional ingredients, the so-called Kitaev model in 1D the symmetry that protects the topological features is much more physical than the chiral symmetry. In fact, it is the particle-hole symmetry that plays a crucial role here, is also inherent to the mean field description of superconductors. Thus, the Kitaev chain is a more realistic model to access the robust zero energy edge modes, the so-called Majorana zero modes. A brief description will be included at the end of this chapter to discuss their physical realizability in experiments, and their possible applications in using the degenerate quantum states at zero energy for storing quantum information.

[6] The edge modes are robust as long as the chiral symmetry is intact. These mean field solutions correspond to the Bogoliubov-de Gennes equations which yield an equivalent description as the BCS theory for conventional superconductors.

The system comprises a 1D p-wave superconductor introduced by Kitaev (2003) where the superconducting correlations occur between spinless (or spin-polarized) fermions at neighboring sites in the chain, as opposed to the onsite pairing discussed in the context of the more familiar BCS (s-wave) superconductors. In this model, the Majorana fermions (MF) which arise as real solutions to the Dirac equation and are their own antiparticles, emerge in a simple and intuitive fashion. The search for Majorana particles remained inconclusive as they have never been observed in nature. Thus, the Kitaev model also serves as a platform for realizing Majorana particles in condensed matter systems. In the following, let us begin with the description of a two-site system which essentially contains all the information that we need.

3.5.2 Two-site Kitaev chain

In this section, we shall study the edge and the bulk properties of a two-site Kitaev chain with a view to exploring the topological properties of the excitation spectrum. At each site, there is a single spinless fermion which is coupled to a p-wave superconductor. It may be noted that by virtue of the fermions being spinless (or spin polarized), a conventional s-wave pairing for fermions belonging to the same site is not possible. The Hamiltonian for such a system is written as

$$\mathcal{H}_{2\text{site}} = -\mu c_1^\dagger c_1 - \mu c_2^\dagger c_2 - t(c_1^\dagger c_2 + \text{h.c.}) + \Delta(c_1^\dagger c_2^\dagger + \text{h.c.}), \tag{3.70}$$

where μ denotes the chemical potential, t is the hopping term among the neighboring sites, and Δ is the p-wave-superconducting order parameter. We assume that all of these energy scales are real and positive. In fact, Δ is usually a complex quantity with an amplitude and a phase. However, taking it either real or complex does not have much consequence in our discussion.

The Hamiltonian can be written on the basis spanned by $\{c_1^\dagger, c_1, c_2^\dagger, c_2\}$. Explicitly write it in this basis, we get,

$$\mathcal{H}_{2\text{site}} = \begin{pmatrix} c_1^\dagger & c_1 & c_2^\dagger & c_2 \end{pmatrix} \begin{pmatrix} -\mu & 0 & -t & \Delta \\ 0 & \mu & -\Delta & t \\ -t & -\Delta & -\mu & 0 \\ \Delta & t & 0 & \mu \end{pmatrix} \begin{pmatrix} c_1 \\ c_1^\dagger \\ c_2 \\ c_2^\dagger \end{pmatrix}. \tag{3.71}$$

In the Dirac notation, we can use the particle-hole representation at each site by writing, $c_1^\dagger = |1e\rangle$, $c_1 = \langle 1h|$, $c_2^\dagger = |2e\rangle$ and $c_2 = \langle 2h|$ where 1 and 2 refer to the sites and e and h refer to the electron (particle) and the hole states, respectively. The electron-hole description that is natural to a superconductor becomes apparent on this basis, and the Hamiltonian can be written as

$$\begin{aligned} \mathcal{H}_{2\text{site}} = &-\mu(|1e\rangle\langle 1e| + |1h\rangle\langle 1h| + |2e\rangle\langle 2e| + |2h\rangle\langle 2h|) \\ &- t(|1e\rangle\langle 2e| - |1h\rangle\langle 2h| + \text{h.c.}) + \Delta(|1e\rangle\langle 2h| - |2e\rangle\langle 1h| + \text{h.c.}). \end{aligned} \tag{3.72}$$

The diagonalization of the above 4×4 Hamiltonian yields the eigensolutions. The eigenvalues are given by

$$E = (t + \sqrt{(\Delta^2 + \mu^2)}),\ (t - \sqrt{(\Delta^2 + \mu^2)}),\quad -(t - \sqrt{(\Delta^2 + \mu^2)}),\ -(t + \sqrt{(\Delta^2 + \mu^2)}). \tag{3.73}$$

The corresponding eigenvectors are

$$\begin{pmatrix} \frac{(t+\sqrt{(\Delta^2+\mu^2)})}{\Delta} - \frac{(\mu+t)}{\Delta} \\ 1 \\ \frac{(\mu+t)}{\Delta} - \frac{(t+\sqrt{(\Delta^2+\mu^2)})}{\Delta} \\ 1 \end{pmatrix}, \begin{pmatrix} \frac{(t-\sqrt{(\Delta^2+\mu^2)})}{\Delta} - \frac{(\mu+t)}{\Delta} \\ 1 \\ \frac{(\mu+t)}{\Delta} - \frac{(t-\sqrt{(\Delta^2+\mu^2)})}{\Delta} \\ 1 \end{pmatrix},$$
$$\begin{pmatrix} \frac{(-t+\sqrt{(\Delta^2+\mu^2)})}{\Delta} - \frac{(\mu-t)}{\Delta} \\ 1 \\ \frac{(-t+\sqrt{(\Delta^2+\mu^2)})}{\Delta} - \frac{(\mu-t)}{\Delta} \\ 1 \end{pmatrix}, \begin{pmatrix} \frac{(-t-\sqrt{(\Delta^2+\mu^2)})}{\Delta} - \frac{(\mu-t)}{\Delta} \\ 1 \\ \frac{(-t-\sqrt{(\Delta^2+\mu^2)})}{\Delta} - \frac{(\mu-t)}{\Delta} \\ 1 \end{pmatrix}. \tag{3.74}$$

Writing down the Hamiltonian in the form as it appears as in Eq. (3.72) has the advantage that it can be trivially extended for a chain containing N sites and can be written as in the following:

$$\begin{aligned} \mathcal{H} = \sum_n [&-\mu\,(|n,e\rangle\langle n,e| + |n,h\rangle\langle n,h|) \\ &- t\,(|n,e\rangle\langle n+1,e| - |n,h\rangle\langle n+1,h| + \text{h.c.}) \\ &+ \Delta\,(|n,e\rangle\langle n+1,h| - |n+1,e\rangle\langle n,h| + \text{h.c.})] \end{aligned} \tag{3.75}$$

$$\begin{aligned} = \sum_n [&-\mu\,(|n,e\rangle\langle n,e| + |n,h\rangle\langle n,h|) \\ &+ (-t+\Delta)\,(|n,e\rangle\langle n+1,e| + \text{h.c.}) \\ &+ (t-\Delta)\,(|n,h\rangle\langle n+1,h| + \text{h.c.})]. \end{aligned} \tag{3.76}$$

The above form appears as a tight binding Hamiltonian with two atoms per unit cell. A Fourier transform can be performed using

$$c_k = \sum_j e^{ikj} c_j; \quad c_k^\dagger = \sum_j e^{-ikj} c_j^\dagger, \tag{3.77}$$

where k is the wave vector defined within the Brillouin zone (BZ) $[-\pi : +\pi]$. Using the Fourier transformed operators and noting that

$$-e^{-ik} c_k^\dagger c_{-k}^\dagger = +e^{-ik} c_k^\dagger c_{-k}^\dagger$$

one gets in the $\{c_k, c_k^\dagger\}$ basis,

$$\begin{aligned} \mathcal{H} &= (-\mu - 2t\cos k)\sigma_z - (2\Delta \sin k \sigma_y) \\ &= \mathbf{d}(k).\sigma, \end{aligned} \tag{3.78}$$

where $\boldsymbol{\sigma}$ denotes the components of the Pauli matrix, and

$$\mathbf{d}(k) = (0, -2\Delta \sin k, -\mu - 2t\cos k)$$

The form is similar to a 2D massless Dirac Hamiltonian (similar to that of graphene). However, it needs to be noted that the components of $\boldsymbol{\sigma}$ do not represent spin degrees of freedom (recall that we

are considering spinless fermions), and instead denotes particle-hole degrees of freedom. It is now easy to diagonalize the 2×2 Hamiltonian whose eigensolution is given by

$$E_k = \pm\sqrt{(-\mu - 2t\cos k)^2 + 4\Delta^2 \sin^2 k} \tag{3.79}$$

with the eigenvectors as

$$\begin{pmatrix} u(k) \\ v(k) \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{4t^2\cos^2 k + 4\mu t\cos k + 4\Delta^2\sin^2 k + i\mu^2}}{2\Delta\sin k} - \frac{(\mu + 2t\cos k)i}{2\Delta\sin k} \\ 1 \end{pmatrix};$$
$$\begin{pmatrix} \frac{-\sqrt{4t^2\cos^2 k + 4\mu t\cos k + 4\Delta^2\sin^2 k + i\mu^2}}{2\Delta\sin k} - \frac{(\mu + 2t\cos k)i}{2\Delta\sin k} \\ 1 \end{pmatrix}. \tag{3.80}$$

It may be noted that the eigenvectors are in the standard format of the mean field solution of a superconductor. Instead of solving the model for its topological characteristics, we shall discuss the symmetry properties first in the following.

3.5.3 Particle-hole symmetry of the Kitaev model

Generally speaking, particle-hole symmetry implies that for a state defined by energy and momentum (E, k), there is always a partner with $(-E, -k)$. This is reflected in the symmetry of the diagonal elements as seen in the Hamiltonian written in the k-space, namely in Eq. (3.78) where the cosine term is symmetric under the transformation of $k \to -k$. This is also an artefact of the p-wave-superconducting term written in the mean field form which ensures particle-hole symmetry of the model.

Let us discuss the particle-hole symmetry of the Kitaev model in detail. The operator that tests the particle-hole symmetry, which we denote by $\Sigma = \sigma_x\mathcal{K}$ where σ_x is the x-component of the Pauli matrix (that acts on the particle-hole degrees of freedom) and $\mathcal{K}$ is the complex conjugation operator, such that $\mathcal{K}i\mathcal{K} = -i$. Σ is an antiunitary operator which transforms the Hamiltonian in Eq. (3.78) as

$$\Sigma\mathcal{H}\Sigma = (\sigma_x\mathcal{K})\mathcal{H}(k)(\sigma_x\mathcal{K}) = -\mathcal{H}, \tag{3.81}$$

where $(\sigma_x\mathcal{K})^2 = 1$. For closer introspection, let us consider a wavefunction, ψ at certain given values of energy and momentum, (E, k). As per the discussion above, there will be a particle-hole partner ψ' with energy and momentum as $(-E, -k)$. Thus,

$$\psi' = \sigma_x\mathcal{K}\psi. \tag{3.82}$$

We can write,

$$(\sigma_x\mathcal{K})\mathcal{H}(k)\psi = E\sigma_x\mathcal{K}\psi. \tag{3.83}$$

By inserting $(\sigma_x\mathcal{K})(\sigma_x\mathcal{K})$ $(= 1)$ between $\mathcal{H}(k)$ and ψ, one gets,

$$(\sigma_x\mathcal{K})\mathcal{H}(k)(\sigma_x\mathcal{K})(\sigma_x\mathcal{K})\psi = E\sigma_x\mathcal{K}\psi. \tag{3.84}$$

Substituting ψ' from Eq. (3.82),

$$(\sigma_x\mathcal{K})\mathcal{H}(k)(\sigma_x\mathcal{K})\psi' = E\psi'. \tag{3.85}$$

Table 3.1
Different choices for the parameters, μ, t, and Δ for showing different phases of the Kitaev model.

μ	t	Δ
0	1	1
1	1	0
1	0	1
1	1	1
1	0.5	0
1	0.3	0
1	0.5	1
1	0	0

Or,

$$\sigma_x \mathcal{H}^*(k)\sigma_x \psi' = E\psi'. \tag{3.86}$$

Since $\sigma_x \mathcal{H}^*(k)\sigma_x = -\mathcal{H}(-k)$,

$$\mathcal{H}(-k)\psi' = -E\psi'. \tag{3.87}$$

This proves the existence of a time reversed partner under the particle-hole transformation.

It may be noted from above that corresponding to $E > 0$, the partners are orthogonal, that is, $\langle\psi|\psi'\rangle = 0$ since they have different eigenvalues. However, for $E = 0$, the situation is more interesting. ψ and ψ' are now identical. These are called the Majorana modes for which $\sigma_x \mathcal{K}\psi = \psi$.

For the sake of convenience and concreteness, we list the parameters μ, t and Δ in Table 3.1 for discussing the topological and the trivial cases.

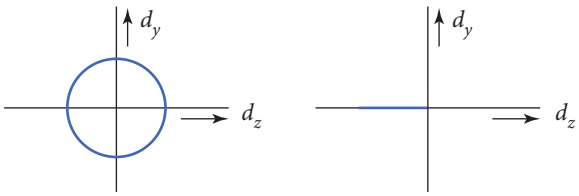


FIG. 3.16
The *d*-vectors corresponding to the trivial and the topological phases are plotted.

This brings us to the notion of topology in the Kitaev model that is familiar in systems where the boundary behaves differently compared to the bulk of the system. This will become clear if we look at the $\mathbf{d}(k)$ vector defined in the yz-plane where the components are $d_y(k) = -2\Delta \sin k$ and $d_z(k) = -\mu - 2t\cos k$. The tip of the $\mathbf{d}$ vector may encircle the origin depending on the values of the parameter yielding an integer winding number.[7] There arises two situations governed by the value of d_z, namely, (i) $d_z =$ positive which results from $\mu > -2t$, and (ii) d_z to be negative for $\mu < -2t$. Representations from each of the cases may be taken as

a. $\mu = 0, t = \Delta = 1$ (a priori, this is the topological limit),
b. $\mu = 1, t = \Delta = 0$ (this is the trivial phase).

Both are insulating phases, however it is easy to see that Fig. 3.16(a) corresponds to winding of the $\mathbf{d}$-vector in the yz-plane surrounding the origin, while for Fig. 3.16(b), $\mathbf{d}$ is a constant vector in the trivial case, and hence cannot wind.

Furthermore, the energy band structure (E *vs* k) for various choices of the parameters, μ, t, and Δ as they appear in Table 3.1 are shown in Fig. 3.17. Both the top left panel and the bottom right panel of Fig. 3.17 denote flat bands and finite spectral gap everywhere in the BZ; however, the former denotes a topological gap, while the latter represents a trivial scenario. Dispersions corresponding to the other parameter values are included for academic interest.

[7] Winding number denotes how many times the origin is wound around by a vector.

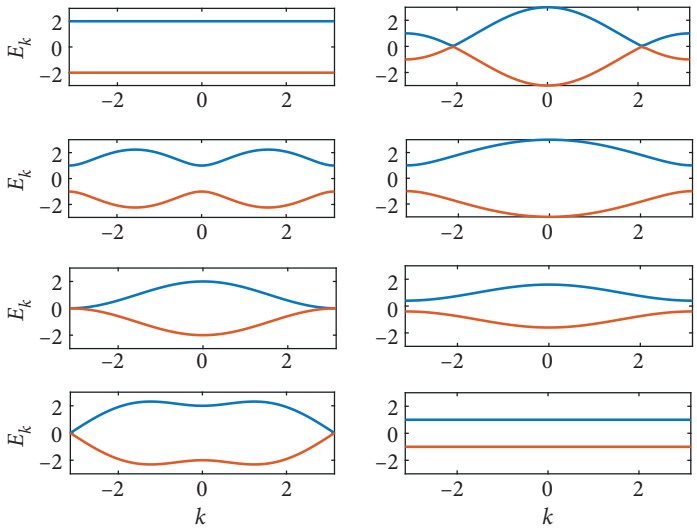


FIG. 3.17
The eigenvalues for a two-site Kitaev model are plotted. Different choices of parameters for all the plots are to be correlated with the values of $\mu, t,$ and Δ given in Table 3.1. The top left and the bottom right correspond to the topological and the trivial phases of the model. Other parameter values are for academic interest.

3.5.4 Winding number

We have already discussed that the Kitaev model is invariant under particle-hole transformation, which states that an eigenstate defined by energy and momentum (E, k) will always have a partner with $(-E, -k)$. We have not made another symmetry explicit earlier that there is also time reversal symmetry, $\mathcal{T}$ defined by $\mathcal{K}$, which is simply the complex conjugation operator.[8] This means that the product of the two, that is, the particle-hole and the time reversal operations, namely, $\Sigma\mathcal{T}$, known as the chiral symmetry, is a valid symmetry operation for the Kitaev model. Further, square of all of these symmetry operations, namely, Σ^2 and $\mathcal{T}^2$ result in $+1$, and hence, according to the tenfold classification introduced by Altland *et al.* and others, the symmetry belongs to the BDI class (Altland and Zirnbauer, 1997; and Ryu *et al.*, 2010).

Possessing a chiral symmetry enables defining a topological invariant called the winding number (we have seen this for the SSH model) which is written as

$$\nu = \frac{1}{2\pi}\int_{BZ} dk|\frac{d}{dk}h(k)|, \tag{3.88}$$

[8] In presence of spin degree of freedom, $\mathcal{T}$ has a more complicated form, namely, $\mathcal{T} = i\sigma_y\mathcal{K}$ where σ_y denotes the real spin degree of freedom.

where

$$h(k) = \tan^{-1}\left(\frac{d_z}{d_y}\right) = \tan^{-1}\left(\frac{\mu + 2t\cos k}{2\Delta \sin k}\right).$$

It can be easily checked that corresponding to the topological case, namely, $t = \Delta = 1$ and $\mu = 0$, the winding number, ν is equal to 1, while for the trivial case $\mu = 1$, $t = \Delta = 0$, $\nu = 0$. The same inferences have already been drawn from the behavior of the d-vector, that is, it winds around the origin for the topological case, and is merely a constant vector for the trivial case. A non-trivial topological (non-zero ν) phase is resilient against perturbations, such as disorder, defect, etc., that does not violate the chiral symmetry of the Hamiltonian.

3.5.5 Majorana fermions in the Kitaev model

E. Majorana postulated their existence in 1937 (Majorana, 2008). It is conjectured that the neutrinos denote an example of such an elementary particle whose antiparticles are thought to be the same as these particles. However, no signature of these particles has been realized in experiments so far. On the other hand, the research on Majorana fermions have gained momentum in the field of condensed matter physics and in particular, topological superconductors, semiconductors, etc.

Before we discuss how Majorana fermions enter the Kitaev model, let us understand what Majorana fermions are. One can write fermion operators, namely, c and $c^\dagger$ in terms of two Majorana fermion operators in the following way,

$$c_i = \frac{1}{2}(\gamma_{1i} - i\gamma_{2i}); \quad c_i^\dagger = \frac{1}{2}(\gamma_{1i} + i\gamma_{2i}) \tag{3.89}$$

where γ_1 and γ_2 are two Majaorana fermions defined at each site i. They are the conjugates of their own, namely,

$$\gamma_1 = \gamma_1^\dagger; \quad \gamma_2 = \gamma_2^\dagger. \tag{3.90}$$

The above properties can easily be verified by inverting the relations in Eq. (3.89). Thus, the Majorana fermion operators represent them to be their own antiparticles. Because of this property, a single Majorana mode is never "filled" (for example $|1\rangle$) or "empty" (that is $|0\rangle$), unlike the way we define usual fermions. This will become clearer as we go along.

The other properties of the Majorana modes can be stated as in the following.

$$\{\gamma_1, \gamma_2\} = \{\gamma_1^\dagger, \gamma_2^\dagger\} = 0 \tag{3.91}$$

$$\gamma_1^2 = \gamma_2^2 = 1 \tag{3.92}$$

for all sites i. As one can see, even though the anticommutation relations are similar to that of fermions, they are quite distinct from the usual fermion operators and hence special. For example, the same relationship for the fermion operators yield,

$$\{c, c^\dagger\} = 1 \tag{3.93}$$

$$c^2 = (c^\dagger)^2 = 0. \tag{3.94}$$

Besides these, there are two other properties which we wish to highlight. The complex conjugation of the Majorana operators can be stated as

$$\mathcal{K}\gamma\mathcal{K} = \gamma; \quad \mathcal{K}\gamma^{\dagger}\mathcal{K} = -\gamma^{\dagger}. \tag{3.95}$$

Next, we discuss the parity of the Majorana zero modes. A few details on these modes will be discussed shortly afterwards. Consider a Kitaev chain that is sufficiently long, such that one can safely ignore any overlap between the Majoranas that exist at the two ends of the chain. The energy spectrum has two zero energy-bound states. The corresponding quasiparticle operators are, say, γ_1 and γ_2. Let us combine them to form a fermions operator, using

$$c = \frac{1}{2}(\gamma_1 + i\gamma_2). \tag{3.96}$$

Inverting the above relation yields

$$\gamma_1 = (c + c^{\dagger}); \quad \gamma_2 = \frac{1}{i}(c - c^{\dagger}). \tag{3.97}$$

Thus, the Majorana operators are a superposition of the fermion creation and annihilation operators. The corresponding fermionic state can be occupied or unoccupied. The bound states have zero energy and hence these are degenerate. These two states differ by fermion number parity. The Kitaev model, being a mean field model for superconductors, does not conserve the particle number [$U(1)$ gauge symmetry]; however, the fermion number parity remains conserved since the pairing term (that is, Δ) adds or removes particles in pairs. In a normal superconductor, the ground state always has even parity for the fermion number. With an odd number of fermions, an unpaired electron results in an energetically less favorable state.

The Kitaev chain is quite unlike a normal superconductor where the non-trivial phase hosts two ground states, each with a different parity, namely an even fermion number parity and an odd fermion number parity. We can define the fermion number parity operator as

$$P_f = 2c^{\dagger}c - 1 = i\gamma_1\gamma_2 \tag{3.98}$$

which has eigenvalues ± 1 for the two zero energy modes.

For N such as Kitaev chains, there are $2N$ Majorana bound states. One can group them into N pairs, where each pair of Majorana operators γ_{2j-1} and γ_{2j} can be combined into a fermion operator c_j. The ground state now has a degeneracy of 2^N. Among these 2^N states, half of the them will have one parity (say, even), and the other half will have the opposite parity (odd). Now, the fermion parity operator for the N-chain system is the product of the operators $2c_j^{\dagger}c_j - 1$ for each of the pairs, such that

$$P_f = i^N\gamma_1\gamma_2 \ldots\ldots\ldots\gamma_{2N}. \tag{3.99}$$

Thus, one has $\frac{1}{2}(2^N) = 2^{N-1}$ states of even and odd parity. This can be easily demonstrated in the following way. One can compute the expectation value of the fermion number parity operator $i\gamma_1\gamma_2$ within the ground state $|\phi_g\rangle$, which yields,

$$\begin{aligned} \langle\phi_g|\, i\gamma_1\gamma_2\, |\phi_g\rangle &= 1 \quad \text{for } c^{\dagger}c = 1 \\ &= -1 \quad \text{for } c^{\dagger}c = 0. \end{aligned} \tag{3.100}$$

Suppose we take a simple one-site Hamiltonian, $\mathcal{H} = \mu c^\dagger c$. Writing in terms of the Majorana operators yield, $\mathcal{H} = \frac{1}{2}\mu(1 + i\gamma_1\gamma_2)$. So this tells us that each fermionic site has to accommodate two Majorana fermions, as opposed to one (or zero) fermion.

Let us extend our discussion to a two-site Kitaev chain. The Hamiltonian is,

$$\mathcal{H} = -\Delta(c_i^\dagger c_{i+1} + c_{i+1}^\dagger c_i + c_i c_{i+1} + c_{i+1}^\dagger c_i^\dagger), \tag{3.101}$$

where we have taken $\Delta = t$ and $\mu = 0$. There are two eigenstates of even parity, and two states of odd parity corresponding to the two eigenvalues Δ and $-\Delta$ respectively. For example, the eigensolutions, $|u^e\rangle$, λ^e with even parity,

$$|u_\pm^e\rangle = \frac{1}{\sqrt{2}}(1 \pm c_{i+1}^\dagger c_i^\dagger)\,|0\rangle; \quad \text{for } \lambda_\pm^e = \mp\Delta. \tag{3.102}$$

Similarly, for the odd solutions, $|u^o\rangle$, λ^o,

$$|u_\pm^o\rangle = \frac{1}{\sqrt{2}}(c_i^\dagger \pm c_{i+1}^\dagger)|0\rangle; \quad \text{for } \lambda_\pm^o = \mp\Delta. \tag{3.103}$$

If we translate it into the language of Majorana fermions,

$$\mathcal{H} = -2i\Delta\gamma_{2,i}\gamma_{1,i+1}. \tag{3.104}$$

These Majorana fermions can again be combined to write down the new fermion operators, namely, the d-operators[9] defined by

$$d_i = \frac{1}{\sqrt{2}}(\gamma_{1,i+1} + i\gamma_{2,i}); \quad i\gamma_{2,i}\gamma_{1,i+1} = 1 - d_i^\dagger d_i. \tag{3.105}$$

This allows the Hamiltonian to be written in terms of these d-operators.

$$\mathcal{H} = 2td_i^\dagger d_i - t = 2t\left(d_i^\dagger d_i - \frac{1}{2}\right) = 2t\left(d_i^\dagger d_i - \frac{1}{2}\right) + 0.a^\dagger a, \tag{3.106}$$

where $a = \frac{1}{2}(\gamma_{2,1} + i\gamma_{1,N})$ which comprises the degrees of freedom that are ignored in the Hamiltonian. Note that the ground state is doubly degenerate, since the states with $\langle a^\dagger a\rangle = 0$ and 1 have the same energy. Suppose N is odd, the sate $|0\rangle$ has even parity, which means that it has even number of a fermions. In the same way, the other ground state, namely, $|1\rangle$, has one fermion and is an odd-parity state, Thus, these two states are quite intriguing, in the sense that they can be considered as the superposition of two Majaorana fermions that reside at the two edges of the chain. The occurrence of these Majorana zero modes and their degenerate nature are artefacts of the topological properties of the model. Evidently, the Hamiltonian commutes with $\gamma_{1,i}$ and $\gamma_{2,i+1}$, that is,

$$\gamma_{1,i}\,|u_+^e\rangle = \frac{i}{2}(c_i^\dagger + c_{i+1}^\dagger)\,|0\rangle; \;\; \gamma_{2,i+1}\,|u_+^e\rangle = \frac{1}{2}(c_{i+1} + c_i)\,|0\rangle = -i\gamma_{1,i}\,|u_+^e\rangle \tag{3.107}$$

which says that the states are identical and differ only by a phase factor.

Let us now consider a chain of N sites (in Fig. 3.18 we show $N = 5$). At each site there is a fermion, and hence there are two Majorana modes, namely, γ_{2j-1} and γ_{2j}. These are called the domino tiles of the

[9] The d-operators here will have to be distinguished from the **d**-vectors discussed earlier.

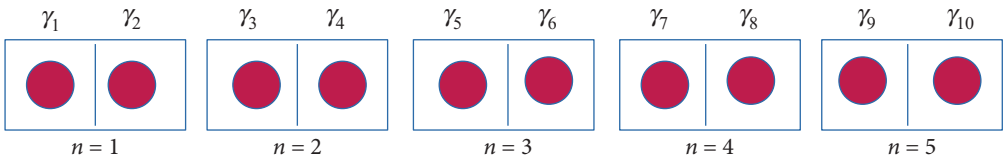


FIG. 3.18
The dominoe model is sketched.

Majorana. There are 10 Majorana modes denoted by $\gamma_1 \ldots\ldots \gamma_{10}$. The Kitaev Hamiltonian written in terms of the Majorana fermions,[10]

$$\mathcal{H} = -\frac{i\mu}{2}\sum_{j=1}^{N}(1 + i\gamma_{1,j}\gamma_{2,j}) - \frac{i}{4}\sum_{j=1}^{N-1}\left[(\Delta + t)\gamma_{2,j}\gamma_{1,j+1} + (\Delta - t)\gamma_{1,j}\gamma_{2,j+1}\right] \tag{3.108}$$

This has a formal similarity with the two coupled SSH chains with an onsite energy, $-\frac{i\mu}{2}$, and hopping terms as $i(\Delta + t)$ and $i(\Delta - t)$. However, we shall not explore this symmetry further.

One way of pairing the Majoranas is to pair them in the same block, and the other would be to pair them across each block. In the first case, the Hamiltonian has only onsite energy that is, μ for the fermions. This yields a Hamiltonian,

$$\mathcal{H} = -\frac{i\mu}{2}\sum_{j=1}^{N}\gamma_{1,j}\gamma_{2,j}. \tag{3.109}$$

This situation clearly subscribes to the trivial case, where $\Delta = t = 0$ and only $\mu \neq 0$. Thus, there are no unpaired Majaorana and the spectrum is gapless with excitation energies $\pm\mu$. Correspondingly, it has no edge modes and the system is a trivial insulator. However, the other scenario, which pairs two Majoranas across the domino tiles, yields a Hamiltonian,

$$\mathcal{H} = it\sum_{j=1}^{N-1}\gamma_{2,j}\gamma_{1,j+1}. \tag{3.110}$$

Clearly, this leaves the two Majorana at the edges to be excluded in $\mathcal{H}$. Also, we may recall that this Hamiltonian corresponds to $t = \Delta = 1$ and $\mu = 0$ which denotes the topological limit of the Kitaev model.

The first case discussed here corresponds to the trivial topology (see upper panel of Fig. 3.19), where the intra-cell pairing of the Majoranas yields no zero modes. However, the latter hosts Majorana zero modes, and thus denotes a topological phase of the system (depicted in the lower panel of Fig. 3.19). It is rewarding to realize that these two cases for the pairing of Majoranas in a domino model correspond to zero modes in the Kitaev model that we have discussed earlier.

[10] We discontinue usage of the symbol "i" for denoting the site index because of the imaginary i being present in the Hamiltonian, and use j instead.

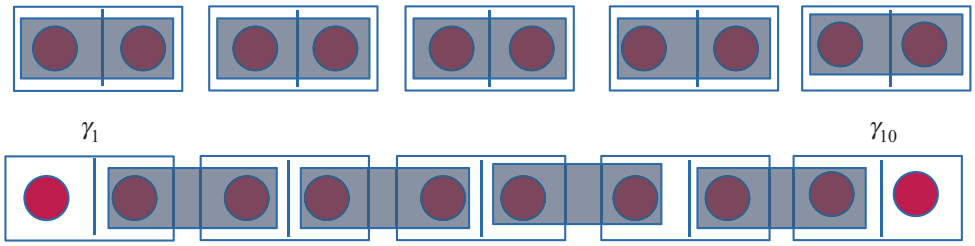


FIG. 3.19
The trivial and the topological dominoe model are demonstrated.

An intuitive way to arrive at the zero energy Majorana modes in the Kitaev chain can be seen by exploring the Kitaev Hamiltonian. For that purpose, we shall write down the Kitaev Hamiltonian once again in real space.[11] We begin with a small k version of the Hamiltonian, where the $-2t\cos k$ in the diagonal term (see Eq. (10)) is replaced by $2t$ (this also necessitates the off-diagonal terms to be written as $2i\Delta k$). Hence, we take a small deviation of the chemical potential, μ from its value where the phase transition occurs from a topological to a trivial phase, namely, $\mu = \pm 2t$. As a specific case, assume

$$\mu = -2t + m \tag{3.111}$$

where m is a small and positive quantity. This yields the $\mathbf{d}$-vector to has a form,

$$\mathbf{d}(k) = (0, 2\Delta k\sigma_y, m\sigma_z). \tag{3.112}$$

We write k as $\frac{\partial}{\partial x}$ (the chain is placed along the x-direction) and assume m to be a smoothly varying quantity in x, that is, $m(x)$ which changes sign across the transition. This allows us to write the Hamiltonian as

$$\mathcal{H} = 2i\Delta\sigma_y\frac{\partial}{\partial x} + m(x)\sigma_z. \tag{3.113}$$

The time independent Schrödinger equation corresponding to $E = 0$ for the Hamiltonian can be written as

$$\left(2i\Delta\sigma_y\frac{\partial}{\partial x} + m(x)\sigma_z\right)\phi(x) = 0 \tag{3.114}$$

where $\phi(x)$ is the eigenfunction. Solving the above equation, one gets, two solutions (corresponding to the 2×2 structure of the Pauli matrices) as

$$\phi_\pm(x) = \exp\left(\pm\int\frac{m(x)}{2\Delta}dx\right)\phi_\pm(0) \tag{3.115}$$

[11] This real space representation is distinct from the one written in Eq. (3.70), and is obtained by discretizing the momentum operator on a chain. We have written it for two sites in Eq. (3.70).

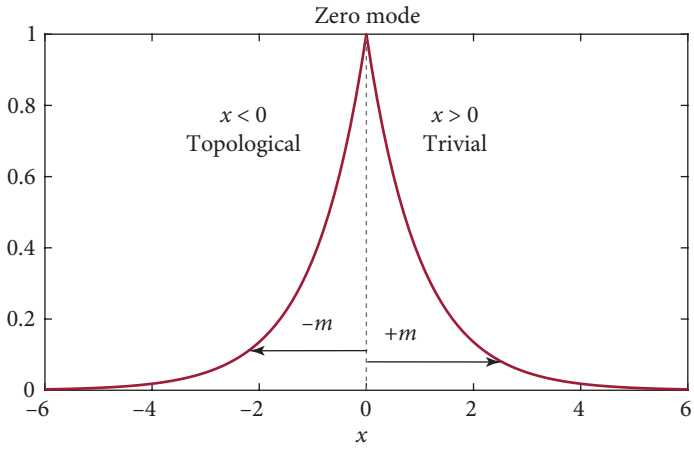


FIG. 3.20
The spatial dependence of the wavefunction is shown. It falls off exponentially on either side of $x = 0$, the left of which denotes a topological phase, while the right represents a trivial phase.

where $\phi_{\pm}(0)$ refers to the eigenfunction of σ_x, that is,

$$\phi_{\pm}(0) = \begin{pmatrix} 1 \\ \pm 1 \end{pmatrix}.$$

The $\pm$-sign yields an unphysical situation as it diverges at $x = \infty$. However, there is a solution corresponding to $m(x) = 0$. This says that the zero energy mode decays exponentially in either direction with respect to $x = 0$. We can invoke any functional behavior of $m(x)$, however, the behavior of the zero mode remains unaltered. Suppose we assume $m(x)$ to change the sign at $x = 0$ and flat otherwise (see Fig. 3.20), the bound state, $\phi_-(x)$ assumes the form,

$$\phi_-(x) = \exp\left(-|m|x\sigma_x\right)\phi_-(0). \tag{3.116}$$

3.5.6 Energy spectrum of *N*-site Kitaev model

In order to obtain the energy spectrum of the Kitaev model, we can use Eq. (3.79) which was originally derived for a two-site Kitaev chain. However, for a N-site system, the same expression can still be used with N modes, that is, N k-values equally spaced between $-\pi$ and $+\pi$.

It is also possible to solve Eqs. (3.109) and (3.110) on the basis of Majorana Fermions in real space. For a chain consisting of N-sites, the Hamiltonian in Eq. (3.109) [or Eq. (3.110)] yields N eigenvalues which can be mapped on to N distinct k-values that belong to the one-dimensional Brillouin zone. The spectrum corresponding to the topological case [Eq. (3.110)] is plotted in Fig. 3.21 as a function of the

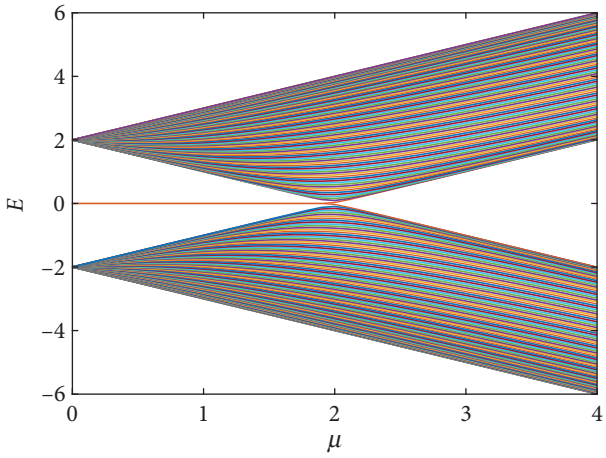


FIG. 3.21
The spectrum for the Kitaev model is plotted for $t = \Delta = 1$. The zero modes persist till $\mu = 2t$, beyond which a trivial band gap opens up.

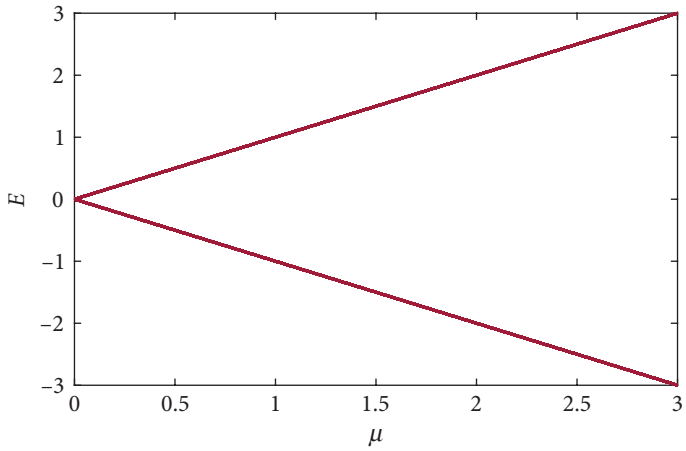


FIG. 3.22
The spectrum for the Kitaev model is plotted for $t = \Delta = 0$. A band gap be seen for all values of μ.

chemical potential, μ. There are two degenerate zero modes which persist till $\mu = 2t$, beyond which these modes merge into the bulk. Hence, for $\mu > 2t$, one obtains a trivial phase. For the trivial case, the spectrum is gapped for all values of μ (see Fig. 3.22). The topological protection of the zero modes is discussed below.

3.6 TOPOLOGICAL PROPERTIES OF THE MAJORANA MODES

It may seem that the topological phase of the Kitaev chain is an artefact of setting the chemical potential, μ, equal to zero, thereby disconnecting the two Majorana modes at the two edges. So a natural question is if we change μ slightly from a value zero, do the Majorana modes disappear by coupling to the rest of the chain? If this is true, then it results from a very carefully controlled model. In reality, this is not true. If we increase μ, it can be easily checked that till $\mu \leq 2t$, the zero modes (two of them) stay together, and they only split for $\mu > 2t$. One can see the same behavior for negative values of the chemical potential, that is, the zero modes stick together at $E = 0$ for $\mu \geq -2t$. At $\mu = \pm 2t$, the system becomes critical and the bulk gap closes, whence the system ceases to be topological. Thus, the Majorana modes are protected as long as there is a gap in the bulk spectrum.

The Majorana modes are indeed protected by the particle-hole symmetry of the Hamiltonian. For $\mu = 0$ (and $\Delta = t \neq 0$), there are an equal number of particle-like (corresponding to $E < 0$), and hole-like states (for $E > 0$). Further, two states at zero energy which are inseparable, as if that is untrue, for example, if they are separated, then that will cause an imbalance in the number of states below and above $E = 0$ (Fermi energy). The only possibility is to couple the two unpaired Majorana modes to each other, however such a coupling is impossible because of the large spatial distance between them. The only way to facilitate a coupling is to close the bulk gap, which precisely happens for $\mu = \pm 2t$, where the zero modes disappear into the bulk.

3.7 EXPERIMENTAL REALIZATION OF THE KITAEV CHAIN

From the discussion so far, it is clear that there are intriguing possibilities of topological features emerging, depending on parameters of the Hamiltonian. The zero energy mode in the topological regime is a coherent superposition of the two Majoranas and has opposite fermion number parity. However, this system may look rather unphysical. Most crucial hindrance is provided by freezing out the spin degrees of freedom. How do we do that? For argument's sake, we can include a strong magnetic field that will polarize all the spins of the electrons in the same direction of the field. However, that brings us to the conflict of superconductivity coexisting with a strong magnetic field. Furthermore, the availability of p-wave superconductors in nature is infrequent. Furthermore, not to forget that a mean field theory of superconductivity is susceptible to large fluctuations in one dimension, which may prevent the stabilization of the superconducting state.

Nevertheless, the seminal works of Fu and Kane (2008) and Fu and Kane (2009) have made it abundantly clear that a Kitaev chain can be experimentally realized in a variety of systems. The key ingredients are proximity-induced superconductivity, spin-orbit coupling (SOC) where the latter boosts the experimental search for exploring SOC in superconductors. From undergraduate level quantum mechanics, it is known that the orbital angular momentum ceases to be a good quantum

number in the presence of SOC, and thus it raises the possibility of p-wave superconductivity (as an admixture to the s-wave pairing) which entails a spin-polarized pairing, and can possibly be induced by proximity effects. For a more detailed discussion, may refer to the papers in Lutchyn *et al.* (2010) and Oreg *et al.* (2010).

As discussed earlier that the Majorana bound states are the superposition of particle and hole excitations at $E = 0$ which are trapped in the bulk gap. Thus, two Majorana bound states encode a non-local qubit that is robust against local perturbations, such as decoherence and, thus, constitute essential elements for topological quantum computation. The Majorana bound states can be experimentally probed by Andreev reflection. Suppose in an experiment with a semiconducting wire placed on a p-wave superconductor (or even an s-wave superconductor is deposited on the surface of a topological insulator whose surface states are denoted by the likes of Dirac fermions), metallic lead couples to the Majorana modes by tunneling of electrons. The Majorana fermions can induce resonant Andreev reflection[12] from the lead to the superconductor (Nilsson *et al.*, 2008; and Law *et al.*, 2009). Usual Andreev reflection converts an electron into a hole in the same lead, however a crossed Andreev reflection non-locally converts electron excitation into hole excitation into different lead. Equivalently, it splits a Cooper pair and distributes them into two leads, which happens at very low excitation energies.

3.8 TOPOLOGY IN 2D: GRAPHENE AS A TOPOLOGICAL INSULATOR

Having studied a prototype model Hamiltonian in 1D, we turn our focus toward 2D, now with the lens on graphene. Particularly, we shall explore whether graphene possesses the credibility of becoming a topological insulator. That may happen, provided by some means, we are able to open a spectral gap at the Dirac cones. Since a non-zero Berry phase can be a smoking gun for non-trivial properties, let us first look at the Berry phase of graphene.

3.8.1 Berry phase

To compute the Berry phase, let us consider the low energy Hamiltonian of graphene given by[13]

$$\mathcal{H} = \hbar v_F(\tau_z \sigma_x q_x + \sigma_y q_y) \tag{3.117}$$

where τ_z denotes the valley degree of freedom, that is, $\tau_z = 1$ for K-point, while it is -1 for the K'-point. As usual, σ denotes the sublattice degree of freedom. The Berry connection was obtained as

$$\mathcal{A} = \langle \psi_- | \nabla | \psi_- \rangle. \tag{3.118}$$

[12] Andreev reflection is briefly discussed in Chap. 7.
[13] We have discussed the electronic properties of graphene in Chap. 3.

To remind ourselves, the Chern number is defined by integrating the Berry curvature over BZ.

$$n = \frac{1}{2\pi} \oint_{BZ} \mathcal{F} d^2K = C \quad \text{(a notation we have used earlier)}$$

We rewrite the Dirac Hamiltonian as

$$h(\mathbf{q}) = \mathbf{q} \cdot \sigma \quad \text{(the velocity term is dropped).} \tag{3.119}$$

In the polar coordinates $\mathbf{q}$ and $h(\mathbf{q})$ can be represented as

$$\mathbf{q} = |\mathbf{q}| \begin{pmatrix} \cos\phi \\ \sin\phi \end{pmatrix} = q \begin{pmatrix} \cos\phi \\ \sin\phi \end{pmatrix} \tag{3.120}$$

and

$$h(\mathbf{q}) = q \begin{pmatrix} 0 & \cos\phi - i\sin\phi \\ \cos\phi + i\sin\phi & 0 \end{pmatrix} = q \begin{pmatrix} 0 & e^{-i\phi} \\ e^{i\phi} & 0 \end{pmatrix}. \tag{3.121}$$

The normalized eigenvectors are,

$$|\psi_-\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} -e^{-i\phi} \\ 1 \end{pmatrix} \text{ and}$$

$$|\psi_+\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{-i\phi} \\ 1 \end{pmatrix}. \tag{3.122}$$

Next, we calculate the Berry connection $\mathcal{A}$, and to remind ourselves that only filled bands are to be taken into account. So we shall consider $|\psi_-\rangle$ in the definition of $\mathcal{A}$.

$$\mathcal{A} = i\langle\psi_-|\nabla_q|\psi_-\rangle. \tag{3.123}$$

The gradient operator, ∇_q in the polar coordinate is given by

$$\nabla_q = \left(\frac{\partial}{\partial q}\hat{q} + \frac{1}{q}\frac{\partial}{\partial\phi}\hat{\phi} \right). \tag{3.124}$$

Note that $|\psi_-\rangle$ does not depend upon q. If we now introduce a band index, n, the Chern number corresponding to a band index, n can be written as C_n. The total Chern number is obtained from the contribution from all the bands, namely,

$$C = \sum_n C_n$$

$$C_n = \frac{1}{2\pi} \int_S \mathcal{F}_n dS, \tag{3.125}$$

where the S is the surface that encloses the loop. With $\mathcal{A} = \frac{1}{2q}$, $\nabla \times \mathcal{A} = 0$. So $\mathcal{F} = 0$, and hence $C = 0$ which is not a surprise, as for time reversal invariant systems, the Chern number should vanish.

The Berry phase around the Dirac points ($\mathbf{K}$ and $\mathbf{K}'$) is nothing but the winding number multiplied by π, which is then either $+1$ or -1. This introduces a measure of the topological *charge* for the Dirac points in the k-space which tells us how the wavefunctions wind around these singular points in k-space

differently with respect to each other. The **K** point carries the topological charge +1 (a vortex) and the **K**′ points carry a topological charge −1 (an anti-vortex). With the Dirac Fermion sitting at **K** carries a Berry phase, $\Phi_B^{\mathbf{K}} = \pi$, and the Dirac Fermion at **K′** has a Berry phase, $\Phi_B^{\mathbf{K}'} = -\pi$. The overall Berry phase, Φ_B is zero, that is, $\Phi_B = 0$.

3.8.2 Symmetries of graphene

It is fairly well known to the readers by now that graphene is represented by the nearest neighbor tight binding model on a honeycomb lattice with a two sublattice basis, namely, A and B sublattices. Carbon (C) atoms occupy both the sublattices. The situation is slightly different in Boron Nitride, which, in spite possessing the same crystal structure, the sublattice symmetry is broken by boron and nitrogen occupying the A and the B sublattices. Thus, graphene is a prototype of a system possessing sublattice symmetry which renders the Hamiltonian block off diagonal written on the sublattice basis. The low energy physics of this model is denoted by the massless Dirac Hamiltonian that we have seen at length earlier. Here, for the sake of completeness, we recall that the low energy Hamiltonian of graphene at both the Dirac points, namely, **K** and **K**′ is written as

$$\mathcal{H}_0(\mathbf{k}) = \hbar v_F (k_x \sigma_x \tau_z + k_y \sigma_y), \tag{3.126}$$

where the pseudospins σ_x, σ_y denote sublattice degrees of freedom, and τ_z (again the z-component of the Pauli matrix) distinguishes the valleys at **K** and **K**′. Needless to say that here is that the Hamiltonian is independent of the real spin, which will continue to be a valid description until spin-orbit coupling is included. Now, consider the inversion (or the sublattice) symmetry, which, along with switching the two sublattices, also changes the momentum **k** to −**k** (remember $\mathbf{p} = m\frac{d\mathbf{r}}{dt}$) which again implies that the two valleys are switched under inversion. The corresponding operator which performs this operation is given by

$$\mathcal{P} = \sigma_x \tau_x$$

Under this inversion operator, the Hamiltonian transforms as

$$\mathcal{P} H(\mathbf{k}) \mathcal{P}^{-1} = \hbar v_F \, \sigma_x \, \tau_x \, (k_x \sigma_x \tau_z + k_y \sigma_y) \, \sigma_x \tau_x = \mathcal{H}_0(-\mathbf{k}). \tag{3.127}$$

The above relation can be proven by using product rules of the Pauli matrices, and it ensures inversion symmetry of the Dirac Hamiltonian.

Now we shall discuss time reversal symmetry. In the case of graphene, time reversal symmetry implies changing the momentum vector **k** to −**k**, followed by complex conjugation of the operator (as explained earlier). Under the time reversal symmetry operation, one Dirac point (say, **K**) goes to another one (say, **K**′), and thus, the two Dirac cones are exchanged. Thus, taking the time reversal operator, $\mathcal{T}$ to be a complex conjugation operator should have been sufficient. However, as discussed earlier, the time reversal symmetry in graphene also implies a transformation from one valley to another, that is **K** changing over to **K**′. This makes us settle for

$$\mathcal{T} = \tau_x \mathcal{K}, \tag{3.128}$$

where τ_x is the x-component of the Pauli matrix. Note that here $\mathcal{T}^2 = 1$ as we are dealing with spinless fermions.[14] To check for the invariance of the Hamiltonian under the operation of $\mathcal{T}$, one needs to prove,

$$\mathcal{T}\mathcal{H}_0(\mathbf{k})\mathcal{T}^{-1} = \hbar v_F\, \tau_x\, (k_x\sigma_x\tau_z + k_y\sigma_y^*)\tau_x = \mathcal{H}_0(-\mathbf{k}). \tag{3.129}$$

The above relation ensures the invariance of the Dirac Hamiltonian under the time reversal operation.

It may also be mentioned that the above symmetries put together, that is, the product of the sublattice (or inversion), and the time reversal symmetries yield further discrete symmetry, known as the charge-conjugate symmetry, usually denoted by $\mathcal{C}$. It can be checked that $\mathcal{H}_0$ is invariant under a combination of these two symmetries.

To summarize, in the context of graphene, we have seen the emergence of three discrete symmetries, namely, the sublattice symmetry (or, the inversion symmetry, denoted by $\mathcal{P}$), the time reversal symmetry (denoted by $\mathcal{T}$), and finally, a combination of the two, that is, the charge-conjugation symmetry ($\mathcal{C}$). They indeed have different properties, such as $\mathcal{P}$ is a unitary operator and anticommutes with the Hamiltonian, $\mathcal{T}$ is an antiunitary operator which commutes with the Hamiltonian, while the charge-conjugation operator $\mathcal{C}$ is antiunitary (since it is a combination of $\mathcal{P}$ and $\mathcal{T}$) which also anticommutes with the Hamiltonian.

Having discussed the fundamental symmetries of graphene, let us return to its prospects of being a topological insulator. In Haldane's own submission (see https://topocondmat.org/w4_haldane/haldane_model.html), there can be simple efforts to tweak the Hamiltonian to achieve topological properties. Thus, the goal is to transform a sheet of graphene into a quantum Hall-like state with conducting edges and insulating bulk. Furthermore, the edge modes have a chiral character, which means that the current is carried in opposite directions at the two edges. There could be two ways of doing this; either break the inversion symmetry, keeping the time reversal symmetry intact, or break the time reversal symmetry, retaining the inversion symmetry. In the following, we show that, while the first option does not yield a topological phase, the second one indeed does. Nevertheless, we shall discuss both, which are, respectively known as the Semenoff insulator (obtained via breaking the inversion symmetry), and a Haldane (or a Chern) insulator (obtained via breaking the time reversal symmetry).

3.8.3 Semenoff insulator

In order to break the inversion symmetry, consider a staggered onsite potential of the form,

$$\mathcal{H}' = \varepsilon_{\mathrm{A}} \sum_{\mathbf{r}_{\mathrm{A}}} c_{\mathrm{A}}^{\dagger}(\mathbf{r}_{\mathrm{A}})c_A(\mathbf{r}_{\mathrm{A}}) + \varepsilon_{\mathrm{B}} \sum_{\mathbf{r}_{\mathrm{B}}} c_{\mathrm{B}}^{\dagger}(\mathbf{r}_{\mathrm{B}})c_{\mathrm{B}}(\mathbf{r}_{\mathrm{B}}), \tag{3.130}$$

where ε_{A} and ε_{B} are onsite potentials at the sites A and B, respectively. For $\varepsilon_{\mathrm{A}} \neq \varepsilon_{\mathrm{B}}$, the inversion symmetry (or the sublattice symmetry) is broken, as is the case for the hexagonal boron nitride (h-BN),

[14] For spin-full systems, $\mathcal{T}^2 = -1$.

where the sites occupied by the C atoms in graphene are occupied by boron (B) and nitrogen (N) at the A and B sublattice sites, thereby causing the onsite energies to be unequal. Thus, including a term that has an equal and opposite magnitude at the two sublattices, the low energy Hamiltonian becomes,

$$\mathcal{H}(\mathbf{q}) = \mathcal{H}_0(\mathbf{q}) + m_I\sigma_z \tag{3.131}$$

where the m_I term makes the massless Dirac particle massive. Here, $m_I = (\varepsilon_A - \varepsilon_B)/2$ can be called the Semenoff mass. $m_I = 0$ for $\varepsilon_A = \varepsilon_B$. Further σ_z anticommutes with $\mathcal{H}_0(\mathbf{k})$, that is,

$$\{\sigma_z, \mathcal{H}_0\} = 0.$$

The spectrum is given by

$$E(\mathbf{q}) = \pm\sqrt{\hbar^2 v_F^2 q^2 + m_I^2}. \tag{3.132}$$

In a compact notation, one may write it as

$$E_\mu(q) = \mu\sqrt{\hbar^2 v_F^2 q^2 + m_I^2} \tag{3.133}$$

where $\mu = \pm 1$ and each sign refers to a valley index.

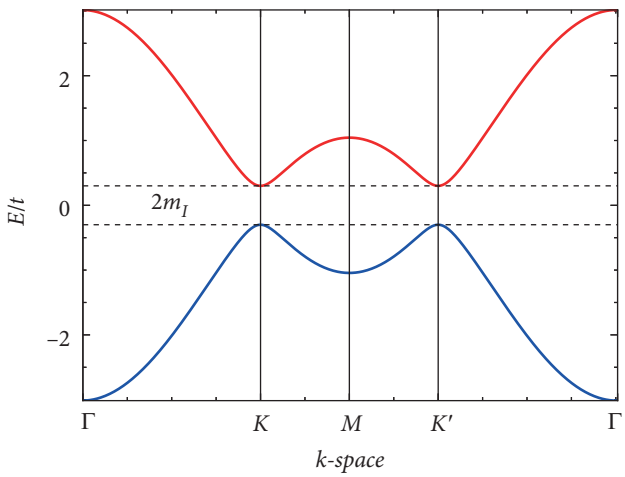


FIG. 3.23
The energy band dispersion with a Semenoff mass, m_I. A gap opens up at the Dirac points of magnitude $2m_I$.

The spectrum is plotted in Fig. 3.23. Spectral gaps open up of magnitude $2m_I$ at each of the Dirac points. This gap earned the name Semenoff insulator. However, the nature of the gap is a trivial one in the following sense. The gap vanishes as $m_I \to 0$. Besides the wavefunction plotted for a graphene nanoribbon of size L_xL_y.[15] There is no trace of the edge states being present. Besides the Berry phase and the Chern number also vanish as we shall show below, thereby certifying the trivial nature of the energy gap in the spectrum. This eliminates the possibility of any topological properties of the model induced by the inclusion of m_I.

To gain a bit of details on the Semenoff insulator, let us complete the mandatory calculations. The eigenfunctions can be written as

$$\Psi^\mu(\mathbf{q}) = \frac{1}{\sqrt{2}}\begin{pmatrix} \sqrt{1 + m_I/E^\mu} \\ \mu\sqrt{1 - m_I/E^\mu}\, e^{i\theta_q} \end{pmatrix}. \tag{3.134}$$

[15] This is called a semi-infinite ribbon. It is finite in y-direction and very large (taken to be infinitely large) along the x-direction ($L_x \gg L_y$). We shall shortly discuss this below.

The corresponding Berry curvature is

$$\Omega^{\mu} = \frac{v_F^2 m_I}{2\mu\left[v_F^2 q_x^2 + v_F^2 q_y^2 + \beta^2\right]^{\frac{3}{2}}} \tag{3.135}$$

which eventually gives the Berry connection as [see Eq. (3.118)],

$$\mathcal{A}^{\mu} = \frac{\tau_z}{2}\left(1 + \mu\frac{m_I}{|E^{\mu}|}\right)\frac{\hat{\theta}_q}{q}. \tag{3.136}$$

Finally, the Berry phase is obtained as

$$\Phi_B = \pi\tau_z\left(1 + \lambda\frac{m_I}{|E^{\mu}|}\right).$$

Thus, the Berry phase for a massless Dirac equation is thus renormalized by the Semenoff mass, m_I. One regains the corresponding result for graphene by putting, $m_I = 0$.

A further (and more robust) check on the trivial nature of the spectral gap can be achieved by computing the dispersion for a graphene nanoribbon. A nanoribbon is a system which is infinite along one direction (say, x-direction), and finite along the other direction (y-direction). Usually, graphene ribbons are recognized by their edges along the x-axis, for example, with zigzag and armchair edges, and are referred to as the zigzag graphene nanoribbons (abbreviated as ZGNR) and armchair graphene nanoribbon (AGNR). There is an important difference between the two. ZGNR is always metallic with gapless edge states, while AGNR is conditionally metallic in the following sense. AGNR has been conducting edge states when $N = 3M - 1$, where N is the number of lattice sites in the y-direction, and M is an integer.

We have taken a ZGNR, as shown in Fig. 3.24, with the total number of lattice sites along the y-axis as 256, that is, $N = 256$ (so number of unit cells is 128) along the y-direction and a width given by $\left(\frac{3N}{2} - 1\right)a$ (a: lattice constant = 1.42Å), which upon putting $N = 256$ yields $383a$ or 543.86 Å. We finally write down the equation of motion, that is solving the Schrödinger equation, $\mathcal{H}\psi = E\psi$ for the amplitudes at the A and B sublattice sites below,

$$E_k a_{k,n} = -\left[t\left\{1 + e^{(-1)^n ik}\right\}b_{k,n} + tb_{k,n-1}\right] + m_I a_{k,n} \tag{3.137}$$

$$E_k b_{k,n} = -\left[t\left\{1 + e^{(-1)^{n+1} ik}\right\}a_{k,n} + ta_{k,n+1}\right] - m_I b_{k,n}. \tag{3.138}$$

Along the x-direction, the ribbon is infinite, which is implemented in our numeric computation by assuming the momentum along the x-direction, that is, k_x to be a good quantum number. The above equations are numerically solved. We show the results in Fig. 3.25 which clearly show the absence of zero modes, which precludes its prospects as a candidate for a topological insulator.

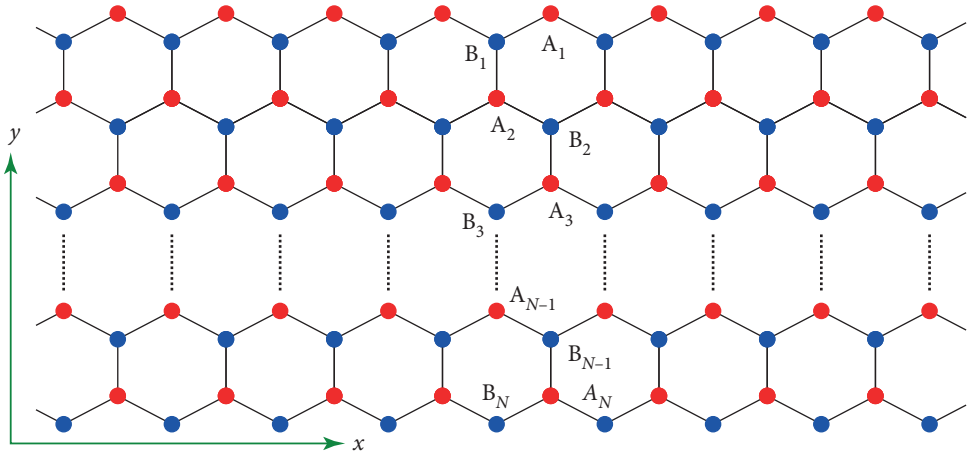


FIG. 3.24
A schematic diagram of a semi-infinite nanoribbon is show. We perform our numeric computation of the edge modes on a geometry such as this.

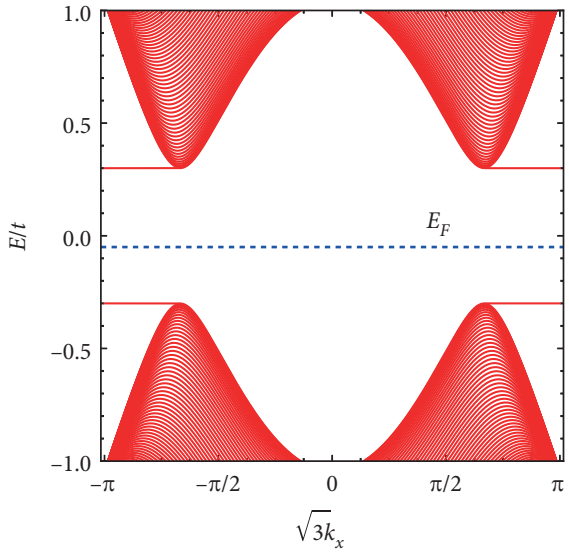


FIG. 3.25
The energy dispersion for a Semenoff insulator in a semi-infinite nanoribbon. A (trivial) gap is visible in the spectrum.

3.8.4 Haldane (Chern) insulator

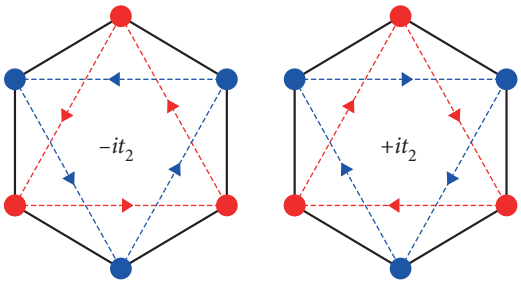


FIG. 3.26
Complex next nearest neighbour hopping in the Haldane model.

The second option of breaking the time reversal invariance is more subtle, and yields a success in obtaining a topological state. The idea involves including an imaginary second neighbour hopping that assumes opposite signs depending on the direction of hopping. For example, if an anticlockwise hopping (shown by the blue arrow in Fig. 3.26) is assumed with a positive sign, then the clockwise hopping (shown by red in Fig. 3.26) acquires a negative sign. A formal way of writing this term is via

$$\mathcal{H}'' = t_2 \sum_{\langle\langle ij \rangle\rangle} e^{i\nu_{ij}\phi} c_i^\dagger c_j, \tag{3.139}$$

where the sum runs over the next nearest neighbor (NNN) sites (double angular bracket $\langle\langle ij\rangle\rangle$ imply NNN sites). ν_{ij} denotes the chiral nature of the hopping term where $\nu_{ij} = -\nu_{ji}$ depends on the direction of the hopping. The convention is $\nu_{ij} = +1$ for clockwise hopping between the NNN sites, while $\nu_{ij} = -1$ for anticlockwise hopping (see Fig. 3.26). The phase $e^{i\phi}$ or $e^{-i\phi}$ depends upon the direction of the hopping. Such a complex direction-dependent hopping breaks the time reversal invariance, since the time reversal flips the direction of hopping. Only the imaginary part of the phase, ϕ is interesting. Thus, to set the real part to zero, we may choose $\phi = \frac{\pi}{2}$. This is known as the Haldane model (Haldane, 1988), which Haldane had prescribed for achieving an anomalous quantum Hall state. It is anomalous in the sense that the Hall effect is realized without an external magnetic field, or equivalently, without the Landau levels. As we have seen earlier, and again shall see shortly that broken time reversal symmetry implies a finite Chern number, a reason why these insulators are known as Chern insulators. The complex phases can be realized by applying a staggered magnetic field pointing at opposite directions at the center of the honeycomb lattice relative to that at the vertices.

Introducing the NNN vectors, $\mathbf{b}_i$, as earlier, where, $\mathbf{b}_1 = \boldsymbol{\delta}_2 - \boldsymbol{\delta}_3$, $\mathbf{b}_2 = \boldsymbol{\delta}_3 - \boldsymbol{\delta}_1$, and $\mathbf{b}_3 = \boldsymbol{\delta}_1 - \boldsymbol{\delta}_2$ where $\boldsymbol{\delta}_i$ denotes the vectors connecting NN sites, the Hamiltonian can be written as (the NN tight binding term is also there, but not written here),

$$\mathcal{H}'' = t_2 \sum_{i=1}^{3} \left[e^{i\phi} \sum_{\mathbf{r}_A} c_{\mathrm{A}}^\dagger(\mathbf{r}_{\mathrm{A}}) c_{\mathrm{A}}(\mathbf{r}_{\mathrm{A}} + \mathbf{b}_i) + e^{-i\phi} \sum_{\mathbf{r}_{\mathrm{B}}} c_{\mathrm{B}}^\dagger(\mathbf{r}_{\mathrm{B}}) c_{\mathrm{B}}(\mathbf{r}_{\mathrm{B}} + \mathbf{b}_i) \right]. \tag{3.140}$$

In the momentum space, the full tight-binding Hamiltonian reads

$$\mathcal{H}''(\mathbf{k}) = 2t_2 \left[\cos\phi \sum_{i=1}^{3} \cos(\mathbf{k}\cdot\mathbf{b}_i)\mathbb{1} + \sin\phi \sum_{i=1}^{3} \sin(\mathbf{k}\cdot\mathbf{b}_i)\sigma_z \right]. \tag{3.141}$$

Until this point, the full NNN Hamiltonian is dispersive, that is, it depends upon the $\mathbf{k}$-vector. However, the low energy Hamiltonian, that is near the $\mathbf{K}$ and $\mathbf{K}'$ points, is independent of $\mathbf{k}$ at the leading order,

where the Hamiltonian can be shown to assume the form,

$$\mathcal{H}''_{\pm\mathbf{K}} = m_H\, \tau_z\, \sigma_z, \tag{3.142}$$

where we have combined the forms at the two Dirac points by using τ_z where $\tau_z = \pm 1$,

$$m_H = -3\sqrt{3}\, t_2\, \sin\phi. \tag{3.143}$$

The above form can be easily obtained by noting that

$$\sum_{i=1}^{3} \cos(\mathbf{k}\cdot\mathbf{b}_i) = -\frac{3}{2} \quad \text{and} \quad \sum_{i=1}^{3} \sin(\mathbf{k}\cdot\mathbf{b}_i) = \mp\frac{3\sqrt{3}}{2},$$

where $\mathbf{k}\cdot\mathbf{b}_i = \mathbf{K}$. The readers are encouraged to fill up a few steps of algebra.

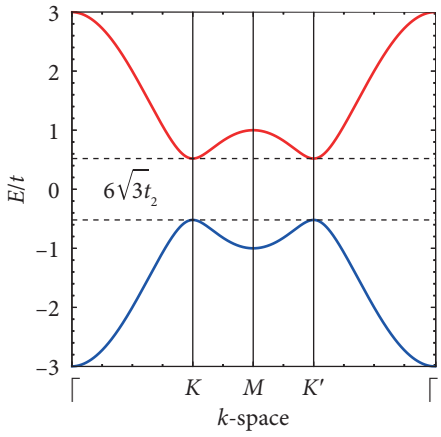


FIG. 3.27
The energy spectrum of the Haldane model.

Therefore, in the leading order $\mathcal{H}''$ is independent of the momentum $\mathbf{k}$. The last term in Eq. (3.142) breaks the time reversal symmetry. σ_z does not change sign, but τ_z being the valley degree of freedom does. Thus, the energy spectrum opens up a gap at the Dirac points for specific values of the complex second neighbor hopping t_2. In fact, adding a small t_2 yields a situation similar to adding a small a Semenoff mass, m_I. However, when t_2 exceeds a value of $\pm m_H/3\sqrt{3}$, the energy closes at one of the two Dirac points (either $\mathbf{K}$ or $\mathbf{K'}$), and opens up at the other Dirac point for one of the signs mentioned above, say $t_2 = m_H/3\sqrt{3}$. The reverse happens for $t_2 = -m_H/3\sqrt{3}$ where the gap closes at the former Dirac point, while it opens at the other. We show this in Fig. 3.27.

In order to elucidate the topological properties, we repeat the calculations same as that of the Semenoff insulator. The eigenfunctions for the Haldane model can be written as

$$\Psi^{\mu}(\mathbf{q}) = \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{1+\beta/E^{\mu}} \\ \mu\sqrt{1-\beta/E^{\mu}}\, e^{i\theta_q} \end{pmatrix}, \tag{3.144}$$

where $\mu = \pm 1$, with the energy spectrum given by

$$E^{\mu} = \mu\sqrt{v_F^2 q_x^2 + v_F^2 q_y^2 + \beta^2}.$$

This further yields a Berry curvature which can be shown to have the form,

$$\Omega^{\mu} = \frac{v_F^2 \beta}{2\mu\left[v_F^2 q_x^2 + v_F^2 q_y^2 + \beta^2\right]^{\frac{3}{2}}}. \tag{3.145}$$

The corresponding Berry connection is hence given by

$$\mathcal{A}^{\mu} = \frac{\tau_z}{2}\left(1 + \mu\frac{\beta}{|E|^{\mu}}\right)\frac{\hat{\theta}_q}{q} \tag{3.146}$$

where $\beta = 3\sqrt{3}t_2$. The Berry phase, Φ_B using

$$\Phi_B = \int \mathcal{A}^\mu \cdot d\mathbf{q}$$

yields

$$\Phi_B = \pi\tau_z\left(1 + \lambda\frac{\beta}{|E^\mu|}\right). \tag{3.147}$$

Finally, the Chern number can be obtained by integrating the Berry curvature over the BZ,

$$C = \oint \Omega^\mu(q)d^2q.$$

For the topological phase, that is for, $|t_2| > m_H/3\sqrt{3}$, one obtains a non-zero Chern number. Owing to a non-zero Chern number (C), the model has earned the name "Chern insulator." The Chern number is the topological invariant which distinguishes the Semenoff insulator from a Chern insulator. For the Semenoff insulator, $C = 0$. We show the phase diagram in Fig. 3.28, where the topological phases are shown via the red ($C = 1$) and the blue ($C = -1$) colors, respectively, while the trivial region ($C = 0$) is shown with a white color.

How do we know that this gap is topological in nature, instead of a trivial one as seen for a Semenoff insulator? This is a valid question since the nature of the gaps looks fairly the same in Figs. 3.25 and 3.27, except that the spectral gaps carry the energy scales proportional to their *masses*, that is, m_I for

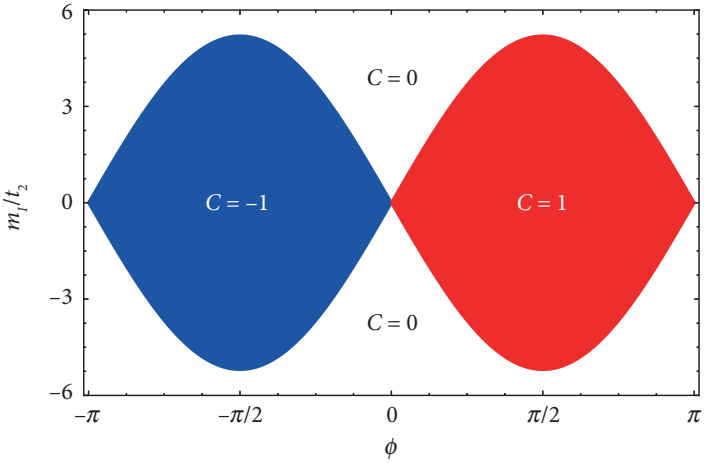


FIG. 3.28
The Chern number phase diagram for the Haldane model. The red region corresponds to $C = 1$, while the blue one denotes $C = -1$. The white region outside the lobes refers to a trivial insulator with $C = 0$.

the Semenoff insulator, and m_H for the Chern insulator. In the following, we check for the chiral edge modes in a semi-infinite graphene nanoribbon.

In a similar fashion as discussed in the context of a Semenoff insulator, the equations of the motion for the amplitudes at the A and B sublattice sites can now be written as

$$E_k a_{k,n} = -\left[t\left\{1 + e^{(-1)^n ik}\right\} b_{k,n} + t b_{k,n-1}\right] - 2t_2\left[\cos(k+\phi)a_{k,n} + e^{(-1)^n \frac{ik}{2}} \cos\left(\frac{k}{2} - \phi\right)\{a_{k,n-1} + a_{k,n+1}\}\right] \tag{3.148}$$

$$E_k b_{k,n} = -\left[t\left\{1 + e^{(-1)^{n+1} ik}\right\} a_{k,n} + t a_{k,n+1}\right] - 2t_2\left[\cos(k-\phi)b_{k,n} + e^{(-1)^{n+1} \frac{ik}{2}} \cos\left(\frac{k}{2} + \phi\right)\{a_{k,n-1} + a_{k,n+1}\}\right]. \tag{3.149}$$

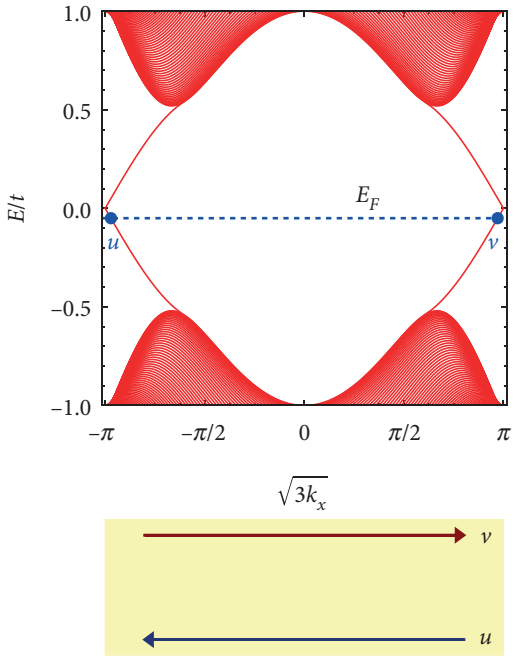


FIG. 3.29
The energy dispersion for a Chern insulator in a semi-infinite nanoribbon. The edge states are shown via the red lines that split from the bulk. In the yellow panel below, we show the chiral edge currents that flow in opposite directions along these edge modes.

In Fig. 3.29 we show the appearance of the edge modes as t_2 crosses $\pm m_H/3\sqrt{3}$ which are absent at small values of t_2. The appearance of the edge modes implies the emergence of a topological phase in the model, and there occurs a phase transition from a topological-insulating phase to that of a band insulator. Thus, we get a quantum Hall-like state, with conducting edge modes (and insulating bulk), albeit without an external magnetic field.

3.8.5 Quantum anomalous Hall effect

Finally, we shall present the results of the Hall conductivity. Here, non-zero Berry curvature yields a finite conductance. To remind ourselves, the full tight binding Hamiltonian [including the NN term which we have excluded earlier in Eq. (3.141)] is written as (Mondal and Basu, 2021)

$$\begin{aligned}\mathcal{H} = &-t\left[\cos(\mathbf{k}\cdot\boldsymbol{\delta}_1) + \sum_{i=2}^{3}\cos(\mathbf{k}\cdot\boldsymbol{\delta}_i)\right]\sigma_x \\ &- t\left[\sin(\mathbf{k}\cdot\boldsymbol{\delta}_1) + \sum_{i=2}^{3}\sin(\mathbf{k}\cdot\boldsymbol{\delta}_i)\right]\sigma_y \\ &+ \left[\Delta - 2\,t_2\,\sin\phi\sum_{i=1}^{3}\sin(\mathbf{k}\cdot\boldsymbol{\nu}_i)\right]\sigma_z \\ &+ \left[2t_2\,\cos\phi\sum_{i=1}^{3}\cos(\mathbf{k}\cdot\boldsymbol{\nu}_i)\right] I \\ = &\, h_x\sigma_x + h_y\sigma_y + h_z\sigma_z + h_0 I,\end{aligned} \tag{3.150}$$

where h_x, h_y and h_z represent the coefficients of the Pauli matrices σ_i. The low energy expansion of this Hamiltonian is convenient for our purpose. In fact, the computation of the Berry curvature is much easier for the low energy Hamiltonian, than it is for the full tight binding one. Arriving at the low energy Hamiltonian involves expanding the sine and the cosine functions to their leading order in the vicinity of the Dirac points. Applying these simplifications, one arrives at

$$\mathcal{H} = \mathbf{d} \cdot \sigma, \tag{3.151}$$

where the components of the $\mathbf{d}$ differ from those of $\mathbf{h}$, and up to linear in k_x and k_y are given by

$$d_x(k_x, k_y) = \frac{3}{2}k_x, \quad d_y(k_x, k_y) = \frac{3}{2}k_y, \quad \text{and} \quad d_z(k_x, k_y) = -3\sqrt{3}$$

The above form facilitates computation of the Hall conductivity using the following form for the Berry connection (Xiao *et al.*, 2010),

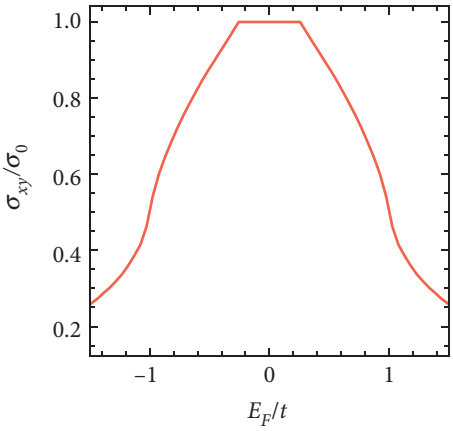


FIG. 3.30
The anomalous Hall conductivity is shown as a function of the Fermi energy. There is a distinct plateau in the vicinity of the zero-Fermi energy.

$$\Omega(E_k) = \frac{\mathbf{d}}{2|\mathbf{d}|^3}\left(\frac{\partial \mathbf{d}}{\partial k_x} \times \frac{\partial \mathbf{d}}{\partial k_y}\right). \tag{3.152}$$

Finally, the Hall conductivity is obtained via

$$\sigma_{xy} = \frac{e^2}{h}\int \frac{d\mathbf{k}}{(2\pi)^2} f(E_k)\Omega(E_k), \tag{3.153}$$

where the integral is taken over the BZ, and $f(E_k)$ is the Fermi distribution function. Since our calculations were at zero temperature, we set $f(E_k) = 1$. The Hall conductivity as a function of the Fermi energy is plotted in Fig. 3.30. A plateau at e^2/h is clearly visible, which enunciates the quantization of the Hall conductivity. Further, the presence of only one plateau is confirmed by the value of the Chern number being 1 (or -1), and also that there is only a pair of gapless edge modes. Thus, an anomalous version of the Hall conductivity is indeed distinct from the usual Hall effect (in presence of an external magnetic field). However, there are experimental realizations of systems with higher values of the Chern number, besides being backed up by a library of theoretical proposals.

3.9 QUANTUM SPIN HALL INSULATOR

Let us set aside the complex second neighbor hopping due to Haldane for a moment, the Dirac points in graphene are protected by time reversal and the inversion symmetries. The complex second neighbor hopping among sites of the same sublattice breaks the time reversal symmetry as we have seen in the preceding section. Kane and Mele (2005) demonstrated that it is possible to restore the time reversal symmetry in the Haldane model if we include (real) spin in the Hamiltonian, thereby making two

copies of the Haldane model, one for each spin. The inclusion of the spin opens up the possibility of a spin-orbit coupling, which however, does not violate any of the fundamental symmetries that we have discussed above. Moreover, the resulting insulating phase is absolutely new, and is referred to as the quantum spin Hall (QSH) phase. It should be clarified that spin-orbit coupling is not an essential ingredient for the realization of the QSH phase. However, to achieve a spin-polarized transport in a material, which shall aid its usage for spintronic applications, spin-orbit coupling is essential. We shall return to this discussion shortly.

Similar to the quantum Hall phase, the QSH phase is distinct from the trivial insulators by the presence of the conducting states at the edges, which are typically protected by the $\mathcal{Z}_2$ topological invariant. However, these edge states are non-chiral, unlike the quantum Hall states. In fact, they are called helical edge states, in the sense that there are two counter propagating edge modes at each edge, one for each spin (see Fig. 3.31). Such conducting modes are immune to the single particle back scattering from defects, disorders, or impurities as they are protected by time reversal symmetry. Thus, as long as there

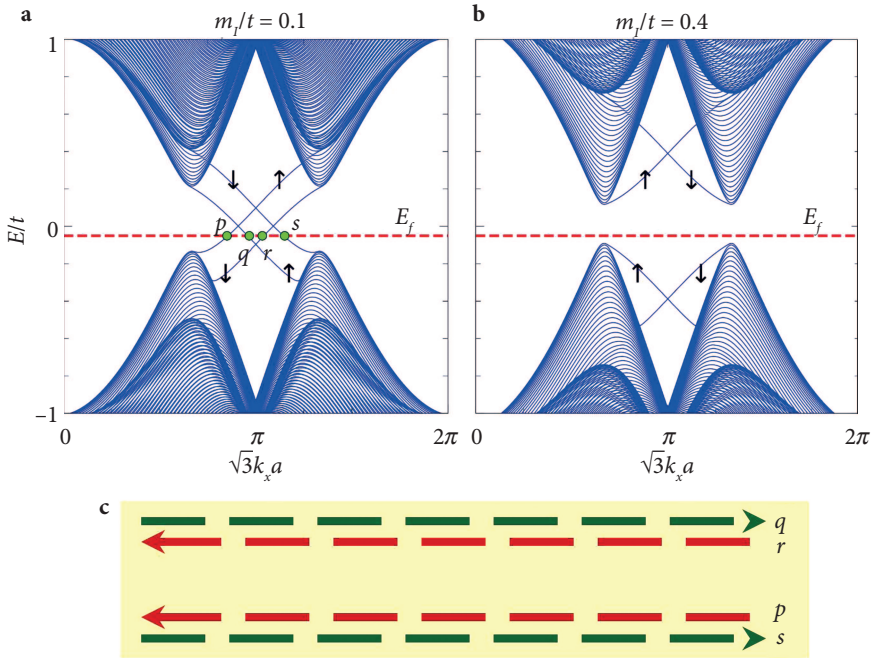


FIG. 3.31
The helical edge modes for the Kane– Mele model on a nanoribbon showing (a) topological, and (b) trivial phases. The yellow panel in (c) shows the spin-polarized helical modes carrying current.

is no time reversal symmetry breaking term, such as a magnetic impurity, the helical edge states are robust, and the QSH phase persists.

It was initially thought that graphene would host a QSH-like phase, however it is almost impossible to realize such a phase owing to an extremely weak spin-orbit coupling (Min *et al.*, 2006). However, a theoretical proposal of a QSH phase happened soon after when Bernevig, Bernevig *et al.* (2006) predicted that quantum wells made of CdTe/HgTe/CdTe host a QSH phase for a certain critical width of the HgTe layer where the band inversion occurs. The corresponding Hamiltonian is called the BHZ model (after Bernevig, Hughes, and Zhang). Quite fortunately, immediately afterwards, Molenkamp and co-workers (König, 2007) experimentally achieved such a scenario where an inverted band structure occurs, followed by the realization of the helical (instead of chiral) edge states.

In the following, we shall describe the Kane–Mele model for graphene, which serves as a toy model for a QSH phase that hosts counter propagating edge modes one for each spin at each of the edges. Further, these modes are found to be robust in the presence of a special type of spin-orbit coupling, known as the Rashba spin-orbit coupling (RSOC). Here, as we shall show below, owing to the restoration of the time reversal symmetry, the Chern number is zero. However, the helical edge modes are still protected by bulk $\mathbb{Z}_2$ topological invariant, which is a consequence of the Kramer's theorem applicable to band properties of fermions in a time reversal invariant system.

3.9.1 Kane–Mele model

The Dirac points have been shown to be protected by the time reversal and the inversion symmetries. Breaking any one of them opens a gap at the Dirac points by splitting the degeneracy. Throughout our discussion on graphene thus far, the spin of the electron never played a role.[16] Kane and Mele included the spin, thereby writing two Haldane Hamiltonians, one for spin-↑ fermions, and the other for spin-↓ fermions. This results in

$$\begin{aligned}\mathcal{H} &= -t\sum_{\langle ij\rangle,\alpha} c^{\dagger}_{i\alpha}c_{j\alpha} + it_2\sum_{\langle\langle ij\rangle\rangle,\alpha\beta} \nu_{ij}c^{\dagger}_{i\alpha}(S_z)_{\alpha\beta}c_{j\beta}\\ &= \mathcal{H}_0 + \mathcal{H}_{\mathrm{KM}},\end{aligned} \tag{3.154}$$

where $\mathcal{H}_0$ is the usual NN tight binding term, and $\mathcal{H}_{\mathrm{KM}}$ is like the Haldane term, now summed over the (real) spins. Traditionally, this term is called the *intrinsic spin-orbit coupling*. S_z is the z-component of the spin of the electrons, and (α, β) denote the spin indices. S_z is indeed the z-component of the Pauli matrices $\boldsymbol{\sigma}$. However, to distinguish it from the sublattice and the valley indices, we have written it with S_z. The NNN hopping term (the second term) describes a spin-orbit coupling that couples the chirality of the electrons, described by ν_{ij} with the z-component of spin ($S_z = \pm 1$). It is as if the orbital angular momentum vector $\mathbf{L}$ is associated with the chirality in a familiar $\mathbf{L}\cdot\mathbf{S}$ term, thus justifying its identification as the spin-orbit coupling.

[16] Chern number which is again analogous to the RHS of the Gauss-Bonnet theorem, namely the Euler-Poincare characteristics. Hall angle is the angle that the resultant of the applied electric field vector (E_x) and the Hall field (E_H) makes with E_x.

The second term, even though it resembles the Haldane term, respects all the symmetries of graphene. Time reversal flips the direction of hopping, that is, reversing the motion, but simultaneously it also flips the spin, thereby yielding another negative sign. This term respects all the symmetries of graphene. It may also be noted that the term does not involve spin flip, and hence, the two bands of the Haldane model (one for each spin) behave distinctly. In a mathematical sense, it means that the Hamiltonian retains a block diagonal form, and hence, is easy to deal with. The bands corresponding to the up-spin electrons are identical to the Haldane model (Chern insulators) discussed earlier. That is, they correspond to the phase of the complex NNN hopping to be $\phi = \pi/2$, and hence have opposite masses at the **K** and the **K′** (remember the term $m_H \tau_z \sigma_z$ in the low energy limit of the Haldane model). Furthermore, it has a Chern number $C_\uparrow = +1$. Please note that we have brought in a spin index to the Chern number. For the down-spin electrons for which $\phi = -\pi/2$, there will be an extra negative sign, which implies reversed signs for m_H at the **K** and the **K′** points as compared to the situation for the up spin. This yields $C_\downarrow = -1$. Thus, the total Chern number, $\sum_\sigma C_\sigma = 0$, which is a consequence of the time reversal symmetry.

A simple way of seeing that the Kane–Mele model is two copies of the Haldane model is that

$$[\mathcal{H}_{KM}, S_z] = 0, \tag{3.155}$$

which signifies that $\mathcal{H}_{KM}$ decouples into Hamiltonian one for each spin. This situation is similar to a Haldane flux $\phi = \frac{\pi}{2}$ for one type of spin, and $\phi = -\frac{\pi}{2}$ for the other. The low energy Kane–Male Hamiltonian can be shown to have a form (readers are encouraged to complete the derivation), written as

$$\mathcal{H}_{KM} = m_H \sigma_z \tau_z S_z$$

where the amplitude $m_H = -3\sqrt{3}\, t_2$ is the Haldane mass as stated earlier.

Let us convince ourselves that time reversal is indeed a valid symmetry operation for the Kane– Mele model. For spinor particles, we have seen that the time reversal operator $\mathcal{T}$ is written as

$$\mathcal{T} = i\sigma_y K$$

where K is a complex conjugation operator. Here we write it as

$$\mathcal{T} = iS_y K.$$

However, since the time reversal transformations from one valley to another, an operator that can be implemented by incorporating a τ_x, we can write,

$$\mathcal{T} = \tau_x\, iS_y K.$$

It is fairly trivial to see that $\mathcal{H}_{KM}$ is even under time reversal. Please recall that the time reversal inflicts complex conjugation, flips the real spin, reverses the valley degree of freedom, and reverses the direction of the momentum $\mathbf{k} \to -\mathbf{k}$. While the last one is not relevant, since the low energy Hamiltonian is independent of **k**, the first two yields under the time reversal,

$$\tau: \quad \tau_z \to -\tau_z, \quad s_z \to -s_z.$$

However, σ_z does not change the sign as it denotes the sublattice degree of freedom. Hence, two negative signs cancel out and we get $\mathcal{H}_{KM}$ to be even under $\mathcal{T}$.

To remind ourselves on the other fundamental symmetry, that the inversion symmetry $\mathcal{P}$, which yields

$$\mathcal{P}: \quad \sigma_z \to -\sigma_z, \quad \tau_z \to -\tau_z, \quad s_z \to s_z.$$

Hence, $\mathcal{H}_{KM}$ respects all symmetries of graphene as claimed earlier.

Let us look at the topological phase transition in a little more detail. For this, it is instructive to look at only one spin at a time, for example, $S_z = +1$ (that is, up spin). The Hamiltonian, including a Semenoff mass becomes,

$$\mathcal{H}_{KM}(\mathbf{k}) = \hbar v_F \left(k_x \sigma_x \tau_z + k_y \sigma_y\right) + (m_I + m_H \tau_z)\sigma_z \tag{3.156}$$

Explicitly writing the above Hamiltonian for the two valleys,

$$\mathcal{H}^{\mathbf{K}}_{KM}(\mathbf{k}) = \hbar v_F (k_x \sigma_x + k_y \sigma_y) + (m_I + m_H)\sigma_z \tag{3.157}$$

$$\mathcal{H}^{\mathbf{K}'}_{KM}(\mathbf{k}) = \hbar v_F (-k_x \sigma_x + k_y \sigma_y) + (m_I - m_H)\sigma_z. \tag{3.158}$$

Now consider two possibilities: (i) $m_I > m_H$ and (ii) $m_H > m_I$. In the first case, consider the extreme limits (for convenience), that is $m_I \gg m_H$ where we have a trivial band insulator. Now consider the other case where $m_H > m_I$: Nothing happens to $\mathcal{H}^{\mathbf{K}}_{KM}(\mathbf{k})$ in Eq. (3.157), but for $\mathcal{H}^{\mathbf{K}'}_{KM}(\mathbf{k})$ in Eq. (3.158), the gap closes and reopens. Thus, the insulating phase with $m_H > m_I$ is distinct from that of a band insulator by a *gap closing* phase transition, which by definition is a topological phase transition.

The situation for $S_z = -1$ is identical, except that the sign of m_H changes which results in a similar phase transition at the other Dirac point, that is, at the **K** point. Now defining,

$$\widetilde{m} = m_I - m_H \tag{3.159}$$

yields, corresponding to,

$$\begin{aligned} &\widetilde{m} < 0, \quad C = 1 \\ &\text{and} \quad \widetilde{m} > 0, \quad C = 0. \end{aligned} \tag{3.160}$$

3.10 BULK-BOUNDARY CORRESPONDENCE

Bulk-boundary correspondence (BBC) yields a guide to the phenomenology of topological insulators. The topological invariants computed from the bulk properties corresponding to a particular phase of the system uniquely reflect the conducting edge modes. Let us try to answer the question that we have posed above, that is, how is $\widetilde{m} < 0$ fundamentally different from that of $\widetilde{m} > 0$? Again, consider a semi-infinite nanoribbon, that is, infinite in x-direction and finite in the y-direction. The Schrödinger equation with the Hamiltonian given earlier can now be solved for a semi-infinite system, as we have discussed earlier.

Before we discuss the numerical solution for a nanoribbon, let us explore an analytic solution. We can assume that the The Hamiltonian has an edge at $y = 0$, so that the system exists for $y < 0$ and a vacuum for $y > 0$. In addition, let us assume a particular value of k_x, namely, $k_x = 0$ (remember k_x is a good quantum number owing to translational invariance in the x-direction). Hence, we can write down the Hamiltonian,

$$\mathcal{H}(y) = -iv_F\sigma_y\frac{\partial}{\partial y} + (m_I - m_H)\sigma_z, \tag{3.161}$$

where, $\hbar = 1$ and $m_I - m_H = \widetilde{m}(y)$. The RHS of Eq. (3.161) resembles y-dependent potential energy in a 1D-free Hamiltonian. Further, let us insist on

$$\begin{aligned} &\widetilde{m}(y) < 0, \quad \text{for } y < 0 \\ &\widetilde{m}(y) > 0, \quad \text{for } y > 0. \end{aligned} \tag{3.162}$$

Thus, as if there is a physical boundary between the topological and trivial states. Let us look at the zero energy solutions.

Now make an ansatz for the y-dependent wavefunction (like variational wavefunction)

$$\psi(y) = i\sigma_y\, e^{f(y)}\, \phi \tag{3.163}$$

where ϕ is a 2-component spinor. Putting Eq. (3.163) in Eq. (3.161)

$$\left(iv_F\frac{df}{dy} + \widetilde{m}(y)\sigma_x\right)\phi = 0 \qquad (\text{using} \quad \sigma_y\sigma_z = i\sigma_x). \tag{3.164}$$

The formal solution for $f(y)$ is obtained as

$$f(y) = -\frac{1}{v_F}\int_0^y dy'\, \widetilde{m}(y') \tag{3.165}$$

where ϕ is assumed to be the eigenstate of σ_x with eigenvalue $+1$.

Also, the effect of $i\sigma_y = e^{i\frac{\pi}{2}\sigma_y}$ is rotated by π around the y-axis.

$$\psi(y) = \exp\left(-\frac{1}{v_F}\int_0^y \widetilde{m}(y')dy'\right)|\sigma_x = -1\rangle. \tag{3.166}$$

The $\exp(-\frac{1}{v_F}\int_0^y \widetilde{m}(y')dy')$ factor allows it to fall off the inside of the sample. So $\psi(y)$ the edge is maximum at the edges. Also, it is an eigenstate of σ_x as it has to mix the two sublattices by hopping along the boundary. For the other Dirac point, the state traverses in the other direction. At larger energies, $\epsilon(k_x) = -v_F k_x$, so that,

$$v_F \text{ (or } v) = \frac{\partial\varepsilon(k_x)}{\partial x} = -v. \tag{3.167}$$

For $\widetilde{m} \to -\widetilde{m}$, we have an electron traversing in the opposite direction at the other cone.

Finally, we show the numeric computation of the edge modes in the Kane–Mele nanoribbon by solving the following sets of equations.

$$E_k a_{k,n} = \left[t\left\{1 + e^{(-1)^{n+1}ik}\right\} b_{k,n} + t b_{k,n+1}\right] s_0 + m_I a_{k,n} s_0 + 2t_2 \left[a_{k,n} \sin k + e^{(-1)^{n+1}\frac{ik}{2}} \sin\frac{k}{2}\{a_{k,n-1} + a_{k,n+1}\}\right] s_z + i\lambda_R \left[\left\{-\frac{1}{2}\left(1 + e^{(-1)^{n+1}ik}\right) b_{k,n} + b_{k,n+1}\right\} s_y - \left\{(-1)^n \frac{\sqrt{3}}{2}\left(1 - e^{(-1)^{n+1}ik}\right) b_{k,n}\right\} s_x\right] \tag{3.168}$$

$$E_k b_{k,n} = \left[t\left\{1 + e^{(-1)^{n}ik}\right\} a_{k,n} + t a_{k,n-1}\right] s_0 - m_I b_{k,n} s_0 + 2t_2 \left[b_{k,n} \cos k + e^{(-1)^{n}\frac{ik}{2}} \cos\frac{k}{2}\{a_{k,n-1} + a_{k,n+1}\}\right] s_z + i\lambda_R \left[\left\{\frac{1}{2}\left(1 + e^{(-1)^{n}ik}\right) b_{k,n} + b_{k,n-1}\right\} s_y - \left\{(-1)^{n+1} \frac{\sqrt{3}}{2}\left(1 - e^{(-1)^{n}ik}\right) a_{k,n}\right\} s_x\right]. \tag{3.169}$$

Fig. 3.31(a) clearly shows the existence of spin-filtered edge modes in the topological phase, while they are absent in Fig. 3.31(b). The presence of the helical modes carrying spin resolved currents at each edge are shown in the yellow panel below [in Fig. 3.31(c)]. Each of the conducting modes denotes a channel for each spin. Thus, the model supports spin-polarized conduction via the edge modes, while the bulk remains gapped. It must be kept in mind that the Chern number is identically equal to zero in this case owing to the time reversal symmetry being intact. However, the topological invariant here is the $\mathcal{Z}_2$ index, which is non-zero.

3.11 SPIN HALL CONDUCTIVITY

We have seen that even though the individual conducting edge states have a non-zero Chern number, the total Chern number, C still vanishes owing to the time reversal symmetry being present. Thus, the charge Hall conductivity vanishes, that is, $\sigma_{xy} = 0$. However, the spin Hall conductivity survives.

In order to calculate the spin Hall conductivity, let us rewind the Corbino disc argument due to Laughlin. When a quantum of flux Φ_0 is added to the inner edge of the disc, an electron is transferred from the inner to the outer edge of the disc. Say this happens for the up spin leading to a e^2/h (charge) Hall conductivity. For the down-spin sector, in the presence of Φ_0, an electron is transferred backward, that is, from the outer edge to the inner one. Including both the spins, the total Hall conductivity is zero, as demanded by the time reversal invariance. However, in the process, a net spin is transferred from the inner to the outer edge. The corresponding spin Hall conductance is given by

$$G_s = \frac{\hbar}{2e}\left(\frac{e^2}{h} + \frac{e^2}{h}\right) = \frac{e}{2\pi}. \tag{3.170}$$

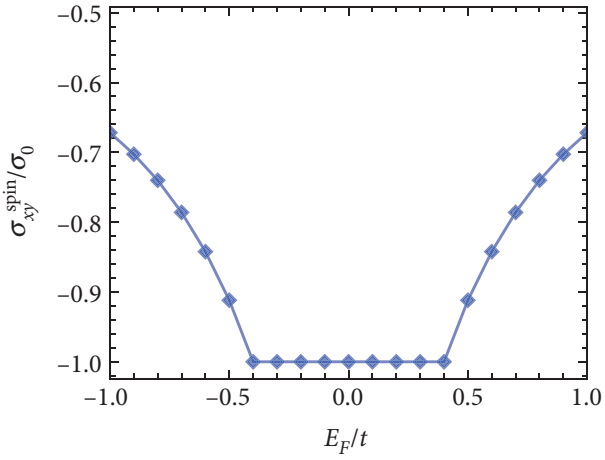


FIG. 3.32
The spin Hall conductivity as a function of the Fermi energy is plotted. A quantization plateau in the vicinity of the zero-Fermi energy is clearly visible.

Thus, G_s is quantized in the unit of $e/2\pi$.

The spin Hall conductance as a function of the Fermi energy is presented for graphene in Fig. 3.32. A plateau is in the vicinity of the zero-Fermi energy (zero bias), a signature of the quantized nature of the spin Hall conductivity, is visible.

There is a subtle point about the intrinsic spin-orbit coupling that deserves special mention, namely, the second term in Eq. (3.155) KM Hamiltonian commutes with the Hamiltonian, that is $[\mathcal{H}_{KM}, S_z] = 0$. However, usually SOC terms are spin non-conserving which means they can mix different spins. Thus, in addition to the intrinsic SOC, other types of SOC can also be present. One such SOC is the Rashba spin-orbit coupling (RSOC), which, as we shall see different spin components.

3.11.1 Rashba spin-orbit coupling

In solids, free (or nearly free) electrons do not feel the strong attraction of the nucleus of their host atoms. However, the electrons may still experience an electric field or a potential gradient due to internal effects. As we know, if the electrons experience a strong electric field of potential gradient, then there is a possibility of emerging of a spin-orbit coupling. So if a potential gradient exists across the interface due to the structural inversion asymmetry, there will be a spin-orbit coupling, and named after its discoverer, E.I. Rashba, that is, Rashba spin-orbit coupling (RSOC) (Bychkov and Rashba, 1984). The importance of the RSOC lies in the fact that asymmetry in the confinement potential can be varied by electrostatics means, allowing one to tune the RSOC strength by an external gate voltage. The strength of the RSOC also depends on the crystal structure in quantum wells and is largest for narrow gap III-V semiconductors, such as InAs and InGaAs. In the following subsection, we shall describe the RSOC in a continuum model. Later on, we shall extend our discussion on graphene.

RSOC yields couples the wave vector with the spin degrees of freedom of the electrons. Further, it leads to the orientation of spins which point perpendicular to the direction of the electron propagation wave vector. The free particle Hamiltonian including RSOC is described by

$$\mathcal{H}_R = -\mu \cdot \mathbf{B} = -\mu \cdot \frac{\mathbf{v} \times \mathbf{E}}{c^2} = \frac{eE}{mc^2}\mathbf{S} \cdot (\mathbf{v} \times \hat{\mathbf{z}}) = \frac{eE\hbar^2}{8\pi^2 m^2 c^2}\sigma \cdot (\mathbf{k} \times \hat{\mathbf{z}}) = \alpha_R(\hat{\mathbf{z}} \times \mathbf{k}) \cdot \sigma, \qquad (3.171)$$

where $\alpha_R = \frac{eE\hbar^2}{8m^2\pi^2c^2}E$ is the strength of the RSOC, $\boldsymbol{\sigma}$ is a vector of the Pauli spin matrix $\mathbf{E} = -\nabla V$ is the electric field along $\hat{\mathbf{z}}$ direction. α_R can be tuned using an external gate voltage. In the absence of any Zeeman coupling, assuming elastic scattering and for $\hat{n} = \hat{z}$ (as per convention), the total Hamiltonian for the electron is given by

$$\mathcal{H} = \frac{p^2}{2m} + \alpha(\mathbf{p} \times \sigma) \cdot \hat{z} = \frac{p^2}{2m} + \alpha(\sigma_x p_y - \sigma_y p_x). \tag{3.172}$$

This Hamiltonian yields the following energy spectrum,

$$E(k) = \frac{\hbar^2 k^2}{2m} \pm \alpha\, \hbar|k|, \tag{3.173}$$

where $|k|$ is the modulus of electron momentum with the plus and the minus signs denote two possible spin directions. The associated wave functions are given by

$$\Psi_{\pm}(x, y) = e^{i(k_x x + k_y y)} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \pm i e^{-i\theta} \end{pmatrix}, \tag{3.174}$$

where $\theta = tan^{-1}(k_y/k_x)$. It is easily understood that the spin states are always perpendicular to the direction of motion [Eq. (3.174)]. If an electron moves along the x-direction, the spinor part of the eigenvector becomes $(1, \pm i)$, that is, the spin up and the spin down are locked in y-direction. In contrast, if the electron moves along the y-direction, the eigenvectors become $(1, \pm 1)$, that is, the spin up and the spin down states are constrained in the x direction (see Fig. 3.33).

In Figs. 3.33(c)–3.33(e), the energy spectrum as a function of momentum, k_y (keeping k_x constant) for a 2DEG are plotted corresponding to the following situations. Fig. 3.33(c) is related to a free electron in 2DEG where the spin degeneracy is present. Figure 3.33(d) represents the energy spectrum for an electron in the presence of a magnetic field $\mathbf{B}$, the spin degeneracy is lifted by the Zeeman splitting, and the gap-separating spin up and spin down bands is equal to $g\mu_B B$ where g is the Bohr magneton. Figure 3.33(e) presents a one-dimensional view of the energy spectrum for an electron in the presence of RSOC. The spin degeneracy is lifted up, except for $k_y = 0$. In this situation, the degeneracy is removed without opening any gap. At $k_y = 0$, the spin spectra are degenerate.

3.11.2 Rashba spin-orbit coupling in graphene

Writing the low energy Hamiltonian in the vicinity of the Dirac points for graphene

$$\mathcal{H}_R = \lambda_R(S_x q_y - S_y q_x) \tag{3.175}$$

which mixes up and down spins, which is why $[\mathcal{H}_R, S_z] \neq 0$. We prefer to call the strength as λ_R here, instead of α_R which was earlier used by us. The corresponding spectrum is given by

$$E_{\gamma\delta}(\mathbf{q}) = \mu\sqrt{\mathrm{q}^2 + (m_H + \delta\lambda_R)^2} + \delta\lambda_R, \tag{3.176}$$

where the indices $\mu = \pm 1$ and $\delta = \pm 1$ yield the conduction and the valence band spectra at the $\mathbf{K}$ and the $\mathbf{K}'$ points. As we have already realized that the spectrum is gapped in the presence of

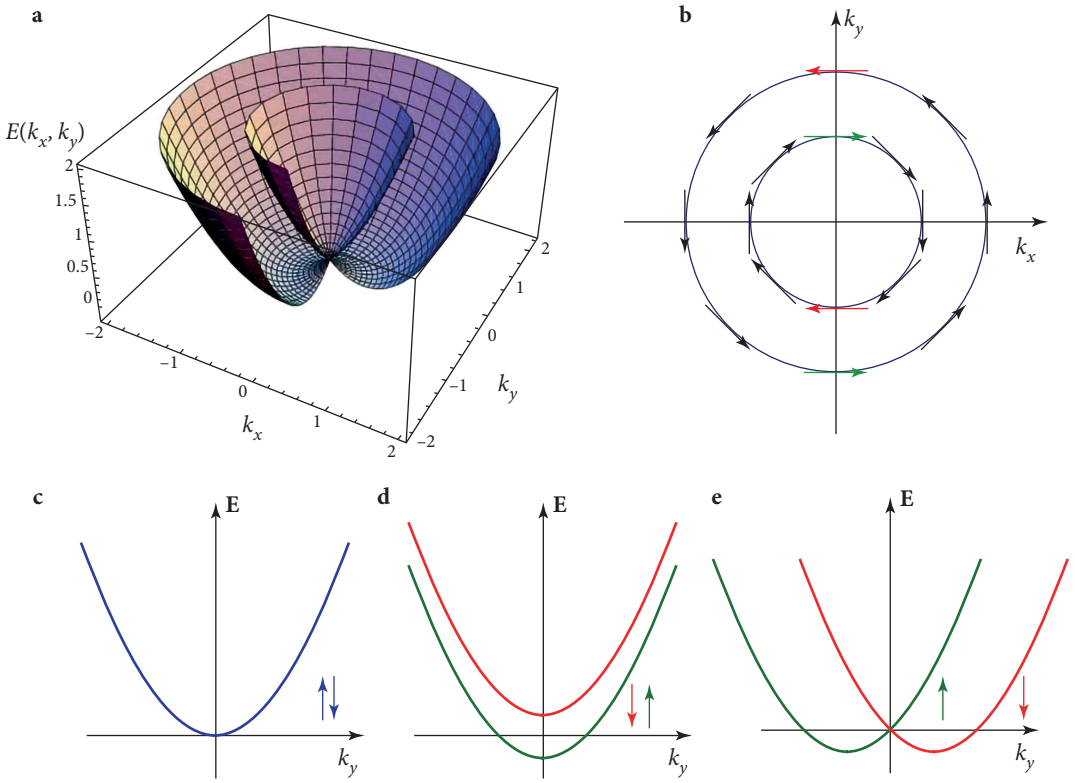


FIG. 3.33
(a) Three-dimensional energy spectrum of the Hamiltonian $\mathcal{H}$ [Eq. (3.172)]. (b) Fermi energy contours for the Hamiltonian $\mathcal{H}$. (c) Energy spectrum for a free electron. (d) Energy spectrum for an electron in the presence of a magnetic field (Zeeman splitting). (e) Energy spectrum for an electron in the presence of Rashba spin-orbit coupling.

the intrinsic spin-orbit coupling (namely, the Haldane term) itself, and now when λ_R is included, it will start competing with m_H when $\delta = -1$. Also, with increasing λ_R, the energy gap decreases. At $m_H = \lambda_R$, the spectrum consists of a Dirac cone with two gapped parabolic bands.

For the sake of completeness, we reiterate that the strength of RSOC is too weak to yield any observable effects. For example, $\lambda_R \sim 10^{-3}$ K in graphene, while the kinetic energy is much larger. There are techniques to enhance RSOC by using heavier adatoms, using an external gate voltage or bend the graphene layer. The basic idea is to create a strong gradient of the electric potential. We shall not discuss this any further and suggest more specialized reviews on the subject.

3.11.3 $\mathbb{Z}_2$ Invariant

We have seen that the Kane–Mele model preserves the time reversal symmetry, and hence the Chern number, which is a $\mathbb{Z}$ invariant, would be zero. This requires us to look for a new topological invariant, namely, the $\mathbb{Z}_2$ invariant, which we shall discuss below. Now, on more general grounds, we need to understand how the presence (or absence) of different discrete symmetries affects the topological invariant of a system. In the appendix, we include a discussion on the "*ten fold*" classification scheme that will aid us in decoding the nature of the topological invariant for a given system.

Since we shall include the Rashba SOC term in the KM Hamiltonian (see Eq. (3.154), and that it respects all symmetries of graphene (for example, Chern number equal to zero), a new topological invariant has to emerge. A priori, it is the $\mathbb{Z}_2$ invariant that we are talking about, however we refer to the topological classification by Altland and Zirnbauer in Altland and Zirnbauer (1997) and Ryu *et al.* (2010).

Let us discuss the topological invariant relevant here, namely, the $\mathbb{Z}_2$ index that characterizes the topological properties of the system. We shall only talk about an inversion symmetric system. For the calculation of the $\mathbb{Z}_2$ index, one may consider the Bloch wave functions, $u_i(\mathbf{k}_i)$ of the occupied bands corresponding to a pair of points $\mathbf{k}_1$ and $\mathbf{k}_2$ in the Brillouin zone. These two points denote the locations of the band extrema (minima for the conduction band and maximum for the valence band) in the BZ. The wave function at one of these points can be obtained by time reversing the wave function corresponding to the other one, that is, $|u_i(\mathbf{k}_1)\rangle = \mathcal{T}\,|u_i(\mathbf{k}_2)\rangle$, and vice versa where $\mathcal{T}$ denotes the time reversal operator. Since the Hamiltonian is time reversal invariant, we can decompose the Hamiltonian, $\mathcal{H}(\mathbf{k})$ and its corresponding occupied band wave functions, $|u_i(\mathbf{k})\rangle$ into even and odd subspaces. The even subspace has the property that $\mathcal{T}\,|u_i(\mathbf{k})\rangle$ is equivalent to $|u_i(\mathbf{k})\rangle$ upto a $U(2)$ rotation. Whereas, the wave functions corresponding to the odd subspace have the property that the space spanned by $\mathcal{T}\,|u_i(\mathbf{k})\rangle$ is orthogonal to that of $|u_i(\mathbf{k})\rangle$. Now, the $\mathbb{Z}_2$ invariant can be calculated by considering the momenta which belong to the odd subspace. We compute the expectation value of the time reversal operator between $|u_i(\mathbf{k})\rangle$ and $|u_j(\mathbf{k})\rangle$, namely, $\langle u_i(\mathbf{k})|\,\mathcal{T}\,|u_j(\mathbf{k})\rangle$. This yields a matrix which is antisymmetric. Hence, we have

$$\langle u_i(\mathbf{k})|\mathcal{T}|u_j(\mathbf{k})\rangle = \epsilon_{ij}P(\mathbf{k}), \tag{3.177}$$

where ϵ_{ij} is the Levi-Civita symbol and $P(\mathbf{k})$ is the Pfaffian of the matrix defined as

$$P(\mathbf{k}) = \mathrm{Pf}\left[\langle u_i(\mathbf{k})|\Theta|u_j(\mathbf{k})\rangle\right]. \tag{3.178}$$

For a 2×2 antisymmetric matrix A_{ij}, the Pfaffian picks up the off-diagonal component. Now the absolute value of this Pfaffian is unity in the even subspace, while it is zero in the odd subspace. Therefore, we dissect the BZ into two halves, such that the points $\mathbf{k}_1$ and $\mathbf{k}_2$ lie in different halves. Thus, the $\mathbb{Z}_2$ index can be computed by performing the integral,

$$\mathbb{Z}_2 = \frac{1}{2\pi i}\oint_C \mathrm{d}\mathbf{k}\cdot\boldsymbol{\nabla}\log\left(P(\mathbf{k}) + i\delta\right), \tag{3.179}$$

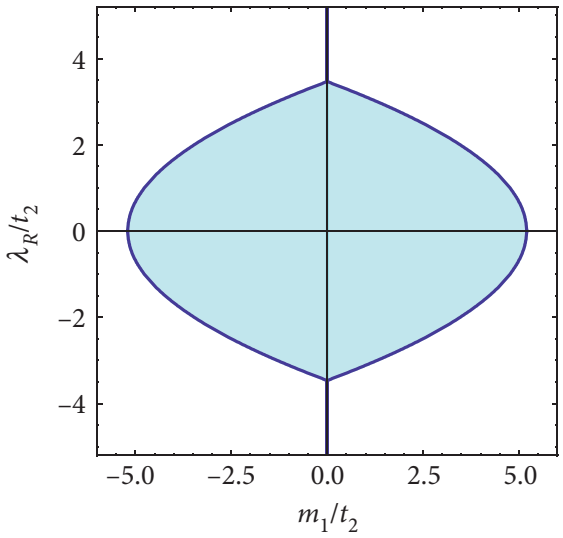


FIG. 3.34
The phase diagram for the Kane–Mele model.The light blue region corresponds to $\mathbb{Z}_2 = 1$, which denotes a quantum spin Hall insulator. The white region outside the lobe refers to a trivial insulator with $\mathcal{Z}_2 = 0$.

where δ is the convergence factor and the contour C is the circumference of the halved BZ discussed above.

The variation of the $\mathbb{Z}_2$ index is shown in the parameter plane defined by the Rashba coupling, λ_R and the Semenoff mass, m_I (both scaled by the NNN hopping t_2), which is shown to have a value 1 in the light blue region, and vanishes outside in Fig. 3.34. The region with non-zero $\mathbb{Z}_2$ invariant will host spin filtered chiral edge modes, and will denote a quantum spin Hall insulator, and the region outside, denotes a trivial insulator. Figure 3.34 denotes the phase diagram of a quantum spin Hall insulator, in the same spirit as Fig. 3.28 denotes the phase diagram for a Chern insulator. To remind the readers, the Chern number vanishes here due to the presence of the time reversal symmetry, thereby ruling out the possibility of a quantum Hall-like state. However, a new topological phase emerges, known as the QSH phase.

3.12 SPIN HALL EFFECT

To shed light on possible applications, we give a brief description of the spin hall effect, and cursorily on the subject of spintronics. The spin Hall effect (SHE) is the generation of the spin current perpendicular to the applied charge current. This leads to the accumulation of spins of opposite kinds at the edges of the sample. The spin selection can be facilitated by a strong spin-orbit coupling (SOC). Strong SOC may be intrinsic to doped semiconductors. The proposal has triggered intense investigation of the phenomenon, and the lead has been taken by the first observation of SHE in *n* doped semiconductors (Kato *et al.*, 2004), and 2D hole gases (Wunderlich *et al.*, 2005). Both the experiments measure directly the spin accumulation induced at the edges of the sample through different optical techniques. However, more quantitative and accurate estimates can be obtained by measuring the Hall angle. An excellent review of the family of the spin Hall effects, comprising of SHE (discussed briefly above), inverse SHE in which a pure spin current generates a charge current, and even an anomalous Hall effect (AHE) in which a charge current generates a polarized transverse charge current in a ferromagnetic material can be found in the review by Sinova *et al.* (2015).

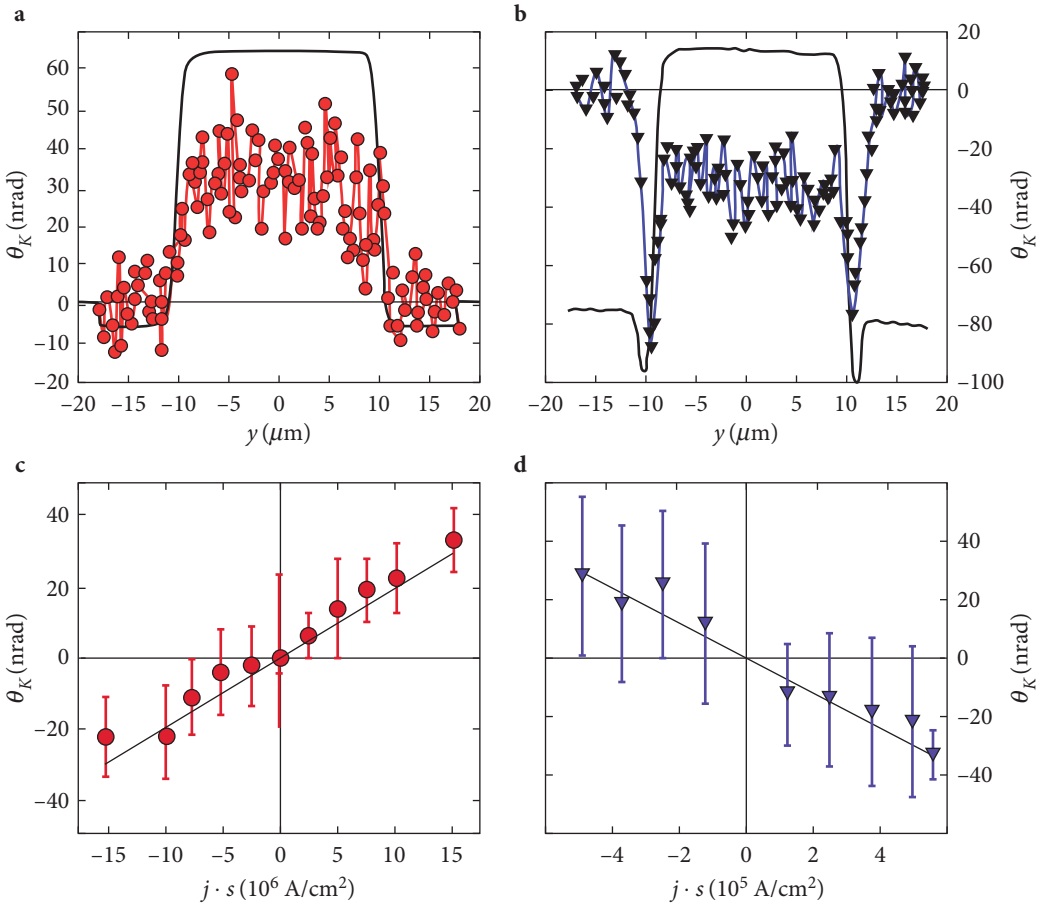


FIG. 3.35
Kerr rotation spectroscopy studies [taken from Stamm *et al.* (2017)] are shown. In (a) and (b), the Kerr rotation angles θ_K as a function of y-scan (in μm) are shown for certain values of the current density, j (j in 10^6 A/cm^2) corresponding to thin films of Pt and W, respectively. (c) and (d) show the linear variation of θ_K with j for Pt and W. The slopes are opposite in the two cases.

Let us briefly look at the early experimental determination of SHE. In the magneto-optical detection of SHE in thin films of platinum (Pt) and tungsten (W), the generation of spin currents from charge currents in the presence of spin-orbit coupling is shown via Kerr rotation spectroscopy (Zhong and Forukas, 2008) in Stamm *et al.* (2017). In Fig. 3.35, the measured Kerr rotation angles, θ_K as a function of line scan (y) are shown for Pt and W samples of 10nm-15nm width at applied current densities in a

particular range of values (see Stamm *et al.* (2017) for details). The Kerr rotation signals are discernible for both samples, as shown in Figs. 3.35(a) and 3.35(b). Apart from the irregularities due to reflection from the edges, θ_K remains approximately constant with mutually opposite signs for Pt and W. In Figs. 3.35(c) and 3.35(d), θ_K as average values of the line scan are shown to vary linearly with the current density j. The results are conclusive in demonstrating SHE using Kerr rotation spectroscopy studies via spatial evolution of the spin dynamics. Possibly, the corresponding data for extrinsic SHE (in presence of an external field) should have a stronger signal, but the above results still confirm the possibility of realizing intrinsic SHE in experiments.

3.12.1 Spin current

Measuring the spin current is central to the study of SHE and hence to the emerging field of spintronics. The spin current has to be contrasted with the charge current. The charge current density, $\mathbf{j}_{el}(\mathbf{r}, t)$ is given by

$$\mathbf{j}_{el}(\mathbf{r}, t) = \mathrm{Re}\left[\psi^{\dagger}(\mathbf{r}, t)(e\mathbf{v})\psi(\mathbf{r}, t)\right] \tag{3.180}$$

which further obeys a continuity equation of the form,

$$\frac{d\rho^{el}}{dt} + \nabla \cdot \mathbf{j}_{el} = 0 \tag{3.181}$$

where $\mathbf{v}$ is the velocity of the electrons (charges), and $\rho^{el}(\mathbf{r}, t) = e\psi^{\dagger}(\mathbf{r}, t)\psi(\mathbf{r}, t)$ is the charge density. The continuity equation in Eq. (3.180) is the consequence of the invariance of charge. However, in the case of spin current, there is an ambiguity that arises from the fact that the spin is not an invariant quantity in spin transport owing to the presence of the spin-orbit coupling (Sun and Xie, 2005) Usually the spin current is defined as $\langle \mathbf{v} \cdot \mathbf{s} \rangle$ which is a non-conserved quantity. However, in the classical sense, just as the charge current density, the spin current density, $\mathbf{j}_s$ (the subscript s refers to the spin) can be written as

$$\mathbf{j}_s = \mathrm{Re}\left[\psi^{\dagger}(\mathbf{r}, t)(\mathbf{v} \cdot \mathbf{s})\psi(\mathbf{r}, t)\right]. \tag{3.182}$$

For a generic Hamiltonian with spin-orbit coupling, it can be shown that,

$$\mathrm{Re}(\psi^{\dagger} v_{\alpha} s_{\beta} \psi) = \mathrm{Re}(\psi^{\dagger} s_{\alpha} v_{\beta} \psi),$$

where α and β refer to the components in an orthogonal coordinate system. From $\mathbf{j}_s(\mathbf{r}, t)$ one can get the total spin current using,

$$\mathrm{I}_{s\alpha}(t) = \int dA\hat{\alpha} \cdot \mathbf{j}_s(\mathbf{r}, t) = \int dA \left[\psi^{\dagger}(\mathbf{r}, t)\frac{1}{2}(\mathbf{v} \cdot \mathbf{s} + \mathbf{s} \cdot \mathbf{v})\psi(\mathbf{r}, t)\right], \tag{3.183}$$

where dA is the elemental area and $\hat{\alpha}$ denotes a certain direction [that is, $\hat{\alpha} \in (\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}})$].

This clearly tells us that the spin current density operator, $\mathbf{j}_s$ is an anticommutator of $\{s_{\alpha}, v_{\beta}\}$ multiplied by a factor of 1/2. In terms of the Pauli matrices,

$$j_s^{\alpha\beta} = \frac{1}{4}\{s_{\alpha}, v_{\beta}\} \quad (\hbar = 1). \tag{3.184}$$

Quite strikingly, unlike the charge current, which is odd under time reversal, the spin current is invariant under time reversal operation. However, Ohm's law ($\mathbf{j} = \sigma \mathbf{E}$) holds for both $\mathbf{j}_{el}$ and $\mathbf{j}_s$. Since the electric field, $\mathbf{E}$ is even under time reversal, the charge conductivity, σ_{el} is odd, while the spin conductivity, σ_s is even.

For concreteness, let us specialize in a particular case, where we choose $\alpha = z$ and $\beta = y$. The y-component of the velocity is obtained from Hamilton's equation of motion,

$$v_y = \frac{\partial \mathcal{H}}{\partial p_y}. \tag{3.185}$$

Considering a Hamiltonian for a 2D electron gas with Rashba spin-orbit coupling,

$$\mathcal{H} = \frac{p^2}{2m} - \lambda_R \sigma \cdot (\hat{\mathbf{z}} \times p) \quad \hbar = 1, \tag{3.186}$$

which yields

$$v_y = \frac{p_y}{m} + \lambda_R \sigma_x. \tag{3.187}$$

Finally, the spin current density assumes the form,

$$j_s^{yz} = \frac{1}{2m} \sigma_z p_y. \tag{3.188}$$

Just as an electric current induces a magnetic field, a pure spin current induces an electric field. The magnetic moment due to the spin of the electron generates a current (Fisher, 1971) This can be understood as follows. Since the motion of a single magnetic moment is equivalent to an electric dipole, which creates an electric field in its vicinity, there will be an electric field due to the motion of a magnetic moment. An estimation of the electric field can be made as follows. Consider two equal and opposite magnetic charges $\pm q_m$ separated by a small distance d moving in opposite directions. Such "*moving*" magnetic dipoles whose magnetic moments are given by $\mathbf{m} = (q_m d)\hat{\mathbf{r}}$ ($\hat{\mathbf{r}}$ denotes the polarization direction) constitute a spin current. Each member of the group, that is a single magnetic moment will generate a magnetic field. This magnetic field in turn generates an electric field, which is given by

$$\mathbf{E} \sim \frac{\mu}{4\pi} \int dV j_s \times \frac{1}{R^3} \left(\hat{\mathbf{r}} - \frac{3\mathbf{R}(\mathbf{R} \cdot \hat{\mathbf{r}})}{R^2} \right), \tag{3.189}$$

where dV is an elemental volume. This electric field is quite tiny in magnitude, yet can produce measurable effects (Shen, 2008)

Just as a current carrying wire experiences a force in a magnetic field ($\sim \mathbf{j} \times \mathbf{B}$), spin current experiences a force $\sim \mathbf{j}_s \times \mathbf{E}$. In spite being small, it is able to control the motion of a spin, including zitterbewegung (jittery motion) of the Dirac electrons. Thus, in semiconductors, where the spin-orbit coupling can be fairly strong, or can even be enhanced by external means, electrons with opposite spins are deflected along the opposite edges of the sample. Thus, a spin unpolarized (paramagnetic) system can yield a pure spin current perpendicular to the direction of the electric field.

Over the last decade and a half, studies concerning the spin current and its application to spintronics in terms of efficiently generating, manipulating and detecting the spin accumulation phenomena have received a plethora of attention. Some progress has also occurred from the device fabrication perspective, via techniques such as spin injection. A major advantage in dealing with the spin current lies in the non-dissipative (or very less dissipation) nature which arises owing to the time reversal invariance of the spin current. This property is in direct contrast with that of the charge current. A simple way to understand the role of time reversal invariance in the phenomenon of dissipation that can be understood with the aid of a damped harmonic oscillator, whose Hamiltonian may be written as

$$\mathcal{H} = \frac{p^2}{2m} + \frac{1}{2}kx^2 + \alpha\dot{x},$$

where the $\alpha\dot{x}$ denotes the damping and breaks the time reversal symmetry. Without this term, the time reversal invariance holds and the scenario is non-dissipative. Thus, a time reversal invariant system presents a non-dissipative scenario, which is precisely the main advantage of the spin transport phenomena.

3.13 APPENDIX

3.13.1 Periodic table of topological materials: Ten fold classification

The ten fold symmetry classification of topological matter involves categorizing topological insulators and superconductors on the basis of the presence or absence of three discrete symmetries. These symmetries include time-reversal symmetry, particle-hole (or equivalently charge conjugation) symmetry and chiral symmetry. In order to decide which class a material falls into, we investigate which of the aforementioned symmetries, the material hosts.

It is important to understand the inherent physical meaning of these non-spatial symmetries in mathematical terms. We have done this earlier in Sec. 3.2, but wish to repeat here some of the essential properties for the discussion to be self-sufficient in this appendix. The time reversal operator has the effect of reversing the motion of a particle, i.e., it transforms $\mathbf{k}$ to $-\mathbf{k}$. If the evolution of a system under the action of the time reversal operator remains the same, the system is called time reversal symmetric. Mathematically, the time reversal operator is represented by K (for spinless) or $i\sigma_y K$ (for spinor systems). Here, K represents the complex conjugation operator and σ_y is the y-component of the Pauli matrix. The charge-conjugation operator, as the name suggests, physically represents replacing a particle with its conjugate within the system. If under such a replacement, the system remains invariant, we call it to be particle-hole symmetric. Importantly, the time reversal symmetry and the particle-hole symmetry are antiunitary symmetries which square to ± 1. Lastly, there are systems which are neither symmetric under time reversal nor particle-hole symmetries. However, if the motion within the

system is reversed and the particles are replaced by their conjugates simultaneously, invariance is re-established. Such systems are purely chiral systems. Chiral symmetry is represented as the product of the time reversal and the particle-hole symmetry. It is difficult to understand the physical implications of chiral symmetry. However, in models where a sublattice structure is apparent, we can associate chiral symmetry with the sublattice symmetry. This means that if the two sublattices of the system are interchanged, the physics remains invariant.

3.13.2 Mathematical representation of the symmetries

We consider a Hamiltonian $\mathcal{H}(k)$ in the k-space. Let us recall what various symmetries do to the Hamiltonian, $\mathcal{H}$. The time reversal symmetry implies that $\mathcal{T}\mathcal{H}(k)\mathcal{T}^{-1} = \mathcal{H}(-k)$. For a particle-hole symmetric system, $\mathcal{C}\mathcal{H}(k)\mathcal{C}^{-1} = -\mathcal{H}(-k)$. Lastly, for a chiral symmetric system $\mathcal{S}\mathcal{H}(k)\mathcal{S}^{-1} = -\mathcal{H}(k)$. To understand the physical picture behind time reversal, it could be interesting to study its effect on the time evolution operator $e^{i\mathcal{H}t}$, $\mathcal{H}$ being the Hamiltonian for the system. We know that time reversal has the effect of complex conjugation, on spinless systems, that is,

$$\mathcal{T}i\mathcal{T}^{-1} = -i. \tag{3.190}$$

Hence, the time evolution operator looks like

$$\mathcal{T}e^{i\mathcal{H}t}\mathcal{T}^{-1} = e^{-i\mathcal{H}t} = e^{i\mathcal{H}(-t)}. \tag{3.191}$$

From this expression, it is apparent why the time reversal operator has the name that it does. It changes t to $-t$ that result in the particle moving in the opposite direction. Another way to study the symmetry operators could be to talk in the language of second quantization, or more specifically, the creation-annihilation operators. Charge-conjugation causes a creation operator to transform into a superposition of annihilation operators and vice versa. Time reversal on the other hand only has the effect of complex conjugation of the coefficients.

$$\begin{aligned}
\mathcal{T}\psi_A\mathcal{T}^{-1} &= \sum_B (U^*_{\mathcal{T}})_{A,B}\psi_B \\
\mathcal{T}\psi^\dagger_A\mathcal{T}^{-1} &= \sum_B (U_{\mathcal{T}})_{B,A}\psi_B \\
\mathcal{C}\psi_A\mathcal{C}^{-1} &= \sum_B (U^*_{\mathcal{C}})_{A,B}\psi^\dagger_B \\
\mathcal{C}\psi^\dagger_A\mathcal{C}^{-1} &= \sum_B (U_{\mathcal{C}})_{B,A}\psi_B.
\end{aligned} \tag{3.192}$$

The classes are differentiated on the basis of whether the time reversal and charge-conjugation operator square to 1, -1 or 0. This gives a total of 9 separate classes. However, as previously discussed, there is another class having only the chiral symmetry. This corresponds to purely chiral systems and constitutes the tenth class in the classification scheme. Two broad sub-groups are apparent from this table. The first one is the purely chiral AIII group, which only hosts chiral systems. The second group

Table 3.2
Table for ten fold symmetry classification. Different symmetry classes are indicated in the bottom row (Ryu *et al.*, 2010).

	BDI	CI	AI	BIII	CII	AII	D	C	A	AIII
$\mathcal{T}$	1	1	1	−1	−1	−1	0	0	0	0
$\mathcal{C}$	1	−1	0	1	−1	0	1	−1	0	0
$\mathcal{S} = \mathcal{TC}$	1	1	0	1	1	0	0	0	0	1

is the D group, which only has charge-conjugation symmetry. This group hosts systems which are superconducting. Also, there are classes like BDI which hosts both chiral and superconducting systems.

Next, we venture upon understanding how this classification originated and wherein lies its importance. The idea behind the tenfold classification is entirely mathematical and was introduced by Elie Cartan, way back in the 1920s. Cartan's classification of $N \times N$ Hermitian matrices into 10 different groups bears one-one correspondence with this 10-fold fermionic symmetry classification. The form of classification that we use today was introduced by Altland and Zirnbauer (1997) The importance of this classification lies in the fact that corresponding to every class and pertaining to the dimension of our system, we may or may not have a distinct topological invariant. This invariant is a signature of the bulk-boundary correspondence and hints towards the presence of edge characteristics that make our system topologically non-trivial, and hence interesting.

The $\mathbb{Z}$ type invariant refers to an integer classification, whereas the $\mathbb{Z}_2$ invariant refers to a binary classification. A zero (0) at any position in the table refers to the fact that no topological insulator or superconductor can be found corresponding to that class and dimension. It would be interesting to associate a few known examples with the given table of invariants. A 2D system belonging to the class "A" resembles our well-known integer quantum Hall system, thereby making the Chern number be a $\mathbb{Z}$ invariant. From the table, we find that it has a $\mathbb{Z}$ type topological invariant. This, as we know, is indeed true and the $\mathbb{Z}$ invariant can be directly associated with the conductivity of the system. Another example would be the quantum spin Hall system, which belongs to the class AII in 2D. It is characterized by a $\mathbb{Z}_2$ topological invariant. Similarly, a 1D SSH chain belongs to the BDI class. A 1D spinless Kitaev chain with p-wave superconducting correlations time reversal, chiral and particle-hole (or charge conjugation) symmetries and hence belongs to the BDI class as well with $\mathbb{Z}$ being the topological invariant.

Table 3.3
Table for the topological invariants in different symmetry classes (Ryu *et al.*, 2010).

AZ/d	1	2	3
A	0	$\mathbb{Z}$	0
AIII	$\mathbb{Z}$	0	$\mathbb{Z}$
AI	0	0	0
BDI	$\mathbb{Z}$	0	0
D	$\mathbb{Z}_2$	$\mathbb{Z}$	0
DIII	$\mathbb{Z}_2$	$\mathbb{Z}_2$	$\mathbb{Z}$
AII	0	$\mathbb{Z}_2$	$\mathbb{Z}_2$
CII	$\mathbb{Z}$	0	$\mathbb{Z}_2$
C	0	$\mathbb{Z}$	0
CI	0	0	$\mathbb{Z}$

The tenfold classification of symmetry shows certain interesting trends when studied carefully. In the given Table 3.4, we find that, corresponding to every dimension, there are only 5 classes which host topologically non-trivial materials. Furthermore, for every class, there exists a periodicity of 8. The invariants repeat

Table 3.4
Periodic table of topological insulators and superconductors.

AZ/d	0	1	2	3	4	5	6	7	8	9
A	$\mathbb{Z}$	0	$\mathbb{Z}$	0	$\mathbb{Z}$	0	$\mathbb{Z}$	0	$\mathbb{Z}$	⋯
AIII	0	$\mathbb{Z}$	0	$\mathbb{Z}$	0	$\mathbb{Z}$	0	$\mathbb{Z}$	0	⋯
AZ/d	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
AI	$\mathbb{Z}$	0	0	0	$\mathbb{Z}$	0	$\mathbb{Z}_2$	$\mathbb{Z}_2$	$\mathbb{Z}$	⋯
BDI	$\mathbb{Z}_2$	$\mathbb{Z}$	0	0	0	$\mathbb{Z}$	0	$\mathbb{Z}_2$	$\mathbb{Z}_2$	⋯
D	$\mathbb{Z}_2$	$\mathbb{Z}_2$	$\mathbb{Z}$	0	0	0	$\mathbb{Z}$	0	$\mathbb{Z}_2$	⋯
DIII	0	$\mathbb{Z}_2$	$\mathbb{Z}_2$	$\mathbb{Z}$	0	0	0	$\mathbb{Z}$	0	⋯
AII	Z	0	$\mathbb{Z}_2$	$\mathbb{Z}_2$	$\mathbb{Z}$	0	0	0	$\mathbb{Z}$	⋯
CII	0	$\mathbb{Z}$	0	$\mathbb{Z}_2$	$\mathbb{Z}_2$	$\mathbb{Z}$	0	0	0	⋯
C	0	0	$\mathbb{Z}$	0	$\mathbb{Z}_2$	$\mathbb{Z}_2$	$\mathbb{Z}$	0	0	⋯
CI	0	0	0	$\mathbb{Z}$	0	$\mathbb{Z}_2$	$\mathbb{Z}_2$	$\mathbb{Z}$	0	⋯

themselves after every 8th dimension. Additionally we find that, if the topological invariant for a particular class in the *nth* dimension is known, we know the invariant for the $(n+1)th$ dimension in the successive class. The origin of the above periodicities can be understood from a discussion on group theory.

3.13.3 Summary and outlook

Quantum Hall states are the first examples of topological insulators which demonstrate completely contrasting electronic behavior between the bulk and the edges of the sample. The bulk of the system is insulating, while there exist conducting states at the edges. Moreover, the Hall conductivity is quantized in units of a universal constant, e^2/h. It became clear later on that the quantization is actually related to a topological invariant known as the Chern number. The geometric interpretation of this invariant is provided by the Gauss-Bonnet theorem, which relates the integral of the Gaussian curvature over a closed surface to a constant which simply counts the number of "genus" (or holes) of the object. In solid state physics, the closed surface is the Brillouin zone and the Gaussian curvature is analogous to a quantity known as the Berry curvature. This brings us to the study of Berry-o-logy where the topological invariants are defined in terms of the Berry phase (a geometric phase picked by a particle during a complete revolution), Berry connection (analogous to the vector potential in electrodynamics), and the Berry curvature (similar to a magnetic field). The machinery is applied to a simple case, in both one and two dimensions, such as an SSH model and graphene. The topology in the SSH model is induced by a dimerized hopping with two atoms per unit cell, and is stabilized by the chiral symmetry of the Hamiltonian. As long as the chiral symmetry is intact, and when the intra-cell hopping amplitude is greater than the inter-cell one, the model displays localized zero modes at the edges. The topological

phase is further characterized by a finite value of the winding number. A similar scenario is presented by the Kitaev model, which involves spinless fermions coupled by p-wave-superconducting correlations on a tight binding chain. The topological property of the model is protected by the particle-hole symmetry that is inherent to superconductors. Further, similar to the SSH model, the topological phase is characterized by zero modes at the edges of an open chain, and a finite value for the winding number. Interestingly, the zero modes is two-fold degenerate, which are inseparable and differ by fermion parity. Moreover, these states have a formal similarity with the Majorana fermions which correspond to their own antiparticles. Further, the symmetry aspects, such as the inversion symmetry (parity), time reversal symmetry are discussed with a view to explore topological properties of graphene. The bulk-boundary correspondence and the existence of the edge states in a nanoribbon geometry are investigated, which serve as an acid test for the topological state. Eventually, following Haldane's conjecture, a topological state emerges by breaking the time reversal symmetry where the system acquires a topological gap at the Dirac points. The presence of such a non-trivial gap is confirmed via the presence of chiral edge states in a ribbon. Further progress is reported in terms of the proposal of a scenario in which the broken time reversal symmetry is revoked using two copies of the Haldane model for each type of spin of the carriers. This is called the Kane–Mele model, which respects all the symmetries that graphene has. Yet there is an important difference which can be brought about by adding the Rashba spin-orbit coupling, which may be weak, but inherent to the two-dimensional systems. The Rashba term leaves the time reversal symmetry intact. Thus, the Kane–Mele model in the presence of the Rashba spin-orbit contribution yields yet another distinct topological state of matter, namely, the quantum spin Hall phase. Owing to the time reversal symmetry is intact. the Chern number vanishes, although the system is characterized by a new topological invariant, known as the $\mathcal{Z}_2$ invariant, which yields the spin Hall conductivity to be non-zero. The prospects of manipulating the spin degree of freedom give birth to an emerging field known as spintronics. The fact that the spin current obeys time reversal symmetry, one gets non-dissipative transport and thus holds the prospect of transmitting information with no (or very little) decay. Thus, the spin transport mechanism does not have any associated Joule heating phenomena. Finally, we have included a brief discussion on the classification of the topological materials according to the schemes introduced by Altland and Zirnbauer. This aid us in identifying the class of the topological systems discussed in this chapter and the corresponding topological invariants.

REFERENCES

Altland, A. and Zirnbauer, M. R., Phys. Rev. B **55**, 1142 (1997).
Bernevig, B. A., Hughes, T. L., and Zhang, S. C., Science **314**, 1757 (2006).
Berry, M. V., Proc. R. Soc. London Ser. A **392**, 45–57 (1984).
Bychkov, Yu. A. and Rashba, E. I., Sov. Phys. JETP Lett. **39**, 78 (1984).
Fisher, G. P., Am. J. Phys. **39**, 1528 (1971).
Fu, L. and Kane, C. L., Phys. Rev. Lett. **100**, 096407 (2008).

Fu, L. and Kane, C. L., Phys. Rev. B **79**, 161408(R) (2009).
Haldane, F. D. M., Phys. Rev. Lett. **61**, 2015 (1988).
Kane, C. L. and Mele, E. J., Phys. Rev. Lett. **95**, 226801 (2005). Phys. Rev. Lett. **95**, 146802 (2005).
Kato, Y. K., Myers, R. C., Gossard, A. C., and Awschalom, D. D., Phys. Rev. Lett. **93**, 176601 (2004).
Kitaev, A.Yu., Annals Phys. **303**, 2 (2003).
Kittel, C., *Introduction to Solid State Physics* (John Wiley and Sons, 1986).
König, M., Science **318**, 766 (2007).
Law, K. T., Lee, P. A., and Ng, T. K., Phys. Rev. Lett. **103**, 237001 (2009).
Lutchyn, R., Sau, J. D., and Das Sarma, S., Phys. Rev. Lett. **105**, 077001 (2010).
Majorana, E., Il Nuovo Cimento (1924–1942) **14**, 171 (2008). ISSN 1827–6121.
Min, H., Hill, J. E., Sinitsyn, N. A., Saha, B. R., Kleinman, L., and McDonald, A. H., Phys. Rev. B **74**, 165310 (2006).
Mondal, S., and Basu, S., J. Phys.: Condens. Matter **33**, 225504 (2021).
Nilsson, J., Akhremov, A. R., and Beenakker, C. W., Phys. Rev. Lett. **101**, 120403 (2008).
Oreg, Y., Rafael, G., and Oppen, F. v., Phys. Rev. Lett. **105**, 177992 (2010).
Perelman, G., Science **314**, 1848 (2006).
Ryu, S. *et al.*, New J. Phys. **12**, 065010 (2010).
Shen, S.-Q., AAPPS Bull. **18**, 29 (2008).
Sinova, J., Valenzuela, S. O., Wunderlich, J., Back, C. H., and Jungwirth, T., Rev. Mod. Phys. **87**, 1213 (2015).
Stamm, C. *et al.*, Phys. Rev. Lett. **119**, 087203 (2017).
Sun, Q.-F. and Xie, X. C., Phys. Rev. B **72**, 245305 (2005).
Thouless, D. J., Kohmoto, M., Nightangle, M. P., and den Nijs, M., Phys. Rev. Lett. **49**, 405 (1982).
Wunderlich, J., Kaestner, B., Sinova, J., and Jungwirth, T., Phys. Rev. Lett. **94**, 047204 (2005).
Xiao, D., Chang, M.-C., and Niu, Q., Rev. Mod. Phys. **82**, 1959 (2010).
Zak, J., Phys. Rev. Lett. **62**, 2747 (1988).
Zhong, Q. and Forukas, J. T., J. Phys. Chem. B **112**, 15529 (2008).

CHAPTER

4 SUPERCONDUCTIVITY

4.1 INTRODUCTION

In solid state physics, many of the phenomena that we are familiar with occur because of the inter-particle interactions among the charge carriers, thereby resulting in an ordered state. The effects of such interactions are best perceptible at low temperature. With an increase in temperature, the thermal motion of the carriers gains prominence, and out do the ordering process. Superconductivity denotes one such ordered phase of matter which stabilizes at low temperature and vanishes when the temperature is increased beyond a certain critical point.

In the year 1908, Kamerlingh Onnes (1911) at the low temperature lab in Leiden, Netherlands successfully liquified Helium (He). At normal atmospheric pressure, the boiling point of helium (He) was found to be 4.2 K, thereby making the exploration of material properties feasible at low temperatures. Studying the electronic conductivity (or resistivity) of metals seemed like a normal choice, as metals are primarily characterized by their electrical resistance. Measurements performed at low temperatures may yield the following possibilities:

i. the resistance vanishes gradually with decreasing temperature.
ii. It may result in a small but finite value at a very low temperature.
iii. It may have a minimum at low temperature, and finally show an upturn before diverging at very low temperatures.

In particular, the last possibility receives support on physical grounds in the sense that at sufficiently low temperature, the carriers are likely to be bound to their respective atoms. Thus, their ability to move around and contribute to the conductivity vanishes.

K. Onnes realized the importance of studying the conductance characteristics of high-purity metals at low temperatures. Initially, he started with gold (Au) and platinum (Pt), mainly because they are noble metals and available in pure form. However, he shifted his attention toward mercury (Hg) which can be obtained in a highly pure form via multiple distillations. Precisely at the boiling point of liquid He, he found that the resistance of an ultra-clean Hg sample sharply vanishes, and at further lower temperatures, the resistance becomes immeasurably small (see Fig. 4.1). The temperature at which the resistance disappears is called the transition temperature, T_c. Below that, a new state of matter emerges, which either completely expels the magnetic field or in some cases, traps magnetic flux having values that are in integer multiple of the flux quantum, Φ_0 ($\Phi_0 = 2.07 \times 10^{-15}$ Wb). This was indeed

https://doi.org/10.1063/9780735422537_004

Modern Perspectives in the Study of Electronic Systems, Saurabh Basu and Sourav Chattopadhyay

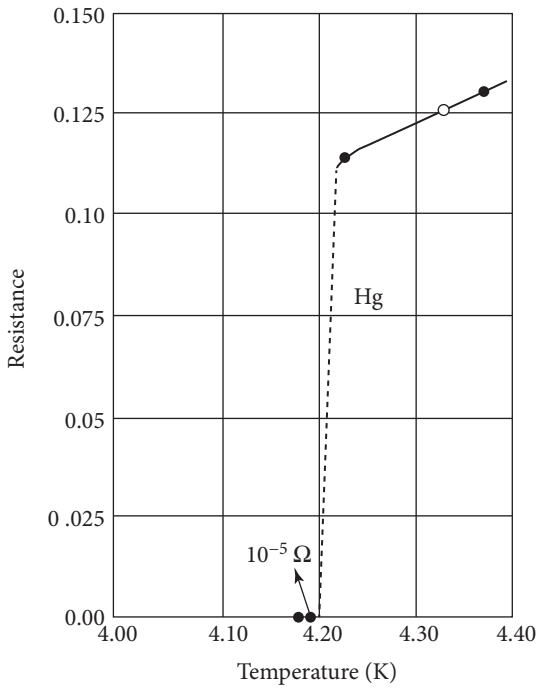


FIG. 4.1
The fall of the electrical resistance of Hg (in Ω) as a function of temperature is shown. Near 4.2 K, there is a sharp fall in resistance, below which the resistance becomes almost zero (After K. Onnes original experiment Kamerlingh Onnes, 1911).

a surprise, since laws of electromagnetic induction predict that ideal conductors retain, and not expel the magnetic field that is trapped inside.

A complete theoretical understanding of the microscopic phenomena had to wait for about half a century until the theoretical derivation by Bardeen, Cooper and Schrieffer (BCS) in 1957 came into existence, which was awarded the Nobel prize in 1972. They realized that at the transition point, the electrons pairwise condense into a new phase, which is a coherent matter wave with a well-defined phase relationship among the pairs. These electrons form pairs mediated via quantized lattice vibrations, namely, the phonons.

For nearly eight decades after the experimental discovery of superconductivity in High in 1908, it remained low temperature phenomena until Bednorz and Müller discovered the onset of superconductivity in the copper oxide planes at larger temperatures. As years passed by, the discovery of a large number of copper oxide superconductors were made with larger and larger critical temperatures, which were thought to have far-reaching consequences in terms of their industrial applications. There was moderate success on that front, however there were many surprises, which arose as the experimental data started coming in on the physical properties of these superconductors. Quite likely, the normal state[1] of these copper oxide materials is very different from those for conventional superconductors, that are completely understood within the BCS theory. Since the beginning point of building up a theory, which in this case is the normal state, is missing, or to the least, is unlike the conventional metallic state, a complete understanding of these superconductors remained elusive.

4.1.1 Historical developments

Empirical rules have been suggested from time to time to enable the identification of possible new superconducting materials. The rules proposed by Hulm and Matthias (1951) in the sixties are especially noteworthy. One of them is based on the number of valence electrons. It is found that the

[1] The normal state refers to the metallic state which is expected as superconductivity is lost.

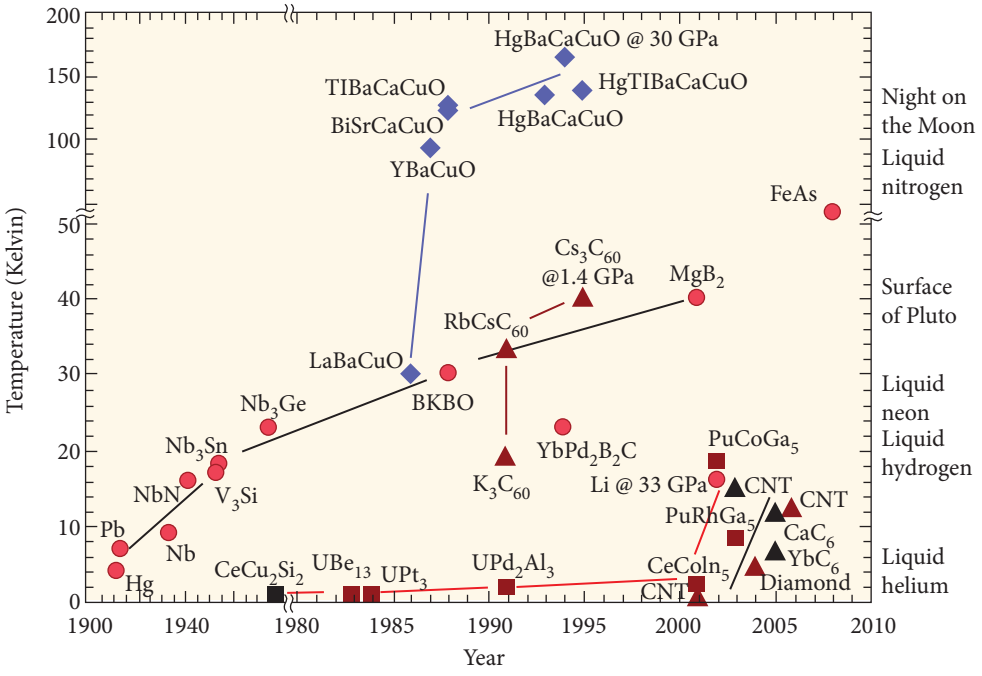


FIG. 4.2
The rise in the superconducting transition temperature, T_c with time since the discovery is indicated in the schematic plot.

materials with on an average 5 to 7 electrons per atom show relatively high T_cs. For example, RuMo with an average of 7 electrons show $T_c = 10.6$ K, whereas, ruthenium (Ru) with an average of 8 electrons show $T_c = 0.5$ K, and molybdenum (Mo) with an average of 6 electrons show $T_c = 1$ K. Later on, a structural dependence is seen on T_c. A-15 structure or the β-tungsten shows high T_c. V_3Si and Nb_3Ge are examples of this structure and they show $T_c = 17.1$ K and 23.2 K, respectively. For quite a few years, this 23.2 K barrier has never been crossed. Among the organic materials, the intercalated compounds of graphite (e.g. C_8 K) become superconducting below 1 K, while K_3C_{60} has $T_c \approx 18$ K. Generally, alkali salts of C_{60} have been found to be superconducting with T_c ranging from 18−35 K (see Fig. 4.2). In Fig. 4.2, a steady rise of T_c over a century has been depicted.

If we refer to the periodic table and consider the top row of the middle section of the periodic table, that is, group IVB (Cr, Mn) to group VIII (Fe, Co, Ni), they have strong magnetic character. While along the same row, Ge, As, Se, or some of the neighboring elements show superconductivity (see Fig. 4.3). Since both the phenomena have electronic origin, and their physical properties are quite distinct, it forces

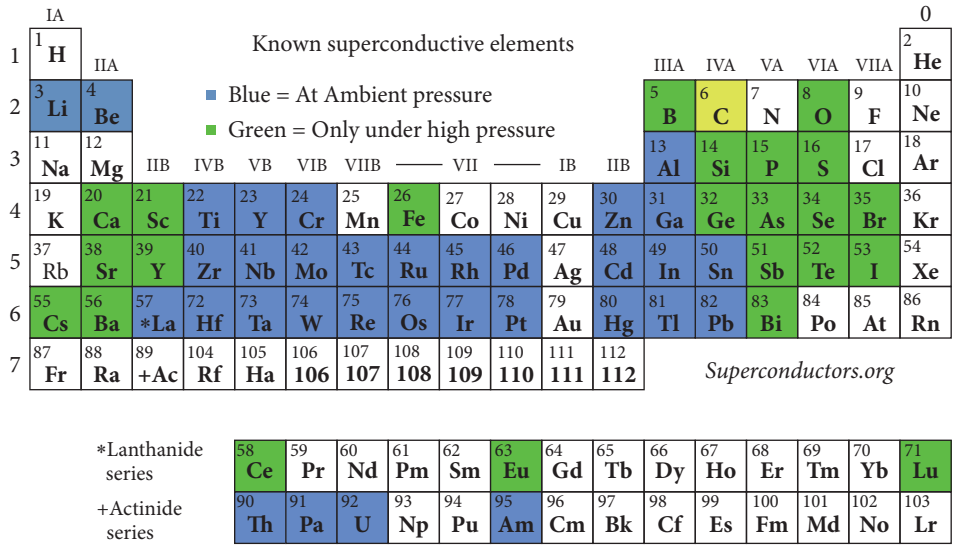


FIG. 4.3
The superconducting elements are shown in the periodic table. The ones shown in blue demonstrate superconductivity at ambient pressure, while those in green become superconducting at high pressure.

us to think about what could be the genesis of this difference. This probably lies in the extent of the wavefunction of the electrons: for magnets, the wavefunction has a small spatial extent, that is spread over a smaller number of atoms, while for superconductors, the spread is large, thereby promoting pairing phenomena between the carriers.

4.1.2 Physical properties

Before we proceed to discussing the physical properties of superconductors, we wish to make the plans clear for our upcoming discussion. We shall mainly mention the salient features of superconductors that make them worthy of studies. However, most of the ideas and concepts are introduced with a minimal amount of derivation. For a thorough understanding of some of the properties, one has to wait for the BCS theory, a topic that we discuss somewhat elaborately afterwards.

Let us rewind the ongoing discussion for more detail. The resistance of a metal drops to zero very sharply, within a temperature window of $\Delta T \sim 10^{-5}$ K below a certain critical temperature, T_c.The superconducting state is characterized by zero electrical resistance. In addition, there is no change in the crystal structure, as verified by x-ray diffraction above and below the transition temperature. Thus, to list out the properties, the state is characterized by

i. infinite electrical conductivity, that is, $\sigma \rightarrow \infty$.
ii. Finite current density, $j =$ finite.
iii. The specific heat as a function of temperature shows a jump at $T = T_c$, and thus the transition to the superconducting state involves latent heat, and hence, the transition from the superconducting state to the normal state is a second-order phase transition.
iv. Vanishing values of the electric field, $E \rightarrow 0$ inside the superconducting sample.
v. Constant magnetic field, $B \rightarrow$ constant inside a superconductor.

Clearly, the last two statements defy the laws of classical electrodynamics. This can be understood in the following way. Using In Ohm's law, one can write, $\mathbf{j} = \sigma \mathbf{E}$ for j having a finite value and $\sigma \rightarrow \infty$, E must be zero. Hence, using Maxwell's equation,

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad \text{or,} \quad \mathbf{B} = \text{constant}.$$

4.1.3 Meissner effect

The complete and sudden vanishing of the electrical resistance below a certain temperature raised a lot of questions. Some physical phenomena that were operative to render electrical resistance above T_c, were suddenly becoming ineffective. Meissner and Ocshenfeld (1933) meanwhile discovered that a superconductor completely excludes an external magnetic field. They measured the magnetic field distributed outside the superconducting materials [such as lead (Pb) or tin (Sn)] which are cooled below their respective transition temperatures while the magnetic field is switched on. The corresponding results could not be explained by superconductors being just resistance-less metals.

Above a certain critical magnetic field (of the order of a few Oersted), there is no expulsion of the magnetic flux, where superconductivity disappears and the material reverts to its normal resistive state and the magnetic field fully penetrates through it. This also endorses a close link between magnetism and superconductivity.

The exclusion of the magnetic field from a superconductor takes place regardless of whether the material becomes superconducting before or after the external field is applied. At equilibrium, the external field is canceled in its interior by screening fields produced by the skin current.

The total exclusion of the magnetic field from inside superconductors assigns a property known as *perfect diamagnetism*. It can be understood as follows:

$B = \mu_0(H_{ext} + M)$.

Since $B = 0$ inside a superconductor, the induced magnetization M cancels the external field, H_{ext}. Thus, the magnetic susceptibility, χ becomes,

$$\chi = \frac{M}{H_{ext}} = -1.$$

Thus, no known material is more diamagnetic than a superconductor. The typical diamagnetic susceptibility of metals is about $10^{-5} - 10^{-6}$. Thus a usual diamagnetic material only expels a small fraction

(about 10^{-5}) of the external field, while a superconductor does it completely. However, even if we claim that the magnetic field is totally expelled from a superconductor, the fact is that it enters only up to a certain distance, called the penetration depth.

4.1.4 Perfect conductors and superconductors

In order to differentiate a perfect conductor (which may show a gradual vanishing of the resistance as $T \to 0$), and a superconductor, we resort to field cooled (FC) and zero field cooled (ZFC) techniques that are schematically shown in Figs. 4.4 and 4.5.

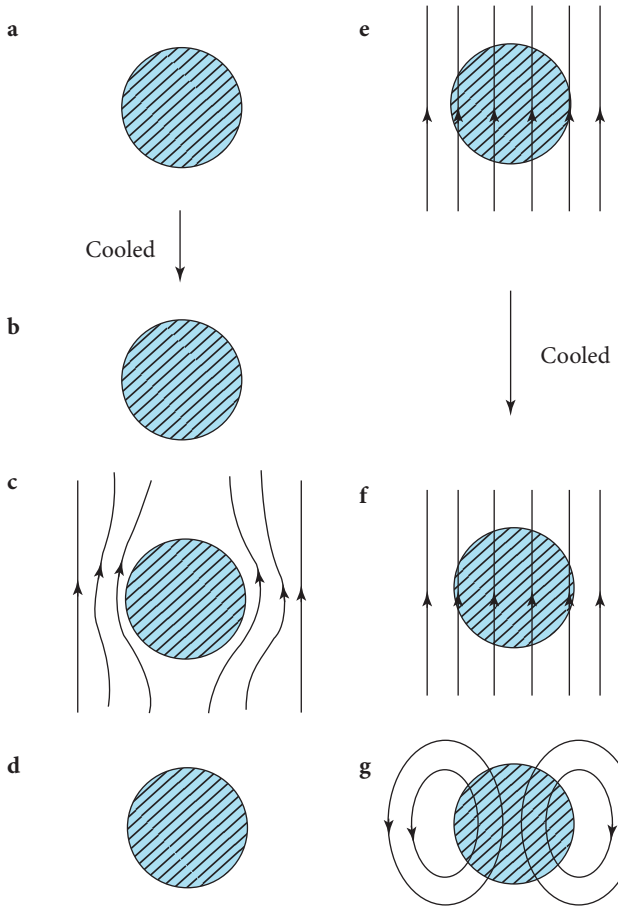


FIG. 4.4
The zero field cooled (left panel) and field cooled (right panel) cases for a perfect conductor are shown schematically.

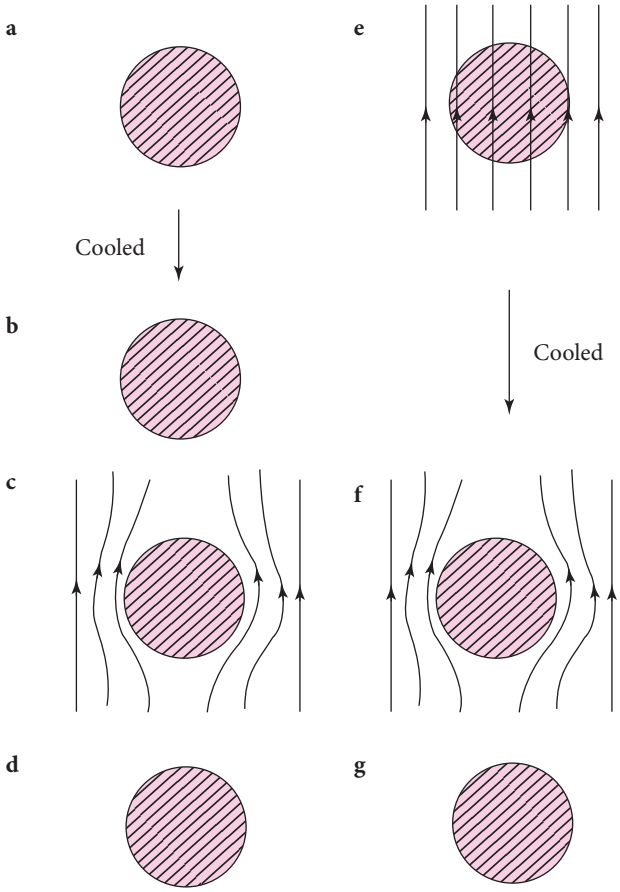


FIG. 4.5
The zero field cooled (left panel) and field cooled (right panel) cases for a superconductor conductor are shown schematically.

If we consider cooling perfect conductors and superconductors in the absence of a magnetic field (left panels of Figs. 4.4 and 4.5), they behave identically when an external magnetic field is applied. That is, they expel the field by developing a surface current, which prohibits the field lines from penetrating beyond a certain distance (discussed below). However, in the presence of an external field, the behavior of a conductor and that of a superconductor are quite distinct. When a conductor is cooled to a perfect conducting state in the presence of an external field, the flux lines get trapped inside, while that for a superconductor, the flux lines are expelled. Upon the withdrawal of the field [panel (g) in Figs. 4.4 and 4.5], a superconductor is left with no memory, where for a perfect conductor, the flux lines form a loop, and thus retains the field in its memory.

4.1.5 Electrodynamics of superconductors: London theory

Before the discovery of the Meissner effect in superconductors, Becker *et al.* (1933) analyzed the electrodynamic behavior of perfect conductors using a simple free electron model. According to this, the electrons accelerate under the application of an electric field, **E**. The argument for resistance-less motion is that whenever the electrons faced a resistance, the applied electric field would accelerate them steadily according to Newton's laws,

$$m^* \frac{d\mathbf{v}}{dt} = e^* \mathbf{E} \tag{4.1}$$

where m^*, e^* are the effective mass and the net charge, respectively. It was found later that $e^* = 2e$ and $m^* = 2m$ which are subsequently called super-electrons or Cooper pairs. If there are n_s numbers of super-electrons moving with a mean local velocity v_s, then the super-current density is given by

$$\mathbf{J}_s = n_s e^* \mathbf{v}_s = -n_s |e^*| \mathbf{v}_s \tag{4.2}$$

The minus sign signifies that the super-current moves in an opposite direction to that of the super-electrons. Substituting Eq. (4.2) in Eq. (4.1) we get

$$\mu_0 \lambda_L^2 \frac{\partial \mathbf{J}_s}{\partial t} = \mathbf{E}, \tag{4.3}$$

where, $\lambda_L = \sqrt{\frac{m^*}{\mu_0 n_s (e^*)^2}}$. λ_L has the dimensions of length. Equation (4.3) is called the first London equation. Taking the curl of both sides of Eq. (4.3),

$$\mu_0 \lambda_L^2 \frac{\partial}{\partial t} (\nabla \times \mathbf{J}_s) = \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}. \tag{4.4}$$

Using the following Maxwell equation:

$$\nabla \times \mathbf{B} = \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} + \mu_0 \mathbf{J}$$

we have

$$\begin{aligned} &\lambda_L^2 \frac{\partial}{\partial t} (\nabla \times (\nabla \times \mathbf{B})) = -\frac{\partial \mathbf{B}}{\partial t} \\ &\lambda_L^2 \nabla^2 \left(\frac{\partial \mathbf{B}}{\partial t} \right) = \frac{\partial \mathbf{B}}{\partial t}. \end{aligned} \tag{4.5}$$

This implies that the rate of the **B** field (or the **B** field itself) will fall off exponentially inside a superconductor to a trapped field B_0. This clearly contradicts the Meissner effect, that is, the magnetic field inside the specimen is zero, irrespective of the initial condition.

Thus, the brothers London and London (1935) suggested that since the macroscopic theory of a perfect conductor makes correct predictions about superconductor for the special case $B_0 = 0$. Thus, it might be reasonable to assume that the magnetic behavior of a superconductor may be correctly described according to the Meissner effect not only to $\frac{d\mathbf{B}}{dt}$ but also to **B** itself, that is

$$\begin{aligned} &\lambda_L^2 \nabla^2 \mathbf{B} = \mathbf{B} \quad [\nabla \times \mathbf{B} = \mu_0 \mathbf{J}_s] \\ &\text{which implies,} \quad \mathbf{B} = -\mu_0 \lambda_L^2 \nabla \times \mathbf{J}_s. \end{aligned} \tag{4.6}$$

By introducing a magnetic vector potential defined by $\nabla \times \mathbf{A} = \mathbf{B}$ and choosing a proper gauge, such as a transverse gauge, namely, $\nabla \cdot \mathbf{A} = 0$, the current density can be written as

$$\mathbf{J}_s = -\frac{1}{\mu_0 \lambda_L^2}\mathbf{A} \quad [\mathbf{B} = \nabla \times \mathbf{A}]. \tag{4.7}$$

This is called the second London equation, and is reminiscent of Ohm's law. This is of course true for simply connected superconductors. The gauge condition demands

$$\nabla \cdot \mathbf{A} = 0 \implies \mathbf{A} \cdot \hat{n} = 0$$

on the boundary surface of the superconductor, $\hat{n}$ being the unit drawn normally.

Now, using Maxwell's equation,

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$$

$$\nabla \times \left(\mathbf{E} - \mu_0 \lambda_L^2 \frac{\partial \mathbf{J}_s}{\partial t}\right) = 0 \tag{4.8}$$

$$\text{or,} \quad \mathbf{E} - \mu_0 \lambda_L^2 \frac{\partial \mathbf{J}_s}{\partial t} = \nabla \phi$$

ϕ is a scalar. One important comment is due here. There is no component of $\nabla\phi$ in the direction of $\mathbf{J}_s$, that is,

$$\nabla \phi \cdot \mathbf{J}_s = 0 \tag{4.9}$$

as $\mathbf{J}_s \cdot \mathbf{E} = 0$ from first London Eq. (4.3). This is correct as otherwise it would mean energy dissipation in a constant magnetic field, which we know to be incorrect as it violates the conservation of energy.

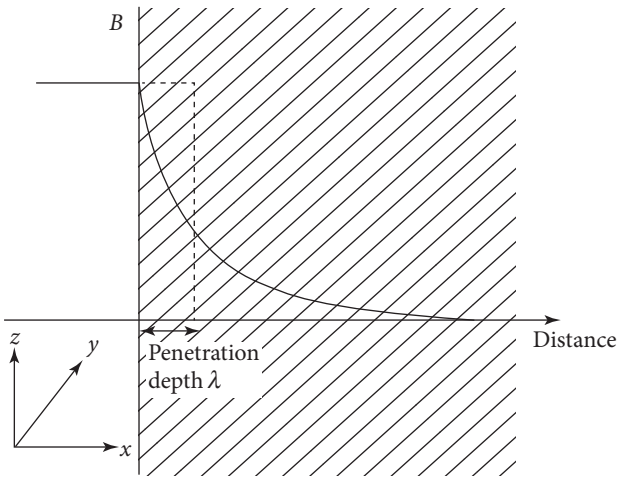


FIG. 4.6
Schematic demonstration of the behavior of an external magnetic field is presented. The field strength falls by a factor 1/*e* inside the superconductor, which is defined as the penetration depth.

4.1.6 Penetration depth

Let us prepare a setup to solve Eq. (4.5) for a superconductor. Refer to Fig. 4.6 where a semi-infinite superconducting sample is placed along the y-axis. We are interested in studying the variation of the magnetic field only in the x-direction along which it has a finite width. The axis is shown along with the figure.

In Fig. 4.6 the variation is expected to be along x-direction, we can convert it into a one-dimensional equation, that is by assuming,

$$\frac{\partial B}{\partial y} = \frac{\partial B}{\partial z} = 0$$

which yields,

$$\frac{dB_x}{dt} = \alpha \frac{\partial^2}{\partial x^2}\left(\frac{dB_x}{dt}\right).$$

Integrating over time yields,

$$B_x = \alpha \frac{\partial^2}{\partial x^2} B_x$$
$$B(x) = B_{ext} e^{-x/\lambda_L}$$

where an exponentially growing solution, even though mathematically admissible, is dropped, and

$$\alpha = \lambda_L^2$$

λ_L is of the order of 500 Å for commonly known superconductors. This is how the magnetic field varies inside a superconducting sample. At $x = 0$, the value of the magnetic field just outside the surface is B_{ext}, which decays exponentially inside the sample.

4.1.7 Flux quantization

In the mid-1930s (London and London, 1935) it was realized that superconductivity is inherently a quantum phenomenon that manifests at macroscopic scales. The description thus requires a wavefunction which can be written as

$$\psi(\mathbf{r}, t) = |\psi| e^{i\theta} \tag{4.10}$$

where θ is a real scalar function representing the phase of the wavefunction and $|\psi|^2 = n_s$, where n_s is the density of super-electrons. The canonical momentum for a particle of charge e^* and mass m^* in a magnetic field is given by

$$\mathbf{p} = m^* \mathbf{v}_s + e^* \mathbf{A}. \tag{4.11}$$

Putting $\mathbf{p} = i\hbar\nabla$ and writing down the Schrödinger equation,

$$\hbar \nabla\theta \psi = (m^* \mathbf{v}_s + e^* \mathbf{A})\psi \tag{4.12}$$

yields,

$$\frac{\hbar}{e^*} \nabla\theta = \mu_0 \lambda_L^2 \mathbf{j}_s + \mathbf{A}, \tag{4.13}$$

where $\lambda_L = \sqrt{\frac{m^*}{\mu_0 n_s e^{*2}}}$ is the London penetration depth (described later), and $\mathbf{j}_s$ is the super-current density defined via, $\mathbf{j}_s = n_s e^* \mathbf{v}_s$. Since $\mathbf{A}$ enjoys gauge freedom, the (local) phase of the wavefunction can be changed in the following manner, so as to keep the velocity unchanged.

$$\theta \rightarrow \theta' = \theta + \frac{e^*}{\hbar}\mathbf{A}.$$

It may be noted that Eq. (4.13) is consistent with Maxwell's equation, which can be checked by taking a curl of the equation, and noting that $\mathbf{B} = \nabla \times \mathbf{A}$. Now taking a derivative of Eq. (4.13) with respect

to time, one gets,

$$\frac{\hbar}{e^*}\nabla\left(\frac{\partial\theta}{\partial t}\right) = \mu_0\lambda_L^2\frac{\partial \mathbf{j}_s}{\partial t} + \frac{\partial \mathbf{A}}{\partial t}. \tag{4.14}$$

Using the first of the London equations, namely,

$$\frac{\partial \mathbf{j}_s}{\partial t} = \frac{\mathbf{E}}{\mu_0\lambda_L^2}$$

one arrives at

$$\mu_0\lambda_L^2\frac{\partial \mathbf{j}_s}{\partial t} = \mathbf{E}. \tag{4.15}$$

Hence, with

$$\mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t}$$

one gets,

$$\frac{\hbar}{e^*}\frac{\partial\theta}{\partial t} = -\phi. \tag{4.16}$$

Thus, the time variation of the local phase denotes a scalar potential ϕ.

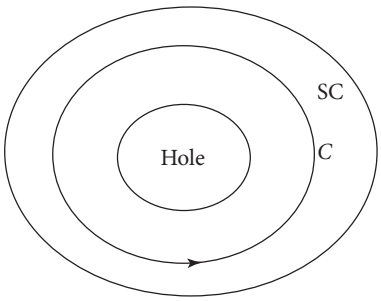


FIG. 4.7
A simply connected superconductor is shown with a "*hole*" within. *C* denotes a contour of integration in Eq. (4.13).

There is a feature of superconductors that deserves special attention, which is quantization of the magnetic flux. Consider a multiply connected superconductor, and consider a closed loop C lying entirely in the superconducting regime, and encircling the hole (marked in Fig. 4.7). Thus, any open surface bounded by C contains regions that are partially superconducting, and partially in a non-superconducting (normal) region.

Applying Stoke's theorem on Eq. (4.13) one obtains,

$$\frac{\hbar}{e^*}\oint_C \nabla\theta\cdot d\mathbf{l} = \Phi_s + \oint_C \mu_0\lambda_L^2\mathbf{j}_s\cdot d\mathbf{l} \tag{4.17}$$

Φ_s is the magnetic flux defined by

$$\Phi_s = \oint \mathbf{B}\cdot d\mathbf{s}. \tag{4.18}$$

The requirement that ψ is a single valued function at each point, that is,

$$\phi e^{i\theta(\mathbf{r},t)} = \phi e^{i\{\theta(\mathbf{r},t)+2n\pi\}}$$

indicates that the LHS of Eq. (4.17) has the form,

$$\frac{\hbar}{e^*}\int_C \nabla\theta\cdot d\mathbf{l} = n\Phi_0 \qquad n = 0, \pm1, \pm2, \ldots \tag{4.19}$$

where Φ_0 is called the flux quantum

$$\Phi_0 = \frac{2\pi\hbar}{|e^*|} = 2.07 \times 10^{-15} \text{ Wb} \tag{4.20}$$

$n = 0$ corresponds to a simply connected superconductor.

Hence integrating over a closed loop C should be the same after one complete rotation from where one has started, except for a phase of 2π, or a multiple of 2π. A flux quantum has this value for $e^* = 2e$. So the super-current is carried by a pair of electrons. In fact,

$$\oint \nabla\theta \cdot d\mathbf{l} = 2\pi n = -\frac{2e}{\hbar}\Phi.$$

Thus, the magnetic flux penetrating through a superconductor can only appear in multiples of the superconducting flux quantum Φ_0, which is often referred to as fluxoid (or fluxon). It may be noted that the value of the fluxoid is extremely small. Experimentally, the flux quantization has been measured by Deaver and Fairbank in 1961 (Deaver and Fairbank, 1961). Importantly, the involvement of Planck's constant ($\hbar$) is a signature of quantum effects existing at the macroscopic scales.

On the right-hand side of Eq. (4.17), we can introduce a fluxoid Φ_L,

$$\Phi_L = \int_s \mathbf{B} \cdot d\mathbf{s} + \oint_c \mu_0 \lambda_L^2 \mathbf{j}_s \cdot d\mathbf{l} \tag{4.21}$$

Thus, within approximation of the London theory,

$$\frac{\partial \Phi_L}{\partial t} = \int_s \left(\frac{\partial \mathbf{B}}{\partial t} + \nabla \times \mathbf{E} \right) \cdot d\mathbf{s} = 0, \tag{4.22}$$

where the last step is obtained using Maxwell's equation. Thus, the fluxoid Φ_L is constant with time. It also does not depend upon the exact shape of the contour C, as long as it encompasses the hole, but only once. If there is no hole, then the fluxoid is zero.

4.1.8 Non-local electrodynamics

The current density equation in electrodynamics, namely,

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

is a local equation in the sense that it relates the current density to a point $\mathbf{r}$ to the velocity of the charge carriers at that point. Thus, it clearly ignores the spatial structure of the electric field. An improvement to this can be made by considering the current density at $\mathbf{r}$ due to the variation of the electric field within a shell around the point $\mathbf{r}$, with a finite radius. On similar grounds, Pippard argued that the current density at a point $\mathbf{r}$ should depend on $\mathbf{E}(\mathbf{r}')$. Further, he conjectured that the wavefunction of the electrons must have a characteristic dimension, namely, ξ which is called the coherence length. A simple uncertainty calculation may be sufficient to yield an estimate of the value of ξ in typical superconductors. We know that electrons within an energy range $k_B T_c$ play a dominant role in the

pairing phenomena. The momenta of these electrons have uncertainty given by

$$\Delta p \simeq \frac{\Delta E}{v_F} \simeq \frac{k_B T_c}{v_F}.$$

Thus, their position uncertainty, using Heisenberg's uncertainty principle is

$$\Delta x \simeq \frac{\hbar}{\Delta p} \simeq \frac{\hbar v_F}{k_B T_c}.$$

This yields

$$\Delta x \simeq \xi = a\frac{\hbar v_F}{k_B T_c}$$

a is usually of the order of unity. In particular, $a = 0.8$ in BCS theory (Tinkham, 1973).

Hence, Pippard suggested that the simple equation for the current density written above should be modified as

$$\mathbf{J}_s(\mathbf{r}) = -(const)\int \frac{\mathbf{R}[\mathbf{R}\cdot\mathbf{A}(\mathbf{r}')]}{R^4} e^{-R/\xi} d\mathbf{r}', \tag{4.23}$$

where, $\mathbf{R} = \mathbf{r} - \mathbf{r}'$, with $\mathbf{r}$ being the source point and $\mathbf{r}'$ being the field point, or where the observation is made.

4.2 MAGNETIC PHASE DIAGRAM OF SUPERCONDUCTORS

When a superconductor is used to form a circuit with a battery, and a steady state is established, all the currents passing through the superconductor are super-currents. Normal currents due to the motion of charged particles contribute zero since no voltage difference can be sustained in a homogeneous superconductor (because otherwise a current will induce a B which is zero). Experiments show that all super-currents flow near the type-I superconductor's surface within a thin layer characterized by the penetration depth. These surface super-currents run so that the B field is canceled in the interior of the conductor.

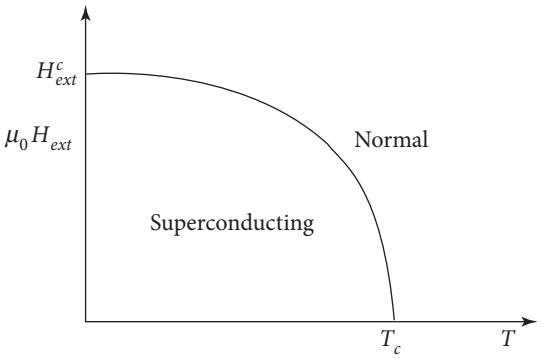


FIG. 4.8
The dependence of the critical field, H_C as a function of temperature T is shown.

Thus, let us review the magnetic phase diagram of superconductors. It is clear by now that the superconducting phase is a stable phase in a certain range of magnetic field and temperatures. For higher fields and temperatures, the normal metallic state becomes more stable. The critical magnetic field (H_c) beyond which the material cannot be superconducting even if it is cooled below T_c (see Fig. 4.8). The value of H_c depends

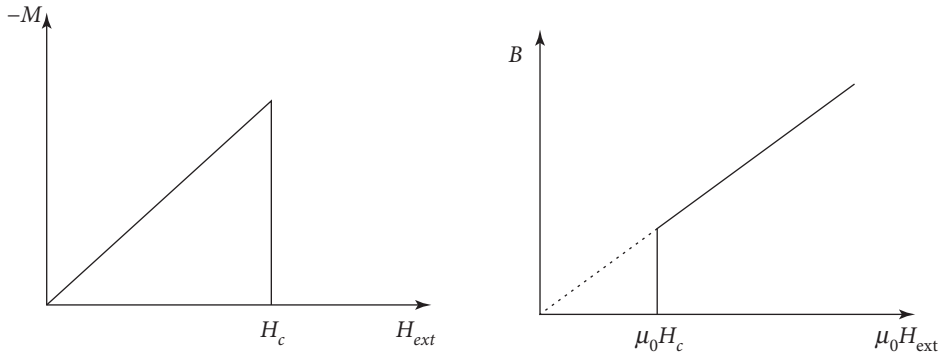


FIG. 4.9
Magnetic phase diagrams for type-I superconductors are shown. There is a direct transition from the superconductor to the normal phase in type-I superconductors.

on temperature, since it is easier to quench a superconductor at higher temperatures, than at lower temperatures. An empirical relation for the same can be written as

$$H_c = H_0\left[1-\left(\frac{T}{T_c}\right)^2\right] \quad \text{where,} \quad H_0 = H_{ext}(T=0). \tag{4.24}$$

The corresponding M vs H phase diagrams are shown in Fig. 4.9. There is a sharp fall of the (negative) magnetization $(-M)$ as a function of the external field, and a sharp rise of the magnetic induction denotes a phase transition from a superconducting to a normal phase.

In type-II superconductors, the phase diagram is more complex. When small magnetic fields are applied, surface currents develop that screen the magnetic field from penetrating homogeneously into the sample. This is termed as the Meissner state, which persists until a certain critical value, called the first critical magnetic field, H_{c1}. Beyond this value, the system enters a mixed state, where the magnetic flux lines, in the form of vortices, penetrate the superconducting sample. Thus, there is a partial screening of the magnetic field.

Beyond H_{c1}, more and more vortices appear insider the superconductor, and eventually when the external field reaches a second critical value, namely, H_{c2} the material becomes a normal metal (Fig. 4.10). The mixed phase is called the Abrikosov phase, where it is energetically favorable for the vortices to self organize into a hexagonal array. The scanning tunneling microscopy (STM) measurements[2] confirm

[2] The working principle of STM is based on the measurement of electronic tunneling current that depends exponentially on the distance between the tip and the sample. The position of a tip on top of a flat surface of the sample can be controlled with subatomic resolution.

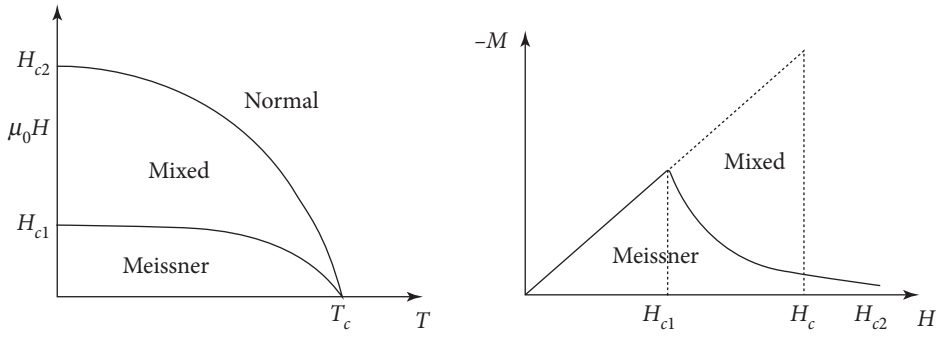


FIG. 4.10
Magnetic phase diagrams for type-II superconductors are shown.In type-II materials, there is an intermediate phase that interrupts a direct transition, where the flux lines penetrate.

the existence of such a vortex array, which minimizes energy by arranging into a hexagonal structure in which the vortices line up. These structures are called Abrikosov lattices.

The magnetic susceptibility of a superconductor is -1 below H_c and has a value 10^{-5} above H_c. This is shown in the following figures.

From here we get a characteristic length called coherence length ξ. ξ is defined as the length scale over which the wavefunction or the order parameter does not vary much. The ratio of these two characteristic length scales is given by

$$\kappa = \frac{\lambda}{\xi}. \tag{4.25}$$

For typical superconductors, $\lambda = 500$ Å, $\xi \approx 5000$ Å. Thus, $\kappa < 1$.

In 1957, Abrikosov found that for some classes of superconductors $\kappa > 1$. This is used as a distinguishing criterion for the type I and type II superconductors, in the following sense.

$\kappa < 1$ type I superconductors
$\kappa > 1$ type II superconductors.

For type II superconductors, the flux lines penetrate the sample until a certain threshold magnetic field. A threshold value of κ for which such flux penetration starts to occur is $\kappa = \kappa_c \approx \frac{1}{\sqrt{2}}$. At this value of κ, the flux penetration starting at a lower critical field H_{c1} and reaching $B_{ext} = H_{ext}$ at $H_{ext} = H_{c2}$.

Even experiments on the high-T_c superconductors confirmed many of the features in the magnetic phase diagram as discussed above. For example, the quantization of vortices ($\Phi_0 = \frac{h}{2e}$), Abrikosov phase etc. are observed in experiments. In the type-II superconductors, the identification of the upper critical field has been particularly controversial, quite unlike conventional superconductors, where the

onset of strong diamagnetism and superconducting critical currents occur at the same H_{c2}. Whereas, in the high-T_c superconductors, they are found to occur at quite different values. There are several other features which are distinct, possibly because of a defect induced pinning mechanism, etc. We shall not continue this discussion here any further and will briefly come back at the end of the chapter. However, certain contrasting properties of these superconductors compared to the conventional ones, will be mentioned during the course of the discussion.

4.2.1 Thermodynamics of superconductors

Thermodynamics route is usually the easiest route to understand the phase transition. Consider Gibbs free energy of a superconductor. If its magnetization is M and magnetic induction H, then the work done in bringing the superconductor into a region with magnetic induction H_{ext} from far away (where $H_{ext} = 0$) is given by

$$W = -\mu_0 \int_0^{H_{ext}} MdH = \mu_0 \frac{H_{ext}^2}{2} \tag{4.26}$$

as $M = -H$ for superconductor.

This is the extra Gibbs free energy of a superconductor. Let us call this g_s (per unit volume)

$$g_s(T, H_{ext}) = g_s(T, 0) + \frac{\mu_0 H_{ext}^2}{2} \tag{4.27}$$

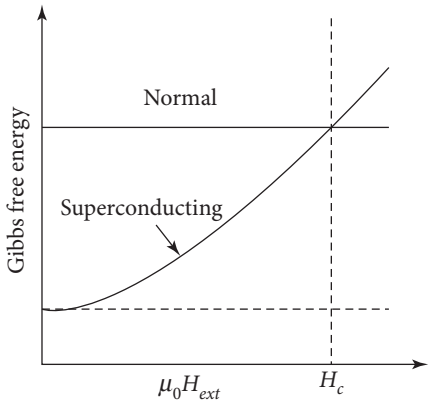


FIG. 4.11
The Gibbs free energy is indicated schematically as a function of the external magnetic field. The energies for the superconducting and the normal states intersect at some value of the magnetic field, beyond which the normal state stabilizes.

or,

$$g_s(T, B_{ext}) = g_s(T, 0) + \frac{B_{ext}^2}{2\mu_0} \tag{4.28}$$

In the normal state, the magnetization is very small value, thus the magnetic work term is vanishingly small. Hence

$$g_n(T, H_{ext}) = g_n(T, 0) \tag{4.29}$$

$g_s(T, H_{ext})$ increases quadratically with H_{ext}. For some H_{ext}, the normal and superconducting states have the same free energy. Beyond this critical value, the normal state is more stable.

Equating Eqs. (4.27) and (4.29) at $H_{ext} = H_c$, (see Fig. 4.11)

$$g_n(T, 0) = g_s(T, 0) + \frac{\mu_0 H_c^2}{2}. \tag{4.30}$$

The RHS is positive, and hence a superconducting state is more stable below H_c. For example, $H_c = 0.08$T for Pb at $T = 0$ (Ramakrishnan and Rao, 1999). Thus, at $T = 0$, the superconducting state is stabilized by 4.25×10^{-25} Joule/mole. This is a really small amount, however, interestingly, this small number stabilizes the superconducting state.

4.2.2 Specific heat

The study of specific heat is in general interesting because it provides a good measure of the range of applicability of phonon-mediated superconductivity. Above the transition temperature, the specific heat, C_n follows Debye's theory. In general, for a normal metal, at temperatures $T < \theta_D$, the specific heat consists of

i. Electronic contribution, γT.
ii. Phonon contribution, AT^3.
iii. Schottky contribution, a/T^2.

Schottky contribution can mostly be neglected and hence,

$$C_n = \gamma T + AT^3. \tag{4.31}$$

Thus, $C_n/T = \gamma + AT^2$. C_n/T versus T^2 thus denotes a linear plot.

At $T = T_c$, the contribution of electronic and phonon specific heat is,

$$\frac{C_{ph}}{C_e} = \frac{1}{\beta}\left(\frac{AT_c^2}{\gamma}\right)$$

where β is the ratio of the number of conduction electrons to that of atoms. From free electron theory,

$$A = 234\frac{R}{\theta_D^3}.$$

Kittel (2004) Remember,

$$C_{ph} = \frac{12\pi^4}{5}R\left(\frac{T}{\theta_D}\right)^3 = 234R\left(\frac{T}{\theta_D}\right)^3$$

Whereas,

$$\gamma = 4.93\frac{R}{T_F}$$

and,

$$C_e = \frac{1}{2}\pi^2 R\left(\frac{T}{T_F}\right) = 4.93R\left(\frac{T}{T_F}\right).$$

Thus,

$$\frac{C_{ph}}{C_e} = \left(\frac{47.5}{\beta}\cdot\frac{T_F}{\theta_D^3}\right)T_c^2$$

β can be set to be of the order of 1. For most of the conventional superconductors, $T_c \ll \theta_D$ So the $C_e \gg C_{ph}$.

If the conduction electrons have an effective mass, m^*, which is defined by the inverse of the curvature of the band structure, and that differs from bare mass m, then the conduction electron-specific heat is

given by

$$\gamma = \left(\frac{m^*}{m}\right)\gamma_0$$

where γ_0 is the bare electron counterpart of γ.

$$\gamma_0 = \frac{\pi^2 R}{2T_F}.$$

Thus

$$\frac{m^*}{m} = \frac{\gamma}{\gamma_0} = \frac{2\gamma T_F}{\pi^2 R}.$$

It is worth mentioning here that for a typical high temperature superconductor, T_F is pretty low, $T_F \sim 4000$ K, at least an order lower than conventional elemental superconductors, such as for Cu, $T_F \sim 80\,000K$. This discrepancy is probably due to the large values of the effective masses.

With C_{es} numerically obtained, we can integrate to find the change in internal energy $U(T)$ as we decrease temperature from T_c. At $T = T_c$ it must be the same as the normal value $U_{en}(0) + \frac{1}{2}\gamma T_c^2$. So

$$U_{es}(T) = U_{en}(0) + \frac{1}{2}\gamma T_c^2 - \int_T^{T_c} C_{es} dT.$$

From the internal energy and entropy, we may compute the free energy using,

$$F_{es}(T) = U_{es}(T) - TS_{es}(T).$$

The critical field is determined by

$$\mu_0 \frac{H_c^2(T)}{2} = F_{en}(T) - F_{es}(T)$$

where,

$$F_{en}(T) = U_{en}(0) - \frac{1}{2}\gamma T^2.$$

The internal energy, U, and the free energy, F, are shown across a superconductor-normal metal transition in Fig. 4.13. To compute the thermodynamic quantities, we explicitly consider the temperature dependence of the specific heat of a superconductor and explore its behavior at the normal-superconductor phase transition. There is a jump in the specific heat at $T = T_c$ as we have seen earlier.

The Gibb's free energy per unit volume for a superconductor in a magnetic field is written as

$$G = U - TS - HM \tag{4.32}$$

neglecting usual PV term. Since,

$$dU = TdS + HdM. \tag{4.33}$$

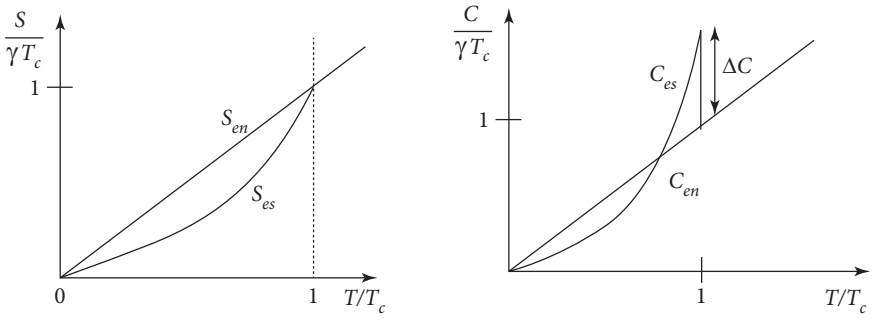


FIG. 4.12
Schematic plots of entropy (S) and specific heat (C) scaled by γT_C (γ defined in text) are presented as a function of reduced temperature, $\frac{T}{T_c}$. $T/T_c < 1$ denotes the superconducting state, while $T/T_c > 1$ depicts the normal state.

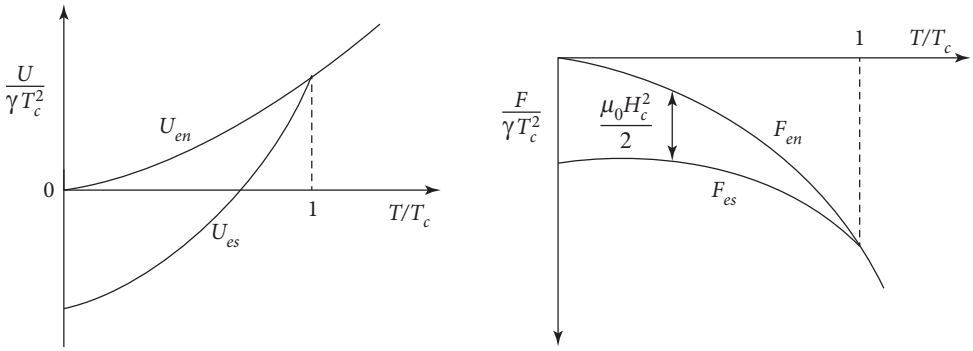


FIG. 4.13
Schematic plots of internal energy (U) and Helmholtz-free energy (F) scaled by γT_C (see definition of γ in text) are presented as a function of reduced temperature, $\frac{T}{T_c}$. $T/T_c < 1$ denotes the superconducting state, while $T/T_c > 1$ depicts the normal state.

This is the change in internal energy in the presence of a magnetic field, from (4.32) and (4.33),

$$dG = -SdT - MdH. \tag{4.34}$$

Thus, integrating (4.34),

$$G_s(H) = G_s(0) + \frac{H^2}{2}. \tag{4.35}$$

Along the critical curve where the SC and the normal metal are in equilibrium,

$$G_n = G_s(0) + \frac{H^2}{2}. \tag{4.36}$$

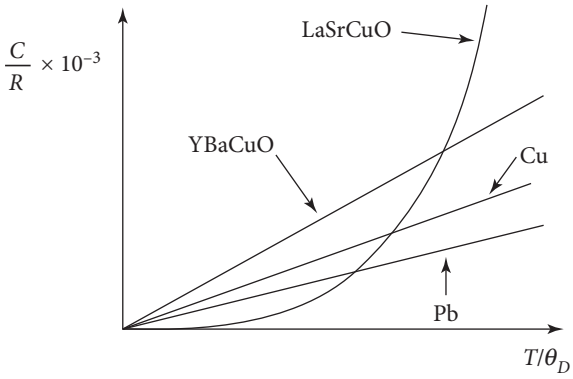


FIG. 4.14
The behaviour of the specific heat is schematically shown as a function of the reduced temperature T/θ_D for both conventional (Cu and Pb) and unconventional (cuprate) superconductors.

where G_s and G_n are the Gibb's free energy for the superconducting and the normal states. From Eq. (4.34)

$$\left(\frac{\partial G}{\partial T}\right)_H = -S. \tag{4.37}$$

Thus, at equilibrium,

$$S_n - S_s = -H_c \frac{dH_c}{dT}. \tag{4.38}$$

S_s denotes the entropy of a SC in a zero field. Since $\frac{dH_c}{dT}$ is always negative, the entropy of the normal state is always greater than that of the superconducting state. Makes sense. Thus,

$$\begin{aligned}\Delta C = C_s - C_n &= -T\frac{d}{dT}(S_s - S_n) \\ &= TH_c\frac{d^2H_c}{dT^2} + T\left(\frac{dH_c}{dT}\right)^2.\end{aligned} \tag{4.39}$$

At $T \to T_c$, $H_c \to 0$,

$$\Delta C = T_c\left(\frac{dH_c}{dT}\right)^2. \tag{4.40}$$

From (4.38) at $T = T_c$, $H_c = 0$, there is no latent heat of transition $\Delta S = 0$ (see left panel of Fig. 4.12), but according to (4.40) $\Delta C =$ finite (see right panel of Fig. 4.12). So the transition is second order (at $T = T_c$). However, away from T_c, the phase transition has a latent heat and is a first-order phase transition.

However, this is not true for high temperature superconductors. In fact for $(La_{0.9}Sr_{0.1}CuO_{4-\delta})$ and $(YBa_2Cu_3O_{7-\delta})$, $AT_c^3 \gg \gamma$, so the phonon term dominates at $T = T_c$ (see Fig. 4.14).

4.2.3 Density of states

The Bogoliubov quasiparticle spectrum, E_k is easily seen to have a minimum of Δ_k for $\xi_k = 0$ (ξ_k is positive definite including zero). Thus, in addition to playing the role of the order parameter Δ_k is also the energy gap in the single particle spectrum. To see this explicitly, we can make a change of variables from ξ_k (for the normal state energies) to E_k (denoted as the quasiparticle energies), namely,

$N(E)dE \to N_n(\xi)d\xi$

Since, the density of states should remain conserved, as the number of carriers remain conserved,

$$N(E)dE = N_n(\xi)d\xi. \tag{4.41}$$

Since, the gap Δ is much smaller than the energy range ($\Delta \ll E$) over which the normal state DOS varies, we can replace,

$$N_n(\xi) \approx N_n(0) = N_0 \tag{4.42}$$

The quasiparticle energies are given by

$$E_k^2 = \Delta^2 + \xi_k^2. \tag{4.43}$$

Dropping the subscripts,

$$E^2 = \Delta^2 + \xi^2$$
$$2EdE = 2\xi\, d\xi$$
$$\frac{d\xi}{dE} = \frac{E}{\xi} = \frac{E}{\sqrt{E^2 - \Delta^2}}.$$

Here,

$$N(E)dE = N_0 d\xi$$
$$\frac{N(E)}{N_0} = \frac{d\xi}{dE}.$$

Thus,

$$\frac{N(E)}{N(0)} = \begin{cases} \frac{E}{\sqrt{E^2-\Delta^2}} & \text{for E } > \Delta \\ 0 & \text{for } E < \Delta. \end{cases}$$

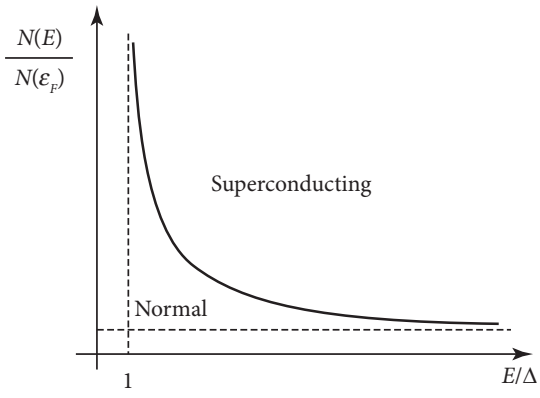


FIG. 4.15
The density of states (DOS) as a function of the scaled energy, E/Δ is shown. The right of the bold line denotes the superconducting phase, while on the left, it is the normal state.

Excitations with all momenta k, even those whose ξ_k fall in the gap have their energies increased above Δ. Moreover, we obtain a divergent DOS for $E \to \Delta$ above.

Considering that across the phase transition, the density of states remains constant, i.e.

$$N_s(E)dE = N_n(\xi)d\xi.$$

Since we are interested in an energy range of a few meV from the Fermi energy, we take $N_n(\xi) = N(\varepsilon_F) =$ a constant.

$$\frac{N_s(E)}{N(\varepsilon_F)} = \frac{d\xi}{dE} = \frac{E}{\sqrt{E^2 - \Delta^2}} \quad \text{for } E > \Delta.$$

The behavior of the DOS is plotted in Fig. 4.15. It is zero for $E < \Delta$, and falls off as E becomes larger than the gap, Δ. Furthermore, the behavior of the quasiparticle energy, E_k as a function of the single particle energy, ξ_k is shown in Fig. 4.16, where a superconducting gap of magnitude Δ is indicated in the figure.

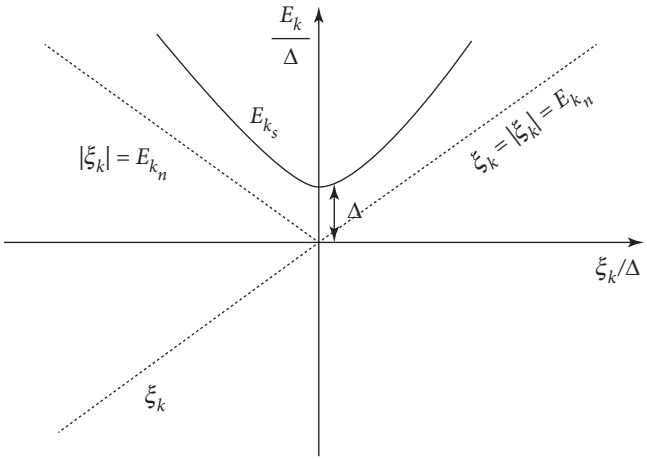


FIG. 4.16
The scaled quasiparticle energy, E/Δ is plotted as function of scaled single particle energy, ξ_k/Δ. The region above the parabola is superconducting (SC) and the region below the dotted lines denote metallic phases.

4.3 BCS THEORY

4.3.1 Introduction

Even though a lot of experimental progress has occurred in the field of superconductivity, a complete understanding of the pairing phenomena and consequently the formation of the superconducting state is lacking. In the following, we describe a microscopic theory that addresses all of these concerns. The The BCS theory of superconductivity was devised by Bardeen Cooper and Schrieffer in 1957. This accurately describes many of the properties of weak - coupling superconductors. The word "*weakly coupled*" refers to the strength of the electron phonon coupling that is considered as the origin of pairing in this class of superconductors, as we shall see below. The basic idea of the BCS theory is that the electrons in metal form bound pairs below a certain critical temperature called the transition temperature. Not all the electrons take part in this pairing process, only the ones that reside close to the Fermi surface, and within a distance of the debye energy, ω_D. The paired state of all of these electrons needs a much body description.

4.3.2 Isotope effect

Fröhlich (1950) realized that the electrons could interact by exchanging phonons, and in such a process, the interaction could be attractive. That was the first proposal that superconductivity can originate from electron-phonon interactions. But how do the phonons come into the picture? The dependence

on phonon parameters was experimentally demonstrated by the fact that the transition temperature, T_c is a function of the ionic mass for different isotopes of the same metal (that becomes a superconductor at lower temperatures). The effect of T_c is shown to have the form,

$$\Delta T_c/T_c = -\frac{1}{2}\left(\frac{\Delta M}{M}\right) \tag{4.44}$$

where M denotes the ionic mass, and ΔM is the difference in masses between the isotopes. This is called the isotope effect. It played a significant role in unraveling the issue concerning the attractive interaction among the charge carriers that eventually leads to superconductivity. Thus, we shall see, the contribution of the lattice excitations or the phonons is central to the formation of bound pairs. It will be shown later that physical quantities, such as the energy of the bound pairs, T_c involve phonon frequencies, and thus should display an isotope effect.

The isotope effect is more conveniently represented in terms of the isotope coefficient, α via the relation,

$$T_c \sim M^{-\alpha} \tag{4.45}$$

where α is given by

$$\alpha = -\frac{M}{\Delta M}\frac{\Delta T_c}{T_c}. \tag{4.46}$$

In general, α has a value 0.5; however, it may vary from one material to another.

4.3.3 Origin of attractive interaction

We consider the interaction between two electrons, one with ($\mathbf{k}$, $\uparrow$) and the other ($-\mathbf{k}$, $\downarrow$). One may consider a spin independent interaction, but since it is not relevant to the context, we can neglect the spin indices for future discussions. There are two mechanisms by which two electrons can interact: they can interact directly or via a third party, which in this case are phonons. We shall explore these two cases in the following.

4.3.3.1 Case I

One obvious process is a direct electron-electron interaction mediated via the Coulomb forces. Let us consider the initial state as $|i\rangle$ (before scattering) and the final state $|f\rangle$ after scattering. Let us denote this interaction term as H^{dir}_{e-e}. The matrix elements of such a term between the states can be written as

$$\langle i|H^{dir}_{e-e}|f\rangle = \int e^{i\mathbf{k}\cdot\mathbf{r}}U_c(\mathbf{r})e^{-i\mathbf{k}'\cdot\mathbf{r}}d^3\mathbf{r}. \tag{4.47}$$

We assume that both $|i\rangle$ and $|f\rangle$ are plane wave states. This term cannot yield a negative contribution and hence does not contribute to the formation of a bound state. The energies involved in this case can be enumerated as

i. Initial: $\epsilon_i = 2\xi_{\mathbf{k}} \quad \xi_{\mathbf{k}} = \frac{\hbar^2k^2}{2m} - \mu.$
ii. Final: $\epsilon_f = 2\xi_{\mathbf{k}'} \quad \mathbf{k}' = \mathbf{k} + \mathbf{q}$

4.3.3.2 Case II

Now let us consider an indirect process. Consider the scattering of two electrons via the exchange of a phonon. Thus, there is an intermediate process involved, and hence, it is a second-order process. There are two ways that such a second-order process can occur.

i. Electron 1 with momentum $\mathbf{k}$ emits a phonon, of frequency ω_q which is later re-absorbed by electron 2 with momentum $-\mathbf{k}$ (see left panel of Fig. 4.17).
ii. Electron 2 with momentum $-\mathbf{k}$ emits a phonon with frequency ω_q which is later re-absorbed by electron 1 with momentum $\mathbf{k}$ (see the right panel of Fig. 4.17).

Since both these processes are equally probable, they need to be incorporated with equal weight. The corresponding energies involved in the second case are,

i. Initial state (before scattering): $\epsilon_i = 2\xi_{\mathbf{k}'}$.
ii. Final state (after scattering): $\epsilon_f = 2\xi_{\mathbf{k}'}$.

Even if the initial and the final states are the same as in case I, the intermediate state is different. However, the corresponding energies are still the same, namely,

$$\epsilon_{int}^{(i)} = \xi_{\mathbf{k}'} + \xi_{\mathbf{k}} + \hbar\omega_q$$
$$\epsilon_{int}^{(ii)} = \xi_{\mathbf{k}} + \xi_{\mathbf{k}'} + \hbar\omega_q.$$

Now, the matrix element of $H_{e-e}^{indirect}$ can be calculated as

$$\begin{aligned}\langle i|H_{e-e}^{indirect}|f\rangle &= \sum_{int} \frac{\langle i|H_{e-e}^{indirect}|int\rangle\langle int|H_{e-e}^{indirect}|f\rangle}{(E_{i,f} - E_{int})^2} \\ &= \sum_{int} \langle i|H_{e-e}^{indirect}|int\rangle \frac{1}{2}\left[\frac{1}{E_f - E_{int}} - \frac{1}{E_i - E_{int}}\right]\langle int|H_{e-e}^{indirect}|f\rangle, \end{aligned} \tag{4.48}$$

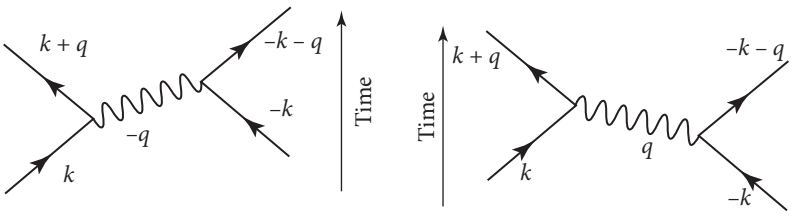


FIG. 4.17
Feynman diagrams show for the electron-electron interaction mediated via phonons. The left panel shows electron 1 emits a phonon, which is captured by electron 2 later. The right panel shows electron 2 emits a phonon which is captured by electron 1 later.

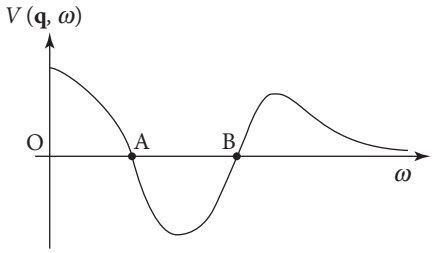


FIG. 4.18
A schematic plot of the interaction energy $V(\mathbf{q}, \omega)$ plotted as a function of the electron energy ω. The region between A and B is where the interaction becomes attractive.

where the summation is over the intermediate states. Note that the energy denominators can be written as,

$$\frac{1}{E_f - E_{int}} = \frac{1}{2\xi_{\mathbf{k}'} - \xi_{\mathbf{k}} - \xi_{\mathbf{k}'} - \hbar\omega_q}$$
$$\frac{1}{E_i - E_{int}} = \frac{1}{2\xi_{\mathbf{k}} - \xi_{\mathbf{k}'} - \xi_{\mathbf{k}} - \hbar\omega_q} \tag{4.49}$$

where, $\xi_{\mathbf{k}'} - \xi_{\mathbf{k}} = \hbar\omega$. Thus,

$$\langle i|H_{e-e}^{indirect}|f\rangle = \frac{1}{\omega - \omega_q} - \frac{1}{\omega + \omega_q}|\widetilde{V}_c(q)|^2 = \frac{2\omega_q}{\omega^2 - \omega_q^2}|\widetilde{V}_c(q)|^2. \tag{4.50}$$

The matrix elements are positive definite, so we are interested to know the sign of the energy denominator. For $\omega < \omega_q$, it is negative, and hence attractive. This is known as Cooper's instability. In real physical situations, the Coulomb term gets weakened by screening, which eventually takes the form,

$$V(\mathbf{q}, \omega) = \frac{4\pi e^2}{q^2 + k_0^2} + \frac{4\pi e^2}{q^2 + k_0^2}\frac{\omega_q}{\omega^2 - \omega_q^2}. \tag{4.51}$$

A schematic plot of the interaction term as a function of the electron energy, ω is plotted in Fig. 4.18. In a small region, between the points A and B marked in Fig. 4.18, $V(\mathbf{q}, \omega)$ becomes negative and hence attractive.

4.3.4 The BCS ground state

The ground states of a free Fermi gas correspond to the filled states with energy ϵ below ϵ_F and all states empty beyond $\epsilon > \epsilon_F$. This Fermi gas of electrons becomes unstable against the formation of at least one bound pair, regardless of how weak the interaction is, so long as it is attractive (Cooper, 1956).

It is well known that the binding does not ordinarily occur in a two body problem in two or three dimensions, until the strength of the potential exceeds a finite threshold value (Sakurai and Napolitano, 2011). So, in order to see how this binding comes about, we consider a simple model of two electrons added to the Fermi sea at $T = 0$. The restriction imposed here is that these two electrons interact with each other, but not with those inside the Fermi sea, except, of course due to the exclusion principle.

Thus, we seek a two-particle wavefunction for our purpose. Bloch argued that the lowest energy state for a two-particle system has zero total momentum. Thus, the two electrons have equal and opposite momenta. This means that the orbital wavefunction, $\psi_0(\mathbf{r_1}, \mathbf{r_2})$ should be of the form,

$$\psi_0(\mathbf{r}_1, \mathbf{r}_2) = \sum_k g_k e^{i\mathbf{k}\cdot\mathbf{r}_1} e^{-i\mathbf{k}\cdot\mathbf{r}_2}. \tag{4.52}$$

Considering the asymmetry of the total wavefunction with respect to the exchange of two electrons, ψ_0 contains even functions, such as $\cos[\mathbf{k}\cdot(\mathbf{r_1}-\mathbf{r_2})]$ with an antisymmetric spin part of the form, $[|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle]$, or a sum of products of $\sin[\mathbf{k}\cdot(\mathbf{r_1}-\mathbf{r_2})]$ with a symmetric triplet function of the form, $[|\uparrow\uparrow\rangle, |\downarrow\downarrow\rangle]$ and $[|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle]$. In a two-particle system, we expect that the singlet to have a lower energy because the cosinusoidal dependence of the orbital wavefunction $\mathbf{r_1}-\mathbf{r_2}$ gives a larger probability of the electrons to are near each other. Thus, we zero in on a two electron singlet wavefunction of the form,

$$\psi_0(\mathbf{r}_1,\mathbf{r}_2) = \left[\sum_{k>k_F} g_k \cos[\mathbf{k}\cdot(\mathbf{r}_1-\mathbf{r}_2)]\right][|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle]\,. \tag{4.53}$$

This can be plugged into the Schrödinger equation to yield,

$$(E-2\epsilon_{\mathbf{k}})g_{\mathbf{k}} = \sum_{k'>k_F} V_{\mathbf{k}\mathbf{k}'}g_{\mathbf{k}'}. \tag{4.54}$$

The above equation is true for any value of k. ϵ_k are the single particle energies (of the plane wave type), and $V_{kk'}$ are the matrix elements of the interaction potential,

$$V_{\mathbf{k}\mathbf{k}'} = \frac{1}{V}\int V(\mathbf{r})e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}}d\mathbf{r}, \tag{4.55}$$

where $\mathbf{r}$ denotes the spatial distance between the two electrons. $V_{\mathbf{k}\mathbf{k}'}$ describes the scattering of a pair of electrons with momenta $(\mathbf{k}'_\uparrow, -\mathbf{k}'_\downarrow)$ to $(\mathbf{k}_\uparrow, -\mathbf{k}_\downarrow)$. We have to solve for the amplitude, $g_{\mathbf{k}}$ such that the total energy $E < 2\epsilon_F$, such that a bound pair exists.

It is hard to carry out calculations for a general form of $V_{\mathbf{k}\mathbf{k}'}$. Cooper introduced a notable simplification,

$$\begin{aligned} V_{\mathbf{k}\mathbf{k}'} &= -V \quad \text{for } \epsilon_F < \epsilon_{k,k'} < \epsilon_F + \hbar\omega_D \\ &= 0 \quad \text{otherwise,} \end{aligned} \tag{4.56}$$

where V denotes the strength of the electron-electron interaction. This implies,

$$g_{\mathbf{k}} = V\sum_{\mathbf{k}'}\frac{g_{\mathbf{k}'}}{2\epsilon_k - E} \tag{4.57}$$

summing over the momentum index k,

$$\frac{1}{V} = \sum_{k>k_F}\frac{1}{2\epsilon_k - E}. \tag{4.58}$$

Note that this summation over k is for k values that are larger than the Fermi wavevector. We replace the summation with integration, by using the DOS at the Fermi level, namely, $N(\epsilon_F)$, which is constant for a particular material, and hence, can be brought out of the integral. Hence,

$$\frac{1}{V} = N(\epsilon_F)\int_{\epsilon_F}^{\epsilon_F+\hbar\omega_D}\frac{d\epsilon}{2\epsilon - E} = \frac{1}{2}N(\epsilon_F)\ln\left(\frac{2\epsilon_F - E + 2\hbar\omega_D}{2\epsilon_F - E}\right). \tag{4.59}$$

For conventional superconductors,

$N(\epsilon_F)V < 0.3$.

This allows the applicability of the weak coupling approximation valid for $N(\epsilon_F)V \ll 1$. This yields,

$$\begin{aligned} \frac{2}{N(\epsilon_F)V} &= \ln\left(\frac{2\epsilon_F - E + 2\hbar\omega_D}{2\epsilon_F - E}\right) \\ \text{or,} \quad \frac{2\epsilon_F - E + 2\hbar\omega_D}{2\epsilon_F - E} &= e^{2/N(\epsilon_F)V} \\ 2\epsilon_F - E + 2\hbar\omega_D &= (2\epsilon_F - E)e^{2/N(\epsilon_F)V} \\ E &= 2\epsilon_F - 2\hbar\omega_D e^{-2/N(\epsilon_F)V}. \end{aligned} \tag{4.60}$$

This is the expression for the total energy of the system of two electrons in the vicinity of the Fermi sea. Thus, with respect to the Fermi surface of two electrons, the energy is negative, and hence it corresponds to a bound state. A bound state of two electrons is thus formed, even for an infinitesimal interaction strength. Also, the form says that the binding energy is not analytic at $V = 0$, so it cannot be expanded in powers of V. Thus, any finite order perturbation theory would not be able to yield this result, and thus the theory is non-perturbative.

As for the wavefunction, it is dependent on the relative coordinate, $\mathbf{r} = \mathbf{r_1} - \mathbf{r_2}$, and is proportional to

$$\sum_{k>k'} \frac{\cos(\mathbf{k}\cdot\mathbf{r})}{2\xi_{\mathbf{k}} + E'}$$

where, $\xi_{\mathbf{k}} = \epsilon_k - \epsilon_F$, and $E' = 2\epsilon_F - E > 0$. E' can now be called the binding energy relative to $2\epsilon_F$. A few comments are in order.

i. The amplitude of the wavefunction, namely, $(2\xi_k + E')^{-1}$ has its maximum value $1/E'$ when $\xi_k = 0$, that is, for electrons at the Fermi level. It falls off at positive values of ξ_k. Thus, it elucidates the dominant role of the electrons at the Fermi energy in the pairing mechanism.
ii. The electronic states within an energy range E' about ϵ_F are involved in the formation of the bound state.
iii. Note that $E' < \hbar\omega_D$ for $N(\epsilon_F)V < 1$, this makes sure that the detailed behavior of $V_{kk'}$ is not important, and the assumption of the form in Eq. (4.56) will suffice.
iv. Also, the small energy range allows estimation of the range of the Cooper pairs via the uncertainty principle, namely, $\xi_0 = \frac{a\hbar v_F}{kT_c}$. Thus, the pairs are highly overlapping.

4.3.5 Statistical description of the BCS ground states

What we have learned so far is that the filled Fermi sea becomes unstable to the formation of a bound pair. In *s*-wave superconductivity, a large number of bound pairs form and a system condenses into the ground state below the same transition temperature, T_c. Let us postulate a paired many body paired

ground state of the form (since it is a ground state, we assume it to be a singlet state),

$$|\psi_0\rangle = \sum_{k>k_F} g_k c^\dagger_{k\uparrow} c^\dagger_{-k\downarrow}|0\rangle \tag{4.61}$$

where $|0\rangle$ does not denote vacuum, rather it represents a filled Fermi sea. Suppose there are M electrons, and we have to involve N electrons of those to form $\frac{N}{2}$ pairs, the number of ways $\frac{N}{2}$ pairs can be formed from M electrons is,

$$\frac{M!}{(M-\frac{N}{2})!(\frac{N}{2})!} \approx (10^{20})^{20} \quad \text{for} \quad M = 10^{23}. \tag{4.62}$$

Thus, a solution of the problem demands that $(10^{20})^{20}$ number of g_k to be solved, which is certainly an impossible task. Thus, a statistical description must be employed here. Further, since the number of particles is large, no serious error will be done if one works with a grand canonical ensemble, where only the average number of particles, that is, $\overline{N}$ is fixed.

To circumvent a significantly difficult problem, the BCS theory postulated a ground state of the form,

$$|\psi_G\rangle = \prod_{\mathbf{k}=\mathbf{k_1},\mathbf{k_2},\ldots\mathbf{k_M}} (u_{\mathbf{k}} + v_{\mathbf{k}} c^\dagger_{k\uparrow} c^\dagger_{-k\downarrow})|\psi_0\rangle. \tag{4.63}$$

The probability of a paired state $(\mathbf{k}_\uparrow, -\mathbf{k}_\downarrow)$ is occupied is denoted by $|v_k|^2$, while it is unoccupied is given by $|u_k|^2$ with the condition,

$$|u_k|^2 + |v_k|^2 = 1 \quad \forall\, k. \tag{4.64}$$

Take an example of two states $\mathbf{k_1}$ and $\mathbf{k_2}$, then the wave function amplitudes represent,

$u_{k_1}, u_{k_2} \rightarrow$ 0 pair
$u_{k_1}, v_{k_2} \rightarrow$ 1 pair in $(\mathbf{k_2}, -\mathbf{k_2})$
$v_{k_1}, v_{k_2} \rightarrow$ 2 pairs in $(\mathbf{k_1}, -\mathbf{k_1}), (\mathbf{k_2}, -\mathbf{k_2})$.

We can calculate the average number of particles in the following way:

$$\overline{N} = \langle \hat{N}_{OP}\rangle = \left\langle \sum_{\mathbf{k},\sigma} \hat{n}_{\mathbf{k},\sigma} \right\rangle = \langle\psi_G| \sum_{\mathbf{k}} c^\dagger_{k\uparrow} c_{k\uparrow} + c^\dagger_{k\downarrow} c_{k\downarrow} |\psi_G\rangle = 2\sum_k \langle \psi_G c^\dagger_{k\uparrow} c_{k\uparrow} \psi_G\rangle$$
$$= 2\sum_k \langle\psi_0|(u^*_k + v^*_k c_{-k\downarrow} c_{k\uparrow}) c^\dagger_{k\uparrow} c_{k\uparrow} (u_k + v_k c^\dagger_{k\uparrow} c^\dagger_{-k\downarrow}) \prod_{l\neq k} (u^*_l + v^*_l c_{-l\downarrow} c_{l\uparrow})(u_l + v_l c^\dagger_{l\uparrow} c^\dagger_{-l\downarrow})|\psi_0\rangle.$$

In principle u_k and v_k are complex quantities. We have also used $\langle A\phi|\psi\rangle = \langle\phi|A^\dagger|\psi\rangle$. Let us look at various terms.

First, let us look at the term for $l \neq k$, namely,

$$|u_l|^2 + u^*_l v_l c^\dagger_{l\uparrow} c^\dagger_{-l\downarrow} + v^*_l u_l c_{-l\downarrow} c_{l\uparrow} + |v_l|^2 c_{-l\downarrow} c_{l\uparrow} c^\dagger_{l\uparrow} c^\dagger_{-l\downarrow} \tag{4.65}$$

where we have taken the expectation with $|\psi_0\rangle$. The middle two terms (second and the third terms) yield zero as they change the occupancy of the state with l-th pair (first term creates and the second

term destroys). The last term only yields a normalization, because it creates and destroys a pair. Thus, the $l \neq k$ term yields,

$$|u_l|^2 + |v_l|^2$$

which, by normalization, is equal to 1.

The same procedure, when applied to $l = k$ term, we shall have two terms surviving with $u_k v_k$ will cancel out. Further,

$$|u_k|^2 c^\dagger_{k\uparrow} c_{k\uparrow} |\phi_0\rangle = 0$$

since there are no states to annihilate for $k > k_F$. Thus, we are left with only $|v_k|^2$.

Thus, the average number of particles yields,

$$\overline{N} = 2 \sum_k |v_k|^2.$$

The factor 2 appears for the summation over spin. Furthermore, our assumption that this reinforces $|v_k|^2$ indeed denotes the probability of occupied states.

To calculate some more statistical quantities that may be of physical relevance, we estimate the RMS fluctuations of $\overline{N}$, we calculate

$$\langle (N - \overline{N})^2 \rangle = \langle N^2 - 2N\overline{N} + \overline{N}^2 \rangle = \langle N^2 \rangle - \overline{N}^2 = 4 \sum_k u_k^2 v_k^2. \tag{4.66}$$

This is a positive definite quantity. In fact, v_k as a function of k goes from 1 to zero and u_k goes from $0 \to 1$ in an energy range given by $k_B T_c$. Thus, the sum above is proportional to $\frac{T_c}{T_F}\overline{N}$.

4.4 THE VARIATIONAL CALCULATION

We write down a generic Hamiltonian that encodes the kinetic energy and a two-particle interaction term as in the following:

$$\mathcal{H} = \sum_{k,\sigma} \xi_{k\sigma} c^\dagger_{k\sigma} c_{k\sigma} + \sum_{k,l} V_{kl} c^\dagger_{k\uparrow} c^\dagger_{-k\downarrow} c_{-l\downarrow} c_{l\uparrow} = \hat{K} + \hat{V}, \tag{4.67}$$

where $\hat{K}$ and $\hat{V}$ denote the kinetic and the interaction energies, and $\xi_k = \epsilon_k - \mu$. k, l are the momenta indices. As exact solutions are hard to arrive at, and perturbation theory will not be of relevance for reasons discussed earlier, we shall resort to a variational calculation, that is, minimize the energy with respect to the tunable parameters, which are the coefficients u_k and v_k. That is,

$$\delta \left[\langle \psi_G | \sum_{k,\sigma} \xi_k n_{k\sigma} + \sum_{k,l} V_{kl} c^\dagger_{k\uparrow} c^\dagger_{-k\downarrow} c_{-l\downarrow} c_{l\uparrow} | \psi_G \rangle \right]. \tag{4.68}$$

Let us look at the kinetic energy term. We have already computed it earlier, namely,

$$\langle \hat{K} \rangle = \langle \psi_G | \sum_{k,\sigma} \xi_k n_{k\sigma} | \psi_G \rangle = 2 \sum_k |v_k|^2 \xi_k. \tag{4.69}$$

Similarly, for the potential energy term,

$$\langle \hat{V} \rangle = \sum_{k,l} V_{kl} u_k^* v_l^* u_l v_k, \tag{4.70}$$

where we have used the normalization constraint, that is, $|u_k|^2 + |v_k|^2 = 1 \; \forall \; k$.

Doing a variational calculation with two variables is a difficult job, which, however, is simplified owing to the normalization condition. This yields an identification of (u_k, v_k) with $(\sin\theta_k, \cos\theta_k)$ which facilitates the Hamiltonian to be dependent only on one parameter, namely, θ_k. In particular, we take

$$u_k = \sin\theta_k, v_k = \cos\theta_k. \tag{4.71}$$

Thus, we can minimize energy with respect to θ_k, which implies

$$\frac{\partial}{\partial \theta_k} \left[\sum_{k'} \xi_{k'} (1 + \cos 2\theta_{k'}) + \frac{1}{4} \sum_{k',l} V_{k'l} \sin 2\theta_{k'} \sin 2\theta_l \right] = 0 \tag{4.72}$$

$$2\xi_k \sin 2\theta_k + \sum_l V_{kl} \cos 2\theta_k \sin 2\theta_l = 0. \tag{4.73}$$

Now define,

$$\Delta_k = -\frac{1}{2} \sum_l V_{kl} \sin 2\theta_l = -\frac{1}{2} \sum_l V_{kl} \sin 2\theta_l. \tag{4.74}$$

Using this in the above equation can be written as

$$2\xi_k \sin 2\theta_k = \frac{1}{2} \sum_l V_{kl} \cos 2\theta_k \sin 2\theta_l = -2\Delta_k \cos 2\theta_k. \tag{4.75}$$

This yields

$$\tan 2\theta_k = -\frac{\Delta_k}{\xi_k}, \quad \text{implying that,} \quad \frac{\sin 2\theta_k}{\cos 2\theta_k} = -\frac{\Delta_k}{\xi_k}. \tag{4.76}$$

Using the definitions,

$$\begin{aligned} 2u_k v_k &= \sin 2\theta_k = \frac{\Delta_k}{\sqrt{\xi_k^2 + \Delta_k^2}} \\ v_k^2 - u_k^2 &= \cos 2\theta_k = -\frac{\xi_k}{\sqrt{\xi_k^2 + \Delta_k^2}}. \end{aligned} \tag{4.77}$$

We have

$$v_k^2 - u_k^2 = -\frac{\xi_k}{E_k},$$

where, $E_k = \sqrt{\xi_k^2 + \Delta^2}$, and, of course, the normalization condition, namely, $u_k^2 + v_k^2 = 1$. Thus, the definition for the coherence factors can be written as

$$v_k^2 = \frac{1}{2}\left(1 - \frac{\xi_k}{E_k}\right) = \frac{1}{2}\left[1 - \frac{\xi_k}{\sqrt{\Delta^2 + \xi_k^2}}\right]$$
$$u_k^2 = \frac{1}{2}\left(1 + \frac{\xi_k}{E_k}\right) = 1 - v_k^2. \tag{4.78}$$

This is a choice which seems to fit all definitions.

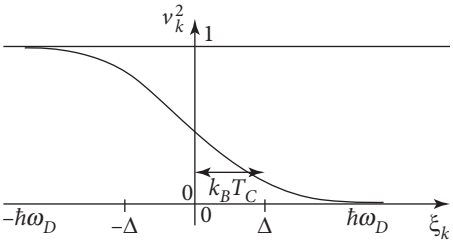


FIG. 4.19
The behaviour of v_k^2 (probability of states occupied by a pair of momentum **k** and $-$**k**) is shown as a function of ξ_k $(= \epsilon_k - \mu)$.

There is a close resemblance between the v_k^2 for the BCS ground state at $T = 0$, and the free Fermi distribution at $T = T_F$ (see Fig. 4.19). Also, v_k^2 falls off as $1/\xi_k^2$ for $\xi_k \gg \Delta$. If one remembers, g_k^2 falls off in a similar fashion in the simple treatment of a single Cooper pair presented earlier. Also, we have chosen cos $2\theta_k$ with a negative sign, because if ξ_k is large (that is, $\epsilon_k \gg \mu$) v_k should go to zero.

Hence, Δ_k assumes the form,

$$\Delta_k = -\frac{1}{2}\sum_l V_{kl} \sin 2\theta_l$$
$$\Delta_k = -\frac{1}{2}\sum_l V_{kl} \frac{\Delta_l^2}{\sqrt{\Delta_l^2 + \xi_l^2}}. \tag{4.79}$$

A trivial solution of the above equation is $\Delta = 0$. This implies, $2u_k v_k = 0$, and $v_k^2 - u_k^2 = -1$. Since, $u_k^2 + v_k^2 = 1$, it implies that $v_k^2 = 0$, which, in turn, means a normal state and hence, there are no pairs. The above discussion clearly indicates that Δ_k can be used as an order parameter for the superconducting transition, which when finite implies a superconducting state, and $\Delta_k = 0$ denotes a normal state.

Let us recall Cooper's original assumption on the nature of the interaction, that is,

$V_{kl} = -V$ if $|\xi_k - \xi_l| < \hbar\omega_D$.

Hence, putting it in Eq. (4.74),

$$\Delta_k = \frac{1}{2} V \sum_l \sin 2\theta_l \quad \text{for } |\xi_k - \xi_l| < \hbar\omega_D = 0 \quad \text{otherwise.} \tag{4.80}$$

From the above equation, it is clear that the order parameter, Δ is independent of the momentum index, k. Substituting sin $2\theta_l$ from Eq. (4.77), one may cancel Δ from both sides to get

$$1 = \frac{V}{2}\sum_l \frac{1}{\sqrt{\Delta^2 + \xi_l^2}}$$

$$\frac{1}{V} = \int_0^{\hbar\omega_D} \frac{N(\varepsilon_F)d\xi}{2\sqrt{\Delta^2+\xi^2}}$$
$$\frac{2}{N(\varepsilon_F)V} = \int_0^{\hbar\omega_D} \frac{d\xi}{\sqrt{\Delta^2+\xi^2}}. \tag{4.81}$$

Integrating and solving for Δ,

$$2\sinh^{-1}\left(\frac{\xi}{\Delta}\right)_0^{\hbar\omega_D} = \frac{2}{N(\varepsilon_F)V}$$
$$2\sinh^{-1}\left(\frac{\hbar\omega_D}{\Delta}\right) = \frac{2}{N(\varepsilon_F)V}. \tag{4.82}$$

Reorganizing,

$$\Delta = \frac{\hbar\omega}{\sinh(1/N(\varepsilon_F)V)}. \tag{4.83}$$

For a weak coupling superconductor, $N(\varepsilon_F)V \ll 1$, one arrives at

$$\Delta = 2\hbar\omega_D e^{-1/N(\varepsilon_F)V}. \tag{4.84}$$

This is the expression for the energy difference between the superconducting state and the normal state. Thus, this much energy has to be supplied in order to break a superconducting state and achieve a normal (metallic) state. Suppose a thermal excitation facilitates such a transition, which yields $\Delta \sim k_B T_c$; A priori, from BCS theory one gets (will be shown immediately afterwards)

$$k_B T_c \approx 1.13\,\hbar\omega_D \exp[-1/N(\varepsilon_F)V], \tag{4.85}$$

which only changes the factor "2" in Eq. (4.84) to 1.13 in Eq. (4.85). Having found Δ, we may simply compute the coefficients u_k and v_k which should specify the optimal BCS wavefunction.

4.4.1 Temperature dependence of the gap

We have introduced that the energy gap (which also acts as the order parameter) in a superconductor is defined by

$$\Delta_k = -\sum_{k'} V_{kk'} \langle c_{-k'\downarrow} c_{k'\uparrow} \rangle \tag{4.86}$$

Δ_k allows us to write the mean field BCS Hamiltonian,

$$\mathcal{H} = \sum_k^{\sigma} \xi_k c^\dagger_{k\sigma} c_{k\sigma} - \sum_k (\Delta_k c^\dagger_{k\uparrow} c^\dagger_{-k\downarrow} + \Delta^*_k c_{-k\downarrow} c_{k\uparrow}). \tag{4.87}$$

A constant term is dropped from the above equation. This Hamiltonian can be diagonalized by Bogoliubov Valatin transformation, where the creation and the annihilation operators are expressed in terms of the quasiparticle operators as

$$c_{k\uparrow} = u^*_k \gamma_{k_0} + v_k \gamma^*_{k_1} \tag{4.88}$$
$$c^\dagger - k\downarrow = -v^*_k \gamma_{k_0} + u_k \gamma^*_{k_1}. \tag{4.89}$$

Being a linear combination of c_k and $c_k^\dagger$, γ_k obeys fermionic anticommutator relations and thus the relationships are canonical. This along with the normalization condition yields

$$\begin{aligned}\mathcal{H} = &\sum_k \xi_k \left[(|u_k|^2 - |v_k|^2)(\gamma_{k_0}^\dagger \gamma_{k_0} + \gamma_{k_1}^\dagger \gamma_{k_1}) + 2|v_k|^2 + 2u_k^* v_k^* \gamma_{k_1}\gamma_{k_0} + 2u_k v_k \gamma_{k_1}^\dagger \gamma_{k_0}^\dagger\right] \\ &+ \sum_k (\Delta_k u_k v_k^* + \Delta_k^* u_k^* v_k)(\gamma_{k_0}^\dagger \gamma_{k_0} + \gamma_{k_1}^\dagger \gamma_{k_1} - 1) + (\Delta_k v_k^{*2} - \Delta_k^* u_k^{*2})\gamma_{k_1}\gamma_{k_0} \\ &+ (\Delta_k^* v_k^2 - \Delta_k u_k^2)\gamma_{k_0}^\dagger \gamma_{k_1}^\dagger.\end{aligned} \tag{4.90}$$

Finally, one gets

$$\mathcal{H} = \sum_k E_k(\gamma_{k_0}^\dagger \gamma_{k_0} + \gamma_{k_1}^\dagger \gamma_{k_1}) + \sum_k (\xi_k - E_k + |\Delta_k|^2) \tag{4.91}$$

where $E_k = \sqrt{\xi_k^2 + |\Delta|^2}$. The second term is just a constant. So, the Hamiltonian is diagonal on the basis of the quasiparticle operators.[3]

For the gap function, Δ, in terms of the quasiparticle operators after a little bit of algebra yields

$$\Delta_k = -\sum_{k'} V_{kk'} u_{k'}^* v_{k'} \langle 1 - \gamma_{k_0'}^\dagger \gamma_{k_0'} - \gamma_{k_1'}^\dagger \gamma_{k_1'}\rangle. \tag{4.92}$$

At zero temperature, $\langle \gamma_{k_0}^\dagger \gamma_{k_0}\rangle = 1$; however, at non-zero temperatures,

$$\langle \gamma_{k_0}^\dagger \gamma_{k_0}\rangle = f(E_k) = \frac{1}{e^{\beta E_k} + 1}$$

where $f(E_k)$ denotes the Fermi distribution function for the quasiparticles. Since the expectation value of each of the number operators yields a Fermi function, we get,

$$\langle 1 - \gamma_{k_0}^\dagger \gamma_{k_0} - \gamma_{k_1}^\dagger \gamma_{k_1}\rangle = 1 - 2f(E_k).$$

Thus, for the gap function,

$$\Delta_k = -\sum_{k'} V_{kk'} u_{k'}^* v_{k'}(1 - 2f(E_k)) = -\sum_{k'} V_{kk'} \frac{\Delta_{k'}}{2E_{k'}} \tanh\left(\frac{\beta E_{k'}}{2}\right).$$

With $V_{kk'} = -V$,

$$\frac{1}{V} = \frac{1}{2}\sum_k \frac{\tanh(\beta E_k/2)}{E_k}. \tag{4.93}$$

The above equation holds a clue to the temperature dependence of the energy gap, $\Delta(T)$. Converting the sum into an integral by introducing the DOS,

$$\begin{aligned}\frac{1}{N(\epsilon_F)V} &= \int_0^{\beta_c \hbar\omega_D/2} \frac{\tanh x}{x} dx, \quad \text{where, } x = \frac{\beta E_k}{2} \\ \int_0^{\beta_c \hbar\omega_D/2} \frac{\tanh x}{x} dx &= \ln\left(\frac{2e^\gamma}{\pi}\beta_c \hbar\omega_D\right)\end{aligned} \tag{4.94}$$

[3] In order for the quasiparticle operators (γ) to diagonalize the idea is to make coefficients of $\gamma\gamma$ or $\gamma^\dagger\gamma^\dagger$ equal to zero and retain $\gamma^\dagger\gamma$. This happens when $2\xi_k u_k v_k + \Delta_k^* v_k^2 - \Delta_k u_k^2 = 0$.

where γ denotes the Euler constant, which has a value, $\gamma = 0.577\cdots$ and thus $\frac{2e^{\gamma}}{\pi} \simeq 1.14$. This yields an expression for the transition temperature, T_c above which superconductivity vanishes,

$$k_B T_c = 1.13\hbar\omega_c e^{-1/N(\epsilon_F)V}. \tag{4.95}$$

One may recall that earlier we have obtained an expression for the energy gap, namely,

$$\Delta = 2\hbar\omega_D e^{-1/N(\epsilon_F)V}$$

which was subsequently equated to $k_B T_c$. Also, a relationship between the energy gap at zero temperature to that of T_c is obtained as

$$\frac{\Delta(0)}{k_B T_c} = \frac{2}{1.13} = 1.764. \tag{4.96}$$

Thus, the gap at $T = 0$ is indeed comparable to the magnitude of $k_B T_c$. The numerical factor has been tested in many experiments and found to be reasonable. The value of $2\Delta(0)$ varies between $3k_B T_c$ and $4.5k_B T_c$ for the conventional (that is, BCS) superconductors, and is mostly clustered around the BCS value $3.5k_B T_c$.

Finally, to arrive at an explicit temperature dependence of the gap function, we consider the integral form of the gap equation, namely,

$$\frac{1}{N(\varepsilon_F)V} = \int_0^{\hbar\omega_D} \frac{\tanh\left[\frac{1}{2}\beta(\xi^2 + \Delta^2)^{1/2}\right]}{(\xi^2 + \Delta^2)^{1/2}} d\xi. \tag{4.97}$$

Now $\Delta(T)$ needs to be computed numerically by self-consistently solving the above equation. For weak coupling superconductors, for which $\hbar\omega_D \gg k_B T_c$, $\Delta(T)/\Delta(0)$ is a universal function of (T/T_c) which decreases monotonically from 1 at $T = 0$ to zero at $T = T_c$. With $\Delta(T)$ determined, the quasiparticle energies can be written as

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

In the vicinity of $T = 0$, the temperature variation is exponentially slow because of $e^{-\Delta/k_B T} \approx 0$ (see Fig. 4.20). Hence, the tan hyperbolic term is insensitive to T and stays very close to 1. Physically speaking, Δ remains nearly a constant until a significant number of quasiparticles are thermally excited.

On the other hand, for $T \sim T_c$, $\Delta(T)$ vanishes with a vertical slope approximately as

$$\frac{\Delta(T)}{\Delta(0)} \approx 1.74\left[1 - \left(\frac{T}{T_c}\right)\right]^{1/2} \text{ for } T \to T_c \text{ from below.}$$

This square root dependence of the Δ on T is a hallmark feature of all mean field theories. For example, the magnetization vanishes in a similar fashion to temperature in molecular field theory of ferromagnetism. The dependence of $\Delta(T)$ presented in Fig. 4.20 closely matches with the experimental data on the elemental superconductors, such as Tin (Sn), Tantalum (Ta), and Niobium (Nb) presented in Fig. 4.21.

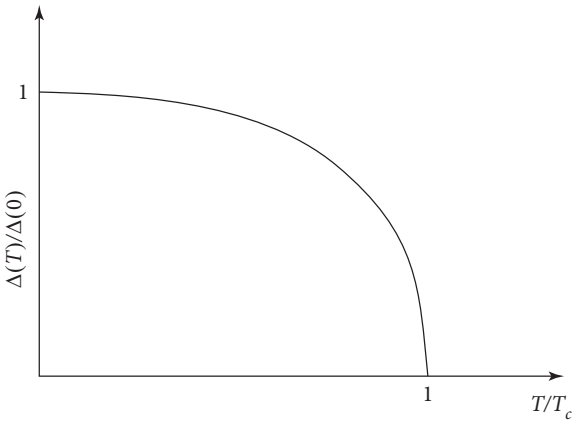


FIG. 4.20
The scaled gap function $\Delta(T)/\Delta(0)$ is numerically obtained and plotted as a function of the reduced temperature, T/T_c. The area enclosed by the curve denotes the superconducting phase.

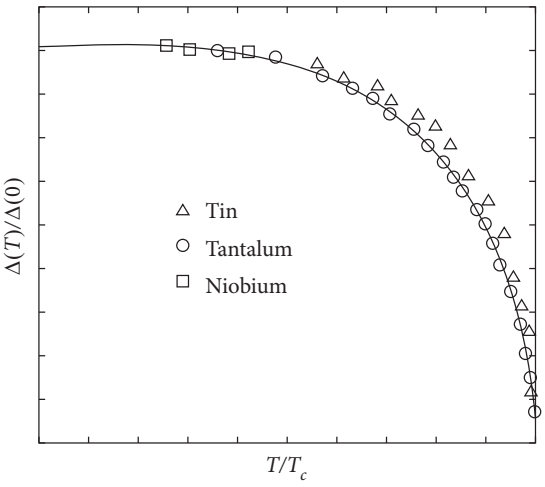


FIG. 4.21
The scaled energy gap is shown as a function of reduced temperature, T/T_c for Tin, Tantalum and Niobium. The figure is taken from Blatt (1992). The continuous lines depict the result of BCS theory.

4.4.2 Thermodynamics from BCS theory

The energy of a superconductor is obtained by explicitly solving the BCS Hamiltonian, which yields

$$E = 2\sum_{k>k_F} \xi_k v_k^2 + 2\sum_{k<k_F} |\xi_k| u_k^2 + \sum_{kk'} V_{kk'} u_k v_k u_{k'} v_{k'} \tag{4.98}$$

where,

$$v_k^2 = \frac{1}{2}\left(1 - \frac{\xi_k}{\sqrt{\xi_k^2 + \Delta_k^2}}\right)$$

with $\xi_k = \varepsilon_k - \mu$, and $\Delta_k = -\sum_{k'} V_{kk'} u_{k'} v_{k'}$. Furthermore, the form of the attractive interaction is

$$V_{kk'} = \begin{cases} V & \text{for both } |\xi_k| \text{ and } |\xi_{k'}| < \hbar\omega_D \\ 0 & \text{otherwise.} \end{cases}$$

Also, we have obtained that

$$\Delta \simeq 2\hbar\omega_D \exp[-1/N(\epsilon_F)V].$$

Thus, the energy difference between the superconducting and the normal states is given by

$$E_s - E_n = -\frac{1}{2}N(\varepsilon_F)\Delta^2\left[1 - \exp\left(\frac{-2}{N(\epsilon_F)V}\right)\right], \tag{4.99}$$

which in the weak coupling limit assumes the form,

$$E_s - E_n = -\frac{1}{2}N(\varepsilon_F)\Delta^2$$

To enumerate the jump in the specific heat from BCS theory, we consider the entropy for the SC state given by Reif (1965)

$$S_s = -2k_B\sum_k\left[(1-f_k)\ln(1-f_k) + f_k\ln f_k\right]. \tag{4.100}$$

Furthermore, in terms of the probability of a microstate being occupied, the entropy has a form,

$$S = -k_B\sum p_i \ln p_i. \tag{4.101}$$

Following which the specific heat can be determined by

$$C_{es} = T\frac{dS_s}{dT} = T\frac{dS_s}{d\beta}\cdot\frac{d\beta}{dT} = T\frac{dS_s}{d\beta}\cdot\left(-\frac{1}{k_BT^2}\right) = -\beta\frac{dS_s}{d\beta}$$

now,

$$\begin{aligned} C_s &= 2\beta k_B\sum_k\frac{\partial f_k}{\partial\beta}\ln\frac{f_k}{1-f_k} = -2\beta^2k_B\sum_k E_k\frac{\partial f_k}{\partial\beta} = -2\beta^2k_B\sum_k E_k\frac{df_k}{d(\beta E_k)}\left(E_k + \beta\frac{dE_k}{d\beta}\right) \\ &= -2\beta k_B\sum_k\left(\frac{df_k}{dE_k}\right)\left(E_k^2 + \frac{1}{2}\beta\frac{d\Delta^2}{d\beta}\right). \end{aligned} \tag{4.102}$$

The first term is the contribution of the quasiparticles to the specific heat, while the second term is due to the temperature-dependent gap function. The features of this specific heat are as follows:

i. C_{es} is exponentially small at $T \ll T_c$, where the excitation energy Δ is much greater than k_BT. This explains the exponential dependence at low temperatures, namely, $C_s \sim T_ce^{-T/T_C}$.
ii. Near T_c, the gap, $\Delta(T) \to 0$, when one can replace E_k by $|\xi_k|$, the first term reduces to

$$C_s(T) = \gamma T = \frac{2\pi^2}{3}N(\varepsilon_F)k_B^2T,$$

which is continuous at $T = T_c$.
iii. The second term is finite below T_c, where $\frac{d\Delta^2}{dT}$ is large but is zero above T_c, thereby giving rise to a discontinuity, ΔC in the electronic specific heat at T_c. This discontinuity is characteristic of a second-order phase transition.

The jump in the specific heat is enumerated as

$$\Delta C = (C_s - C_n)\,|_{T=T_c} = N(\varepsilon_F)k_B\beta^2\frac{d(\Delta^2)}{d\beta}\int_{-\infty}^{\infty}\left(\frac{df}{d|\xi|}\right)d\xi = N(\varepsilon_F)\left(-\frac{d\Delta^2}{dT}\right)_{T=T_c}. \tag{4.103}$$

Here, $\frac{df}{d|\xi|} = \frac{df}{d\xi}$ as f is an even function of ξ. Furthermore, using[4]

$$\Delta(T) = 1.74\Delta(0)\left(1 - \frac{T}{T_c}\right)^{1/2}$$

the magnitude ΔC can be computed to yield

$$\Delta C = 9.4N(\varepsilon_F)k_B^2 T_c,$$

where $\Delta(0) = 1.76\, k_B T_c$. The normalized discontinuity, usually measured in experiments, is given by

$$\frac{\Delta C}{C_{en}} = \frac{9.4}{2\pi^2/3} \approx 1.43$$

$C_{es}(T)$ can be found numerically from the parent expression [Eq. (4.102)]. $C_{es}(T)$ can be integrated to obtain $U_{es}(T)$. The clue is that at $T = T_c$, $U_{es}(T)$ must be same as the normal value $U_{en}(0) + \frac{1}{2}\gamma T_c^2$.

4.5 ELECTROMAGNETIC CONSIDERATIONS

Let us consider a system of N electrons subjected to a magnetic field, $\mathbf{B}$ which is described by a vector potential. The momentum for each of these particles described by $i = 1, \ldots, N$ under an external vector potential, $\mathbf{A}(\mathbf{r}, t)$ can be replaced by

$$\mathbf{p}_i \rightarrow \mathbf{p}_i + e\mathbf{A}(\mathbf{r}).$$

This yields a Hamiltonian,

$$\mathcal{H} = \frac{1}{2m}\sum_{i=1}^{N}(\mathbf{p}_i - e\mathbf{A}(\mathbf{r}_i))^2 + \hat{V} = \mathcal{H}_0 + \mathcal{H}', \tag{4.104}$$

where, $\mathcal{H}_0$ is the many body Hamiltonian with $\mathcal{H}_0 = \sum_i \frac{p_i^2}{2m}$. For the linear order in $\mathbf{A}$, the coupling to the external probe can be written as

$$\mathcal{H}' = -\frac{e}{2m}\sum_{i=1}^{N}\left[\mathbf{p}_i \cdot \mathbf{A}(\mathbf{r}_i) + \mathbf{A}(\mathbf{r}_i) \cdot \mathbf{p}_i\right]. \tag{4.105}$$

With the choice of a transverse gauge, that is, $\nabla \cdot \mathbf{A} = 0$, both the terms yield the same contribution. Hence, we can write,

$$\mathcal{H}' = -\frac{e}{m}\sum_{i=1}^{N}\mathbf{p}_i \cdot \mathbf{A}(\mathbf{r}_i). \tag{4.106}$$

This greatly simplifies the analysis, as will be clear in the following discussion.

[4] This is the numerical solution of the gap function, $\Delta(T)$ shown in Fig. 4.20.

The total current operator can be written as

$$\mathbf{j}(\mathbf{r}) = \frac{e}{2}\sum_{i=1}^{N}(\mathbf{v}_i\delta(\mathbf{r}-\mathbf{r}_i) + \delta(\mathbf{r}-\mathbf{r}_i)\mathbf{v}_i) \tag{4.107}$$

with the velocity operator given by

$$\mathbf{v}_i = \frac{1}{m}(\mathbf{p}_i - \frac{e}{c}\mathbf{A}(\mathbf{r}_i)) \tag{4.108}$$

and the δ-functions denote the positions of the particles. The above form for current can be split into

$$\mathbf{j}(\mathbf{r}) = \mathbf{j}_p(\mathbf{r}) + \mathbf{j}_d(\mathbf{r}), \tag{4.109}$$

where $\mathbf{j}_p$ and $\mathbf{j}_d$ denote the paramagnetic and diamagnetic current densities, and are given by the following:

$$\mathbf{j}_p(\mathbf{r}) = \frac{e}{2m}\sum_{i=1}^{N}(\mathbf{p}_i\delta(\mathbf{r}-\mathbf{r}_i) + \delta(\mathbf{r}-\mathbf{r}_i)\mathbf{p}_i)\,. \tag{4.110}$$

In terms of the fermionic operators,

$$\mathbf{j}_p(\mathbf{q}) = \frac{e\hbar}{m}\sum_{\mathbf{k},\sigma}(\mathbf{k}+\mathbf{q}/2)c^{\dagger}_{k\sigma}c_{k+q,\sigma}\,, \tag{4.111}$$

where $\mathbf{k}$ and $\mathbf{q}$ denote momentum indices. Next, the diamagnetic current density is written as

$$\mathbf{j}_d(\mathbf{r}) = -\frac{e^2}{m}\mathbf{A}(\mathbf{r})\sum_{i=1}^{N}\delta(\mathbf{r}-\mathbf{r}_i), \tag{4.112}$$

which again in terms of the fermionic operators can be expressed via

$$\mathbf{j}_d(\mathbf{q}) = -\frac{e^2}{m}\mathbf{A}_{\mathbf{q}}\sum_{\mathbf{k},\sigma}c^{\dagger}_{k\sigma}c_{k\sigma} = -\frac{ne^2}{m}\mathbf{A}_{\mathbf{q}}. \tag{4.113}$$

Using these definitions, the coupling to the vector potential is expressed as

$$\mathcal{H}' = -\frac{1}{\Omega}\int d\mathbf{r}\,\mathbf{j}_p(\mathbf{r})\cdot\mathbf{A}(\mathbf{r}) = -\frac{1}{\Omega}\sum_{\mathbf{q}}\mathbf{j}_p(-\mathbf{q})\cdot\mathbf{A}(\mathbf{q}), \tag{4.114}$$

where Ω denotes the volume and can be set to unity for convenience. It may be noted that the external probe $\mathbf{A}$ couples only to the paramagnetic current $\mathbf{j}_p$ in the Hamiltonian, $\mathcal{H}'$. $\mathbf{j}_d$ is already linear in $\mathbf{A}$. So it is only needed to evaluate $\mathbf{j}_p$ at the linear response level, and hence add to $\mathbf{j}_d$ to get the total current response, such that,

$$\langle\mathbf{j}\rangle = \langle\mathbf{j}_p\rangle + \langle\mathbf{j}_d\rangle. \tag{4.115}$$

According to linear response theory, one can write

$$\langle j_\alpha\rangle(\mathbf{q},\omega) = \chi_{\alpha\beta}(\mathbf{q},\omega)A_\beta(\mathbf{q},\omega) - \frac{e^2}{m}\left\langle\sum_{k\sigma}c^{\dagger}_{k\sigma}c_{k\sigma}\right\rangle A_\alpha(\mathbf{q},\omega), \tag{4.116}$$

where ω denotes the frequency and the current-current response function is given by

$$\chi_{\alpha\beta} \equiv \chi^{p}_{j_\alpha j_\beta}(\mathbf{q}, \omega)$$

with the density appearing in the second term,

$$n \equiv \left\langle \sum_{k\sigma} c^{\dagger}_{k\sigma} c_{k\sigma} \right\rangle.$$

To obtain the Meissner effect, we are interested in the zero momentum (long wavelength) and zero frequency (static) response of $\chi_{\alpha\beta}$, that is, for $\omega = 0$, $q \to 0$. One finally gets the current density,

$$\langle j_\alpha \rangle(\mathbf{q} \to 0, \omega = 0) = \chi_{\alpha\beta}(\mathbf{q} \to 0, \omega = 0) A_\beta(\mathbf{q} \to 0, \omega = 0) - \frac{ne^2}{m} A_\alpha(\mathbf{q} \to 0, \omega = 0). \tag{4.117}$$

To compute conductivity, we have to consider the finite frequency limit ($\omega \neq 0$).[5] Consider the electric field (the scalar potential is set to zero because of the gauge condition),

$$\mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t} = i\omega \mathbf{A}. \tag{4.118}$$

Thus,

$$\langle j_\alpha \rangle = \sigma_{\alpha\beta} E_\beta = i\omega \sigma_{\alpha\beta} A_\beta, \tag{4.119}$$

which subsequently yields,

$$\sigma_{\alpha\beta} = \frac{1}{i\omega} \left[\chi_{\alpha\beta}(q = 0, \omega) - \frac{ne^2}{m} \delta_{\alpha\beta} \right], \tag{4.120}$$

here,

$$\sigma_{\alpha\beta} = \mathrm{Re}\, \sigma_{\alpha\beta} + \mathrm{Im}\, \sigma_{\alpha\beta}. \tag{4.121}$$

Finally, at the linear response level, the Kubo formula for the conductivity tensor is written as

$$\sigma_{\alpha\beta}(\omega) = \frac{1}{i\omega} \left[\chi_{j_p \alpha j_p \beta}(q = 0, \omega) - \frac{ne^2}{m} \delta_{\alpha\beta} \right]. \tag{4.122}$$

4.5.1 Meissner effect

We demonstrate the expulsion of the magnetic field or the Meissner effect in the following.

$$\mathbf{j}_p = \frac{e\hbar}{m} \sum_k (\mathbf{k} + \mathbf{q}/2)(c^{\dagger}_{k\uparrow} c_{k+q\uparrow} + c^{\dagger}_{-k\downarrow} c_{-(k+q)\downarrow}). \tag{4.123}$$

The creation and the annihilation operators are expressed in terms of the quasiparticle operators via

$$\begin{aligned} c_{\mathbf{k}}\uparrow &= u_{\mathbf{k}} \gamma_{\mathbf{k}\uparrow} + v_{\mathbf{k}} \gamma^{\dagger}_{\mathbf{k}\downarrow} \\ c^{\dagger}_{-\mathbf{k}\downarrow} &= -v_{\mathbf{k}} \gamma_{\mathbf{k}\uparrow} + u_{\mathbf{k}} \gamma^{\dagger}_{\mathbf{k}\downarrow} \end{aligned} \tag{4.124}$$

[5] For the dc conductivity, we can take $\omega \to 0$.

where $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ are the coherence factors. These relationships are canonical, since they preserve the fermionic anticommutation relations. This allows the paramagnetic current to be written as

$$j_p = \frac{e\hbar}{m}\sum_k (\mathbf{k}+\mathbf{q}/2)[(uu'+vv')(\gamma^\dagger_{k\uparrow}\gamma_{k+q\uparrow} - \gamma^\dagger_{-(k+q)\downarrow}\gamma_{-k\downarrow}) + (uv'-vu')(\gamma^\dagger_{k\uparrow}\gamma^\dagger_{-(k+q)\downarrow} - \gamma_{-k\downarrow}\gamma_{k+q\uparrow})], \tag{4.125}$$

where the momentum indices are dropped by redefining,

$$u = u_k \quad v = v_k \; u' = u_{k+q} \; v' = v_{k+q}.$$

Now in the linear regime, the current-current response function is written as

$$\chi_{\alpha\beta}(q,\omega=0) = \sum_{m,n}\frac{e^{-\beta\omega_m}}{Z}\left\{\frac{((j^\alpha_p(q))_{m,n}(j^\beta_p(-q))_{n,m}}{\omega_{nm}} + c.c.\right\}. \tag{4.126}$$

One can import the contribution of j_p from Eq. (4.125). In particular, we look at different terms below.

i. $\gamma^\dagger\gamma$ term with $\sigma = \uparrow$

$$\sum_{k,k'} f(u,v)(k'+q/2)^\alpha (k-q/2)^\beta \langle m|\gamma^\dagger_{k'}\gamma_{k'+q}|n\rangle\langle n|\gamma^\dagger_k\gamma_{k-q}|m\rangle\frac{1}{\omega_{nm}},$$

where $f(u, v)$ is the coherence factor determined later. The only $|n\rangle$ that contributes is given by $|n\rangle = \gamma^\dagger_k\gamma_{k-q}|m\rangle$. Thus,

$$\mathbf{k}'+\mathbf{q} = \mathbf{k}, \quad \text{and}, \quad \omega_{nm} \equiv \omega_n - \omega_m = E_k - E_{k-q}.$$

This yields the above contribution to having the form,

$$\sum_k f(u,v)\frac{(k-q/2)^\alpha(k-q/2)^\beta}{E_k - E_{k-q}} = \sum_m \frac{e^{-\beta\omega_m}}{Z}\langle m|\gamma^\dagger_{k-q}\gamma_k\gamma^\dagger_k\gamma_{k-q}|m\rangle = \langle\gamma^\dagger_{k-q}\gamma_{k-q}(1-\gamma^\dagger_k\gamma_k)\rangle = f_{k-q}(1-f_k). \tag{4.127}$$

The above is the thermal average, we get the Fermi distribution function, f_k in the above expression. Finally, the coherence factor, $f(u, v)$ is

$$= (u_{k'}u_{k'+q} + v_{k'}v_{k'+q})(u_k u_{k-q} + v_k v_{k-q})$$
$$= (u_k u_{k-q} + v_k v_{k-q})^2, \quad \text{since} \quad \mathbf{k}' = \mathbf{k}-\mathbf{q}.$$

One may shift the dummy index, $\mathbf{k} \to \mathbf{k}+\mathbf{q}$ which makes the complex conjugate equal to the term calculated above. This yields the term $\gamma^\dagger\gamma$ for $\uparrow$-spin as

$$= \frac{2e^2\hbar^2}{m^2}\sum_k \frac{(uu'+vv')^2 f(1-f')}{E'-E}(k+q/2)^\alpha(k+q/2)^\beta.$$

ii. Combination of $\gamma^\dagger\gamma$ with $\sigma = \downarrow$
Same as (i) except for a negative sign in front of $\gamma_\downarrow^\dagger\gamma_\downarrow$ comes in $(-1)^2$ in χ and $\mathbf{k} \to \mathbf{k} + \mathbf{q}$ in u's, v's, fs and E's. This yields for the $\gamma^\dagger\gamma$ term corresponding to $\downarrow$-spin,

$$= \frac{2e^2\hbar^2}{m^2}\sum_k \frac{(uu' + vv')^2 f'(1-f)}{E - E'}(k + q/2)^\alpha (k + q/2)^\beta.$$

Subtracting (ii) from (i), that are contributions from spin-$\uparrow$ and spin-$\downarrow$,

$$= \frac{2e^2\hbar^2}{m^2}\sum_k (k + q/2)^\alpha (k + q/2)^\beta (uu' + vv')^2 \frac{(f - f')}{E - E'}.$$

iii. and (iv) The combination of the $\gamma^\dagger\gamma^\dagger$ and $\gamma\gamma$ terms
This comes with a coherence factor $(u_k v_{k+q} - v_k u_{k+q}) \to 0$ as $q \to 0$ so they will not contribute, however, we shall retain them for considering the finite q response. They can be written as

$$\sum_{m,n} \frac{e^{-\beta\omega_m}}{Z}\langle m|\gamma_k^\dagger\gamma_{-(k+q)}^\dagger|n\rangle\langle n|\gamma_{-(k+q)}\gamma_k|m\rangle \frac{(-1)}{\omega_{nm}} = \frac{(-1)\langle\gamma_k^\dagger\gamma_k\gamma_{-(k+q)}^\dagger\gamma_{-(k+q)}\rangle}{-E_k - E_{k+q}} = \frac{ff'}{E + E'}.$$

The negative sign, that is, -1 is the only difference between the $\gamma^\dagger\gamma$ and the $\gamma\gamma$ terms. The other one is

$$= \frac{(-1)\langle\gamma_{-(k+q)}\gamma_k\gamma_k^\dagger\gamma_{-(k+q)}^\dagger\rangle}{E_k + E_{k+q}} = \frac{(-1)(1-f)(1-f')}{E + E'}.$$

Here, the coherence factor is written as

$$(uv' - vu')(u'v - v'u) = -(uv' - vu')^2.$$

Adding contributions of (iii) and (iv)

$$= \frac{(1 - f + f')}{E + E'}(uv' - vu')^2 \times 2.$$

The current-current response function is written as

$$\chi_{\alpha\beta}(\mathbf{q}, \omega = 0) = \frac{2e^2\hbar^2}{m^2}\sum_{\mathbf{k}} (k + q/2)^\alpha (k + q/2)^\beta \left[(uu' + vv')^2 \frac{(f - f')}{E - E'} + (uv' - vu')^2 \frac{(1 - f - f')}{E + E'}\right]. \quad (4.128)$$

Now consider special cases of this general formula, that is, $\mathbf{q} \to 0$ response, which yields

$$(uv' - vu') \to 0 \quad (uu' + vv') \to 1.$$

In which case,

$$\chi_{\alpha\beta}(\mathbf{q} \to 0, \omega = 0) = \frac{2e^2\hbar^2}{m^2}\sum_{\mathbf{k}} k_\alpha k_\beta \left(-\frac{\partial f_k}{\partial E_k}\right). \quad (4.129)$$

The same result can be obtained in a more direct manner Tinkham (1973).

In the large wavelength limit, namely, $\mathbf{q} = 0$

$$\begin{aligned}\mathbf{j}_p &= \frac{e\hbar}{m}\sum_{\mathbf{k}}\mathbf{k}\left(\gamma^{\dagger}_{k\uparrow}\gamma_{k\uparrow} - \gamma^{\dagger}_{-k\downarrow}\gamma_{-k\downarrow}\right)\\ \langle\mathbf{j}_p\rangle &= \frac{e\hbar}{m}\sum_{\mathbf{k}}\mathbf{k}\left(f_{k\uparrow} - f_{-k\downarrow}\right).\end{aligned} \tag{4.130}$$

We must keep track of how the quasiparticle spectrum, $E_{k\uparrow}$ changes under the influence of the external Hamiltonian,

$$\mathcal{H}' = -\frac{1}{c}\sum_{\mathbf{q}}\mathbf{j}_1(-\mathbf{q})\cdot\mathbf{A}(\mathbf{q}) = -\frac{e\hbar}{mc}\sum_{\mathbf{k}}(\mathbf{k}\cdot\mathbf{A}_0)(\gamma^{\dagger}_{k\uparrow}\gamma^{\dagger}_{k\uparrow} - \gamma_{-k\downarrow}\gamma_{-k\downarrow}), \tag{4.131}$$

where the $q = 0$ term is retained in the last step of the equation. This yields an energy spectrum that is reminiscent of spin splitting in a magnetic field. The energies are modified as

$$E_{k\uparrow} \to E_{k\uparrow} - \frac{e\hbar}{mc}\,\mathbf{k}\cdot\mathbf{A}_0$$

Similarly,

$$E_{-k\downarrow} \to E_{-k\downarrow} + \frac{e\hbar}{mc}\,\mathbf{k}\cdot\mathbf{A}_0$$

where $\mathbf{A}_0$ is the vector potential for $q = 0$. Now,

$$\begin{aligned}f_{k\uparrow} - f_{-k\downarrow} &\approx \left(-\frac{\partial f_k}{\partial E_k}\right)\frac{2e\hbar}{mc}\,\mathbf{k}\cdot\mathbf{A}_0\\ \langle\mathbf{j}_p(\mathbf{q}=0)\rangle &= \frac{2e^2\hbar^2}{m^2c}\sum_{\mathbf{k}}(\mathbf{k}\cdot\mathbf{A}_0)\,\mathbf{k}\left(-\frac{\partial f_k}{\partial E_k}\right).\end{aligned} \tag{4.132}$$

Since, by symmetry $\langle\mathbf{j}_p(\mathbf{q}=0)$ is parallel to $\mathbf{A}_0$. Also,

$$\int\frac{d\Omega}{4\pi}(\mathbf{k}\cdot\hat{A}_0)^2 = \frac{k^2}{3}, \quad \text{as } \langle\cos^2\theta\rangle = 1/3.$$

Therefore, $\hat{A}_0$ is the unit vector along A_0,

$$\begin{aligned}\langle\mathbf{j}_p(0)\rangle &= \frac{2}{3}\frac{e^2\hbar^2}{m^2c}\sum_{\mathbf{k}}k^2\left(-\frac{\partial f}{\partial E_k}\right)\mathbf{A}_0\\ \text{Since, } \sum_{\mathbf{k}}k^2\left(-\frac{\partial f}{\partial E_k}\right) &\approx k_F^2\,N(0)\int d\xi_k\left(-\frac{\partial f}{\partial E_k}\right),\end{aligned} \tag{4.133}$$

where $N(0)$ is the density of states at the Fermi level. Moreover, the pre-factor of $\int d\xi_k$ is given by

$$= \frac{2}{3}\frac{e^2\hbar^2}{m^2}\,k_F^2\,\frac{3n}{4\varepsilon_F} = \frac{ne^2}{m}.$$

Hence, the average paramagnetic current is given by

$$\langle \mathbf{j}_p(q=0)\rangle = \frac{ne^2}{m}\int d\xi_k \left(-\frac{\partial f}{\partial E_k}\right)\mathbf{A}_0. \tag{4.134}$$

The normal state can be obtained by putting $\Delta = 0$ in the expression for quasiparticle energies, that are given by $E = \sqrt{\xi^2 + \Delta^2}$. This yield,

$$E = |\xi| \quad \text{Also,} \quad \int d\xi \left(-\frac{\partial f}{\partial \xi}\right) = 1.$$

Finally, one gets

$$\langle \mathbf{j}_p(q=0)\rangle = \frac{ne^2}{m}\mathbf{A}_0. \tag{4.135}$$

Hence, the paramagnetic response, $\langle \mathbf{j}_p\rangle$ and the diamagnetic response $\langle \mathbf{j}_d\rangle$ exactly cancel each other, thereby yielding the total induced current, $\langle \mathbf{j}\rangle = 0$. Consequently, no circulating currents are induced in a normal metal in equilibrium conditions, and, thus, the external magnetic field $\mathbf{H}$ is not shielded.

One can ignore the weak Landau diamagnetic effects, which arise from the fact that,

$$\langle \mathbf{j}_p(\mathbf{q}\to 0)\rangle = \left(\frac{ne^2}{m} - \alpha q^2\right)\mathbf{A}(q) \tag{4.136}$$

where one can write,

$$\mathbf{A}_{total} = \left(\frac{ne^2}{m} - \alpha q^2\right)\mathbf{A}(q).$$

The total current density now becomes

$$\langle \mathbf{j}(\mathbf{q}\to 0)\rangle = -\alpha q^2\,\mathbf{A}_{total} \qquad (\alpha > 0), \tag{4.137}$$

where α is a positive quantity, and $\mathbf{A}_{ind}$ obeys the wave equation,

$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right)\mathbf{A}_{ind} = \mu_0\langle \mathbf{j}\rangle. \tag{4.138}$$

In a static situation, the contribution of the term $\frac{1}{c^2}\frac{\partial^2}{\partial t^2} = 0$. Then,

$$q^2\mathbf{A}_{ind} = -\mu_0 \times q^2\,\mathbf{A}_{total}. \tag{4.139}$$

Using self-consistent fields,

$$\mathbf{A}_{total} = \mathbf{A}_{ind} + \mathbf{A}_{ext}$$

$$\text{or, } \mathbf{A}_{total} = \frac{\mathbf{A}_{ext}(\mathbf{q})}{1+\mu_0\alpha} \qquad (\alpha > 0). \tag{4.140}$$

Hence, the total vector potential derives a contribution from the external field and $\mathbf{A}_{ind}$. Taking curl to obtain the magnetic field,

$$|\mathbf{H}_{total}| = \frac{|\mathbf{H}_{ext}|}{1+\eta}, \tag{4.141}$$

where η is a positive quantity, such that $\mathbf{H}_{total} < \mathbf{H}_{ext}$ which means that the magnetic field is somewhat screened, and the system shows weak diamagnetism.

Let us now discuss the superconducting state, which is characterized by $\Delta \neq 0$ and the cancelation of the paramagnetic response $\langle \mathbf{j}_p \rangle$ and the diamagnetic response $\langle \mathbf{j}_d \rangle$ are not valid any longer. In fact, as we shall see, at $T = 0$, $\langle \mathbf{j}_p \rangle = 0$ and therefore, only the diamagnetic response $\langle \mathbf{j}_d \rangle$ contributes.

To this effect, let us define the ratio of the density of normal electrons to that of the total density, namely, $\frac{n_n}{n}$. This quantity will vanish in the superconducting state, or as $T \to T_c^-$, that is, when T approaches T_c below.

$$\begin{aligned} \frac{n_n}{n} &= \int_{-\infty}^{\infty} d\xi_k \left(-\frac{\partial f_k}{\partial \xi}\right) = 2\int_{\Delta}^{\infty} dE \, \frac{E}{\sqrt{E^2-\Delta^2}} \left(-\frac{\partial f}{\partial E}\right) \\ &= \frac{1}{2T}\int_0^{\infty} d\xi \, \mathrm{sech}^2\left(\frac{\sqrt{\xi^2+\Delta^2(T)}}{2T}\right) = Y(T), \end{aligned} \tag{4.142}$$

where $Y(T)$ is known as the Yosida function. The Yosida function characterizes the response of quantities to which other pairs do not contribute, and thus represents the response of the normal fluid. Let us explore the following limits, namely, $T \ll T_c$ and $T \to T_c^-$.

i. For $T \ll T_c$, $Y(T)$ behaves as $Y(T) \sim e^{-\Delta(0)/T}$.
ii. For $T \to T_c^-$, $Y(T) \approx 1 - 2\frac{(T_c - T)}{T_c}$.

The $q = 0$ components of the paramagnetic and the diamagnetic contributions are, respectively, written as

$$\langle \mathbf{j}_p \rangle_{q=0} = \frac{ne^2}{m}\left(\frac{n_n}{n}\right)\mathbf{A}_0 \tag{4.143}$$

$$\langle \mathbf{j}_d \rangle_{q=0} = -\frac{ne^2}{m}\mathbf{A}_0. \tag{4.144}$$

Thus,

$$\langle \mathbf{j} \rangle = \langle \mathbf{j}_p + \mathbf{j}_d \rangle = -\frac{ne^2}{m}\left(1 - \frac{n_n}{n}\right)\mathbf{A}_0 \equiv -\frac{n_s e^2}{m}\mathbf{A}_0, \tag{4.145}$$

where one can define the super-electron density, that is, $\frac{n_s}{n}$

$$\frac{\rho_s}{\rho} = 1 - \frac{\rho_n}{\rho}.$$

So,

$$\langle \mathbf{j} \rangle = -\frac{n_s e^2}{m}\mathbf{A}. \tag{4.146}$$

Thus, the super-electron density can be described in terms of the Yoshida function as

$$\frac{n_s(T)}{n} = 1 - Y(T).$$

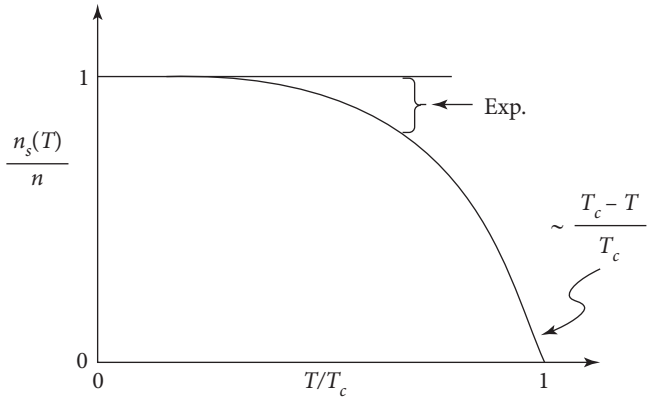


FIG. 4.22
The behaviour of the superfluid density scaled by the total density, $\frac{n_s(T)}{n}$ is shown as a function of reduced temperature, T/T_c. The behavior of the superfluid density close to $T = T_c$ is given by $(1 - \frac{T}{T_c})$.

Hence, the London penetration depth (or the inverse of it) is expressed as

$$\frac{1}{\lambda_L^2(T)} = \frac{4\pi n_s(T)e^2}{m} = \frac{1}{\lambda_L^2(0)}(1 - Y(T)). \tag{4.147}$$

Thus, the temperature dependence of the penetration depth is expressed by the temperature variation of the Yoshida function, $Y(T)$ (see Fig. 4.22).

4.5.2 Electromagnetic response in the transverse gauge

An externally applied vector potential, $\mathbf{A}_{ext}(\mathbf{r}, t)$, induces current in the system $\langle \mathbf{j}(\mathbf{r}, t)\rangle$, which in turn generates an electromagnetic field. Using Maxwell's equation, we have

$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right)\mathbf{A}_{\text{ind}}(\mathbf{r}, t) = \mu_0\langle \mathbf{j}(\mathbf{r}, t)\rangle \tag{4.148}$$

which follows from

$$\mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t} \text{ and } \nabla \times (\nabla \times \mathbf{A}) = -\nabla^2\mathbf{A}$$

in a gauge where $\nabla \cdot \mathbf{A} = 0$. Therefore, Fourier transforming one obtains

$$\left(q^2 - \frac{\omega^2}{c^2}\right)\mathbf{A}_{\text{ind}}(\mathbf{q}, \omega) = \mu_0\langle \mathbf{j}(\mathbf{q}, \omega)\rangle. \tag{4.149}$$

The total vector potential is written as

$$\mathbf{A}_{\text{tot}} = \mathbf{A}_{\text{ext}} + \mathbf{A}_{\text{ind}}$$

and for the current density,

$$\langle \mathbf{j}(\mathbf{q},\omega)\rangle = \left(\chi^T - \frac{ne^2}{m}\right)\mathbf{A}_{\text{tot}} = -\frac{n_s e^2}{m}\mathbf{A}_{\text{ext}} = -\frac{1}{\lambda_L^2}\mathbf{A}_{\text{tot}}. \tag{4.150}$$

Now,

$$\begin{aligned} &\mathbf{B} = \mu_0(\mathbf{H} + 4\pi\mathbf{M}) \\ &\text{also } \mathbf{B} = \nabla \times \mathbf{A}_{\text{tot}} \\ &\text{and } \mathbf{H} = \nabla \times \mathbf{A}_{\text{ext}} \text{ that is, } \quad \mathbf{M} = \frac{1}{\mu_0}\nabla \times \mathbf{A}_{\text{ind}}. \end{aligned} \tag{4.151}$$

In the static limit, that is, $\omega = 0$,

$$\begin{aligned} &\mathbf{A}_{\text{ind}} = \frac{4\pi}{c}\frac{1}{q^2}\langle \mathbf{j}(\mathbf{q},\omega)\rangle \\ &\text{or, } \mathbf{A}_{\text{ind}} = \frac{-1/\lambda_L^2(T)}{q^2 + 1/\lambda_L^2(T)}\mathbf{A}_{\text{ext}}. \end{aligned} \tag{4.152}$$

Thus, the susceptibility can be written as

$$\chi = \frac{\partial M}{\partial H} = \frac{\frac{1}{\lambda_L^2(T)}}{q^2 + \frac{1}{\lambda_L^2(T)}} \tag{4.153}$$

that is, $\lim_{q\to 0}\chi = -1$.

The above equation implies the conditions for perfect diamagnetism. Superconductors are perfect diamagnets. It may be noted that the diamagnetic susceptibility of metals is usually of the order of 10^{-5}–10^{-6}.

4.6 GINZBURG-LANDAU (GL) THEORY

In 1937 Ginzburg and Landau (1950) developed a model to describe second-order phase transition,[6] that is, those transitions which involve latent heat. It is physically meaningful to state that the quantum many body states, ψ (something similar to what we have seen in BCS theory) should be a function that minimizes the free energy, F_s of the superconducting state. Ginzburg and Landau considered F_s to be a functional of the wavefunction ψ, and hence use a variational principle to minimize F_s with respect to ψ, and obtain a set of differential equations for ψ that govern the physical properties of the superconductor in equilibrium conditions. To this end, a concept of the order parameter has been developed, which smoothly vanishes at a temperature $T > T_c$, while it is non-zero for $T < T_c$. The identification of the order parameter is often obvious from the properties of second-order phase transitions. For example, for a ferromagnetic to a paramagnetic transition, the order parameter is the magnetization,

[6] It may be mentioned here that we have introduced the Ginzburg Landau theory very briefly. In no way should it be considered as complete. For a detailed discussion, see *Theory of type-II superconductors* in Fetter and Hohenberg (1969).

which is finite below the magnetic transition temperature, T_c and goes to zero above T_c. Whereas, a superconducting state has an energy gap that separates the ground state and the first excited state, as the order parameter assumes non-zero values below the superconducting transition temperature, T_c which, however, vanishes above T_c.

Since the order parameter evolves continuously from zero below T_c, it is natural to expand the free energy as a power series in the order parameter. The free energy is a scalar, but the order parameter, may in general be a vector, or even a complex quantity. We shall study GL theory by constructing the free energy for the superconductor in terms of the wavefunction for the "*super-electrons*" as the order parameter (instead of the energy gap between the ground and the excited states), $\psi(\mathbf{r})$, the density of the super-electrons is $n_s = |\psi(\mathbf{r})|^2$.

Consider a homogeneous superconductor at zero external magnetic field, that is, $\mathbf{H} = 0$. Also, $\psi(\mathbf{r})$ is independent of $\mathbf{r}$. Let us expand the free energy near the transition from a normal metal to a superconductor, that is, near T_c,

$$F_s = F_n + \alpha|\psi|^2 + \frac{\beta}{2}|\psi|^4, \tag{4.154}$$

where F_s denotes the free energy density of the superconductor, and F_n is the free energy density of the normal state. Furthermore, α, β are the phenomenological expansion coefficients characterizing the material, and they depend on temperature. The minimum of F_s is obtained by minimizing F with respect to the order parameter,

$$\frac{dF_s}{d|\psi|^2} = 0. \tag{4.155}$$

The minimization implies that (writing the equilibrium value of ψ and ψ_0),

$$|\psi_0|^2 = -\alpha/2\beta.$$

Recall that $|\psi_0|^2$ is the density of the super-electrons that minimizes F_s. Now,

$$\frac{dF_s}{d\psi} = 2\alpha\psi + 2\beta\psi^3 = 2\psi[\alpha + \beta\psi^2]. \tag{4.156}$$

Thus, the difference in energies between the normal and the superconducting states is obtained from

$$F_n - F_s = \alpha\left(\frac{\alpha}{\beta}\right) - \frac{\beta}{2}\left(\frac{\alpha}{\beta}\right)^2 = \frac{\alpha^2}{\beta} - \frac{\alpha^2}{2\beta} = \frac{\alpha^2}{2\beta}. \tag{4.157}$$

The temperature dependence of α and β can be deciphered as follows. At the first order in temperature, that is, when T is close to T_c, let us postulate a linear behavior of the form,

$$\alpha \simeq a(T - T_c) \tag{4.158}$$

where a is independent of the temperature. This implies

$$\begin{aligned} &\alpha = 0 \text{ at } (T = T_c) \\ &\alpha < 0 \text{ at } (T < T_c). \end{aligned} \tag{4.159}$$

Thus, the order parameter assumes a form,

$$|\psi| = 0 \text{ for } T > T_c$$
$$|\psi| = \left[\frac{a(T - T_c)}{\beta}\right]^{1/2} \text{ for } T < T_c \tag{4.160}$$

α changes sign across the phase transition, and hence β should not change sign and remains constant, at least for small deviations in temperature from the transition point.

Next, consider a superconductor in an external magnetic field, $\mathbf{H_{ext}}$. The electrons interact with the external field via their momenta transforming as

$$\begin{aligned} &\mathbf{p} \to \mathbf{p} - e\mathbf{A} \text{ where } \mathbf{B}_{ext} = \nabla \times \mathbf{A} \\ \text{which implies,} \quad &\nabla \to \nabla - ie\mathbf{A}. \end{aligned} \tag{4.161}$$

Thus, the free energies have to be written as

$$F_s = F_n + \alpha|\psi|^2 + \frac{1}{2}\beta|\psi|^4 + \frac{\hbar^2}{2m}\left|\left[\nabla - \frac{ie}{\hbar c}A(\mathbf{r})\right]\psi(\mathbf{r})\right|^2 + \mu_0 H_{ext}^2(\mathbf{r}) \tag{4.162}$$

where $\mathbf{H}_{ext}(\mathbf{r}) = \nabla \times \mathbf{A}(\mathbf{r})$. The free energy is a functional of

$$\{\psi(\mathbf{r}), \psi^*(\mathbf{r}), \nabla\psi(\mathbf{r}), \nabla\psi^*(\mathbf{r}), H(\mathbf{r})\}$$

and each one of these is a function of the spatial variable $\mathbf{r}$. Minimizing with respect to $\psi^*(\mathbf{r})$ yields

$$\begin{aligned} \delta F_s = &\left\{-\frac{\hbar^2}{2m}\left(\nabla - \frac{ie}{\hbar}A(\mathbf{r})\right)^2 \psi(\mathbf{r}) + \alpha\psi(\mathbf{r}) + \beta|\psi(\mathbf{r})|^2\psi(\mathbf{r})\right\}\delta\psi^*(\mathbf{r}) \\ &+ \frac{\hbar^2}{2m}\left(\nabla - \frac{ie}{\hbar}A(\mathbf{r})\right)\psi(\mathbf{r})\delta\psi^*(\mathbf{r}). \end{aligned} \tag{4.163}$$

The last term can be written as $\mu_0 H^2$, or $\mu_0(\nabla \times \mathbf{A})^2$. Furthermore, variation with respect to ψ yields the complex conjugate of this equation. Putting $\delta F_s = 0$ for an arbitrary variation of the order parameter, namely, $\delta\psi^*(\mathbf{r})$, we get the first GL equation,

$$-\frac{\hbar^2}{2m}\left(\nabla - \frac{ie}{\hbar}A(\mathbf{r})\right)^2 \psi(\mathbf{r}) + \alpha\psi(\mathbf{r}) + \beta|\psi(\mathbf{r})|^2\psi(\mathbf{r}) = 0 \tag{4.164}$$

where we have neglected the magnetic energy density due to the external field, that is, the last term. Now minimize it with respect to $\mathbf{A}(\mathbf{r})$ yields Ampere's law,

$$\nabla \times \mathbf{H}(\mathbf{r}) = \mu_0 \mathbf{j}(\mathbf{r}) \tag{4.165}$$

where,

$$\mathbf{j}(\mathbf{r}) = -\frac{ie\hbar}{2m}\left[\psi^*(\mathbf{r})\nabla\psi(\mathbf{r}) - \psi(\mathbf{r})\nabla\psi^*(\mathbf{r})\right] - \frac{e^2}{m}|\psi|^2 A(\mathbf{r}). \tag{4.166}$$

This is the second GL equation. These two GL equations in Eqs. (4.164) and (4.166) are the main triumphs of GL theory. They enunciate the variation of the order parameter, $\psi(\mathbf{r})$, and the current density, $\mathbf{j}(\mathbf{r})$.

4.6.1 Coherence length and the Penetration depth

Using the GL equations obtained above, we wish to derive expressions for some physical observables that are measured in experiments, such as the coherence length and the penetration depth. Let us first talk about the coherence length.

Consider an inhomogeneous order parameter for a system generated by the presence of the boundary in the absence of an external magnetic field.[7] The superconducting (SC) region is denoted by $x > 0$. Thus, the order parameter is zero for the interface. For deriving expression for the coherence length, putting $\mathbf{A} = 0$ in the first GL equation Eq. (4.164),

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + \alpha|\psi| + \beta\psi^3 = 0. \tag{4.167}$$

It may be noted that α is negative in the superconducting state, which allows us to set, $\alpha = -|\alpha|$. Now defining the coherence length as

$$\xi^2 = \frac{\hbar^2}{2m|\alpha|}. \tag{4.168}$$

Also writing $(\beta/|\alpha|)\psi^2 = f^2$, we can write Eq. (4.167) as

$$-\xi^2 f'' - f + f^3 = 0. \tag{4.169}$$

Multiply both sides by f' to get

$$\frac{d}{dx}\left[-\frac{\xi^2 f'^2}{2} - \frac{1}{2}f^2 + \frac{1}{4}f^4\right] = 0 \tag{4.170}$$

or, $-\xi^2\frac{f'^2}{2} - \frac{1}{2}f^2 + \frac{1}{4}f^4 = \text{constant}.$

Far from the boundary, f' (or ψ') should be zero, and $f^2 = 1$. Hence, $|\psi|^2 = \frac{|\alpha|}{\beta}$. Equation (4.170) now becomes

$$\xi^2(f')^2 = \frac{1}{2}(1 - f^2)^2. \tag{4.171}$$

The solution of the above equation is given by

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

Hence,

$$\psi = \left(\frac{|\alpha|}{\beta}\right)^{1/2}\tanh\left(\frac{x}{\sqrt{2}}\right)$$

This yields that ξ is the measure of the distance over which the order parameter responds to a perturbation. Also, since $\alpha = a(T - T_c)$, the temperature dependence of the coherence length can be

[7] To derive the coherence length, presence of magnetic field is not essential

obtained as

$$\xi(T) = \frac{\hbar^2}{2maT_c}\left(1 - \frac{T}{T_c}\right)^{-1/2} = \xi_0\left(1 - \frac{T}{T_c}\right)^{-1/2}. \tag{4.172}$$

Thus, $\xi(T)$ diverges as $(T - T_c)^{-1/2}$ as $T \to T_c$ below. Also, the temperature independent length, ξ_0 is given by

$$\xi_0 = \frac{\hbar^2}{2maT_c}.$$

A little manipulation of the above formula yields,

$$\xi_0^2 = \frac{\hbar^2}{2m\epsilon_F}\left(\frac{\epsilon_F}{k_B T_c}\right)^2.$$

Thus ξ_0 is larger than $(\frac{\hbar^2}{2m\epsilon_F})^{1/2}$, which is of the order of interparticle spacing, by a factor of $(\epsilon_F/k_B T_c)^2$. Remember, ϵ_F is usually of the order of a few eV (typically, 5–6 eV), and $k_B T_c$ for superconductors is of the order of a few meV. Thus, ξ_0 is usually of the order of a few thousand lattice spacings.

Next, in order to arrive at an expression for the penetration depth, we examine the current expression from the second GL equation, that is Eq. (4.166). If we neglect the first term in comparison to the second term, then one obtains the London equation (discussed earlier), namely,

$$\mathbf{j}(\mathbf{r}) = \mu_0 \frac{1}{\lambda_L^2}\mathbf{A}(\mathbf{r}), \tag{4.173}$$

where,

$$\lambda_L^2 = \frac{mc^2}{4\pi e^2 |\psi|^2}$$

Again, the temperature dependence of the penetration depth is obtained as

$$\lambda_L = \left(\frac{mc^2\beta}{4\pi e^2 a T_c}\right)\left(1 - \frac{T}{T_c}\right)^{-1/2} \tag{4.174}$$

λ_L has a similar divergence as that of $\xi(T)$. Also, a ratio of the two can be found as

$$\kappa = \lambda_L/\xi = \frac{mc}{e\hbar}\left(\frac{\beta}{2\pi}\right)^{1/2}. \tag{4.175}$$

As discussed earlier, the value of κ demarcates between various types of superconductors, namely, the type-I and type-II superconductors, which have very different magnetic properties.

4.7 COOPER PAIRS WITH FINITE MOMENTUM

In BCS theory, we have seen that the pairing occurs between electrons of opposite spins and equal and opposite momentum, namely, $(k\uparrow, -k\downarrow)$, such that it results in the total momentum to is zero for the

pair. The spectral gap of the order parameters is uniform as well. However, there are other superconductors which can accommodate pairs having non-zero total momentum, and a non-uniform order parameter. Such superconducting properties were predicted independently by Fulde and Ferrell (1964) and Larkin and Ovchinnikov (1964) and were shown to arise in the presence of large magnetic fields. These superconducting states are called Fulde-Ferrell-Larkin-Ovchinnikov (FFLO) states, where the superconductivity is destroyed by "*Pauli pair breaking*" mechanism, rather than by orbital pair breaking that occurs in BCS superconductors. Quasi-2D heavy fermion superconductors, such as $CeCoIn_5$ etc are potential candidates for the formation of FFLO states.

One basic way where such finite momentum pairing can be realized is when the populations of the participating species are unequal. A possible way of creating an imbalance in the densities of spin-↑ and spin-↓ electrons is using a magnetic field which makes the Fermi surfaces (corresponding to ↑ and ↓-spin states) shift away with respect to each other. However, such population imbalances can be detrimental to pairing. This can be understood as follows. The breakdown of superconductivity by such population imbalance of the participating species can be realized by the fact that deep in the Fermi sea, particles cannot form a pair or even scatter off one another because their motion is "frozen" by the exclusion principle. Pairing requires partially empty energy states that can only be found near the Fermi surface. But if the populations of the two different spin components are different, the two Fermi surfaces no longer match up in momentum space, that is, there are no partially occupied states in which both species have opposite momenta and can form a zero-momentum pair. This makes the pairing energetically less favorable and eventually causes superfluidity to break down. However, there is one way that superconductivity can stabilize, which will happen when the pairing between the Zeeman splitted parts of the Fermi surface give rise to an unconventional pairing state, namely, $(k\uparrow, -k + q\downarrow)$, comprising pairs having a finite center of mass momentum, q.

The origin of such unconventional pairing phenomena goes back to as early as the 1960*s* when Maki and Tsuento (1964) motivated by the unusual dependence of the critical temperature as a function of the magnetic field along with the existence of the Clogston-Chandrasekhar limit (Chandrasekhar, 1962; and Clogston, 1962), which showed that the critical field in a certain class of superconductors was set by Pauli paramagnetism. The inference drawn from the above studies was that the second-order phase transition which exists between superconducting and normal phases becomes a first-order transition at a critical value of temperature and magnetic field, making the critical value a tri-critical point.

4.7.1 The FFLO phase

The study of superconductivity in the presence of a magnetic field commenced nearly half a century ago with the works of Clogston (1962) and Chandrasekhar (1962). In their work, they had considered pair breaking only by the Paulli paramagnetic effect and the orbital component was assumed to be negligibly small. Later, it was observed that the uniform state with population imbalance is unstable as the magnetic field is increased (until $H = \frac{\Delta_0}{\sqrt{2}}$, where H is the magnetic field and Δ_0 is the superconducting order parameter) at zero temperature which was confirmed by comparing the free energies of

the normal and superconducting phases. Thus, there is an indication of phase separation of the two spin species at $H > \frac{\Delta_0}{\sqrt{2}}$. The subject was intermittently revived to discuss the bounds on the upper critical field and its effect on the phase boundary where the paramagnetic effect governs the physics. More abounding implications of the presence of an external magnetic field are elucidated by Fulde and Ferrell (1964) and by Larkin and Ovchinnikov (1964) where possibilities of a finite momentum pairing between the different participation species of electrons are explored.

The relevant discussion on the effect of the magnetic field on superconductivity is as follows. The Fermi surfaces of the spin-↑ and ↓ electrons split when an external magnetic field is turned on, resulting in an imbalance between the two electron species. Under such circumstances, along with the conventional BCS state, the spin-polarized state (normal state) and possibly more states compete for the ground state. The imbalance of the two electron species in turn leads to the breaking of a fraction of the Cooper pairs. If the number of broken pairs is small, then the energy gap, Δ_0 is not affected. It may be noted that energy of $2H$ (where H is the strength of the magnetic field) is gained from the new spin orientation at a cost of $2\Delta_0$ in breaking a Cooper pair, while attaining a spin-polarized state. A continuous increase in the number of broken Cooper pairs demand H to be greater than Δ_0 in order to make the spin polarized state eventually energetically more favored over the BCS state. But the spin-polarized state has a lower energy when $H = 1/\sqrt{2}\Delta_0$ which is known as the "Pauli limit"Clogston (1962) and Chandrasekhar (1962). Thus, the BCS state with a large number of broken Cooper pairs becomes unstable, and an alternative solution exists for the ground state. This unusual phase is known as the FFLO phase which is stabilized by a large Zeeman splitting between the ↑ and the ↓-spin electrons which form pairs across the Fermi surface and subsequently condense to give a state which has lower free energy than the normal spin-polarized state. These pairs have non-zero total momentum ($2q = k_\uparrow - k_\downarrow$) as opposed to the traditional zero momentum pairs in the conventional BCS state (refer to Fig. 4.23) since the paired electrons have different momenta ($k_\uparrow$ and $k_\downarrow$). Interestingly, finite momentu m pairing results in a spatially oscillating superconducting order parameter with a

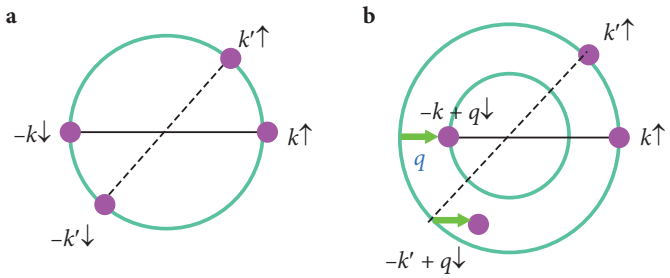


FIG. 4.23
Schematic view of pairing states. (a) BCS pairing state, (b) FFLO pairing state. The inner and outer circles represent the Fermi surface of the spin down and up bounds respectively (Matsuda and Shimahara, 2007).

wavelength of the order of superconducting coherence length, ξ. The FFLO phase thus breaks spontaneously translational symmetry and resembles the unconventional superconductivity in strongly correlated systems, where the order parameter changes sign in the momentum space.

The orbital (diamagnetic) pair breaking is a crucial mechanism that limits the realization of the FFLO state. It is a dominating pair breaking mechanism that destroys superconductivity for magnetic fields much weaker than the Clogston-Chandrasekhar limit (H_c). The significance of the diamagnetic pair breaking is usually described in terms of the Maki parameter $\alpha = \sqrt{2}H_{c2}^{orb}/H_c$, where $H_{c_2}^{orb}$ is the upper critical field calculated without the Zeeman splitting. Thus, the paramagnetic effect (due to the exchange field) responsible for FFLO has to dominate over a competing orbital effect which annihilates superconductivity due to the formation of screen currents arising from vortices. The above requirement demands few stringent conditions which must be satisfied for the realization of the FFLO phase, such as ultra-clean type II superconductor, such that the pairing correlations survive at a finite magnetic field and the electronic mean free path, $l \gg \xi$ where ξ is the superconducting coherence length.

In spite of a robust theoretical possibility for the existence of FFLO, the experimental scenario remained bleak mainly due to the lack of phase space available for pairing to occur. Thus, FFLO occupies a very small space in the phase diagram. There exist two general possibilities to reduce the destructive role of the orbital pair breaking. In the layered superconductors, the formation of Landau orbits should be suppressed with magnetic fields applied parallel to the layers. This may explain the possible observations of the FFLO state in some organic superconductors. The role of the orbital pair breaking should also be limited to systems with narrow energy bands. It is worthwhile to mention at this stage that the FFLO phase is possible in the case of condensation (mass polarized) quarks (color superconductivity). Astrophysical objects, for example, neutron stars or pulsars, can be a good candidate for realizing such a phenomenon. A good introduction to the subject is obtained in reviews by Casalbouni and Narduli (2004) and by Buzdin (2005).

4.7.2 Experimental signatures of the FFLO phase

The experimental studies of FFLO in heavy fermion compounds, for example, Ce and U based materials, started in the early nineties as some of the prerequisite conditions for observing FFLO is met in these systems. Notable of them is $CeRu_2$ which is usually obtained in a metallurgically clean state. Among the heavy Fermion systems, the most explored one is $CeCoIn_5$ which receives strong support as a candidate with an FFLO phase. These are essentially quasi two-dimensional structures and the magnetic field is applied parallel to the *ab* plane. The experiments on $CeCoIn_5$, have shown the heat capacity to undergo two phase transitions, a second order one within the superconducting (SC) state at low field values and a higher field first-order transition at H_{c_2} (Radovan *et al.*, 2003), the intervening regime acquiring a nonuniform nature. With the external magnetic field acquiring an angle with the *ab*-plane, the orbital effect starts playing a role and the large field transition goes away, rendering an absence of the nonuniform or the FFLO state. The enhancement of the penetration depth (as a function of magnetic field) (Martin *et al.*, 2005), anomalous thermal and magneto-thermal conductivity

(a discontinuous jump) in the vicinity of the upper critical field (Capan *et al.*, 2004) and a structural transformation with vortices appearing in the flux line lattice observed via ultrasound measurements (Watanabe *et al.*, 2004) have provided reasonable, though not robust, support for $CeCoIn_5$ in hosting a FFLO phase.

Other experiments, such as anisotropic magneto-thermal measurements (Capan *et al.*, 2004) ultrasound velocity measurements (Watanabe *et al.*, 2004) etc. provide only indirect and cursory evidence supporting the presence of the FFLO state.

More recently, ^{115}In NMR studies on $CeCoIn_5$ with the applied field parallel to *ab* plane revealed dramatic asymmetry in the NMR spectrum for a field greater than the upper critical field when compared with the one less than that (Kakuyanagi *et al.*, 2005). Furthermore, an unusual temperature dependence of the knight shift of ^{115}In is noted in the latter case. These facts are correlated with direct evidence of the FFLO phase and a simulation of the NMR spectrum with a spatially modulated gap function seems to satisfactorily explain the experimental findings (Kumagai *et al.*, 2006). These results were challenged in the light of other NMR studies (Mitrović *et al.*, 2006).

The main factors that act in favor of organic superconductors as possible candidates for observation FFLO is that their layered structures (with the magnetic field parallel to the layers) thereby making the orbital H_{c_2} extremely high and hence, the Pauli paramagnetic effect becomes supremely important. The reduced dimensionality or the planar structure of these materials is conjectured to be a facilitator for observing the FFLO phase. Recent critical field measurements (Nam *et al.*, 1999) in the quasi-two-dimensional organic superconductor $\kappa - (BEDT\text{-}TTF)_2Cu(NCS)_2$ strongly suggest that a state of the FFLO type exists in this material. The agreement between this experiment and existing theories has been successfully verified (Manalo and Klein, 2000) both in view of both the angle-dependence (Shimahara and Rainer, 1997) and the temperature dependence Singleton *et al.* (2000) of the upper critical field. In Manalo and Klein (2000), a comparison between the experimental (Nam *et al.*, 1999) temperature dependence of the plane parallel upper critical field with the theoretical results for $\kappa - (BEDT\text{-}TTF)_2Cu(NCS)_2$ has been reported. This is the first time since the original predictions of the FFLO phase that quantitative agreement between theory and experiment with regard to the FFLO phase boundary has been established.

An increase in the critical field at low temperatures with positive d^2H_c/dT^2 was observed in $(TMTSF)_2PF_6$ (Lee *et al.*, 1997). Such behavior of the critical field behavior is very similar to the behavior that is theoretically obtained in low-dimensional FFLO superconductors (Shimahara, 1994; and Dupuis, 1995). This might suggest the possibility of the FFLO state in this material or similar organic compounds, although this is not conclusive evidence for the existence of the FFLO phase.

4.7.3 Theoretical developments

A theoretical analysis of the FFLO state requires, in general, a self-consistent calculation of the amplitude of the spatially varying order parameter. Such calculations are within the reach of present

computation capacities, but were not possible in the early years of discovery of the FFLO phase. Early theoretical works on this problem were mostly centered around studying the region near the second-order phase transition to the normal state or studied the Fulde-Ferrell state characterized by a spatially modulating order parameter (Takada and Izuyama, 1969),

$$\Delta(r) = \Delta_0 \, e^{i\mathbf{q}\cdot\mathbf{r}} \tag{4.176}$$

Although analytical solutions to both the problems exist, they fail to provide a complete description of the spin imbalanced inhomogeneous phase. For example, the Fulde-Ferrell (FF) state is not the correct minimum energy state, as shown by Larkin and Ovchinnikov (LO) who considered the order parameter to be of the form $\Delta = \Delta_0 \cos(\mathbf{q} \cdot \mathbf{r})$. It may be noted that the amplitude of the order parameter is no longer a constant in real space in the LO phase. Various studies have been performed with decreasing magnetic field which have shown that the LO phase evolves into a state with a set of domain walls. This has been shown in one (Machida and Nakanishi, 1984), two (Burkhardt and Rainer, 1994), three (Matsuo *et al.*, 1998) dimensions, and also for *d*-wave superconductors (Vorontsov *et al.*, 2005). In the limit of small population imbalance, the order parameter is constant in real space with the value being equal to that in the state with no population difference, excepting near the domain walls. These are the locations where the magnetization concentrates and the phase of the order parameter changes by π when the walls are crossed. The physics of these domain walls is closely related to that of the π junctions in superconductor-ferromagnet-superconductor junctions (Buzdin, 2005).

A detailed study of the effect of dimensionality on the exotic FFLO phase was also conducted. In one dimension, it was argued that the ground state of the homogeneous attractive Fermi gases with unequal spin populations is the one-dimensional analog of the FFLO phase (Yang, 2001). A similar study of the scenario in two-dimensional systems has yielded many intriguing results (Zhu *et al.*, 1995; and Pieri *et al.*, 2007). In particular, a quasi-classical analysis in Combescot and Mora (2005) using a Ginzburg-Landau expansion of the free energy in the Fourier components of the superconducting order parameter, had shown that the FFLO transition in two dimensions is continuous at low temperatures.

Various studies have shown that the system must be in the ultra-clean limit where the quasiparticle mean free path l is much longer than the coherence length ξ for the realization of the FFLO phase. This, in turn, demands that the Ginzburg Landau parameter, κ must be much larger than unity. These conditions are readily met in *d*-wave superconductors like high-T_c superconductors and organic superconductors e.g. $\kappa - (ET)_2$ salts and $\lambda - (ET)_2$ salts (Nam *et al.*, 1999). Thus, the discovery of these new classes of superconductors has triggered an extensive search of the FFLO state in *d*-wave superconductors (Vorontsov and Graf, 2006). In this regard, some of the theoretical studies have predicted that the stability of the FFLO state in two dimensional *d*-wave superconductors is much more than in three-dimensional *s*-wave superconductors (Shimahara, 1998; and Yang and Sondhi, 1998). Some other studies have been performed to investigate the effect of disorder on the FFLO state in *d*-wave superconductors. For example, Vorontsov and coworkers (Vrontosov *et al.*, 2008), have investigated

the stability of the FFLO state in dirty *d*-wave superconductors using the quasi-classical theory based on the self-consistent *t*-matrix approximation for impurities. Their results are counter-intuitive as it shows the FFLO state to be robust in "dirty" *d*-wave superconductors. In a separate study by Yanase (2009), the spatial structure of the superconducting order parameter and the magnetic properties in the disordered FFLO state are investigated using the Bogoliubov de Gennes method. The results show that the nature of the superconducting order parameter is strongly dependent on the kind of disorder taken into consideration.

The modulating superconducting order parameter, which yields direct evidence of the presence of the FFLO phase, is very difficult to observe in experiments. In this context, the quasiparticle local density of states (LDOS) is very useful as it can be directly detected using scanning tunneling microscopy (STM). The LDOS for a two dimensional *d*-wave superconductor has been computed by Vorontsov *et al.* (2005) by solving quasi-classical Eilenberger equations. Wang *et al.* (2006) have also investigated LDOS for both *s* and *d*-wave superconductors, along with the superconducting order parameter which is found to be stripe-like for *s*-wave superconductors and a square lattice for *d*-wave superconductors.

Most of the theoretical studies of the FFLO state are based on the mean field approach, which is counterintuitive as one expects the quantum and thermal fluctuation effects to play a significant role in the FFLO phase than in conventional BCS superconductors, as the FFLO phase breaks the translational symmetry. In this regard, Yang (2001) has studied the FFLO phase in quasi-one dimensional superconductors using bosonization for an exact treatment of the intra-chain quantum fluctuations and has shown that the transition from the FFLO phase to BCS is a continuous transition of the commensurate-incommensurate type. The effect of thermal fluctuations has been discussed by Matsuda and Shimahara (2007) which reveals that the mean field FFLO state is destabilized by the enhanced fluctuation effects. In all these studies for treating phase fluctuations for an FFLO superconductor, the most common method is to incorporate a fluctuation in the phase of the order parameter and hence construct a generalized Ginzburg-Landau (GL) action for the phase variable (Shimahara, 1999; and Dalidovich and Yang, 2004). Spatial correlations of the phase variable indicate a rapid suppression of the long-range order (LRO). Furthermore, the fluctuation-driven transition from normal to FFLO is found to be first order that corroborates the experimental evidence in CeCoIn5, however contradicts the mean field results (Dalidovich and Yang, 2004).

The experimental realization of the phase has been illusive mainly because its occurrence requires several stringent conditions to be met simultaneously. In most type-II superconductors, destruction of superconductivity occurs through the orbital pair breaking effect, leading to the emergence of the vortex state. However, such an orbital effect is always detrimental to the FFLO state. Hence, for the FFLO state to occur, the orbital pair breaking effect must be weak relative to the Pauli paramagnetic effect. Moreover, the system needs to be clean, since the FFLO state is readily destroyed by impurities. The cleanliness condition is met when the coherence length, ξ is very small as compared to the mean free path, λ. Thus, the superconducting materials fulfilling these necessary conditions are few. Some

of these conditions are satisfied in heavy fermion superconductors with large orbital-limiting fields (Gloos *et al.*, 1993; and Huxley *et al.*, 1993).

An alternative route to achieve the supremacy of paramagnetic effect is to use a layered structure in a strong magnetic field applied parallel to the layers, thereby undermining the orbital pair breaking effect further and augmenting the parameter space where FFLO can exist (Barzykin and Gorkov, 2002). The organic superconductors strongly fit into these requirements and hence are considered as ideal candidates for the FFLO phase (Shimahara, 1994). Apart from these compounds, signatures of FFLO phases are also observed in neutron stars (Alford *et al.*, 2000).

4.8 EXPERIMENTAL DETERMINATION OF THE ENERGY GAP

The existence of a forbidden energy regime in superconductors has given us a simple way to understand the fact that below certain excitation energies, the bound state of the electrons is robust and they cannot dissociate into single particles. Thus, the current in the superconducting state is mediated via the cooper pairs (also referred to as the *super-electrons*), and the normal current carried by the electrons is zero. More precisely, there are no unpaired electrons in the superconducting state. Here we describe a few of the early experimental methods below to determine the energy gap. We shall mainly focus on three methods that have been widely used in obtaining the magnitude and the pairing symmetry of gaps in the early days of the discovery of superconductors. They are

i. Absorption of electromagnetic radiation.
ii. Ultrasound absorption.
iii. Tunneling experiments.

4.8.1 Absorption of electromagnetic radiation

This technique of exploring the energy gap was reported in 1957 by Tinkham and Glover (1958). They observed infrared transmission in thin-superconducting films. It was earlier proposed that electromagnetic (EM) waves of appropriate frequencies can break the ordered state of a superconductor. Assuming a cooper pair has a binding energy given by $E_B \sim 10^{-3}$ eV, one needs a radiation frequency $\nu \sim E_B/h = 2.4 \times 10^{11}$ Hz or 240 GHz. The wavelength of the waves is about 1 mm, for which neither generation of such frequencies, or detection methods were available in 1930s. Only two decades later, such em waves could be produced. Typical transition temperatures and the magnitude of the superconducting gaps for a few superconductors are indicated in Table 4.1.

Table 4.1
2Δ and T_c are reported for a few conventional superconductors. The gap is of the order of a few meV.

	T_c	2Δ (meV)
Nb_3Sn	18	6.55
MgB_2	40	10
Rb_3C_{60}	30	12
$ErRh_4B_4$	9	3

In the experimental method, EM radiation is guided through a small cavity made into a superconducting material. Within the cavity, the radiation is reflected several times before it emerges through the cavity, and hence is detected. The stronger the absorption of the radiation from the walls of the cavity, lower = the output power. At a fixed temperature, $T < T_c$, that is, when the material is in the

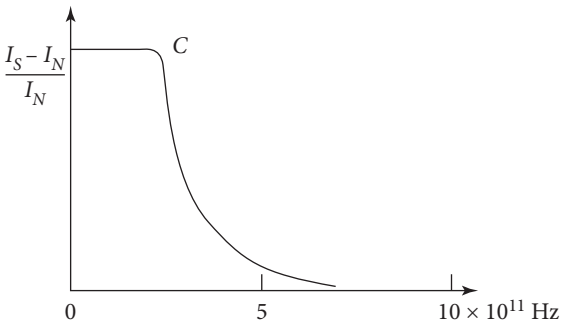


FIG. 4.24
The absorption of electromagnetic radiation is shown as a function of frequency. The frequency at which the intensity almost vanishes yields the superconducting gap 2Δ.

superconducting state, let the power in the superconducting and the normal states be I_s and I_n respectively. $\frac{I_s - I_n}{I_n}$ is plotted as a function of frequency of the EM waves in Fig. 4.24.

At small frequencies, we see a distinct difference between the detected power in the superconducting and the normal states, with the absorption being larger for the superconducting state. At certain threshold frequencies, the difference drops abruptly to zero. At even larger frequencies, the relative difference in the detected power stays at values that are vanishingly small. The abrupt drop sets in as soon as the quantum energy of the radiation is large enough to break the cooper pairs. Thus, there is an absorption edge at the point C in Fig. 4.24. Similar absorption edges are also seen for semiconductors; however, the scale of the energy gap is much higher there, e.g., ~0.8 eV in Ge. As an example, for $h\nu > 2\Delta_0$, the energy gap practically vanishes and so does the absorption spectrum. This aids us in obtaining $2\Delta_0$, which is proportional to T_c. Some typical values for superconductors, and their transition temperatures are tabulated below.

4.8.2 Ultrasound absorption

Sound waves denote the propagation of phonons in a material. Most measurements with sound waves are carried out with typical frequencies of the order of 10 MHz. These frequencies correspond to energies that are much smaller than the superconducting gap. Only very close to T_c, where the energy gap Δ approaches zero, sound energies at the MHz frequency range become comparable to the gap. Unpaired electrons below T_c interact with the lattice excitations leading to damping. Below T_c, as the temperature is decreased, the number of unpaired electrons decreases rapidly, thereby causing reduced damping of the sound intensity. A schematic plot of the ultrasound absorption spectra (in arbitrary units) is presented as a function of the reduced temperature, T/T_c. As T tends to T_c from below, the cooper pairs start dissociating yielding an increasingly large number of unpaired electrons in the system. These electrons absorb the sound waves leading to an enhancement in the absorption intensity (see Fig. 4.25). Since at a given temperature, the number of unpaired electrons depends on the energy gap, the latter can be determined from such absorption experiments.

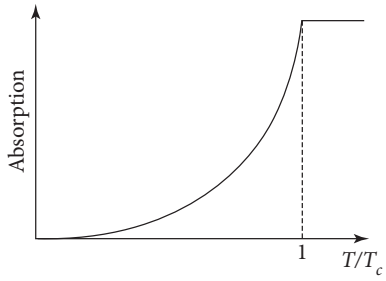


FIG. 4.25
Schematic depiction of the ultrasound absorption intensity is presented as a function of T/T_c. The absorption is low in the superconducting phase owing to the absence of unpaired electrons.

4.8.3 Tunneling experiments

In 1961, Giaever (1961) pointed out the possibility of determining the energy gap by means of tunneling experiments. The experiments involve the observation of tunneling current across a metal-insulator-semiconductor (MIS) junction. The metal and the insulating materials are usually taken as Al and Al_2O_3 respectively. Al_2O_3 is known to be a good insulator which can be fabricated nearly perfectly with a thickness of only a few nanometers.

One can understand the tunneling effect without a detailed calculation. Consider a particle that incidents on a barrier with energy lower than the barrier height. While a classical particle is unable to travel to the other side of the barrier, it is possible for a quantum mechanical particle (because of its wave nature) to do so. Considering the wave character of the particle, when a wave is incident on the surface separating two media into which it cannot enter, the wave must be completely reflected. We understand intuitively that being a wave, it can penetrate up to a small distance into the forbidden region with its amplitude decreasing exponentially. The decrease is faster and complete within the barrier if the width and height of the barrier are large. However, for a sufficiently thin barrier, there is a finite possibility that the wave can penetrate into the barrier (of the order of de Broglie wavelength) and propagate into the medium (lead) on the other side.

If a superconductor is introduced on one or both sides of the barrier, the current as a function of the biasing voltage ($I - V$) characteristic is bound to be different. This happens because an energy gap appears in the superconducting state. So no tunneling current can flow below a certain critical voltage, given by $V_c = \Delta/e$. From the above plot (Fig. 4.26) it is clear that from V_c, the value of Δ_0 can be determined.

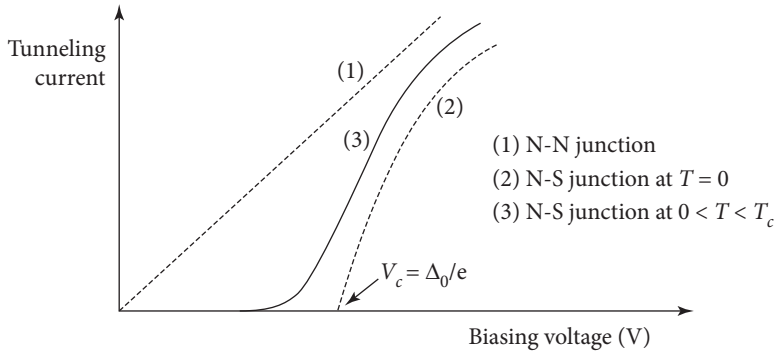


FIG. 4.26
The current-voltage characteristics are shown schematically for N-N and N-S junctions at $T = 0$ and $0 < T < T_c$.

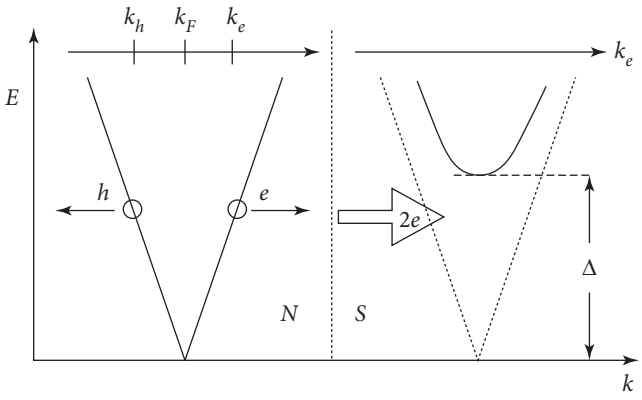


FIG. 4.27
A schematic diagram of the dispersion E vs k corresponding to the Andreev reflection at the normal-superconductor (N-S) junction is shown. A cooper pair is created at the boundary, which can propagate at low bias inside a superconductor. The electron is specularly reflected as a hole on the metallic side.

However, at the normal-superconductor junction there is an additional phenomenon which contributes to the tunneling current. Here especially we are interested in a situation in which the biasing voltage is much smaller compared to the energy gap, Δ in a superconductor. Thus, in the usual circumstances, no tunneling can occur. Thus, for an electron arriving at the interface of a normal-superconductor junction, two possible processes can take place (See Fig. 4.27), namely,

i. a normal reflection in which an electron is simply reflected back, giving rise to no net current.
ii. Andreev reflection which is an anomalous reflection of an electron. What actually happens here is the following. An electron comes at the interface, instead of undergoing a reflection
   a. a cooper pair is added to the superconducting region, it has an energy Δ.
   b. the incident electron is annihilated.
   c. a hole is reflected back along the original path of the electron.

This is known as Andreev reflection. The incident, and the reflected quasi-particles have approximately equal wave vectors, which are given by

$$k^{el} = k_F + \frac{\epsilon}{\hbar v_F}$$
$$k^{n} = k_F - \frac{\epsilon}{\hbar v_F} \tag{4.177}$$

but have opposite directions of motion, as appears from the sign of the group velocity $\frac{1}{\hbar}\frac{\partial \epsilon}{\partial k}$. The momentum is conserved up to the order,

$$\hbar\,|k^{el} - k^{n}| \leq \hbar/\xi$$

where $\xi = \frac{\hbar v_F}{\pi \Delta_0}$, which is the coherence length.

Energy is also conserved in the process. The minimum energy of a cooper pair has the value $2\epsilon_F$, the incident electron has energy $\epsilon_F + \epsilon$, and that of the reflected hole is $\epsilon_F - \epsilon$. In the process, a charge $2e$ is transferred from the normal to the superconducting side of the barrier, which is equivalent to a cooper pair being injected into the superconductor.

We can understand the tunneling problem in the following manner. The number of electrons tunneling per unit time with an energy E from the left (say, a metal) to the right (a superconductor) is

proportional to the number of occupied states on the left, namely, $N_1(E)f(E)$, and that to the unoccupied states on the right, $N_2(E')(1 - f(E'))$. $N_1(E)$ and $N_2(E')$ denote densities of states for the metallic and the superconducting regions, respectively. In the presence of the biasing voltage, V, the energies are given by

$$E' = E + eV$$

Thus, the unoccupied states on the right are denoted by

$$N_2(E + eV)(1 - f(E + eV)) \tag{4.178}$$

The tunneling current is proportional to both these independent processes, which can be written as

$$I_{1\to 2} \propto \int_{-\infty}^{\infty} N_1(\varepsilon)f(\varepsilon)N_2(\varepsilon + eV)(1 - f(\varepsilon + eV))\, d\varepsilon \tag{4.179}$$

This tunneling current is a measurable quantity in experiments.

For a normal metal-superconductor junction, the current, $I_{N\to S}$ can be obtained by replacing $N_1 \to N_N\, N_2 \to N_S$. A schematic plot of the tunneling current for the junctions, namely, and metal-metal $(N - N)$ and metal-superconductor $(N - S)$ both at $T = 0$ and for $T < T_c$ as a function of the bias voltage, are presented in Fig. 4.26. The threshold potential yields the energy gap. A differential conductance, $\frac{dI}{dV}$, which is proportional to the density of states at the Fermi level, is often obtained in experiments.

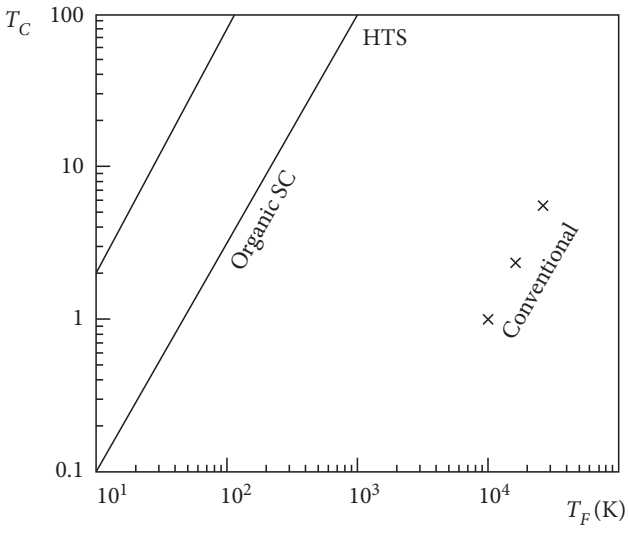


FIG. 4.28
Uemura plot showing the superconducting transition temperature, T_c vs the Fermi temperature, T_F (both in log scale). The positions of different superconductors are shown. The conventional (BCS) superconductors lie far away from unconventional superconductors.

4.8.4 Unconventional superconductivity

There are several superconductors, that have been discovered over the last few decades, which challenge the well-established paradigm of conventional superconductivity. Quite a few of them, including the high-T_c cuprates, pose anomalies owing to features that are unfamiliar in the context of Fermi liquid theory. As a first signature, Uemura presented a plot of T_c vs T_F (both in log scale) that yields a straight line, while the conventional superconductors are far away from this line. This is called the Uemura plot and is presented in Fig. 4.28.

4.9 HIGH-T_C CUPRATES

In 1986, 75 years after the discovery of superconductivity, G. Bednorz and K. Muller at IBM, Zurich demonstrated superconductivity in perovskite structured lanthanum (La) based copper oxide material, which registered a T_c of about 39 K for which the Nobel prize was awarded in 1987. It was a remarkable discovery, as it later allowed chemical substitution in perovskite cuprates to push the T_c well beyond the liquid nitrogen temperature (77 K), which is a much cheaper and easily accessible quantity than liquid Helium. Later in 1987, La was substituted with yttrium (Y) in the form of $YBa_{2-x}Cu_3O_{7-x}$ which showed a T_c of about 92 K. The materials show highest T_c when they are slightly oxygen deficient, that is, when $x \simeq 0.15$. Superconductivity disappears at larger values of x, which is also accompanied by a structural phase transition when YBCO changes from an orthorhombic structure to a tetragonal structure. Subsequently, thallium (Th) and mercury (Hg) based superconductors show even higher transition temperatures. These superconductors are generally type-II superconductors which show a gradual change in temperature as a function of the external magnetic field.

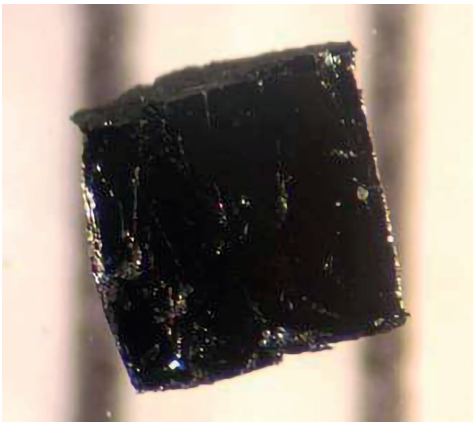


FIG. 4.29
A cube of sides 1 mm of the Bismuth Strontium Calcium Copper Oxide (BSCCO) is shown.

A large volume of research has taken place since its discovery Varma (2020), however, owing to its ill-understood normal state, a comprehensive understanding of the microscopic phenomena remained elusive. A simple and intuitive way to understand the complexity in the physical properties demonstrated by the cuprates is to look at the photograph of the Bismuth-based superconductor in Fig. 4.29. The highlight of the figure is that it despite being a ceramic superconductor, it appears black and without any luster. Conventional ceramic compounds possess a glow which is visibly absent here. It probably indicates toward the lack of a conventional metallic state too, as will be reflected in the discussion that follows below.

The electronic pairing mechanism is itself questionable and possibly other pairing symmetries, other then the conventional s-wave (the latter discussed in the context of BCS theory) play crucial roles. Even after several decades of active research, several anomalies that questioned the established paradigm of superconductivity, were left without the necessary clarification. We discuss some of them below.

4.10 PHYSICAL PROPERTIES

The discovery of the high-T_c superconductors gave an intense push to the research on superconductivity in the direction of achieving larger and larger transition temperatures (T_c). Over the last three

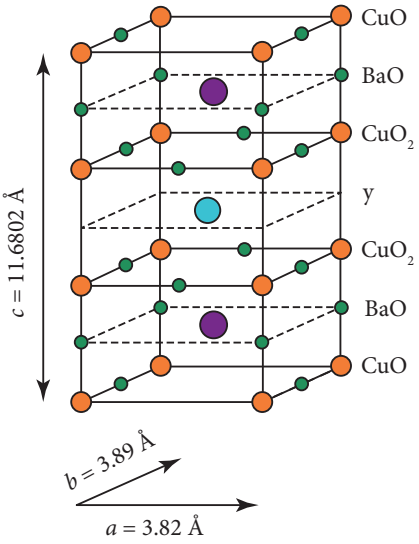


FIG. 4.30
Unit cell of Yttrium Barium Copper Oxide (YBCO) is shown. The lattice parameters are indicated.

decades, the maximum T_c has increased by 140 K and there is a lingering hope of realizing even larger values of T_c. Much of the difficulties in understanding both the normal and the superconducting states in cuprates involves a complex multi-layered crystal structure with a large anisotropy along the c-axis. Besides, there are difficulties in obtaining good quality single crystals. In the following, we shall discuss a few key properties of the cuprates that should serve as an introduction to a beginner and set the platform for building up the knowledge on this class of materials with rather curious properties.

Several families of thermodynamically stable cuprate superconductors are experimentally realized. The fundamental block is the CuO_2 plane where Cu ions in $3d^9$ configuration form a square lattice. Each such Cu ion is connected to another through the oxygen ions. Depending upon the class of materials, the unit cell may contain one, two, or more CuO_2 layers. A typical unit cell of Yttrium Barium Copper Oxide ($YBa_2Cu_3O_{7-\delta}$), with $\delta \approx 0.1 - 0.3$, known by the more familiar name YBCO with a $T_c \approx 90$ K) is shown in Fig. 4.30. The planar and the interplanar lattice parameters are $a = 3.82$ Å, $b = 3.89$ Å, $c \simeq 11.7$ Å. It is generally believed that T_C scales with the number of CuO_2 layers. Many of the unusual properties are attributed to the large volume of the unit cell and the interplanar anisotropy.

The compounds that are of special interest are the $Bi_2CaSr_2Cu_2O_8$ (called as Bi-2122 compounds) and $Tl_2CaBa_2Cu_2O_8$ (called as Tl-2122 compounds). Because of their large T_c (>100 K), they have different numbers of cuprate layers and are denoted by a general formula $A_2Ca_nY_2Cu_nO_{2n+4}$ with A = Bi, Tl; and Y = Sr, Ba; etc. The ones with Bi and Sr are more widely studied and are described by the formula $Bi_2Sr_2Ca_{n-1}Cu_nO_{2n+4+\delta}$ where $n = 1 \rightarrow 3$ and δ denotes the oxygen doping. These are called BSCCO (pronounced as BISKO). $n = 1$, $n = 2$ and $n = 3$ have a T_c of 33 K, 96 K, and 108 K respectively. Similarly, for the Tl family (denoted by $TlBa_2Ca_{n-1}Cu_nO_{2n+3}$) $n = 1, 2$ and 3 correspond to $T_c \sim 14$ K, ~80 K and ~120 K, respectively.

There are numerous examples of anomalies noted in various experiments on cuprate superconductors. For example, consider ultrasound attenuation measurements. In conventional (BCS) superconductors, the ultrasound attenuates exponentially due to the loss of unpaired electrons (in favor of forming cooper pairs). The coefficient of attenuation α is proportional to $e^{-\Delta/K_BT}$ where Δ denotes the superconducting energy gap. The scenario is not too different in the cuprate family. However, the ultrasonic attenuation below T_c is found to be lower than that of the BCS superconductors, possibly owing to the

formation of pseudogap. However, the longitudinal sound velocity and the elastic moduli show discontinuities in most of the cuprate superconductors (except for the Tl-based family). The discontinuity points toward a structural change and "*lattice softening/hardening*" at the superconducting transition. The anomaly in the elastic modulus points toward a collective ordering of the oxygen atoms in the CuO_2 layers. There are many such examples which show significant digression compared to the behavior observed for conventional superconductors. Apart from being beyond the scope of the present discussion, there is a volume of good reviews on the subject, and interested readers are encouraged to look at some of these. In the following we review the most significant feature of the cuprates which eluded a microscopic theory so far, namely, the formation of a pseudogap phase.

4.11 THE PSEUDOGAP PHASE

The nature of the spectral gap in superconductors and how it vanishes (or changes) across a transition can be detected by angle-resolved photoemission spectroscopy (ARPES) experiments. In the context of high T_c superconductors, high energy photons from a synchrotron source with an energy $\approx$20 eV made to incident on the surface of a single crystal. As a result, a photoelectron is ejected at an angle θ normal to the surface of the crystal. The kinetic energy of the ejected electron is measured via an electron spectrometer whose energy resolution is of the order of a few meV. The ARPES data for high T_c cuprates are thoroughly analyzed and reviewed in Shen *et al.* (1993) and Randeria and Campuzano (1997).

The parallel component of the momentum ($\mathbf{k}$) is balanced by the momentum of the excitations created inside the sample. By varying the detecting angle θ between the normal and the position of the detector, it is possible to map out the momentum dependence of the filled states of the crystal. Thus, the intensity of the photoelectron distribution will demonstrate a sharp peak corresponding to the parallel component of the momentum with an energy ω which is given by

$$I(\mathbf{k},\omega) = I_0(\mathbf{k}) f(\omega) A(\mathbf{k},\omega), \tag{4.180}$$

where the energy is measured with respect to the Fermi-level (Randeria and Campuzano, 1997) and $A(\mathbf{k},\omega)$ denotes the spectral function of the excitations that corresponds to the hole that is left behind. $I_0(\mathbf{k})$ is the arbitrary intensity factor that preserves the dimensionality of the equation and $f(\omega)$ is the Fermi distribution function, $f(\omega) = \frac{1}{e^{\beta(\omega-\mu)}+1}$, $\hbar = 1$ and β being the inverse temperature.

The important quantity in the above Eq. (4.180) is the spectral function which contains the information about the electron-electron interaction present in the material via the self-energy $\Sigma(\mathbf{k},\omega)$, which is in general a complex quantity. The spectral function $A(\mathbf{k},\omega)$ is expressed in terms $\Sigma(\mathbf{k},\omega)$ as (Fetter and Walecka, 1971)

$$A(\mathbf{k},\omega) = \frac{\Sigma''(\mathbf{k},\omega)}{(\omega - \epsilon_k - \Sigma'(\mathbf{k},\omega))^2 + \Sigma''(\mathbf{k},\omega)}, \tag{4.181}$$

where Σ' and Σ'' represent the real and the imaginary parts of the self-energy and ϵ_k denotes the single particle energies.

The ARPES spectra measure $A(\mathbf{k}, \omega)$ and its shows a broad background as a function of ω that extends to lower energies and yields the energy loss of the photoelectrons within the crystal. For a recent review, see Yu *et al.* (2020). However, as a result of strong interparticle interactions, which dominate in these cuprate superconductors, this broad background is attributed to the electronic correlations. The energy of the emergent photoelectrons yields a sharp coherent peak at the single particle energies given by ϵ_k broadens rapidly for state below the Fermi level as the detection angle θ is varied. As said earlier, the loss of coherence is due to the presence of strong electronic correlations (Timusk and Statt, 1999) and hints toward the breakdown of the Fermi liquid picture (see appendix). The optical conductivity data support the loss of coherence as well. The sharp peak at the Fermi level as a function of the detection angle θ maps out the Fermi surface as it yields the information on the momentum (and hence the energy) of the emitted photoelectron. The same experiment can be repeated at various doping levels. Mostly in cuprates, at optimal doping and at the overdoped regime, the presence of the sharp peak in the ARPES spectrum confirms the Fermi liquid picture to be valid. However, a noticeable broadening is noted in the underdoped regime at the Fermi level and that implies the breakdown of the Fermi liquid theory (FLT).

Both at $T < T_c$ (superconducting state) and $T > T_c$ (normal state), there are deviations from the FLT which are attributed to the formation of pseudogap in cuprates. The pseudogap is most likely related to and has the same symmetry as the superconducting gap. (believed to be d-wave like). However, unlike a conventional superconducting gap, the magnitude of the pseudogap is temperature independent. ARPES and other experiments indicate that the magnitude of the gap does not change significantly as the temperature is lowered through T_c. Thus, in case of the cuprates, unlike a conventional superconducting gap, the gap starts as a pseudogap in the normal state and evolves into the superconducting gap below T_c. In other words, $\Delta(T)$ does not go to zero at $T = T_c$.

The pseudogap appears in the underdoped regime of the phase diagram and weakens toward optimal doping. A weak pseudogap is still present at optimal doping,[8] but rapidly disappears thereafter. The spectral gap is purely superconducting in the overdoped region. This is illustrated in the phase diagram shown in Fig. 4.31 where T^* approaches T_c just into the overdoped region. T^* is temperature dependent in the underdoped regime. Well into the overdoped region, the pseudogap merges with the superconducting gap, which means that T^* merges with T_c, and there remains no trace of the pseudogap in the normal state.

To wind up the discussion, we note down a few salient features. They are,

i. the pseudogap is different from the superconducting gap in the underdoped phase of cuprates, and gradually evolves into the superconducting gap at larger doping levels.
ii. Both the pseudogap and superconducting gaps are believed to have d-wave symmetry.
iii. The crossover temperature T^* (shown in Fig. 4.31) merges with T_c in the overdoped region of the phase diagram.

[8] Optimal doping corresponds to highest T_c.

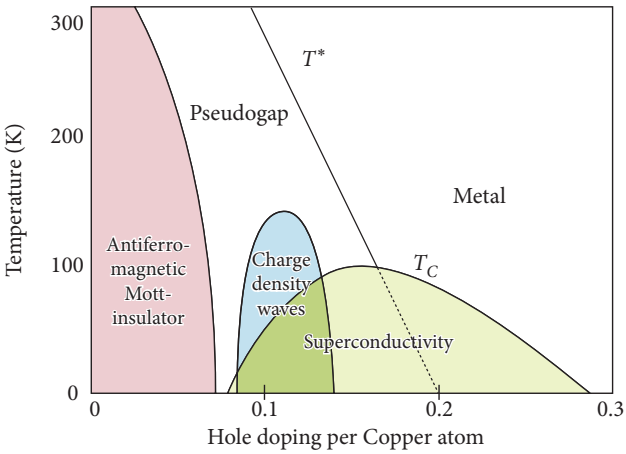


FIG. 4.31
Generic phase diagram showing temperature vs hole doping (per copper atom) for cuprate superconductors. The different phases are shown. The superconducting dome extends upto a doping of 0.3. In the underdoped regime, at high temperatures, one gets a "*strange*" metallic phase, which cannot be explained by the Fermi liquid theory, whereas the overdoped regime shows more familiar metallic features.

Table 4.2
T_cs of some typical cuprate superconductors are shown.

	T_c
$La_{2-x}Ba_xCuO_4$	30 K
$La_{2-x}Sr_xCuO_4$	39 K
$YBa_2Cu_3O_7$	92 K
$Bi_2Sr_2Ca_2Cu_3O_10$	110 K
$Te_2Ba_2Ca_2Cu_3O_10$	125 K
$HgBa_2Ca_2Cu_3O_8$	134 K

The cuprate superconductors have a generic phase diagram (see Fig. 4.31) which itself poses a bunch of surprises. The proximity of the Mott-insulating state and the superconducting phase, a pseudogap phase in the underdoped phase, are among the few. There is a superconducting dome (shown in light green color in Fig. 4.31) which extends from a (hole) doping per Cu atom of about 0.08 to about 0.3. The maximum T_c occurs at a doping $\lesssim 0.2$. Below this is known as the underdoped regime, while above it is denoted as the overdoped phase. The high temperature phase of the underdoped regime is not a "*normal*" metal and shows significant deviations from the FLT, quite likely due to the presence of the pseudogap phase. The overdoped regime is "*well behaved*" Fermi liquid metal. The transition temperatures of a few commonly known high-T_c cuprates are shown in Table 4.2.

We shall not prolong the discussion here, and rather motivate the readers to look at the review mentioned in the footnote and the references therein. A few of the CuO_2 superconductors are listed here along with their transition temperatures.

4.12 IRON BASED SUPERCONDUCTORS

The discovery of the iron (Fe) based superconductors defied the occurrence of superconductivity in compounds with strong magnetic correlations. Further, the discovery of superconductivity in cuprates

created a bias in the community that copper oxide planes are indispensable for achieving higher T_c in materials. However, the Fe-based superconductors do not contain copper oxide planes, and even do not contain oxygen in certain situations. Instead, they have electrons from Fe whose density changes unusually rapidly under doping, thereby showing an interplay between the magnetic phenomenon and superconductivity. Similar to the high-T_c cuprates, the normal state in Fe-based superconductors is unconventional, and thus the mechanism behind the superconducting pairing remains elusive.

The excitement in the study of these Fe-based superconductors started with iron oxypnictides ($LaFeAsO_{1-x}Fe_x$) which shows a superconducting transition at $T_c = 26$ K (Kamihara *et al.*, 2008). It may not constitute the first discovery reported for the family of superconductors, but in all others, such as Th_7Fe_3 ($T_c = 1.8$ K) (Mathias *et al.*, 1961), U_6Fe ($T_c = 3.9$ K) (Chandrasekhar and Hulm, 1958). etc. In fact, Fe itself ($T_c = 1.8$ K) is superconducting under the application of pressure (∼20 GPa) (Shimuzu *et al.*, 2001). However, much higher T_c is observed in fluorine (F) doped LaFeAsO, and the subsequent discoveries, such as Th doped GdFeAsO (Wang, 2008) and many others have portrayed a growth of T_c similar to that in cuprates. This has led to the emergence of a new class of high temperature superconductors known by the name Fe pnictides, (Fe-Pn) and Fe chalcogenides, (Fe-Ch).[9] There is literally an outburst in the number of members in the family of these superconductors and about 150 of them have been realized so far only belonging to the LaFeAsO (called the "1111" structure) family. Then there is the "122" family corresponding to MFe_2As_2 (M: Sc, V, Cr), "111" family corresponding MFeAs, "11" family corresponding to FeSe, and many others. For an extensive review, see Stewart (2011).

The existence of the isotope effect (although the pairing mechanism is primarily *not* phonon mediated) leaves room for further questions. Spin or orbital fluctuations play a crucial role in pushing the normal state toward a superconducting phase. Furthermore, the symmetry of the superconducting gap function is not an isotropic *s*-wave type, and instead contains nodal lines with possibilities of gapless excitations. Also, the specific heat discontinuity ΔC (see discussion in Sec. 4.4.2) instead of scaling linearly with T_c (which is a hallmark signature of conventional superconductors), it scales very differently in Fe-Pn/Ch superconductors. For example, $\Delta C \sim T_c^3$, along with a normal state-specific heat behaving as $C \sim T^3$ is observed.

4.12.1 Properties of Fe-based superconductors

The crystal structures of these Fe-based superconductors possess square lattice as the building blocks, corners of which are occupied by the Fe^{2+} ions. The Pn/Ch are located at the apical sites of a tetrahedron (see Fig. 4.32). The nearest neighbor Fe-Fe distance is approximately 2.8 Å. These Fe-Pn and the

[9] The fifth group elements in the periodic table are called as pnictides (Pn), which are **N, P, As, Sb, Bi**, etc. and the sixth group elements are called as chalcogenides, which are O, S, Se, Te, Po, etc.

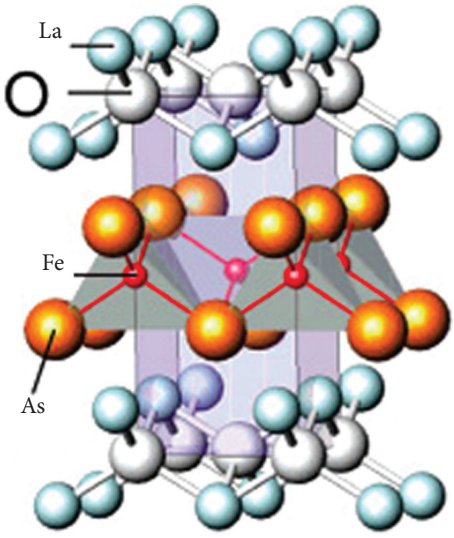


FIG. 4.32
The crystal structure of 1111 LaFeAsO (from Kamihara *et al.* (2008)).

Fe-Ch layers are arranged alternately with other interlayers, thereby making a sandwich-type layered structure.

It is practically an uphill task to depict the crystal structures of all the members of the family of Fe-based superconductors, owing to a huge library of resources available. It is also an arduous task to enumerate the properties of all these materials and hence, it is beyond the scope of this book. Interested readers may consult the review in Stewart (2011). What we shall aim in the following is to provide the readers with useful information on only one class of Fe-based superconductors, namely, the 1111 family, for which the crystal structure is shown here (see Fig. 4.32). In the following, we shall focus on the dependence of T_c on doping, pressure, and the external magnetic field. We shall also include brief discussions on a couple of key thermodynamic and transport properties, such as the specific heat, resistivity, and susceptibility.

Fe plays a key role in the superconducting properties of Fe-based compounds as the parent undoped compounds possess five 3d orbitals at the Fermi surface. This scenario is in contrast with the cuprates which have only one 3d orbital (due to Cu) at the Fermi level. Further, the Fermi surfaces of the Fe-based materials contain several hole- and electron-like cylinders, thereby forming quite a complex band structure. The transition temperature T_c can be increased via oxygen deficiency induced by doping with fluorine (F) as shown by Kamihara in the 1111 Fe-Pn compounds. The transition temperature T_c can be increased via oxygen deficiency induced by doping with fluorine (F) as shown by Kamihara in the 1111 Fe-Pn compounds (Kamihara *et al.*, 2008). The usage of smaller Lanthanide elements in LnFeAsO (Ln = La, Dy, Tb, Gd, Sm, Nd, etc) has been derived from the above discovery when 11% F doping was carried out at the oxygen sites in $LnFeAsO_{1-x}F_x$ that ramped up the T_c from 26 K to 43 K under pressure (Takahashi *et al.*, 2008). We show the temperature vs doping (by F) for $CeFeAsO_{1-x}F_x$ in Fig. 4.33, where at low doping, an antiferromagnetic metallic (AFM) phase is realized. A superconducting phase is obtained at about 5% doping. The result was supplemented by Eisaki *et al.* (2008) where it was shown that T_c is independent of the nature of the Lanthanide elements (Ln), but depends on the planar lattice parameter.

Furthermore, T_c was found to increase with pressure. In the F-doped $LnFeAsO_{1-x}F_x$, $\frac{dT_c}{dP} = 2$ K/GPa for $x = 0.05$ (Takahashi *et al.*, 2008) in the limit of zero pressure. However, for the optimally doped compound ($x \simeq 0.11$), T_c goes up to 43 K at 4 GPa pressure with a slope $\frac{dT_c}{dP} = 3$ K/GPa, and hence steadily decreases to 9 K. T_c vs pressure phase diagram thus generated (Takahashi *et al.*, 2008) seems not so generic to other members of the family. For example, (Lorenz *et al.*, 2008) measured the pressure response of $SmFeAsO_{1-x}F_x$ upto a pressure of 1.7 GPa and found a different behaviour from its

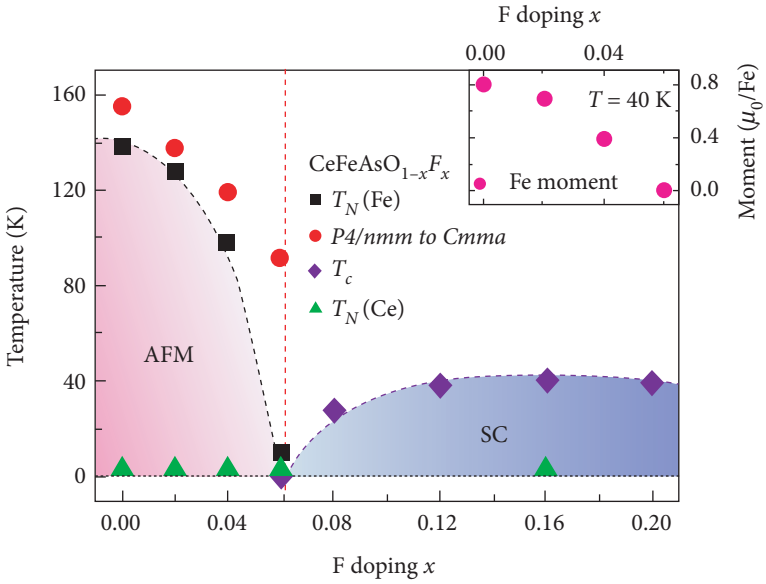


FIG. 4.33
The phase diagram of of 1111 CeFeAsO$_{1-x}$F$_x$ vs F doping (from Zhao *et al.* (2008)) An antiferromagnetic metallic (AFM) phase is realized at low doping. At about 5% doping the material becomes superconducting. The structural phase transition is denoted by the red dots. For more details, see Zhao *et al.* (2008).

La counterpart in the sense that T_c increases with pressure for the undoped material (SmFeAsO), however it decreases for the doped material in the overdosed regime. The 1111 family of superconductors belong to the type-II variety and record very high upper critical magnetic fields, H_{c_2}. Using dc fields up to 45 T, $H_{c_2}(T = 0)$ was reported to be 60T for an optimally doped LaFeAsO$_{0.89}$F$_{0.11}$ (Jarozynski *et al.*, 2008). Furthermore, there observed anisotropies have in the slope of $\frac{dH_{c_2}}{dT}|_{T=T_c}$ along the planar and out-of-plane directions.

To shed light on the transport properties of Fe-based compounds, Kamihara *et al.* (2008) in their pioneering discovery of LaFeAsO$_{1-x}$F$_x$ found that the resistivity of the undoped compound (LaFeAsO) has a temperature independent resistivity with a value 5 mΩ-cm at low temperature. However, it shows an anomaly at ∼150 K with an upturn in its magnitude at 100 K. At optimal doping ($x \simeq 0.11$), the resistivity is metallic at room temperature and falls steadily at lower temperatures. The same group (Kamihara *et al.*, 2008) has also reported data on the magnetic susceptibility. The undoped compound reports a value 0.4 memu/mole and again independent of temperature below 300 K, except for an anomaly at 150 K, and an upturn in its behavior below ∼25 K.

We have cursorily talked about the thermodynamic properties earlier, where the behavior of the specific heat at the transition point (T_c) was discussed. To remind ourselves of the context, in the study of solid state physics, C/T is usually plotted vs T as $T \to 0$[10] where C/T is found to be temperature independent and is represented by γ. γ depends on the material properties via the density of states at the Fermi level ($N(\epsilon_F)$), and the effective mass (m^*) of the carriers. Importantly, C/T shows a discontinuity at $T = T_c$, signaling a second-order phase transition from a metal to a superconductor. In fact, the discontinuity is often considered as the hallmark feature of the onset of superconductivity. However, for the 1111 family of materials, possibly owing to high values of T_c, C/T is not temperature independent, and hence the value of γ cannot be ascertained. Although, γ is much better known for other families of Fe-based superconductors. A general consensus is that $\Delta C/T_c$ behaves as T_c^2 in most of these materials, which is at least true in the underdoped regime, and most probably are signatures of a non-Fermi liquid metallic state[11] at higher temperatures. Possibly the Fe-Pn/Ch superconductors are weak Fermi liquids with strong pair breaking mechanism in place.

4.12.2 Pairing symmetry

The pairing symmetry of the superconducting state is not beyond doubt and has generated a lot of debate for Fe-based superconductors. The electron-phonon interaction is the *pairing glue*, and is hence a key element to conventional (BCS) superconductors. However, it is found to be insufficient, and possibly cannot give rise to such high values of T_c found in experiments conducted on the pnictides or the chalcogenides. Orbital or spin fluctuations may play a role in the pairing mechanism. Theoretical studies suggest an extended-s wave ($\cos k_x + \cos k_y$) pairing symmetry owing to the multi-orbital nature (Kuroki *et al.*, 2009). In addition, $s_{x^2-y^2}$ symmetry has been predicted by Parish *et al.* (2008), and d-wave in some 122 Fe-As compounds (Reid *et al.*, 2012). Further, there are hints of the presence of two superconducting gaps which are inferred from two different values of $2\Delta/k_BT_c$,[12] one with a larger value $\frac{2\Delta}{k_BT_c} \simeq 7$, and the other corresponds to a smaller gap, that is, $\frac{2\Delta}{k_BT_c} < 3$. Pairing correlations with two different gaps indeed indicates an unconventional origin of superconductivity.

Thus, to summarize various features that have made studies of the Fe-based superconductors exciting across the last decade and a half, we enumerate some of them below.

i. There is a prejudice against Fe for it to be detrimental to superconducting pairing.
ii. The interplay of magnetism, superconductivity and orbital degrees of freedom is observed.
iii. The specific heat scales very differently with temperature compared to the conventional superconductors.
iv. The d-electrons of Fe have a complicated band structure that changes drastically upon doping.
v. There are two pairing gaps.

[10] In the limit $T \to 0$, the phonon contribution is suppressed, leaving only the electronic contribution to the specific heat.
[11] A brief introduction to the main results of the Fermi liquid theory is presented in the appendix.
[12] $\frac{2\Delta}{k_BT_c} = 3.52$ for BCS superconductors.

vi. There is an isotope effect of Fe, as opposed to the oxygen isotopes observed in conventional superconductors. Despite the observed isotope effect, electron-phonon coupling is not a dominant player in the pairing phenomena.

With this collection of facts, we hope to have provided a platform for the beginners.

We next move on to the potential applications of superconductors, which have constantly enriched the field of superconductivity throughout a period spanning over the last century.

4.13 APPLICATIONS OF SUPERCONDUCTORS

The chief applications of superconductivity lie in the large scale production of high intensity magnetic fields needed for the functioning of magnetic resonance imaging (MRI) and nuclear magnetic resonance (NMR). Superconducting rails are used to the magnetic levitation principle, which capitalizes on the lack of contact between the vehicle and the rails, and hence there will be no wear and friction, besides aiding to attain far larger velocities compared to ordinary trains. The magnetic levitation technology is quite advanced, but at its core, it uses the phenomenon of the Meissner effect, that is, the expulsion of the magnetic flux from a superconducting sample, as we have seen before.

There is another application of superconductors that we should be concerned about in this text is the Josephson effect. The Josephson effect is a manifestation of long-range quantum coherence in superconductors. The effect is discovered by B.D. Josephson in 1962[13] when he was a PhD student at Cambridge (UK). In mathematical terms, the effect describes the relationship between the current and the voltage across a link in a superconductor-insulator-superconductor junction. Thus, a Josephson junction (JJ) is formed by placing two superconducting materials with an intervening non-superconducting material, where the Cooper pair can tunnel from one superconductor to another through the barrier. Extensive reviews of the Josephson effects can be found in Makhlin *et al.* (2001) and Golubov *et al.* (2004).

4.14 JOSEPHSON EFFECT

In the following, we show different types of JJ with two bulk superconductors connected via a weak link, that is, regions that allow the passage of electrons in a restricted sense (see Fig. 4.34).

To understand the physical phenomenon, consider two isolated superconductors with the superconducting order parameters ψ_1 and ψ_2 ($\psi_1 \neq \psi_2$), which further can be expressed in terms of the amplitudes ($|\psi_{1,2}|$) and phases ($\phi_{1,2}$) as $\psi_1 = |\psi_1|e^{i\phi_1}$ and $\psi_2 = |\psi_2|e^{i\phi_2}$. In the absence of any interaction between the superconductors, the phases ϕ_1 and ϕ_2 are in general different. If these two regions are strongly coupled, the two phases may be equal ($\phi_1 = \phi_2$), which demands that all the Cooper pairs

[13] In 1973, Josephson was awarded Nobel prize for this work.

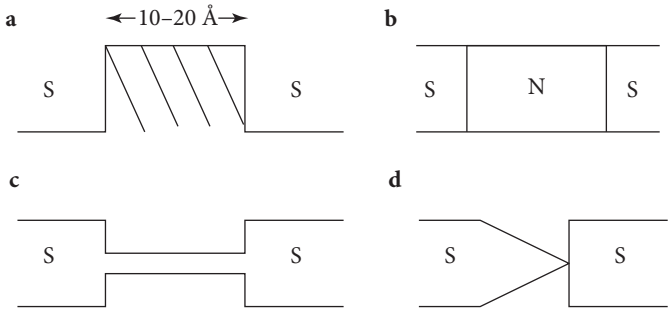


FIG. 4.34
Different types of Josephson junctions, namely, (a) Tunnel junction, superconductor-insulator-superconductor (SIS) junction, (b) Proximity junction, superconductor-normal-superconductor (SNS) junction, (c) Constriction (Microbridge), (d) Point-contact junction.

are in the same state. However, in case of a weak coupling, the phases will differ, and it is possible to maintain this difference by passing a small current through the JJ. This is precisely the Josephson effect.

4.14.1 AC and DC Josephson effects

Josephson predicted that the current and voltage have dependencies given by

$$I = I_c \sin \Delta\phi \tag{4.182}$$

$$\frac{d}{dt}\Delta\phi = \frac{2eV}{\hbar}, \tag{4.183}$$

where the $\Delta\phi$ is the phase difference $\Delta\phi = \phi_1 - \phi_2$, ϕ_1 is the phase of the order parameter for the superconductor on the left, while ϕ_2 is the phase of the order parameter for the one on the right, and V is the voltage across the junction. Equation (4.182) implies the presence of a dissipationless sinusoidal current with a phase difference $\Delta\phi$. For a DC voltage V applied across the junction, the current oscillates as [from Eq. (4.183)],

$$I(t) = I_c \sin\left(\frac{2eVt}{\hbar}\right). \tag{4.184}$$

This is called the AC Josephson effect.

Actually it is rather difficult to generate a constant DC voltage across a junction, and thus a more conventional situation is to talk about a constant current, namely,

$$I(t) = I = \text{constant}. \tag{4.185}$$

For $I < I_c$,

$$\Delta\phi = \text{constant} = \sin^{-1}\left(\frac{I}{I_c}\right)$$

which yields from Eq. (4.183) for $V = 0$ (as $\frac{d\Delta\phi}{dt} = 0$). This is called the DC Josephson effect.

It is possible to discuss the Josephson effect and the related phenomena in terms of flux quantization. Consider a contour C as shown by dotted lines in Fig. 4.35. Assume that the self-inductance effect of the ring is small, that is, $LI_c \ll \Phi_0$, L is the self-inductance of the coil, I_c is the critical current, Φ_0 is the flux quantum ($= \frac{h}{2e}$). For the magnetic field to be zero in the ring (defined by the contour C), the thickness of the ring is required to be larger than the penetration depth of the superconducting sample. The magnetic field should completely die out. Thus, no flux can be present along the contour C (see Fig. 4.35). From the equation of the electric current density (refer to London equation), namely,

$$\mathbf{j}(\mathbf{r}) \propto \nabla\phi(\mathbf{r}) - \frac{2e\mathbf{A}(\mathbf{r})}{\hbar} \tag{4.186}$$

for $\mathbf{j}(\mathbf{r}) = 0$, which is true for superconductors,

$$\nabla\phi(\mathbf{r}) = \frac{2e\mathbf{A}(\mathbf{r})}{\hbar}. \tag{4.187}$$

Now integrate both sides along the contour C,

$$\oint \nabla\phi \cdot d\mathbf{l} = \frac{2e}{\hbar}\oint \mathbf{A}\cdot d\mathbf{l} = \frac{2e}{\hbar}\int_s \mathbf{B}\cdot d\mathbf{s} = \frac{2e\Phi}{\hbar} = 2\pi\frac{\Phi}{\Phi_0}. \tag{4.188}$$

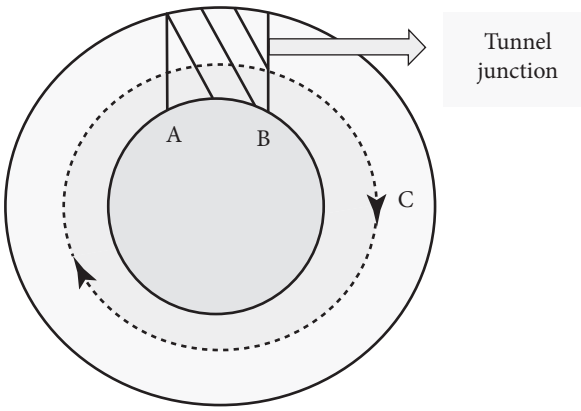


FIG. 4.35
Schematic diagram of a typical tunnel junction is shown. C is an imaginary contour in the superconducting region.

Thus, the phase difference is given by

$$\Delta\phi = 2\pi\frac{\Phi}{\Phi_0}, \tag{4.189}$$

where $\Delta\phi$ denotes the difference between the phases of the cooper pairs at the points A and B in Fig. 4.35. This is a fundamental relationship for a JJ. Now taking a derivative of Eq. (4.188) and using Faraday's law,

$$\frac{d}{dt}\Delta\phi = \frac{2eV(t)}{\hbar} \tag{4.190}$$

The wavefunction at the superconducting electrode is given by

$$\psi(\mathbf{r}, t) = \psi_0 e^{-i\frac{Vt}{\hbar}}, \tag{4.191}$$

where $V(t)$ is the emf developed around the ring, which can be interpreted as a voltage drop across

the junction. where $\theta(t) = \frac{Vt}{\hbar}$ is the phase of the wavefunction. Since the current is continuous, it must thread through the insulating barrier. This is not classically allowed, but the super-electrons can tunnel through the insulating barrier with a zero voltage drops and resulting in a Josephson current.

Let us perform a simple analysis of the above scenario by assuming a rectangular potential barrier of height V_0 and width $2a$ for the insulating region of JJ [see the schematic diagram in Fig. 4.36(a)]. The corresponding Schrödinger equation is written as

$$i\hbar\frac{\partial}{\partial t}\psi(\mathbf{r},t) = \frac{1}{2m^*}(-i\hbar\nabla - e^*\mathbf{A}(\mathbf{r},t))^2\psi(\mathbf{r},t) + e^*\phi(\mathbf{r},t)\psi(\mathbf{r},t) + V(x)\psi(\mathbf{r},t), \tag{4.192}$$

where the energy of the super-electrons is ϵ_0. Let us ignore the electromagnetic field for now,

$$\frac{\hbar^2}{2m^*}\nabla^2\psi = (\epsilon_0 - V_0)\psi(\mathbf{r}) \quad \text{for } |x| \leq a. \tag{4.193}$$

Thus, in the insulating region, the solution can be written as

$$\psi(x) = C_1\cosh\left(\frac{x}{\xi}\right) + C_2\sinh\left(\frac{x}{\xi}\right) \tag{4.194}$$

where

$$\xi = \sqrt{\frac{\hbar^2}{2m^*(V_0 - \epsilon_0)}}. \tag{4.195}$$

The scenario is depicted in Fig. 4.36(b). Now, let us recall the expression for the current,

$$J_s = \frac{2e^*}{m^*}Re(\psi^*\frac{\hbar}{i}\nabla\psi) = \frac{e^*\hbar}{m^*\xi}Im(C_1^*C_2). \tag{4.196}$$

At the boundaries, by matching the wavefunctions, one gets

$$\begin{aligned}\psi(-a) &= \sqrt{n_1}e^{i\theta_1}\\ \psi(a) &= \sqrt{n_2}e^{i\theta_2}.\end{aligned} \tag{4.197}$$

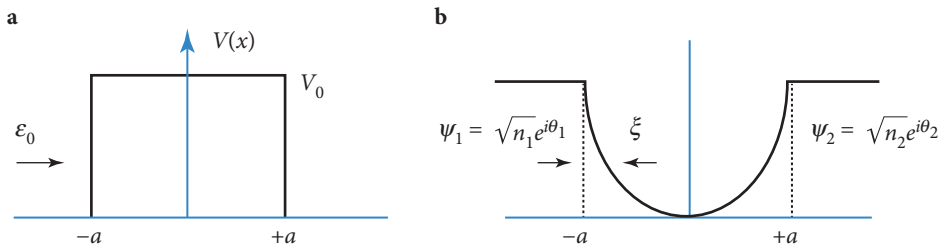


FIG. 4.36
(a) Potential experienced by the super-electrons because of the presence of the insulating barrier. A particle of energy ϵ_0 approaches the barrier from the left. (b) Fall off of the wave functions inside the barrier. ξ denotes the distance over which the wavefunction falls off by a factor $\frac{1}{e}$ of the value at the surface.

The coefficients C_1 and C_2 can be computed as

$$C_1 = \frac{\sqrt{n_1}e^{i\theta_1} + \sqrt{n_2}e^{i\theta_2}}{2\cosh\left(\frac{a}{\xi}\right)},$$
$$C_2 = \frac{\sqrt{n_1}e^{i\theta_1} - \sqrt{n_2}e^{i\theta_2}}{2\sinh\left(\frac{a}{\xi}\right)}. \tag{4.198}$$

Using the definition of the super current density, $\mathbf{J}_s$,

$$\mathbf{J}_s = \mathbf{J}_c \sin(\theta_1 - \theta_2) = \mathbf{J}_c \sin(\phi); \quad \phi = \theta_1 - \theta_2 \tag{4.199}$$

where the critical $\mathbf{J}_c$ is given by

$$|\mathbf{J}_c| = \frac{e\hbar\sqrt{n_1 n_2}}{m\xi \sinh(2a/\xi)}. \tag{4.200}$$

Now, including the electromagnetic field, one gets

$$\phi = \theta_1 - \theta_2 - \frac{2\pi}{\Phi_0}\int_1^2 \mathbf{A}(\mathbf{r}, t)\cdot d\mathbf{l}. \tag{4.201}$$

The rate of change in this phase is

$$\frac{\partial\phi}{\partial t} = \frac{\partial\theta_1}{\partial t} - \frac{\partial\theta_2}{\partial t} - \frac{2\pi}{\Phi_0}\frac{\partial}{\partial t}\int_1^2 \mathbf{A}(\mathbf{r}, t)\cdot d\mathbf{l}. \tag{4.202}$$

The current and the voltage across an SIS junction can be represented as

$$I_s = I_c \sin\phi$$
$$V = \frac{\Phi_0}{2\pi}\frac{d\phi}{dt}. \tag{4.203}$$

The corresponding work done is given by

$$W_J = \int_0^{t_0} IVdt = \int_0^{t_0} I_c \sin\phi'(t)\frac{\Phi_0}{2\pi}\frac{d\phi'(t)}{dt}dt = \frac{\Phi_0 I_c}{2\pi}\int_0^{\phi} \sin\phi'(t)d\phi'(t)$$
$$= \frac{\Phi_0 I_c}{2\pi}(1 - \cos\phi). \tag{4.204}$$

Thus, the current and the work done lag in phase, which is shown in Fig. 4.37.

Let us discuss the alternating current (AC) Josephson effect with a DC voltage bias. Consider $V = V_0$ (see Fig. 4.38), such that,

$$\phi(t) = \phi(0) + \frac{2\pi}{\Phi_0}V_0 t$$
$$I_s = I_c \sin\left(\frac{2\pi}{\Phi_0}V_0 t + \phi(0)\right) \tag{4.205}$$
$$I_s = I_c \sin\left(2\pi f_J t + \phi(0)\right).$$

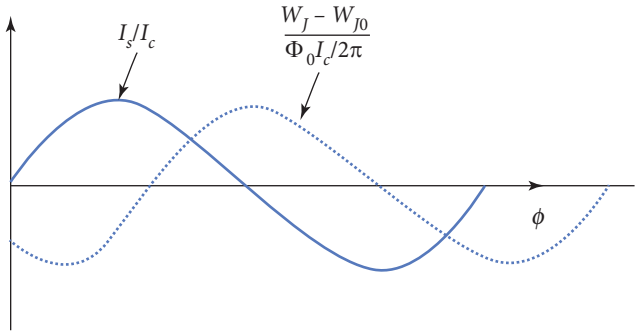


FIG. 4.37
$I_s - \phi$ curve (Bold line) and $W_J - \phi$ curve (Dotted line).

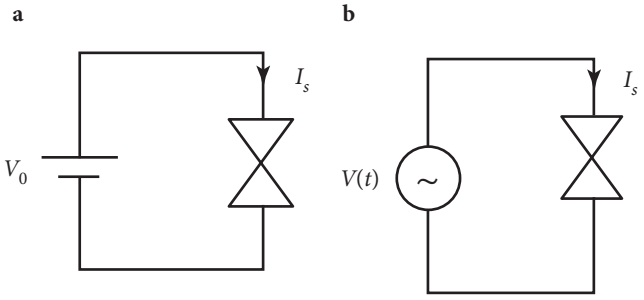


FIG. 4.38
(a) Schematic representation of the DC voltage bias setup. (b) Schematic representation of the AC voltage bias setup.

The Josephson frequency is given by

$$f_J = \frac{V_0}{\Phi_0} = \frac{2e}{h} V_0 = 484 \times 10^{12} V_0 \quad \text{(Hz)}. \tag{4.206}$$

Thus, a DC voltage of 10 μV causes an oscillation of about 5 GHz. This is the principle behind the Josephson microwave oscillator. The oscillator delivers a small power of the order of 10 nW for a typical $I_c = 1$ mA.

The rapid oscillations are due to quantum interference between the two junctions. The period is determined by the magnetic field that is required to generate one flux quantum in the loop. So the critical current maximum occurs at $\Phi/\Phi_0 = 0, \pm 1, \pm 2, \ldots . \pm n$. The observations of these oscillations are analogous to the double slit experiment in optics.

Let us look at the case of AC voltage bias. Consider an AC source (see Fig. 4.38),

$$V(t) = V_0 + V_s \cos \omega_s t$$
$$\phi(t) = \phi(0) + \frac{2\pi}{\Phi_0} V_0 t + \frac{2\pi V_s}{\Phi_0 \omega_s} \sin \omega_s t. \tag{4.207}$$

The current is given by

$$I_s = I_c \sin\left(\phi(0) + \frac{2\pi}{\Phi_0} V_0 t + \frac{2\pi V_s}{\Phi_0 \omega_s} \sin \omega_s t\right). \tag{4.208}$$

Hence, a constant current will occur when

$$2\pi f_J = n\omega_s \tag{4.209}$$

which yields,

$$V_0 = n\left(\frac{\Phi_0}{2\pi}\omega_s\right). \tag{4.210}$$

Thus, an AC voltage of 1 GHz frequency applied across the junction will provide a DC current at $V_0 = 0$, and at integral multiples of 2 μV.

4.15 SQUID

A superconducting quantum interference device (SQUID) measures the magnetic flux and the output voltage signal, which is a periodic function of the flux threading the superconducting loop. The flux variation that SQUID can measure is $10^{-5}\Phi_0$ with $\Phi_0 = 2.07 \times 10^{-15}$ Wb. Interested readers may consult the review article by Greenberg (1998).

The SQUID uses the Josephson effect phenomenon to measure extremely small variation of magnetic flux. Typically, a SQUID is a ring of superconductors interrupted by one or more Josephson junctions. There are usually two kinds of SQUIDs that are used, namely, a one junction device called the RF SQUID, and a device comprising two junctions which is called the DC SQUID.

A bias current I_b (for RF SQUID) or $2I_b$ (for DC SQUID) is applied, putting the operational point midway between the superconducting and the resistive behaviors. Shunt resistors are used to prevent hysteresis behavior in the $I - V$ curve shown in Fig. 4.39(a). A schematic view of the SQUID set up is shown in Fig. 4.39(b).

By fixing the current such that $I_b > I_c$, when an external magnetic flux, $\Phi_{ext} = B_{ext} \cdot A$, is threaded through the Josephson loop, the voltage drop across the Josephson junction will change. As the external flux increases or decreases, the voltage will change in a periodic manner with a period in multiples of flux quantum Φ_0 (see Fig. 4.40). Monitoring the change in voltage allows us to determine the magnetic flux that has been coupled to the SQUID loop. By using an external circuit, it is possible to "lock" the SQUID at a particular point in the $V - \Phi_0$ curve (say at the point P in Fig. 4.40). The feedback current

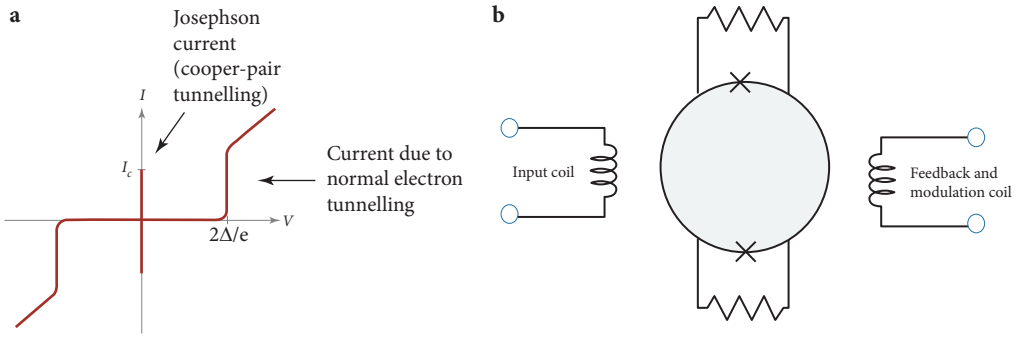


FIG. 4.39
(a) $I - V$ characteristic curve. (b) A typical SQUID setup is shown with an input coil (left) and a feedback coil (right).

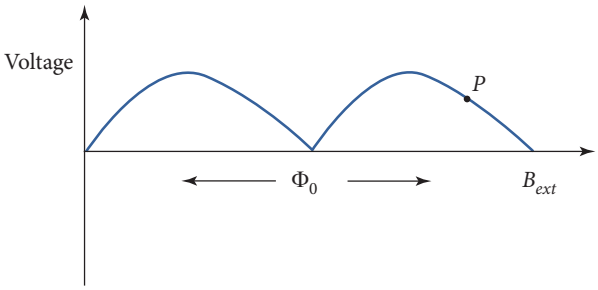


FIG. 4.40
$V - B_{ext}$ characteristic curve.

measured by the feedback loop [in Fig. 4.39(b)] is a measure of the externally applied flux. A SQUID is normally operated at the steepest part of the $V - \Phi_0$ curve to have $\frac{\partial V}{\partial \Phi}$ maximum. Thus, SQUIDs have been a key factor in the development and commercialization of ultrasensitive electric and magnetic measurement systems. In many cases, SQUIDs offer the possibility of measuring very small changes in magnetic flux, where no other methodology is available. Hence, SQUIDs have a wide range of sensing applications.

One disadvantage of using the SQUID is that the area of the detection coil is small (hence, it can thread only a small flux) and so is the associated self-inductance ($\approx 10^{-10}$ H). Increasing the area of the loop increases the self-inductance effects and hence introduces a hysteresis effect in the $I - V$ characteristics. Using a flux transformer coil, as shown in Fig. 4.41 aids in threading larger flux values. Thus, measurement of larger flux is also possible using a SQUID.

To complete our discussion, let us discuss the DC and RF SQUIDs. A DC SQUID differs from the RF SQUID with regard to two key elements, namely, with regard to biasing the Josephson junction and the number of Josephson junctions being two connected in parallel instead of one, as is the case for RF SQUID.

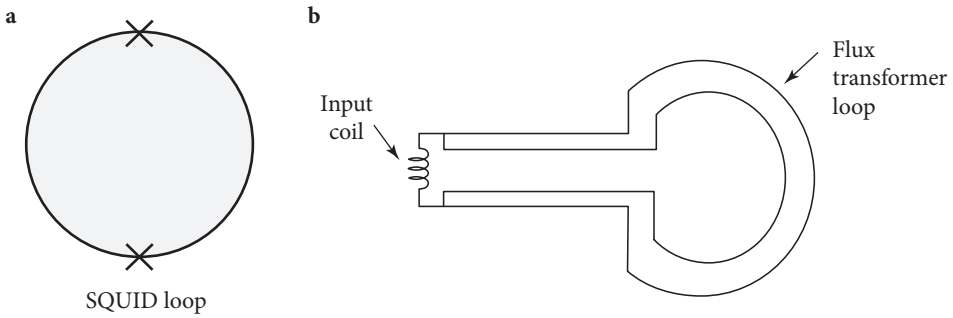


FIG. 4.41
(a) SQUID loop. (b) Flux transformer coil.

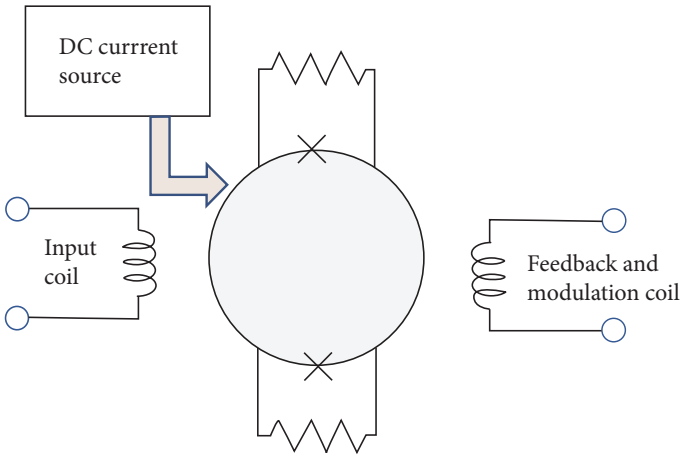


FIG. 4.42
A schematic setup for DC SQUID is shown.

Since there are two junctions which need to be matched within a few percentages, a job achieved by shunt resistors, it may be possible to operate a SQUID with mismatched junctions. But this performance will be sufficiently degraded. Similar to the RF SQUID, the input, the feedback, and the modulation coils are not wound around the SQUID loop, but inductively coupled to it (see Fig. 4.42). It is biased with a DC current approximately twice that of I_c, and thus develops a DC voltage across the junction and the shunt resistors. A change in magnetic flux applied through the SQUID loop induces a phase change of the condensate wave-function that enhances the current through one Josephson

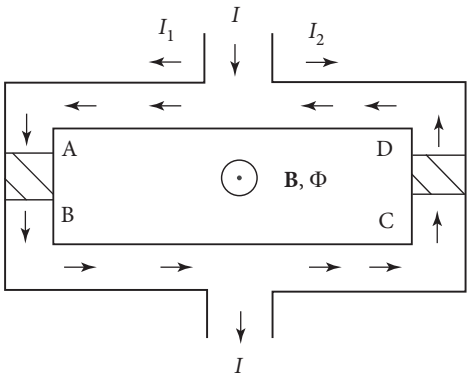


FIG. 4.43
DC SQUID setup with two junctions and a fux, Φ threading the device due to a magnetic field **B** is shown.

junction ($I_{net} = I_c + I_b$), and reduces the current through the other ($I_{net} = I_c - I_b$). As the external flux increases or decreases, the voltage will change periodically with a Φ_0 period. The current in the feedback circuit is the direct measure of flux change applied in the SQUID.

Commercially DC SQUIDs are mostly used. Besides sensors, SQUIDs have large applications in magnetoencephalography (MEG). Tumors or pathological tissues are detected by the change in the magnetic field by the neurons. For example, the magnetic field due to a single neuron can be estimated using the familiar Biot-Savart law in electromagnetics,

$$\mathbf{B} = \frac{\mu_0}{4\pi}\frac{\mathbf{Q}\times\mathbf{r}}{r^3}. \tag{4.211}$$

Healthy tissues produce $|\mathbf{B}|$ about 20 PT at 10 μm, whereas $|\mathbf{B}| \approx 100$ PT is produced by the pathological tissues (that are affected by tumours). Thus, it involves a charge $|\mathbf{Q}| = 20\ fA\cdot m$ for the former and 100 $fA\cdot m$ for the latter.

The DC SQUID setup is shown in Fig. 4.43. Let us assume for simplicity that the two junctions are identical. Also, the condensates are considered same on both sides of the junction. The total current is given by

$$I = I_1 + I_2 = I_c \sin\phi_1 + I_c \sin\phi_2 = 2I_c \cos\frac{\phi_1-\phi_2}{2}\sin\frac{\phi_1+\phi_2}{2}. \tag{4.212}$$

The corresponding phase difference around the loop is

$$\oint \nabla\theta\cdot d\mathbf{l} = (\theta_B-\theta_A)+(\theta_D-\theta_C)+(\theta_C-\theta_B)+(\theta_A-\theta_D) = 2\pi n. \tag{4.213}$$

From the definition, one may obtain the phases at the points A, B, C, and D (see Fig. 4.43) as

$$\theta_B - \theta_A = -\phi_1 - \frac{2\pi}{\Phi_0}\int_A^B \mathbf{A}\cdot d\mathbf{l}. \tag{4.214}$$

$$\theta_D - \theta_C = \phi_2 - \frac{2\pi}{\Phi_0}\int_C^D \mathbf{A}\cdot d\mathbf{l} \tag{4.215}$$

Introducing the supercurrent via the London equation, the phase differences may be written as

$$\theta_C - \theta_B = \int_B^C \nabla\theta\cdot d\mathbf{l} = -\lambda_L\int_B^C \mathbf{J}_s\cdot d\mathbf{l} - \frac{2\pi}{\Phi_0}\int_B^C \mathbf{A}\cdot d\mathbf{l} \tag{4.216}$$

$$\theta_A - \theta_D = \int_D^A \nabla\theta\cdot d\mathbf{l} = -\lambda_L\int_D^A \mathbf{J}_s\cdot d\mathbf{l} - \frac{2\pi}{\Phi_0}\int_D^A \mathbf{A}\cdot d\mathbf{l} \tag{4.217}$$

where λ_L is the penetration depth given by $\lambda_L = \frac{m}{ne^2}$. Adding Eqs. (4.214)–(4.217) and putting in Eq. (4.213) one gets,

$$\phi_2 - \phi_1 = 2\pi n + \frac{2\pi}{\Phi_0}\oint \mathbf{A}\cdot d\mathbf{l} + \lambda_L \int_B^C \mathbf{J}_s \cdot d\mathbf{l} + \lambda_L \int_D^A \mathbf{J}_s \cdot d\mathbf{l}. \tag{4.218}$$

Thus, the phase difference between the condensate wavefunction is given by

$$\phi_2 - \phi_1 = 2\pi n + \frac{2\pi\Phi}{\Phi_0} + \lambda_L \int_c \mathbf{J}_s \cdot d\mathbf{l}. \tag{4.219}$$

If the contour falls within the London penetration depth for a given superconducting sample, then the last term will vanish. Hence,

$$\phi_2 - \phi_1 = 2\pi n + 2\pi\frac{\Phi}{\Phi_0}. \tag{4.220}$$

So the total current is given by

$$I = 2I_c \cos\left(\frac{\pi\Phi}{\Phi_0}\right)\sin\left(\phi_1 + \frac{\pi\Phi}{\Phi_0}\right). \tag{4.221}$$

The flux piercing through the contour becomes

$$\Phi = \Phi_{ext} + LI_{cir} \tag{4.222}$$

where, I_{cir} is the circulating current $= \frac{I_1 - I_2}{2}$ and L is the self-inductance of the coil. The total flux thus can be written as

$$\Phi = \Phi_{ext} + LI_c \sin\left(\frac{\pi\Phi}{\Phi_0}\right)\cos\left(\phi_1 + \frac{\pi\Phi}{\Phi_0}\right). \tag{4.223}$$

For a given external flux, there is a range between I and Φ that satisfies these equations. One needs to know the maximum I that can be made to pass through the SQUID and still have zero voltage (critical current). Figure 4.44(a) describes the variation of the flux (Φ) in the circuit as a function of the external flux, Φ_{ext}. Furthermore, the sinusoidal variation of the maximum current as a function of the external flux is depicted in Fig. 4.44(b).

Next, for simplicity, let us consider the self-inductance of the coil to be zero ($L = 0$). Hence, $\Phi = \Phi_{ext}$, which yields the following expression for the current,

$$I = 2I_c \cos\left(\frac{\pi\Phi_{ext}}{\Phi_0}\right)\sin\left(\phi_1 + \frac{\pi\Phi_{ext}}{\Phi_0}\right). \tag{4.224}$$

The extremum occurs when $\frac{dI}{d\phi_1} = 0$, that is, when

$$\cos\left(\phi_1 + \frac{\pi\Phi_{ext}}{\Phi_0}\right) = 0 \tag{4.225}$$

or,

$$\sin\left(\phi_1 + \frac{\pi\Phi_{ext}}{\Phi_0}\right) = \pm 1. \tag{4.226}$$

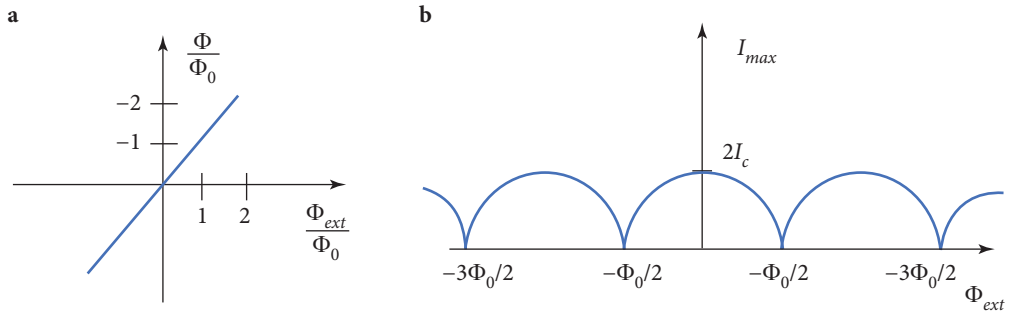


FIG. 4.44
DC SQUID without self inductance. (a) Φ vs Φ_{ext} curve. (b) I vs Φ_{ext} curve.

This yields the maximum current as

$$I_{max} = 2I_c \left| \cos\left(\frac{\pi \Phi_{ext}}{\Phi_0} \right) \right|. \tag{4.227}$$

The current has a maximum of $2I_c$ for $\Phi = n\Phi_0$, with n being an integer. Also, a minimum value of 0 occurs for $\Phi = (n + 1/2)\Phi_0$.

Next, generalize it for the case with self-inductance ($L \neq 0$, refer to Fig. 4.45). The two junctions are connected in parallel on a superconducting loop of inductance L. In this SQUID one applies a constant bias current $I_B > 2I_0$. The voltage across the SQUID oscillates with a period Φ_0 as one alters the external magnetic flux, Φ. There are two condensate pair wave functions which (weakly) interfere with the junction. If a magnetic flux passes through the loop, it changes the relationship between the phase difference across the junction. As a result, the critical current of the SQUID changes. The total current has a sinusoidal variation with the external flux in the argument of the variation.

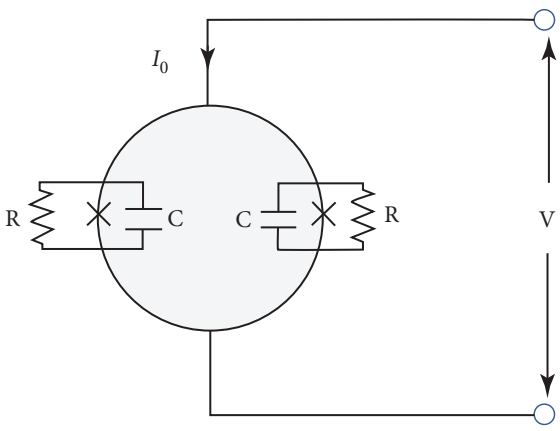


FIG. 4.45
DC SQUID with self inductance.

Finally, let us discuss the RF SQUID. The RF SQUID utilizes a single Josephson junction and flux is inductively coupled into the SQUID loop via an input coil which connects the SQUID to the experimental setup and an "RF" coil that is part of a high- Q resonant tank circuit, to note down the changes of the current in the SQUID loop. See Fig. 4.46. The tank circuit is driven

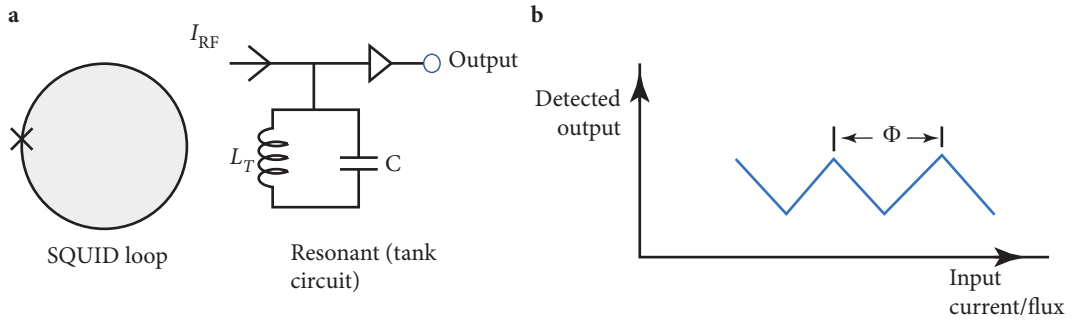


FIG. 4.46
(a) RF SQUID setup. (b) Output current vs input flux for RF SQUID (saw-tooth like pattern).

by an RF current, while the SQUID loop is inductively coupled to the inductance L_T of the LC tank circuit. The resulting RF voltage is periodic in the flux applied to the SQUID with period Φ_0. The output vs the input current and the output flux variations are shown in Fig. 4.46.

RF as a SQUID is actually a misnomer as no interference takes place. One could make a junction by pressing together two pieces of oxidized niobium Nb wire at right angles to each other.

4.16 APPENDIX

4.16.1 Fermi liquid theory

One last thing we wish to pursue very briefly is the essence of the Fermi liquid theory (FLT). The subject demands more attention than what it gets here. Since the motivation is to provide the readers with a bird's eye view of the topic, we introduce the basic concepts here. The mathematical derivations are beyond the purview of the presented discussion here.

The success of the single electron picture of metals rests on the FLT theory by Landau (1965). In the FLT, the weakly interacting electrons in a metal form quasiparticles that follow the usual fermionic properties. The FL theory is remarkably successful, despite a few failures, especially where the interactions between electrons are quite strong and hence cannot be ignored.

In order to understand the conditions of applicability of FLT, we can define the mean electronic separation, r_s and density ρ_e via

$$\rho_e \times \frac{4}{3}\pi r_s^2 = 1.$$

The average Coulomb interaction per electron is given by

$$\langle PE\rangle \sim \frac{1}{2}\frac{e^2}{4\pi\epsilon_0 r_s}.$$

Furthermore, the mean kinetic energy is

$$\langle KE\rangle \sim \frac{\hbar^2}{2m}\frac{1}{r_s^2}$$

$$\frac{\langle PE\rangle}{\langle KE\rangle} \sim \frac{e^2}{8\pi\epsilon_0 r_s}\times\frac{2mr_s^2}{\hbar^2} = \frac{me^2}{4\pi\epsilon_0\hbar^2}r_s = \frac{r_s}{a_0}. \tag{4.228}$$

Thus, the ratio of the mean potential to the mean kinetic energies is of the order of electron-electron separation, which further can be measured in units of the Bohr radius, a_0. Some typical values are

$$r_s/a_0 = 1.9 \quad \text{for Be}$$
$$= 5.6 \quad \text{for Cs}.$$

Thus, the ratio r_s/a_0 is large, and for Wigner crystals, $r_s/a_0 > 20$. This means that the mean electronic separation is larger than the Bohr radius, which in turn, advocates that FLT will provide an appropriate description of electronic systems.

The need for a Fermi liquid theory in the first place comes from the fact that the specific heat of electrons, according to the classical (and non-interacting) theory, should be $3k_B/2$; however, one usually gets a value much lower than this. Also, the susceptibility, χ of the free moments deviates significantly from the classical behavior, namely, $\chi \sim \frac{1}{T}$. Thus, the non-interacting theory is found to be insufficient in a variety of cases.

These puzzles are solved by the FLT, which says that only a small fraction of the electrons near the Fermi surface takes part in contributing to the physical observable, such as the specific heat or the susceptibility etc. These electrons are promoted from just inside the Fermi surface to just outside, and such a phenomenon is known as particle-hole excitations. The same formalism as that of the non-interacting systems will continue to be valid, except for the bare electronic mass, m is to be replaced by the effective mass, m^*. Only the electrons within k_BT of the Fermi energy contribute to the specific heat so that the specific heat is proportional to T and is small. Also, the electrons within an energy slice $\mu_B B$ of the Fermi surface can be magnetized with a moment, μ_B leading to a temperature independent Pauli susceptibility. These dependencies were matched with experiments on metals and were found to have good support.

However, though FLT solved these riddles, there remain questions that are yet to be answered, such as how does a non-interacting theory explain the behavior of systems where the interaction effects are important. The answer provided by Landau rests on the concept of "*adiabatic continuity*." The adiabatic continuity indicates that the labels associated with eigenstates are more robust against perturbation than the eigenstates themselves. To have clarity in understanding, consider a one-dimensional box whose eigenstates are given by sinusoidal functions. More importantly it is expressible in terms of the number of nodes (that is the number of zeros), that is how many zeros it has (see left panel

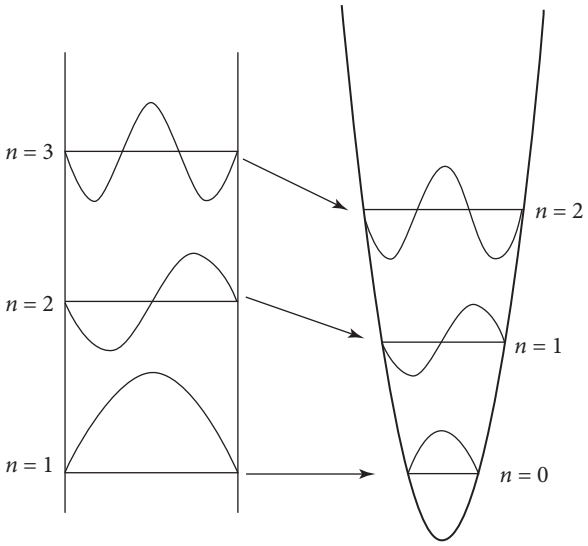


FIG. 4.47
The one-to-one correspondence between the states of a particle in an infinite potential and a particle subjected to a harmonic potential is shown. The arrows denote the corresponding energy states for which the number of nodes remains fixed.

of Fig. 4.47). The larger the energy, greater the number of nodes. Now, if we perturb the system by a small harmonic oscillation potential, $V(x) = \frac{1}{2}\epsilon x^2$ (ϵ is small). The new eigenstates of the system will no longer remain simple sine waves, but involve a mixing of all the eigenstates of the original unperturbed problem. However, the number of nodes remains a good indicator to describe the eigenstates of this more complex (interacting) problem. The correspondence is shown schematically in Fig. 4.47.

Landau applied this idea to the interacting gas of electrons. He imagined that a turning on the interaction effects slowly, and observed how the eigenstates of the system behave. He postulated that there will be a one-to-one mapping of the low energy states of that of the interacting system with that of the non-interacting Fermi gas. The assumption is that the good quantum numbers associated with the excitations of the non-interacting systems will remain valid, even after the interactions are turned on. Just as Pauli's exclusion principle holds for the non-interacting electrons, this would remain so even in the presence of interactions. We can therefore retain the picture of the excitation of particles and holes, with them carrying the same quantum numbers as their electronic counterparts of the free Fermi gas. These excitations are called quasiparticles, whose wavefunctions and the energies are different from those of the corresponding electrons in the non-interacting problem. These quasiparticles are core to the understanding of FLT, and account for the observed temperature dependencies of the specific heat and susceptibility, with the only requirement of having a Fermi surface.

The energies of the quasiparticles are not the same as those of the non-interacting electrons. In Landau's theory, the modified energy appears through two terms. Their origin can be understood as follows.

i. First, when a quasiparticle moves, there will be a backflow in the filled Fermi sea (due to momentum conservation), as the quasiparticles "*push*" the ground state out of the way. This modifies the inertial mass of the quasiparticles, m is replaced by m^*.
ii. Second, the quasiparticle energy must depend on the distribution of other quasiparticles, which Landau had included via some "f" function, where the total energy of the interacting system can be

expanded as a functional of variation in density, namely,

$$E = \sum_{\mathbf{k},\sigma} \frac{p_F}{m^*}(\hbar k - p_F)\delta n_{\mathbf{k},\sigma} + \frac{1}{2}\sum_{\mathbf{k},\mathbf{k}',\sigma,\sigma'} f_{\mathbf{k}\sigma,\mathbf{k}'\sigma'}\delta n_{\mathbf{k},\sigma}\delta n_{\mathbf{k}',\sigma'},$$

where p_F denotes the Fermi momentum, and $\delta n_{\mathbf{k},\sigma}$ ($\delta n_{\mathbf{k}',\sigma'}$) denotes the change in the Fermi distribution function. We skip a detailed discussion of the above and encourage the readers to look at more specialized articles (Kinza, 2018). Also, for a short review, see Neilson (1996).

Let us look at the additional constraints on the validity of Landau FLT. The quasiparticles must be long lived in the vicinity of the Fermi sphere. In fact, the inverse of the QP lifetime, τ_{el} can be evaluated by using Fermi's golden rule,

$$\frac{1}{\tau_{el}} = \frac{2\pi}{\hbar}\sum_f |V_{fi}|^2\delta(\varepsilon - \varepsilon_p). \tag{4.229}$$

Assuming V_{fi} to be a constant and independent of the energy transfer ε, one can convert the above sum into an integral, and hence, integrate over the final (continuum) states. This goes as

$$\frac{1}{\tau_{el}} \sim \varepsilon^2.$$

Also, τ_{el} is found to have the temperature dependence of the form,

$$\frac{1}{\tau_{el}} \sim T^2.$$

This decay rate is important in transport properties of metals, as it yields a T^2 resistivity.

Here we present a few examples, which from a theoretical perspective generate non-Fermi liquid (NFL) characteristics in interacting electronic systems. These are

i. metals close to the quantum critical point, where a phase transition occurs close to $T = 0$, the quasiparticles scatter so strongly that they cease to follow the usual results of FLT.
ii. Metals in one dimension called Luttinger liquid. In 1D, electrons are unstable and decay into two separate particles (spinons and holons) that carry the spin and charge, respectively.
iii. Disordered Kondo metals. Here, the scattering from the magnetic impurities is too strong to allow for stable quasiparticles to form.

4.17 SUMMARY AND OUTLOOK

Let us include a brief recap of the topics discussed. First, and foremost, both the semiconductors and the superconductors possess a spectral gap in their electronic band structure. The question is how are these two gaps distinguished? Evidently, the values differ by three orders of magnitude, that is, a semiconducting gap is usually of the order of an eV, while the superconducting gap typically lies in the regime of a few meV. There is a more important difference, where the ground state of a superconductor

is a coherent state, and devoid of any unpaired electron, where all the (Cooper) pairs have amplitude and phase coherence. The valence band of a semiconductor has no such feature and comprises energy levels that disperse differently across the Brillouin zone.

By and large, we have confined ourselves to the study of conventional superconductors. We have highlighted several distinguishing properties of superconductors, such as the Meissner effect, the isotope effect, distinction between perfect conductors and superconductors, flux quantization, penetration depth, and the coherence length, type-I and type-II superconductors, etc. Hence, we have discussed the magnetic and thermodynamic properties of superconductors, where the former shows a phase diagram as a function of the external magnetic field that encodes a phase transition, either directly from a superconductor to a metal, or the same intervened by a mixed phase that admits the magnetic flux lines (in unit of a constant flux quantum) to penetrate the sample. Most of the well-known conventional superconductors are of the latter variety. The thermodynamic properties, such as the specific heat shows a jump of fixed magnitude at the transition point and are a generic feature of a superconducting phase transition.

Hence, we have embarked on the famous BCS theory, where we have discussed Cooper's instability that leads to the formation of bound pairs of electrons mediated via phonons and is known to provide the microscopic origin of the pairing phenomena. This is followed by a detailed description which yields an appropriate estimation of the temperature dependence of the spectral gap, and hence the transition temperature, where the latter provides an enumeration of different superconductors. Among the crucial properties, we have outlined the calculation of the specific heat, computed the behavior of the specific heat and the jump therein at the transition temperature and discussed the Meissner effect.

We have also included a brief recap of the Ginzburg-Landau (GL) theory, which is a phenomenological theory of superconductivity proposed by Ginzburg and Landau in the early 1940s and is broadly applicable to all second-order phase transitions. Apart from the discussion of the two GL equations, we have estimated two important length scales that characterize a superconductor, namely, the coherence length and the penetration depth. By no means, the above discussion is complete and the readers are encouraged to look up more detailed notes of the subject.

We have also described in brief a few experimental methods for estimating the magnitude of the energy gap in superconductors. These are electromagnetic, ultrasound absorption experiments, and measuring the tunneling spectra in junction systems, involving a normal metal and a superconductor (NS junction), or a metal between two superconductors (SNS junction), etc.

To wind down our discussion on superconductors, we have mentioned very briefly the unconventional superconductivity that has been discovered over the last two and a half decades. In particular, we have touched upon the cuprate superconductors which have created enormous excitement owing to several unconventional features. The lack of a well defined starting point for a microscopic theory to be developed, at least in the underdoped regime, eluded a conclusive understanding of the pairing phenomena

and the superconducting state. About two decades later, Fe-based superconductors were discovered with moderately high transition temperatures (although not as high as the cuprates), whose normal phase again hinted toward a non-Fermi liquid state. Furthermore, their physical properties differ from conventional superconductors on a number of counts. Thus, there remained several issues that necessitate going beyond the conventional (BCS) paradigm. Once again the readers are encouraged to look at literature, including some excellent review articles on the subject of unconventional superconductivity that transcends beyond the cuprates and Fe-based pnictides or chalcogenides.

For completeness, we discuss the applications of superconductivity. In this connection, different types of superconducting junctions are illustrated, which eventually leads us to the topic of the Josephson effect. The AC and DC Josephson effects are described along with a derivation of the current voltage relationships. Practical examples of the Josephson junctions are denoted by the SQUIDs. Different types of SQUIDs and a brief discussion is made on their utility in detecting extremely small values of magnetic flux.

Further, in the appendix, we have included a crisp introduction to the Fermi liquid theory. The main inferences of the theory are clearly violated for the cases of both the Fe-based and cuprate superconductors owing to strong electronic correlations.

REFERENCES

Alford, M. G., Berges, J., and Rajagopal, K., Nucl. Phys. B **571**, 269 (2000).
Barzykin, V. and Gorkov, L. P., Phys. Rev. Lett. **89**, 227002 (2002).
Becker, R., Heller, G., and Sauter, F., "Uber die Stromverteilung in einer supraleitenden Kugel," Z. Phys. **85**, 772–87 (1933).
Blatt, F. J., *Modern Physics* (McGraw Hill, 1992).
Burkhardt, H. and Rainer, D., Ann. Phys. **3**, 181 (1994).
Buzdin, A. I., Rev. Mod. Phys. **77**, 935 (2005).
Capan, C. *et al.*, Phys. Rev. B **70**, 134513 (2004).
Casalbouni, R. and Narduli, G., Rev. Mod. Phys. **76**, 263 (2004).
Chandrasekhar, B. S., Appl. Phys. Lett. **1**, 7 (1962).
Chandrasekhar, B. S. and Hulm, J. K., J. Phys. Chem. Solids **7**, 259 (1958).
Clogston, A. M., Phys. Rev. Lett. **9**, 266 (1962).
Combescot, R. and Mora, C., Eur. Phys. J. B **44**, 189 (2005).
Cooper, L. N., Phys. Rev. **104**, 1189 (1956).
Dalidovich, D. and Yang, K., Phys. Rev. Lett. **93**, 247002 (2004).
Deaver, B. S. and Fairbank, W. M., Phys. Rev. Lett. **7**, 43 (1961).
Dupuis, N., Phys. Rev. B **51**, 9074 (1995).
Eisaki, H. *et al.*, J. Phys. Soc. Jpn **77**(Suppl. C), 36 (2008).
Fetter, A. L. and Walecka, J. D., *Quantum Theory of Many Particle Systems* (McGraw Hill, 1971).

Fetter, L. and Hohenberg, P. C., *Superconductivity*, edited by R. D. Parks, 1st ed. (CRC Press, 1969).
Fröhlich, H., Phys. Rev. **79**, 845 (1950).
Fulde, P. and Ferrell, R. A., Phys. Rev. **135**, A550 (1964).
Giaever, I, See the Nobel lecture by I. Giaever available at https://www.nobelprize.org/prizes/physics/1973/giaever/lecture/.
Ginzburg, V. L. and Landau, L. D., Zh. Eksp. Teor. Fiz. **20**, 1064 (1950). English translation in: L. D. Landau, Collected papers (Oxford: Pergamon Press, 1965), p. 546.
Gloos, K. *et al.*, Phys. Rev. Lett. **70**, 501 (1993).
Golubov, A. A., Yu. Kupriyanov, M., and Ilíchev, E., Rev. Mod. Phys. **76**, 411 (2004).
Greenberg, Ya. S., Rev. Mod. Phys. **70**, 175 (1998).
Hulm, J. K. and Matthias, B. T., Phys. Rev. **82**, 273 (1951); Phys. Rev. **89**, 439 (1953); Phys. Rev. **87**, 799 (1952); B.T. Matthias, Phys. Rev. **92**, 874 (1953).
Huxley, A. D. *et al.*, J. Phys.: Condens. Matter **5**, 7709 (1993).
Jarozynski, J. *et al.*, Phys. Rev. B **78**, 064511 (2008).
Kakuyanagi, K. *et al.*, Phys. Rev. Lett. **94**, 047602 (2005).
Kamerlingh Onnes, H., Research Notebooks Vol. 56, 57; H. Kamerlingh Onnes Archive, Boerhaave Museum, Leiden, the Netherlands. H. Kamerlingh Onnes, Comm. Phys. Lab. Univ. Leiden **122**, 124 (1911), H.Kamerlingh Onnes, Comm. Leiden, I913, Nr.I33atoI33d.
Kamihara, Y., Watanabe, T., Hirano, M., and Hosno, H., J. Am. Chem. Soc. **130**, 3296 (2008).
Kinza, M., *Theory of Fermi Liquids in Metals* (Springer, 2018).
Kittel, C., *Introduction to Solid State Physics* (Wiley, 2004).
Kumagai, K., Kakuyanagi, K., Takashima, S., Nohara, M., Takagi, H., and Matsuda, Y., J. Supercond. Novel. Magn. **19**, 1 (2006).
Kuroki, K. *et al.*, New J. Phys. **11**, 25017 (2009).
Landau, L. D., "The theory of a fermi liquid," in *Collected Papers of L.D. Landau*, edited by D. Ter Haar (Pergamon, 1965).
Larkin, A. and Ovchinnikov, Y. N., Sov. Phys. JETP **47**, 1136 (1964), available at https://www.osti.gov/biblio/4653415.
Lee, I. J., Naughton, M. J., Danner, G. M., and Chaikin, P. M., Phys. Rev. Lett. **78**, 3555 (1997).
London, F. and London, H., Proc. R. Soc. A: Math., Phys. Eng. Sci. **149**, 71 (1935).
Lorenz, B., Sasmal, K., Chaudhury, R. P., Chen, X. H., Liu, R. H., Wu, T., and Chu, C. W., Phys. Rev. B **78**, 012505 (2008).
Machida, K. and Nakanishi, H., Phys. Rev. B **30**, 122 (1984).
Makhlin, Y., Schön, G., and Shnirman, A., Rev. Mod. Phys. **73**, 357 (2001).
Maki, K. and Tsuento, T., Prog. Theor. Phys. **31**, 945 (1964).
Manalo, S. and Klein, U., J. Phys.: Condens. Matter **12**, L471 (2000).
Martin, C. *et al.*, Phys. Rev. B, **71**, 020503(R) (2005).
Mathias, B. T., Compton, V. B., and Corenzwit, E., J. Phys. Chem. Solids **19**, 130 (1961).
Matsuda, Y. and Shimahara, H., J. Phys. Soc. Jpn. **76**, 051005 (2007).
Matsuo, S., Higashitani, S., Nagato, Y., and Nagai, K., J. Phys. Soc. Jpn. **67**, 280 (1998).

Mitrović, V. F., Horavatić, H., Berthier, C., Knebel, G., Lapertot, G., and Floquet, J., Phys. Rev. Lett. **97**, 117002 (2006).
Nam, M. S. *et al.*, J. Phys.: Condens. Matter **11**, L477 (1999).
Neilson, D., Aus. J. Phys. **49**, 79–102 (1996).
Parish, M. M., Hu, J., and Bernevig, B. A., Phys. Rev. B **78**, 144514 (2008).
Pieri, P., Neilson, D., and Strinati, G. C., Phys. Rev. B **75**, 113301 (2007).
Radovan, H. A. *et al.*, Nature **425**, 51 (2003).
Ramakrishnan, T. V. and Rao, C. N. R., *Superconductivity Today* (University Press, 1999).
Randeria, M. and Campuzano, J. C., "Varenna lectures," arXiv:9709107.
Reid, J. P. *et al.*, Phys. Rev. Lett. **109**, 1 (2012).
Reif, F., *Fundamentals of Statistical and Thermal Physics* (McGraw Hill, 1965).
Sakurai, J. J. and Napolitano, J., *Modern Quantum Mechanics*, 2nd ed. (Addison-Wesley, 2011).
Shen, Z.-X *et al.*, Phys. Rev. Lett. **70**, 3999 (1993).
Shimahara, H., Phys. Rev. B **50**, 12760 (1994).
Shimahara, H., J. Phys. Soc. Jpn. **67**, 1872 (1998).
Shimahara, H., Physica B **259**, 492 (1999).
Shimahara, H. and Rainer, D., J. Phys. Soc. Jpn. **66**, 3591 (1997).
Shimuzu, K. *et al.*, Nature **412**, 316 (2001).
Singleton, J. *et al.*, J. Phys.: Condens. Matter **12**, L641 (2000).
Stewart, G. R., Rev. Mod. Phys. **83**, 1589 (2011).
Takada, S. and Izuyama, T., Prog. Theor. Phys. **41**, 635 (1969).
Takahashi, H. *et al.*, J. Phys. Soc. Jpn **77**(Suppl. C), 78 (2008).
Takahashi, H., Igawa, K., Arii, K., Kamihara, Y., Hirano, M., and Hosno, M., Nature **453**, 376 (2008).
Timusk, T. and Statt, B., Rep. Prog. Phys. **62**, 61 (1999).
Tinkham, M., *Introduction to Superconductivity* (Dover Books, 1973).
Tinkham, M. and Glover, R. E., Phys. Rev. **110**, 778 (1958).
Varma, C. M., Rev. Mod. Phys. **92**, 031001 (2020).
Vorontsov, A. B. and Graf, M. J., Phys. Rev. B **74**, 172504 (2006).
Vorontsov, A. B., Sauls, J. A., and Graf, M. J., Phys. Rev. B **72**, 184501 (2005).
Vrontosov, A. B., Vekhter, I., and Graf, M. J., Phys. Rev. B **78**, 180505(R) (2008).
Wang, C., Europhys. Lett. **83**, 67006 (2008).
Wang, Q., Chen, H.-Y., Hu, C.-R., and Ting, C. S., Phys. Rev. Lett. **96**, 117006 (2006).
Watanabe, T. *et al.*, Phys. Rev. B, **70**, 020506 (2004).
Yanase, Y., New J. Phys. **11**, 055056 (2009).
Yang, K., Phys. Rev. B **63**, 140511 (2001).
Yang, K. and Sondhi, S. L., Phys. Rev. B **57**, 8566 (1998).
Yu, T. *et al.*, npj. Quantum. Mater. **5**, 46 (2020).
Zhao, J. *et al.*, Nature Mater. **7**, 953 (2008).
Zhu, X., Littlewood, P. B., Hybertsen, M. S., and Rice, T. M., Phys. Rev. Lett. **74**, 1633 (1995).

INDEX

Modern Perspectives in the Study of Electronic Systems focuses on recent developments in selected critical areas of condensed matter physics. The book blends theoretical developments and experimental discoveries of materials, helping the reader understand different phenomena. It studies the Quantum Hall Effect (integer and fractional) and superconductivity from phenomenological and microscopic perspectives. Magnetism and magnetic phenomena in solids via approximate solutions of the Hamiltonian model are also addressed. The book also discusses the topological aspects of materials that have fundamentally changed the study of condensed matter physics in the recent times.

This timely and practical book:

- Introduces topology that has governed the study of two-dimensional Dirac materials and led to the discovery of three-dimensional materials
- Includes illustrations and plots for better visualization and understanding of the book content
- Uses a hands-on approach to teach subjects that have influenced teaching and research worldwide

Modern Perspectives in the Study of Electronic Systems is an excellent tutorial for senior undergraduate students, graduate students, and early-career researchers. It will be helpful for teachers in teaching courses on specialized topics in condensed matter physics.

Saurabh Basu, Ph.D., joined the Indian Institute of Technology (IIT) Guwahati, India, in 2003, where he is a Senior Professor of Physics. He supervised many Ph.D. students and is the author of over 100 scientific publications across many peer-reviewed journals. Dr. Basu's teaching areas include advanced topics in quantum mechanics, statistical physics, mathematical methods, and condensed matter physics.

Sourav Chattopadhyay, Ph.D., has been an Assistant Professor in the Department of Physics at ICFAI University Tripura since 2022. He completed his master's and doctoral programs at IIT Guwahati. He pursued Physics (Hons.) from the Scottish Church College, Kolkata. His primary research interest is computational statistical mechanics both equilibrium and out-of-equilibrium with an emphasis on magnetism. He has a keen interest in numerical simulations of physics problems and coding.

AIP Publishing

9 780735 422513