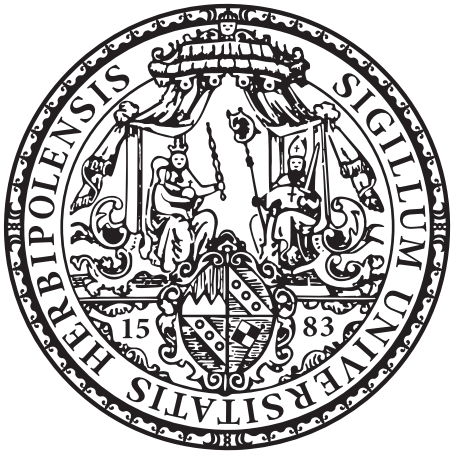
Correlated Topological Materials

Dissertation zur Erlangung des
naturwissenschaftlichen Doktorgrades
der Julius-Maximilians-Universität Würzburg



vorgelegt von
Stefan Jürgens
aus Hamm

Würzburg 2017

Eingereicht am: 10.04.2017 .....
bei der Fakultät für Physik und Astronomie

1. Gutachter: Prof. Dr. Björn Trauzettel .....

2. Gutachter: Prof. Dr. Giorgio Sangiovanni .....

der Dissertation

Vorsitzende: Prof. Dr. Johanna Erdmenger .....

1. Prüfer: Prof. Dr. Björn Trauzettel .....

2. Prüfer: Prof. Dr. Giorgio Sangiovanni .....

3. Prüfer: Prof. Dr. Friedrich Reinert .....

im Promotionskolloquium

Tag des Promotionskolloquiums: 21.07.2017 .....

Doktorurkunde ausgehändigt am: .....

Zusammenfassung

Das Thema dieser Doktorarbeit ist die Kombination von topologisch nicht-trivialen (TnT) Phasen mit Coulomb Wechselwirkungseffekten, die zwischen den Elektronen eines Systems der kondensierten Materie auftreten. Ein Schwerpunkt wird sowohl auf die sich ergebenen Vorteile als auch möglichen Nachteile gelegt, z.B. bezogen auf den topologischen Schutz in der Gegenwart von starker Wechselwirkung.

Die topologischen Effekte in der Physik werden in Kap. 2 vorgestellt. Basierend auf der topologischen Bandtheorie führen wir die topologischen Materialien ein, inklusive Chern Isolatoren, topologischer Isolatoren (TIs) in zwei und drei Dimensionen und Weyl Halbmetallen (WSMs). Die Formalismen für eine kontrollierte Behandlung der Coulomb Korrelationen werden in Kap. 3 präsentiert, beginnend mit der topologischen Feldtheorie. Die Random Phase Approximation (RPA) bietet einen störungstheoretischen Ansatz, während im Bereich der starken Wechselwirkung die Theorie des Quanten-Hall-Ferromagnetismus greift. Wechselwirkende Systeme in einer Dimension sind besonders und werden von uns als Luttinger Flüssigkeit beschrieben. Das Kapitel endet mit einem Überblick über die zu erwartenden Vorteile und Möglichkeiten einer Kombination von Topologie und Korrelationen in Kap. 3.3.

Diese Ideen werden im Forschungsteil weiter ausgeführt. In Kap. II beschäftigen wir uns mit schwach wechselwirkenden, zweidimensionalen (2D) TIs, beschrieben durch das Bernevig-Hughes-Zhang (BHZ) Modell. Dies ist z.B. anwendbar für Quantentrogstrukturen basierend auf HgTe/CdTe oder InAs/GaSb. Die Bandstruktur im Volumen ist hier gegeben durch eine Mischung aus linearen Dirac and quadratischen Schrödinger Fermionen. Wir untersuchen die Anregungen für kleine Energien mittels RPA und finden ein neues Interbandplasmon, das aus der Kombination von Dirac und Schrödinger Physik entspringt und in den jeweiligen Grenzfällen nicht existiert. Während es bereits im undotierten Fall zu finden ist, konkurriert es bei endlicher Dotierung mit dem gewöhnlichen Intrabandplasmon. Die gebrochene Teilchen-Loch Symmetrie in HgTe Quantentrögen ermöglicht eine Trennung der Beiden im Anregungsspektrum, für experimentell zugängliche Parameter in der richtigen Größenordnung für Raman- oder Elektronenspektroskopie. Das wechselwirkende Anregungsspektrum des Bulk zeigt hier klare Unterschiede zwischen dem topologisch trivialen und nicht-trivialen Regime. Ein noch deutlicheres experimentelles Signal erwarten wir von der optischen Leitfähigkeit des Systems, welche somit eine quantitative Möglichkeit bietet, zwischen den topologischen Phasen eines 2D TIs mittels einer Bulk Messung zu unterscheiden.

In Kap. III untersuchen wir stark-wechselwirkende Systeme, die sich in einem geordneten, Quanten-Hall-Ferromagnetischen (QHFM) Zustand befinden. Dieser Zustand kann auch in schwach-wechselwirkenden Systemen in einem starken magnetischen Feld auftreten. In diesem Fall bilden die Elektronen flache Landau-Niveaus mit minimierter kinetischer Energie aus, sodass die Coulomb Wechselwirkung dominiert. Solche Systeme bilden die Klasse der Quanten-Hall topologischen Isolatoren (QHTIs): TnT Zustände bei endlichem Magnetfeld, deren gegenläufige Randzustände nicht durch Zeitumkehr, sondern durch räumliche oder spin Symmetrien geschützt werden. Infrage kommende Materialien sind 2D TIs wie HgTe Heterostrukturen oder Graphen. Unsere Analyse fokussiert sich auf die Umgebung des topologischen Phasen-

übergangs, in der sich das System in dem stark-wechselwirkenden QHFM Zustand befindet. Hier kann die Physik sowohl des Bulks als auch die der Randzustände mittels des nichtlinearen σ-Modells für den Ordnungsparameter beschrieben werden. Wir zeigen, dass eine effektive, kontinuierliche U(1) Symmetrie für den topologischen Schutz sorgt. Ist diese Symmetrie erhalten, bleibt die TnT Phase auch für starke Wechselwirkungen bestehen und die Randzustände bilden eine helikale Luttinger Flüssigkeit. Diese kann durch das magnetische Feld stark beeinflusst werden, sodass die effektive Wechselwirkungsstärke zwischen schwach wechselwirkend für vernachlässigbares Feld, $\mathcal{K} \approx 1$, und stark wechselwirkend am topologischen Phasenübergang, $\mathcal{K} \to 0$, variiert.

Im letzten Kap. IV erforschen wir, ob WSM- and drei-dimensionale TI-Phasen zeitgleich und am selben Ort existieren können, mit einem hybriden Oberflächenzustand an der gemeinsamen Grenzfläche. Ein entsprechender Austausch zwischen den Materialien kann durch Coulomb Wechselwirkung oder eine räumliche Bandüberlagerung realisiert werden. Ein Tunnelkopplungsansatz erlaubt es uns, den hybriden Oberflächenhamiltonian analytisch herzuleiten und ermöglicht so eine detaillierte Analyse der Oberflächendispersionsrelation. Im Fall von spinsymmetrischer Kopplung entstehen weitere Diracpunkte aus der Kombination eines einzelnen Diracpunktes und eines Fermibogens. Bricht man die Spinsymmetrie durch die Kopplung entstehen Bandlücken in der Oberflächendispersion und die ursprünglichen Diracpunkte werden spinpolarisiert. Wir schlagen experimentelle Umsetzungen dieser hybriden Physik vor, z.B. kompressiv verspanntes HgTe oder auch Heterostrukturen aus TI and WSM Materialien.

Summary

The topic of this PhD thesis is the combination of topologically non-trivial (TnT) phases with correlation effects stemming from Coulomb interaction between the electrons in a condensed matter system. Emphasis is put on both emerging benefits as well as hindrances, e.g. concerning the topological protection in the presence of strong interactions.

The physics related to topological effects is established in Sec. 2. Based on the topological band theory, we introduce topological materials including Chern insulators, topological insulators (TIs) in two and three dimensions as well as Weyl semimetals (WSMs). Formalisms for a controlled treatment of Coulomb correlations are presented in Sec. 3, starting with the topological field theory (TFT). The Random Phase Approximation (RPA) is introduced as a perturbative approach, while in the strongly interacting limit the theory of quantum Hall ferromagnetism (QHFMism) applies. Interactions in one dimension are special, and are treated through the Luttinger liquid description. The section ends with an overview of the expected benefits offered by the combination of topology and interactions, see Sec. 3.3.

These ideas are then elaborated in the research part. In Chap. II, we consider weakly interacting two-dimensional (2D) TIs, described by the Bernevig-Hughes-Zhang (BHZ) model. This is applicable, e.g., to quantum well (QW) structures made of HgTe/CdTe or InAs/GaSb. The bulk band structure is here a mixture stemming from linear Dirac and quadratic Schrödinger fermions. We study the low-energy excitations in RPA, where a new interband plasmon emerges due to the combined Dirac and Schrödinger physics, which is absent in the separate limits. Already present in the undoped limit, one finds it also at finite doping, where it competes with the usual intraband plasmon. The broken particle-hole (p-h) symmetry in HgTe QWs allows for an effective separation of the two in the excitation spectrum for experimentally accessible parameters, in the right range for Raman or electron loss spectroscopy. The interacting bulk excitation spectrum shows here clear differences between the topologically trivial (TT) and TnT regime. An even stronger signal in experiments is expected from the optical conductivity of the system. It thus offers a quantitative way to identify the topological phase of 2D TIs from a bulk measurement.

In Chap. III, we study a strongly interacting system, forming an ordered, quantum Hall ferromagnetic (QHFM) state. The latter can arise also in weakly interacting materials with an applied strong magnetic field. Here, electrons form flat Landau levels (LLs), quenching the kinetic energy such that Coulomb interaction can be dominant. These systems define the class of quantum Hall topological insulators (QHTIs): TnT states at finite magnetic field, where the counter-propagating edge states are protected by a symmetry (spatial or spin) other than time-reversal ($\boldsymbol{\mathcal{T}}$). Possible material realizations are 2D TIs like HgTe heterostructures and graphene. In our analysis, we focus on the vicinity of the topological phase transition, where the system is in a strongly interacting QHFM state. The bulk and edge physics can be described by a nonlinear σ-model for the collective order parameter (OP) of the ordered state. We find that an emerging, continuous U(1) symmetry offers topological protection. If this U(1) symmetry is preserved, the TnT phase persists in the presence of interactions, and we find a helical Luttinger liquid at the edge. The latter is highly tunable by the magnetic field,

where the effective interaction strength varies from weakly interacting at zero field, $\mathcal{K} \approx 1$, to diverging interaction strength at the phase transition, $\mathcal{K} \to 0$.

In the last Chap. IV, we investigate whether a WSM and a three-dimensional (3D) TI phase can exist together at the same time, with a combined, hybrid surface state at the joint boundaries. An overlap between the two can be realized by Coulomb interaction or a spatial band overlap of the two systems. A tunnel coupling approach allows us to derive the hybrid surface state Hamiltonian analytically, enabling a detailed study of its dispersion relation. For spin-symmetric coupling, new Dirac nodes emerge out of the combination of a single Dirac node and a Fermi arc. Breaking the spin symmetry through the coupling, the dispersion relation is gapped and the former Dirac node gets spin-polarized. We propose experimental realizations of the hybrid physics, including compressively strained HgTe as well as heterostructures of TI and WSM materials, connected to each other, e.g., by Coulomb interaction.

Contents

Symbols

k_B	Boltzmann constant, $k_B = 1.380\ldots \cdot 10^{-23}$ J/K
$\hbar$	reduced Planck constant, $\hbar = 6.626\ldots \cdot 10^{-34}/(2\pi)$ Js, usually set to 1
e	positive elementary charge, $e = 1.602\ldots \cdot 10^{-19}$ As
e	Euler number
ε_0	dielectric constant, $\varepsilon_0 = 8.854\ldots \cdot 10^{-12}$ As/Vm
B	magnetic field
l_B	magnetic length, $l_B = \sqrt{\frac{\hbar}{eB}}$
$k_\pm$	$k_\pm = k_x \pm ik_y = k_\parallel e^{\pm i\phi_k}$
$k_\parallel$	$k_\parallel^2 = k_x^2 + k_y^2$
μ	chemical potential
$\boldsymbol{\sigma}$, $\boldsymbol{\tau}$	Pauli matrices, $\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$, $\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$
S^1, S^2, T^2	manifolds: ring, sphere, torus
ν	$\mathbb{Z}_2$ invariant or filling factor
$\lvert u_{n,\boldsymbol{k}}\rangle$	Bloch state
E_k	dispersion relation; energy as a function of momentum k
E_k^μ	$E_k^\mu = E_k - \mu$; dispersion relation, normed to chemical potential
$\boldsymbol{G}$	reciprocal lattice vector
a	lattice constant, usually set to 1
$\mathcal{F}$	Berry curvature
$\boldsymbol{A}$	Berry connection or electromagnetic vector potential
γ_c	Berry phase
Θ	time-reversal operator, $\Theta = \mathrm{e}^{-i\pi S_y}K$ for spinful particles, $\Theta = -i\sigma_y K$ ($\Theta = K$) for spin 1/2 (1)
K	complex conjugation operator
P	spatial inversion operator, eigenvalues ξ_m of P give the parity
ξ_m	parity eigenvalues
Λ_a	time-reversal invariant momentum
σ_{xy}	Hall conductance

χ	chirality
$n_\downarrow, n_\uparrow, n_\mathcal{M}$	spin and mirror Chern numbers
$A(k,\omega)$	spectral function
g_s	spin degeneracy factor, $g_s = 2$
ε_r	background dielectric constant
α	effective Dirac fine-structure constant
χ_{nn}	density-density response function
χ_0	Lindhard response function
Π^R	polarization function
$\mathcal{F}^{\lambda,\lambda'}_{\boldsymbol{k},\boldsymbol{k}'}$	overlap function
v_q	bare Coulomb interaction energy, $v(r) = \frac{e^2}{4\pi\varepsilon_0}\frac{1}{r}$, in 2 dimensions $v_q = \frac{e^2}{2\varepsilon_0}\frac{1}{q}$
$\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}, M$	parameters of the BHZ model, Eqs. (2.38) and (4.1)
$\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{M}$	parameter functions of the Bi_2Se_3 model, Eqs. (2.57) and (9.1)
$q_0 = \frac{\mathcal{A}}{\lvert\mathcal{B}\rvert}$	BHZ model intrinsic momentum scale
$E_0 = \mathcal{A}q_0$	BHZ model intrinsic energy scale
ω_p	plasma frequency
$\Omega = \frac{\omega}{E_0}$	dimensionless frequency
$X = \frac{k}{q_0}$	dimensionless momentum
γ	imaginary part of the plasma frequency
$\Gamma = \frac{\gamma}{E_0}$	dimensionless imaginary part of the plasma frequency
$f(E)$	Fermi-Dirac function, $f(E) = \frac{1}{\mathrm{e}^{\beta(E-\mu)}+1}$
$\beta = \frac{1}{k_B T}$	inverse temperature
$\bar{\beta} = \frac{E_0}{k_B T}$	dimensionless inverse temperature
Π	polarization function or canonically conjugate momentum to ϕ in the Luttinger description
$\mathbb{L}$ L	Lagrangian and Lagrangian density
S	action, $S = \int \mathrm{d}t\, \mathbb{L}$
v	Luttinger velocity
$\mathcal{K}$	Luttinger interaction parameter
γ_W	parameter switches between type I and II Weyl semimetal for $\lvert\gamma_W\rvert \lessgtr 1$

Acronyms

TI	topological insulator
TCI	topological crystalline insulator
NI	normal insulator
TT	topologically trivial
TnT	topologically non-trivial
WSM	Weyl semimetal
QSH	quantum spin Hall
QH	quantum Hall
IQH	integer quantum Hall
$\mathcal{T}$	time-reversal
$\mathcal{I}$	inversion
p-h	particle-hole
e-h	electron-hole
TFT	topological field theory
CS	Chern-Simons
TRIM	time-reversal invariant momentum
QW	quantum well
2DEG	two dimensional electron gas
BIA	bulk inversion asymmetry
SIA	structural inversion asymmetry
BHZ	Bernevig-Hughes-Zhang
RPA	Random Phase Approximation
DOS	density of states
SPE	single-particle excitation

QHFM quantum Hall ferromagnetic

QHFMism quantum Hall ferromagnetism

QHTI quantum Hall topological insulator

OP order parameter

SOC spin-orbit coupling

F ferromagnetic

CAF canted antiferromagnetic

LL Landau level

BZ Brillouin zone

1D one-dimensional

2D two-dimensional

3D three-dimensional

1. Introduction

Two of the most useful concepts in physics are the principles of symmetries and scales. Quite reasonably, on different length scales different laws of physics dominate, e.g. quantum mechanics on the microscopic scale, while general relativity becomes relevant on the cosmological scale. The breaking of symmetries is directly related to this concept of scales. According to the Grand Unification Theory and the Standard Model, the universe is expected to have been in an ultra symmetric state at very high energies shortly after the Big Bang. Due to its expansion and corresponding lowering of its average energy density, the universe went through a series of phase transitions in which the symmetries were broken one after the other, arriving finally in the universe we live in today.

The same concepts can be used in many-body physics, see e.g. Ref. [Vol03] for a comparison between condensed matter and high-energy physics. The topic of this thesis can be assigned to the field of mesoscopic physics. Here, we are interested in quantum mechanical effects which occur not in the nanometer range of the individual atom, but on larger scales, usually in the range of micrometers. Naturally, this is the length scale of modern electronics, making this field not only interesting but also very useful for real life applications. One trend is here the shrinking of existing structures down to the nanometer level. For modern transistors with their 14 nm gates, the way down to the single-electron transistor does not seem so far fetched anymore.

In this thesis, we want to consider a different approach, not thinking about the effects that occur when one shrinks large structures to smaller scales, but consider what happens if one adds lots of small objects together. The emergent physics, as described by P. W. Anderson in his article "More Is Different" [And72], can be fundamentally different from the physics of the single particle. In our everyday life, there are only very few of these emergent physical effects based on quantum mechanics, that survive on our macroscopic scale. Magnetism, resulting from the quantum mechanical spin of the particles, is probably the most common and useful one. Superconductivity also works on large length scales as proven in modern particle accelerators, but has problems with higher temperatures. Whether there will be superconducting materials at room temperature and ambient pressure is still an open question. Both of these effects correspond to a broken symmetry, the magnet breaks rotational symmetry due to the fixed direction of its magnetization, while the BCS superconductor breaks the conservation of particles. In this thesis, we will consider topological phases, which might belong to this exclusive class of macroscopic quantum mechanical effects in the near future. Interestingly, they do not break symmetries, but actually rely on certain ones, like time-reversal ($\boldsymbol{\mathcal{T}}$) or inversion ($\boldsymbol{\mathcal{I}}$) symmetry, for their protection.

While the underlying physics of such topological insulators (TIs) is quite profound and will be elucidated below, there is a straight forward working definition. TIs are phases of matter with insulating bulk and conducting surface states. The surface has unique properties, including e.g. spin-momentum locking, see Fig. 1.1, and suppressed backscattering, and has to exist due to topological reasons. It is protected by the symmetries of the system and the size of the bulk band gap, such that there is no fundamental reason why such topological materials

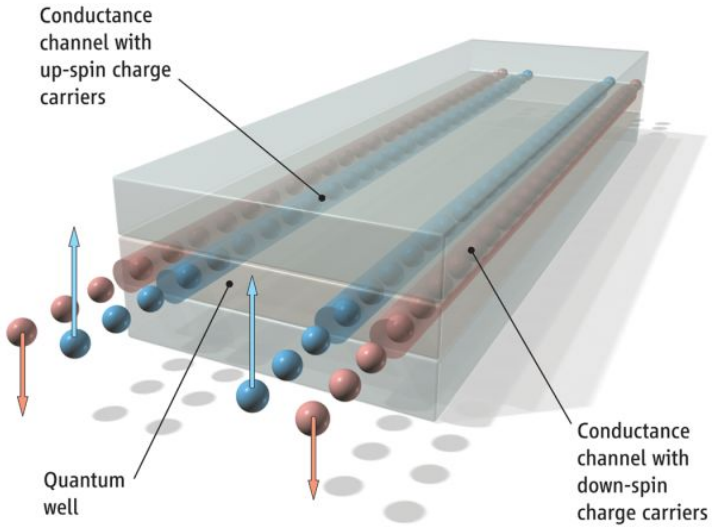


Figure 1.1. Two-dimensional TI (Quantum well) with helical edge states. In these channels, the spin of the electron is correlated with its direction of motion, such that spin-up electrons move clockwise and spin-down electrons move anti-clockwise along the edge. From Ref. [KWB+07]. Reprinted with permission from AAAS.

should not work at room temperature. The perfect conduction of the edge channels as well as their spin polarization could be interesting aspects for the future of information technology.

The physics of TIs and their semimetallic counterparts, the Weyl semimetals (WSMs), is quite well understood by now. Interestingly, although these topological effects clearly belong to the emergent physics as described by Anderson, their theoretical description is done on a single-particle level. In this thesis, we want to extend these theories to include correlation effects. Such interaction effects in condensed matter are notoriously difficult to treat, as kinetic and Coulomb energy are often on the same order of magnitude. This makes the field of correlated topological materials very interesting and challenging. Depending on the considered material and method, the inclusion of correlation effects results in qualitatively new effects or in quantitative changes to the physics of the system. But as Marx noted: merely quantitative differences, beyond a certain point, pass into qualitative changes. In this sense, the combination of topological and correlation effects bears many possibilities where indeed more is different.

Organizational matters

This thesis is structured into four chapters. Chap. I gives an introduction to topological and interacting matter and is based on the literature. At the beginning of each section, a detailed list of the used references is given for further studies. The following Chaps. II-IV are based on the published articles [JMT14a, JMT14b, KJT16, JT17]. These papers have been considerably reworked and restructured for this thesis, but were also partly adapted without extensive reformulation.

Part I.

Topological Matter

The study of topological effects in condensed matter systems is a fruitful endeavor. It began with the investigation of topological insulator (TI) systems based on non-interacting theories, but now also includes gapless topological phases as well as correlated topological materials. Faced with this huge variety, giving a comprehensive description of topological matter is out of the question. In this introductory chapter we will therefore focus on topics that are of direct importance for the following main research chapters II-IV. Still, we have to be brief on each separate topic. For readers interested in additional informations, we always give the specific references the paragraphs are build on at the beginning of each section.

This chapter comprises two main sections. The first one, Sec. 2, gives an introduction to topological effects and classification. It is mostly based on the uncorrelated topological band theory, introducing concepts such as Berry phase, Chern and $\mathbb{Z}_2$ invariants. All topological phases that are relevant for later chapters are established, including two-dimensional (2D) and three-dimensional (3D) TIs as well as Weyl semimetals (WSMs). Additionally the corresponding low-energy models are derived, which will be used in chapters II-IV.

The second part, Sec. 3, gives an introduction to correlation effects and how to treat them. The topological field theory (TFT) is introduced as a complementary theory to the topological band theory, working also in the correlated regime. Theories about how to treat interaction effects in the electron gas in a controlled fashion are covered, again with a focus on the formalisms relevant for the later chapters. This section closes by motivating why the connection of correlation effects with topology is interesting and fruitful. This bridges the gap between the introductory part of this thesis, Chap. I, and the research part in chapters II-IV.

2. Topology and Band Theory

Topological concepts form the basis of this thesis, so it is rather natural to start off by examining how topology enters condensed matter physics. In most of this section, we will rely on the help of a topological band theory. This may seem rather odd for a thesis entitled "Correlated Topological Materials", as band theory, by definition, does not include any correlation effects. Yet it is an easy and intuitive way to understand the connection of topology and condensed matter physics in most of its facets. As the majority of topological materials investigated so far feature only weak or no correlation effects, band theory seems like a reasonable starting point. Correlation effects will then be introduced in Sec. 3, and their influence on topology is the central topic of chapters II-IV.

This section is organized as follows. In Sec. 2.1 we introduce topology as the theory of continuity. We provide examples where topology appears in general life and in physics in particular. A more mathematically solid definition is provided as well. Topological band theory is explained in Sec. 2.2. Topology enters band theory through the geometrical Berry phase. The connection of topology to symmetries and invariants like Chern and $\mathbb{Z}_2$ invariants is highlighted. The effect of bulk topology on the surface physics is explained as well as the topological classification according to the so-called 10-fold way. The following sections are then about how these invariants relate to physical effects in real system. Sec. 2.3 explains the connection of the Chern number to the integer quantum Hall (IQH) effect. Sec. 2.4 introduces TIs in 2 and 3 spatial dimensions. This includes the quantum spin Hall (QSH) insulator, where HgTe quantum wells (QWs) serve as the prime example, as well as weak and strong 3D TIs. Appropriate models are given which will be used in the following chapters. This section closes with establishing WSMs as a new addition to the zoo of topological matter in Sec. 2.5.

2.1. Topology - Theory of Continuity

Topology is a general concept that goes well beyond its application in condensed matter physics. In abstract terms, it is the study of properties of a space, that are preserved under continuous transformations. In condensed matter physics, this space is usually the reciprocal k-space, described by a (topological) band theory. Yet the tools and concepts needed for classifying band structures of materials according to their topological properties appear also in very different contexts. Thus before we dive into the depths of a topological band theory, in this section, we will first motivate the concept. We begin with a short discussion in Sec. 2.1.1 of descriptive examples of topological concepts in everyday life. The idea of topology as a theory of continuity is put on a mathematical basis in Sec. 2.1.2. We close in Sec. 2.1.3 by giving an overview of where in physics topology plays a role and why it appears in condensed matter but not in high-energy physics.

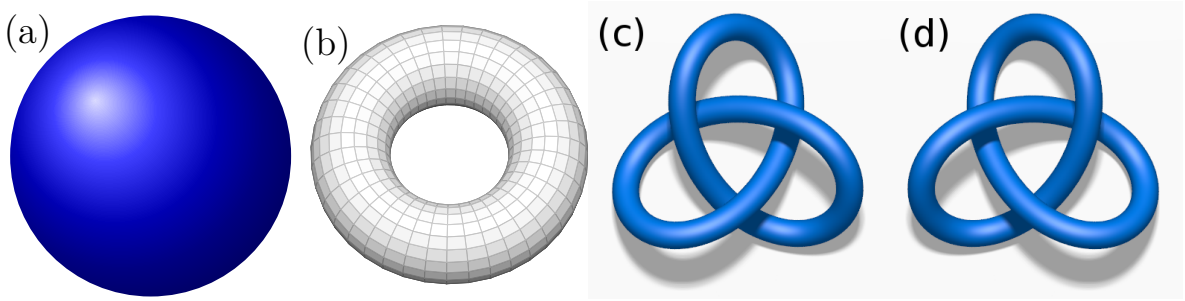


Figure 2.1. (a) Sphere with g=0. (b) Torus with g=1. Also called unknot. (c) and (d) Trefoil knots, left- and right-handed.

2.1.1. Examples from Everyday Life

This section has the purpose of providing an intuition about topological classification without the use of any math. The working principle is the following: two objects are topological equivalent, if they can be transformed into one another by a smooth transformation. This sounds a bit vague, but we can certainly state that for example neither cutting nor gluing are smooth transformations. Equivalent objects have some kind of characteristic feature that stays unchanged if the object is smoothly deformed. Such a resilient property can usually be described by an invariant that does not change in the transformation. We will now have a look at some examples. The section is based on Ref. [FM13] Chap. 1, Sec. 2.1 and Ref. [wik17a].

Sphere and Torus

We begin by taking a bit of modeling clay into our hands and form a sphere out of it like in Fig. 2.1 (a). Now, by smoothly deforming it, we can form a disk or a bowl out of this sphere. Yet, to get a doughnut like in Fig. 2.1 (b) one needs to punch a hole into the middle of the sphere - this is not considered a smooth deformation. So we can conclude that the number of holes in an object is an invariant that does not change under smooth deformations. In mathematics, the number of holes is called the genus g. It is related to the Euler characteristic χ and can be calculated for boundary-less surfaces S with the Gauss-Bonnet theorem

$$g = 1 - \frac{1}{2}\chi, \quad \chi = \frac{1}{2\pi}\int_S K \mathrm{d}A \tag{2.1}$$

as an integral over the Gaussian curvature K.

Annulus and Möbius Strip

As another simple example, take a strip of paper and glue the ends together. After glueing, the outcome should look like the objects shown in Fig. 2.2. The number of twists that one introduced into the strip is an invariant - without cutting the strip again it cannot be changed, just shifted around the loop by smooth deformations.

Unknot and Trefoil Knot

The last example from handicraft work is the tying of knots into a rope. Here actually several invariants can be defined to classify knots, including the already known to us genus as well as

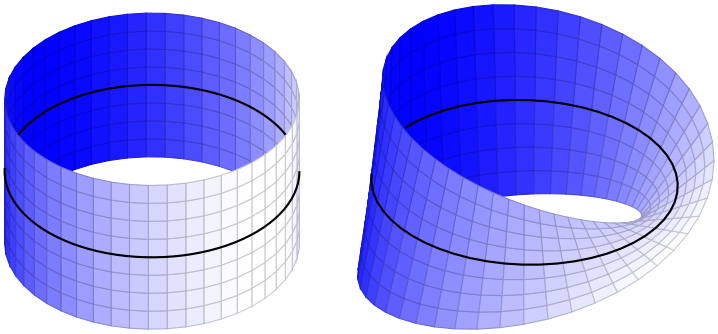


Figure 2.2. (left) Annulus B_0 with no twist and (right) Möbius strip B_1 with a single twist.

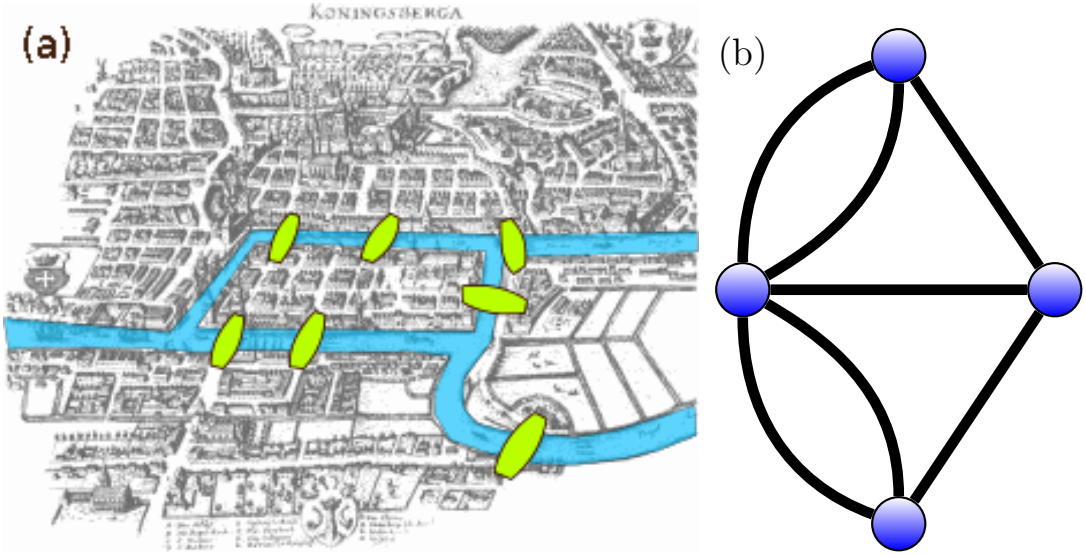


Figure 2.3. (a) Map of the center of Königsberg, indicating the 4 main parts of the city, separated by the river and connected by seven bridges. (b) Connected graph description of the same problem. Figure (a) adapted from Ref. [wik17b].

the number of crossings of the rope. An example is the trefoil knot shown in Fig. 2.1 (c) and (d). It comes in two versions, the left-handed and right-handed one. They are distinguished by their combinations of over and under crossings and are topological distinct. Such a handedness will appear again in the topological band theory, e.g. when we talk about helicity in Sec. 2.4.1 or chirality in Sec. 2.5.

Königsberg Bridge Problem

Our last example is the problem of the seven bridges of Königsberg. This notable mathematics problem was solved by Leonhard Euler in 1736 and can be considered as one of the foundations of graph theory. This section is based on Ref. [wik17b] and Ref. [Eul35].

The problem considers the city of Königsberg. The Pregel River divides the city into different parts, including two large islands, which are overall connected by seven bridges, see Fig. 2.3 (a). The question is whether one can take a walk, during which one crosses each bridge once and only once. Of course, all kinds of cheating, like taking a boat or only half-crossing of

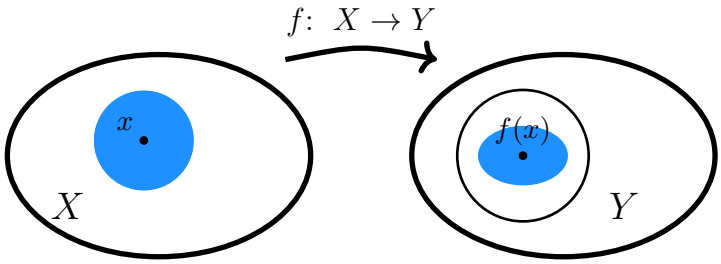


Figure 2.4. Depiction of point x in set X and its image $f(x)$ in set Y. The image of the blue neighborhood of x fits into the neighborhood of Y, indicated by a black circle.

bridges, is ruled out. The start and end point can be arbitrary.

Euler proved that it is not possible, using topological reasoning for his solution. First, we note that the precise path we take in the land mass is irrelevant. Only the sequence in which we cross the bridges is important. Thus one can map the problem to a graph as depicted in Fig. 2.3 (b), where only the nodes and edges are relevant. Now, for non-endpoints the number of entering and leaving the node has to be equal, meaning the number of bridges should be even. For start and end points, we need an additional connection, so here the number of bridges should be odd. Looking at Fig. 2.3, all parts of the city are connected by an odd number of bridges - the walk around the city can not be done by crossing each bridge only ones.

2.1.2. Mathematical Definitions

We now have an intuitive feeling about invariant properties under smooth deformations. In this section, we complement this feeling with the proper mathematical definitions. It is based on Ref. [Bud11] Secs. 1.1f and Ref. [wik17a].

Definition of Topology

Topology is the theory of continuity. Interestingly, one does not need a metric for a proper definition of continuity. We use a neighborhood U to define the continuous function $f(x)$. For a continuous function f, the image of the neighborhood of point x fits into the neighborhood U around the image of $f(x)$. If one considers a neighborhood like a small ball around a point, see Fig. 2.4, this essentially means that a continuous function makes no jumps. A different definition for continuity relies on open sets. An open set is a set which is a neighborhood of each of its points. A function is then called continuous, if the preimage of any open set is open.

With this we come now to the definition of a topology. A topology T on a set X is a family of sets (set of sets) fulfilling three properties:

1. the empty set and X are elements of T
2. a union of elements of T is an element of T
3. an intersection of finitely many elements of T is an element of T

A topological space is defined by (X, T). Elements of T are called open sets. Although one does not need a metric for these definitions, if one has a metric it automatically defines a

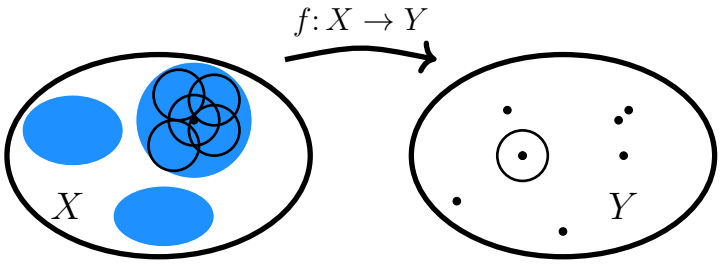


Figure 2.5. Continuous function f that maps different points in X to the same point in Y. X can be divided in disjunct subsets (blue ellipses), covered by neighborhoods (black circles).

topology. So one can think about a topology as a classification of elements of a set. This classification can be made based on invariants.

Let us have a look at an example for such a topological invariant. Consider Fig. 2.5, where many configurations in X are mapped by the continuous function f to the same point in Y. Following our definition of continuity, the preimage of the single point in Y contains an entire ball in X. Covering X with overlapping balls, we conclude that the function has to be locally constant. The constant can be different on different disjunct subsets, depicted by dark blue ellipses in Fig. 2.5. This picture has a direct physical application in the IQH effect, where the conductance is quantized in multiples of $\frac{e^2}{h}$. As conductivity is a continuous function of the system parameters, we conclude that it is locally constant, invariant under smooth changes to the system. The connected regions of parameter space are given by the regions with a mobility gap (between the Landau levels (LLs)). Outside these regions, and at phase transitions between them, the conductivity is not quantized.

One can distinguish between two types of maps, on which a decomposition into equivalence classes can be done:

- homeomorphism: a continuous, bijective map, the inverse map is also continuous. Spaces that are homeomorphic have the same topology. Example: doughnut to cup of tea.
- homotopy: continuous deformation of one map into another, leaving the boundary of the domain fixed. It is not necesseraly invertible. Example: a solid sphere can be shrunken to a point, but this is not invertible, thus not homeomorphic.

As an example, we classify the alphabet according to homeomorphism and homotopy.

- Homeomorphic: {A,R}, {B}, {C,G,I,J,L,M,N,S,U,V,W,Z}, {D,O}, {E,F,T,Y}, {H,K}, {P}, {Q}, {X}
- homotopic: {A,R,D,O,P,Q}, {B}, {C,E,F,G,H,I,J,K,L,M,N,S,T,U,V,W,X,Y,Z}

For classification based on homotopy, only the number of holes of the letter are important. For the homeomorphic classification, the number of tails is important too, as they can not be shrunken to a point and back in an invertible fashion. Properties of topological spaces that are invariant under homeomorphisms are called topological invariants.

Twisted Products and Manifolds

Next we want to apply these concepts to the classification of manifolds. This will be relevant for the application to physics, as topology enters physics via the underlying space of the wave functions. This makes it obvious that topology is a global, not local, property of a system.

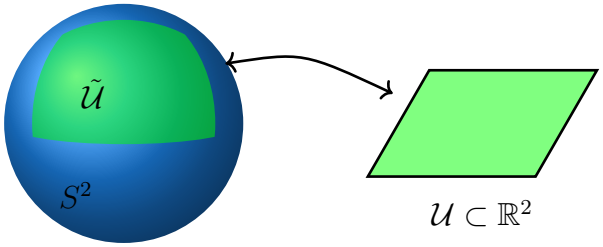


Figure 2.6. Homeomorphic mapping of a region $\tilde{\mathcal{U}}$ of a sphere S^2 to the flat space $\mathcal{U} \subset \mathbb{R}^2$.

A basic example of a trivial product is a ring $S^1 \times \mathbb{1}$ as shown in Fig. 2.2 (a). It is the trivial product of a unit circle S^1, called base manifold, and a unit interval $\mathbb{1}$, the fiber. The total space, $S^1 \times \mathbb{1}$, is called a fiber bundle. A non-trivial fiber bundle is the Möbius strip, Fig. 2.2 (b), which contains an additional twist of π, such that it can not be written as a trivial product of two manifolds. Yet locally it can be parametrized like the trivial ring. Non-trivial fiber bundles need always a base manifold that is not homotopic to a point, S^1 in this example.

Next we have a look at a local parametrization. We take $\tilde{\mathcal{U}}$, a part of S^2, and map it homeomorphically to a flat piece of plane $\mathcal{U} \subset \mathbb{R}^2$, see Fig. 2.6. In order to cover the entire sphere, several patches are needed, as S^2 is not homeomorphic to the plane. The single patches are trivial fiber bundles. The non-triviality is introduced by transition functions in the overlap regions. They model the twist by changing the fiber coordinates in a non-trivial fashion. For the Möbius strip, this corresponds to a flip of the sign of the fiber coordinate.

2.1.3. Physics Overview

Topology has become an interesting part of condensed matter physics. Here we want to give a brief overview why the study of topological properties of a system is interesting, what are the most prominent effects and why they are somewhat restricted to condensed matter systems. This section is based on Ref. [Bud11] Sec. 1, Ref. [FM13] Chap. 1, Sec. 1, Ref. [BH13] Sec. 1 and Ref. [She13] Secs. 1.1f. It serves as a physical introduction for the following, more technical, sections.

An important and successful paradigm in condensed matter is the classification of matter according to its ordered phases. Some phases can be characterized by a local, spontaneous breaking of symmetries. In these cases one defines a local order parameter, which is finite in the ordered phase and zero otherwise. A prominent example is the finite magnetization of a ferromagnetic phase. According to Landau theory of phase transitions, one can expand in the order parameter as it is supposed to be small around the transition. The different orders in the parameter then characterize different sorts of phase transitions.

In contrast to this, there are other phases involving no broken symmetries, which makes them more subtle. An example is the IQH effect, see Sec. 2.3. Its quantized conductance is insensitive to local fluctuations and smooth changes in material parameters. Only a closing of the gap can change the conductivity, and we addressed the link to topology already in Sec. 2.1.2.

This robustness is a key property of topological effects, which makes them interesting for applications in physics. Topology offers a global characterization of systems as it is a property of the underlying space of the wave functions. Compared to microscopic theories which depend on all kinds of material parameters, this provides robustness to topological effects. Topology

is usually protected by global symmetries and different topological states can be classified by global invariants, much like in real space as described in Sec. 2.1.1, but now in reciprocal k-space. This link between topology and global symmetries of a system led to the generation of a periodic table of topological matter, given in Sec. 2.2.5.

A powerful tool for classification of crystals is the topological band theory, discussed in the following Sec. 2.2. The topological information is hidden in the geometric phase called Berry phase associated with each band. This implies that topology is encoded in the eigenstates, not the energy spectrum, of a crystal. Invariants like Chern numbers and $\mathbb{Z}_2$ invariants are given as integrals over this Berry phase and allow for a classification of insulators according to preserved symmetries. The trivial insulator serves here as a reference system. It is defined as an insulator which adiabatically flows into the atomic limit when the hopping is slowly turned off.

Experimentally relevant are topological effects due to their robustness against local perturbations. A perfectly quantized conductance as a basis for a system of units or error protected quantum computing based on nonlocal qubits are two of the more prominent ideas for practical relevance of topology. While the bulk of TIs is still insulating and thus quite boring in practice, the gapless surface states are what makes TIs interesting for applications.

The presence of edge states can be understood from an important, experimentally verified, example, the IQH effect, see Sec. 2.3. A strong magnetic field perpendicular to a two dimensional electron gas (2DEG) leads to LLs with quantized energies $(n+1/2)\omega_c$, where $\omega_c = eB/m$ and m the particle mass. At the edge, the local potential pushes the LLs through the Fermi energy, which leads to an edge conductance proportional to the number of LLs, the filling factor ν. The QSH effect, see Sec. 2.4, can be understood on a similar basis. Instead of a magnetic field, a strong spin-orbit coupling (SOC) is present in the system. Now electrons with different spin feel different effective magnetic fields. This leads to helical edge states where spin and momentum are coupled. The consequence thereof can be a finite spin Hall conductance.

As a final remark of this section, we consider the question why all these effects only appear in condensed matter systems, not in high-energy physics. Thus why is quantum electro dynamics (QED) always topologically trivial, while condensed matter is sometimes not. Wave functions are defined up to a phase, $\left\{\mathrm{e}^{i\phi}\psi\left(x\right)|\mathrm{e}^{i\phi} \in \mathrm{U}\left(1\right)\right\}$ at a point $x \in \mathbb{R}^4$ for QED. $\mathbb{R}^4$ is a trivial, contractible space, which implies that a global gauge can be chosen. In condensed matter, the wave functions live in k-space in the Brillouin zone (BZ). The BZ is periodic, and thus k-space is given by a torus T^d, which is a non-trivial manifold. Wave functions are made up of the Bloch functions $|u_k\rangle$, $\left\{\mathrm{e}^{i\phi}|u_k\rangle|\mathrm{e}^{i\phi} \in \mathrm{U}\left(1\right)\right\}$. Here the gauge field is given by Berrys connection Sec. 2.2.2, which can have a non-trivial character.

2.2. Topological Band Theory

In accordance to the idea of classifying matter by its phase, discussed in Sec. 2.1.3, we are now interested in a topological classification of matter. This is a very complicated task in general. In this section we restrict ourselves to matter describable by band theory which simplifies the classification considerably. The drawback is that all systems in which Coulomb interaction plays a dominant role fall not in this category. For these systems, we will discuss the complementary TFT in Sec. 3.1.

A band insulator can be described in the independent electron approximation, its ground state is represented by a Slater determinant. We will see that two insulators are topologically

equivalent if their Hamiltonians can be changed into one another while the system stays in the ground state. As an insulator has a band gap, the transformations can always be done slow enough as not to excite the system. Such a process is called adiabatic. Topological transitions then need a band closing during the transformation. It is the only way to change topological properties of a system, within the limitations given by band theory. The band gap also ensures that weak interaction effects to not alter the topological classification. Interacting states that can be adiabatically connected to noninteracting electrons have the same topology, as will be shown in Sec. 3.1. The limitation of band theory to noninteracting electrons are thus not as severe as one might fear.

We begin in the following with a short repetition of band theory in Sec. 2.2.1. The topological aspects are introduced in Sec. 2.2.2 via the Berry phase. The presence of time-reversal ($\boldsymbol{\mathcal{T}}$) symmetry deepens the topological classification as discussed in Sec. 2.2.3. Topological systems have gapless edge states at the boundaries, which follow directly from a change in topology at interfaces, see Sec. 2.2.4. This section concludes with the classification table of TIs on the basis of the Altland-Zirnbauer symmetry classes in Sec. 2.2.5.

2.2.1. Band Theory

We recap shortly the basis of band theory, i.e. Bloch theorem. This section is based on Ref. [FM13] Chap. 1, Sec. 2.2 and Ref. [She13] Secs. 3.1f & 4.1.

As a starting point, the considered physical material is assumed to be crystalline and effectively noninteracting. It has translational symmetry, which allows for a labeling of single-particle states by crystal momentum $\boldsymbol{k}$. Bloch theorem implies that the eigenstates can be written as

$$|\psi_{n,\boldsymbol{k}}(\boldsymbol{r})\rangle = \mathrm{e}^{i\boldsymbol{k}\cdot\boldsymbol{r}}|u_{n,\boldsymbol{k}}(\boldsymbol{r})\rangle. \tag{2.2}$$

$|u_{n,\boldsymbol{k}}(\boldsymbol{r})\rangle$ is a cell periodic eigenstate of the Bloch Hamiltonian $H(\boldsymbol{k}) = \mathrm{e}^{-i\boldsymbol{k}\cdot\boldsymbol{r}} H(\boldsymbol{r})\, \mathrm{e}^{i\boldsymbol{k}\cdot\boldsymbol{r}}$, thus

$$H(\boldsymbol{k})\,|u_{n,\boldsymbol{k}}(\boldsymbol{r})\rangle = E_{n,\boldsymbol{k}}|u_{n,\boldsymbol{k}}(\boldsymbol{r})\rangle. \tag{2.3}$$

The eigenvalues $E_{n,\boldsymbol{k}}$ and eigenvectors $|u_{n,\boldsymbol{k}}(\boldsymbol{r})\rangle$ define the band structure. These bands are filled up from the bottom, where the Pauli principle dictates that every state can only be filled once. Insulators are characterized by an energy gap at the Fermi energy, separating the highest occupied band, called valence band, and the lowest empty band, called conduction band.

Translational symmetry implies $H(\boldsymbol{k}+\boldsymbol{G}) = H(\boldsymbol{k})$ with the reciprocal lattice vector $\boldsymbol{G}$. Thus crystal momentum is defined in the periodic BZ up to translations by a multiple of $\boldsymbol{G}$, $\boldsymbol{k} \equiv \boldsymbol{k} + \boldsymbol{G}$. This corresponds to the topology of a torus T^d in d dimensions. The band structure can be seen as a mapping from the BZ torus to the space of Bloch Hamiltonians with an energy gap.

How can one obtain the band structure of a crystal in practice? A powerful method is given by tight binding calculations. One considers electrons to be almost localized around their individual atoms in a lattice. Only the outer electron orbits overlap, providing a possibility for the electrons to jump from one atom to the next, see Fig. 2.7. The overlapping orbitals can then be described by delocalized bands.

Often one is only interested in the physics close to a special point in the BZ. In this case it can make sense to derive an effective continuum model based on an expansion in k-space via the $\boldsymbol{k}\cdot\boldsymbol{p}$ theory. These continuum models capture the essential low-energy physics if the band structure is gapped and thus effectively inert in the rest of BZ. In this thesis, we will

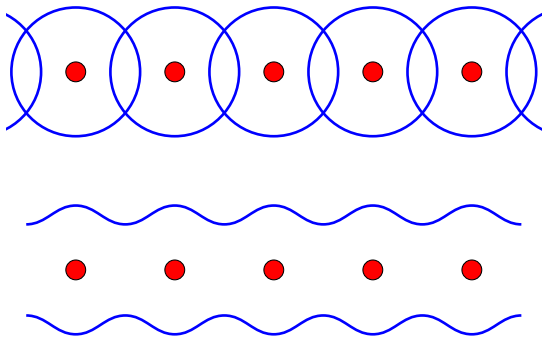


Figure 2.7. Red dots depict the positive ions of a lattice, the blue circles the orbits of the outer electrons. Due to an overlap of the electron orbits, bands form and electrons can jump from one atom to the next.

rely heavily on these continuum models. An easy way to switch from a continuum model to a lattice regularized one is the substitution

$$k \to \frac{1}{a}\sin(ka), \quad k^2 \to \frac{2}{a^2}(1-\cos(ka)) \tag{2.4}$$

with the lattice constant a.

Translational symmetry of a crystal makes it possible to use Bloch theorem. Additional real space symmetries can greatly facilitate the derivation of effective, low-energy models. By classifying orbitals according to there type, s, p, ..., and total orbital momentum, group theory in combination with $\boldsymbol{k}\cdot\boldsymbol{p}$ theory makes it possible to derive continuum models including all symmetry allowed terms. The free parameters are then determined from a fit to experimental or ab-initio band structures. Most of the models used in this thesis, see Sec. 2.4.1, Sec. 2.4.2 and Sec. 2.5, were obtained this way.

2.2.2. Berry Phase and the Chern Invariants

We are now in a position to understand where topological arguments enter band theory. This is done by having a closer look at the phase factors an electron in a Bloch band picks up during time evolution. We introduce the Berry phase and discuss its quantization in terms of Chern invariants. This section is based on Ref. [BH13] Secs. 2 & 3.0, Ref. [She13] Sec. 4.2 & App. A as well as Ref. [FM13] Chap. 1, Sec. 2.4.

A wave function in quantum mechanics is defined up to a phase, such that the substitution $|u_{n,\boldsymbol{k}}(\boldsymbol{r})\rangle \to \mathrm{e}^{i\phi(\boldsymbol{k})}|u_{n,\boldsymbol{k}}(\boldsymbol{r})\rangle$ with the phase $\phi(\boldsymbol{k})$ keeps the eigenenergies, Eq. (2.3), invariant. The eigenenergies are therefore gauge independent, as observables should be.

Let's have a look at the time evolution of the system. Consider a Hamiltonian dependent on one parameter $\boldsymbol{R}(t)$, which follows a slow cyclic evolution from $t=0$ to $t=T$ such that $\boldsymbol{R}(0)=\boldsymbol{R}(T)$, with T the period. In parameter space this corresponds to a closed path C. The instantaneous eigenstate basis of $H(\boldsymbol{R}(t))$ is defined by

$$H(\boldsymbol{R}(t))\,|u_n(\boldsymbol{R}(t))\rangle = E_n(\boldsymbol{R}(t))\,|u_n(\boldsymbol{R}(t))\rangle. \tag{2.5}$$

The phase of the eigenstate is not yet fixed, and it can be $\boldsymbol{R}(t)$ dependent. The phase function should be smooth and single valued along the path. If this is not possible globally, then we

have to define it on several overlaying patches. The full time evolution of the quantum state $|\Psi(t)\rangle$ is governed by

$$i\hbar\partial_t|\Psi(t)\rangle = H(\boldsymbol{R}(t))\,|\Psi(t)\rangle. \tag{2.6}$$

In the adiabatic limit, the system stays in the instantaneous eigenstate. Therefore quantum state and eigenstate can be related by

$$|\Psi(t)\rangle = \mathrm{e}^{i\gamma_c(t)} \exp\left(-\frac{i}{\hbar}\int_0^t \mathrm{d}t' E_n(\boldsymbol{R}(t'))\right) |u_n(\boldsymbol{R}(t))\rangle \tag{2.7}$$

where the geometric phase $\gamma_c(t)$, called Berry phase, is defined by

$$\partial_t\gamma_c(t) = i\langle u_n(t)|\partial_t|u_n(t)\rangle. \tag{2.8}$$

It originates from the fact that not only the eigenenergies change under time evolution, but the instantaneous eigenbasis as well. One can rewrite this as a path integral in k-space like

$$\gamma_c = \oint_C \boldsymbol{A}_n \cdot \mathrm{d}\boldsymbol{k} = \int_S \mathcal{F}_n \cdot \mathrm{d}\boldsymbol{S} \tag{2.9}$$

with the vector potential called Berry connection

$$\boldsymbol{A}_n = i\langle u_{n,\boldsymbol{k}}|\nabla_{\boldsymbol{k}}|u_{n,\boldsymbol{k}}\rangle, \tag{2.10}$$

the surface S and the Berry curvature $\mathcal{F}_n = \nabla \times \boldsymbol{A}_n$. $\boldsymbol{A}_n$ changes under a gauge transformation like $\boldsymbol{A}_n \to \boldsymbol{A}_n - \nabla_{\boldsymbol{k}}\phi(\boldsymbol{k})$, but the Berry phase is a gauge invariant quantity, formally corresponding to a magnetic flux. The Berry curvature is the equivalent magnetic field and can be written in the gauge invariant form

$$\mathcal{F}_n = i\sum_{m\neq n} \frac{\langle u_{n,\boldsymbol{k}}|\nabla_{\boldsymbol{k}}H(\boldsymbol{k})\,|u_{m,\boldsymbol{k}}\rangle \times \langle u_{m,\boldsymbol{k}}|\nabla_{\boldsymbol{k}}H(\boldsymbol{k})\,|u_{n,\boldsymbol{k}}\rangle}{(E_n - E_m)^2}. \tag{2.11}$$

It follows directly from the anticommutativity of the cross product that the sum of Berry curvature over all energy levels is 0. The denominator of Eq. (2.11) goes to 0 for degenerate bands, indicating that the Berry curvature has a monopole in this case.

The physical consequences of the Berry phase depend on the dimension and whether the system is a metal or insulator. According to Eq. (2.9) the Berry phase can be written as an integral over a closed curve (one-dimensional (1D) manifold), which is relevant for insulators in 1D and the Fermi surfaces of 2D metals. For the purposes of this thesis, the representation of the Berry phase as a surface (2D manifold) integral over the Berry curvature is more relevant. We apply this to insulators in 2D, see Sec. 2.4.1, and the Fermi surfaces of 3D metals. As an example for the latter case, in InfoBox 2.1 the Berry phase of a 3D Dirac cone is discussed. One finds that the Chern number, given by

$$n = \frac{1}{2\pi}\int_S \mathcal{F}\cdot\mathrm{d}\boldsymbol{S} \tag{2.15}$$

for a closed surface S, is always quantized.

The Berry phase is not just an abstract tool, it is straightforward to think about an experimental setup where it can be directly measured. Consider a beam of particles prepared in spin state n. The beam is split and runs along two different paths. On one, a constant magnetic field B is applied. On the other, B is constant in magnitude but slowly varies in direction, subtending the total angle γ_c. Combine the beams again, they will have acquired a phase difference that leads to the diffraction pattern in the intensity of

$$\left|1 + \mathrm{e}^{i\gamma_c}\right|^2 = 4\cos^2(\gamma_c/2)\,. \tag{2.16}$$

Consider the two level Hamiltonian

$$H(\boldsymbol{k}) = \boldsymbol{d}(\boldsymbol{k}) \cdot \boldsymbol{\sigma}. \tag{2.12}$$

It corresponds to a 3D Dirac cone for $\boldsymbol{d} = \boldsymbol{k}$. In this case, we have $\nabla_{\boldsymbol{k}} H = \boldsymbol{\sigma}$ and with the help of Eq. (2.11) we find

$$\mathcal{F}_{\pm}^{Dirac} = \mp \frac{\boldsymbol{k}}{2|\boldsymbol{k}|^3}. \tag{2.13}$$

$\pm$ stands for the positive and negative energy part of the Dirac cone. This shows that Dirac points are monopoles of Berry curvature. Integrating over a 3D sphere containing the monopole gives a Berry phase of

$$\gamma_c^{Dirac} = \mp 2\pi \tag{2.14}$$

such that the Chern number, Eq. (2.15), is an integer.

InfoBox 2.1. Berry phase of a 3D Dirac cone

2.2.3. Time-reversal Symmetry and the $\mathbb{Z}_2$ Invariant

Symmetries are essential for the endeavor to classify matter. Besides the spatial symmetries of a crystal, in this section we have a look at the discrete symmetry of time-reversal ($\boldsymbol{\mathcal{T}}$). It will prove a key ingredient in the understanding of topological properties of condensed matter systems. The section is based on Ref. [FM13] Chap. 1, Sec. 5.2 & Chap. 2, Sec. 2.1, Ref. [BH13] Secs. 4 & 10 & 12 as well as Ref. [She13] Sec. 4.8f & App. B. It provides the basis for the discussion of TIs in Sec. 2.4.

Time-reversal Symmetry

The operation $\boldsymbol{\mathcal{T}}$ reverses the arrow of time, $t \to -t$. The corresponding anti-unitary operator will be denoted by Θ, and the Hamiltonian of a $\boldsymbol{\mathcal{T}}$ symmetric system should commute with it,

$$[H, \Theta] = 0. \tag{2.17}$$

A standard way of breaking $\boldsymbol{\mathcal{T}}$ symmetry is by applying an external magnetic field $\boldsymbol{B}$ to the system, which does not reverse its direction under Θ.

In the special case of spin 1 or 1/2 particle, the $\boldsymbol{\mathcal{T}}$ operator can be represented by the anti-unitary operator

$$\begin{aligned} \Theta &= K; \ \textit{spin } 1 \\ \Theta &= \exp(-i\pi S_y/\hbar) K; \ \textit{spin } 1/2. \end{aligned} \tag{2.18}$$

Here S_y is a spin operator and K the complex conjugation operator. Thus spin 1/2 particles gain an additional rotation of their spin by π around the y-axis. One immediately finds $\Theta^2 = 1$ ($\Theta^2 = -1$) for spin 1 (1/2) particles. This is due to the fact that one needs a total rotation of 4π to bring back a spin 1/2 particle to its original state. For $\boldsymbol{\mathcal{T}}$ invariant Bloch Hamiltonians one finds

$$\Theta H(\boldsymbol{k}) \Theta^{-1} = H(-\boldsymbol{k}) \tag{2.19}$$

which can be motivated from the substitution $\boldsymbol{k} = -i\hbar\nabla$. Thus Θ maps the Hamiltonian at k to the one at $-k$.

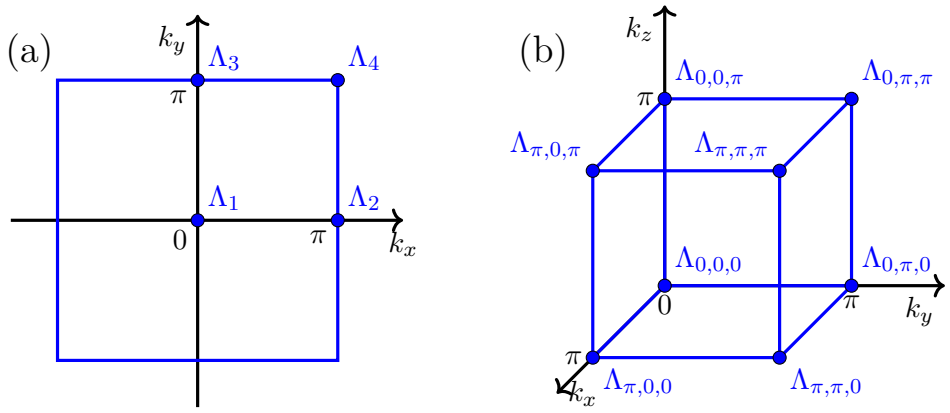


Figure 2.8. Square lattice examples. (a) In 2D, the BZ has 4 TRIMs Λ_a. (b) In 3D, 8 TRIMs can be defined. Due to $\boldsymbol{\mathcal{T}}$ symmetry, only half of the BZ needs to be considered in each principal direction.

In the BZ, there are special time-reversal invariant momenta (TRIMs) where $k = -k$ mod $\boldsymbol{G}$. In these cases, the mapping of Eq. (2.19) leads to the exact same Hamiltonian. For square lattices, which will serve as an example in the following, these points are $k = 0$ and $k = \pm\pi/a$ along each principal axis, see Fig. 2.8. In $\boldsymbol{\mathcal{T}}$ invariant systems, one needs only to look at $1/2^d$ of the BZ, as the other part follows from Eq. (2.19).

Kramers Theorem

Kramers theorem states that eigenstates of a $\boldsymbol{\mathcal{T}}$ invariant Hamiltonian of half-integer spin particles are at least twofold degenerate. This is the case because the state $|\chi\rangle$ is different from state $\Theta|\chi\rangle$. Otherwise one could write $\Theta|\chi\rangle = c|\chi\rangle$ with constant c, which would lead, with $\Theta^2|\chi\rangle = |c|^2|\chi\rangle$, to the contradiction $|c|^2 = -1$. Kramer's partners are usually split according to Eq. (2.19) into a state at k and one at $-k$. Only at TRIMs is the spectrum necessarily degenerate.

Without SOC in the system, Kramer's theorem corresponds just to the spin degeneracy. Yet it is also valid with SOC when no clear spin can be defined, which makes it more interesting. As an immediate consequence of the theorem, scattering between Kramer's partners is forbidden, such that

$$\langle \Theta\psi | H | \psi \rangle = 0 \tag{2.20}$$

for a $\boldsymbol{\mathcal{T}}$ invariant Hamiltonian H. It also follows that a $\boldsymbol{\mathcal{T}}$ invariant insulator needs at least four bands for spin 1/2 electrons, as a pair of bands will always be degenerate at the TRIMs in the BZ.

$\mathbb{Z}_2$ Invariant in 2D

All systems with $\boldsymbol{\mathcal{T}}$ symmetry have zero Chern number $n = 0$. In these systems, however, an additional $\mathbb{Z}_2$ invariant can be defined that serves as a subdivision of the class of $n = 0$ systems. As a physical consequence of a finite $\mathbb{Z}_2 = 1$, one finds a Kramers pair at the edge of a finite system, see Sec. 2.2.4.

We now derive a formula for the $\mathbb{Z}_2$ invariant in 2D systems. We begin by defining a unitary

matrix from occupied Bloch functions like

$$\omega_{mn}(\boldsymbol{k}) = \langle u_{m,\boldsymbol{k}}|\Theta|u_{n,-\boldsymbol{k}}\rangle. \tag{2.21}$$

As Θ is anti-unitary with $\Theta^2 = -1$, it follows that $\omega^T(\boldsymbol{k}) = -\omega(-\boldsymbol{k})$. In 2D there are four special TRIMs, denoted Λ_a, with $\boldsymbol{k} = -\boldsymbol{k}$ at which $\omega(\Lambda_a)$ is antisymmetric. The determinant of an antisymmetric matrix is the square of its pfaffian. The pfaffian itself is gauge dependent for a U(1) gauge transformation. A gauge invariant quantity can be defined by

$$\delta_a = \mathrm{Pf}\left[\omega(\Lambda_a)\right] / \sqrt{\mathrm{Det}\left[\omega(\Lambda_a)\right]} = \pm 1. \tag{2.22}$$

If we define $|u_m(\boldsymbol{k})\rangle$ continuously throughout the BZ, which is possible due to the zero Chern number, the square root can be specified globally. We then define the $\mathbb{Z}_2$ invariant ν as

$$(-1)^\nu = \prod_{a=1}^{4} \delta_a \tag{2.23}$$

with $\nu = 0$ or $\nu = 1$.

The calculation of ν is much simpler in systems with additional symmetries. For a spin symmetry where e.g. S_z is conserved, independent Chern numbers $n_\uparrow$ and $n_\downarrow$ for spin up and down can be defined. $\boldsymbol{\mathcal{T}}$ symmetry dictates that $n_\uparrow + n_\downarrow = 0$ and with $n_\sigma = (n_\uparrow - n_\downarrow)/2$ the $\mathbb{Z}_2$ invariant is given by

$$\nu = n_\sigma \mod 2. \tag{2.24}$$

In crystals that have inversion ($\boldsymbol{\mathcal{I}}$) symmetry, another simplification is possible, following Ref. [FK07]. In such systems, Bloch states are also parity eigenstates at the TRIMs Λ_a. Thus for the $\boldsymbol{\mathcal{I}}$ operator P one finds $P|u_{m,\Lambda_a}\rangle = \xi_m|u_{m,\Lambda_a}\rangle$ with eigenvalues $\xi_m(\Lambda_a) = \pm 1$. The product over m Kramers pairs (the two bands in a Kramers pair have the same parity eigenvalue) of occupied bands

$$\delta_a = \prod_m \xi_m(\Lambda_a) \tag{2.25}$$

can then be used to determine ν via Eq. (2.23).

In all these definitions, SOC has not played an obvious role. Yet it is needed to ensure the presence of a bulk gap. For an $\boldsymbol{\mathcal{I}}$ and $\boldsymbol{\mathcal{T}}$ invariant system without SOC, the bulk gap needs to close somewhere in the BZ.

$\mathbb{Z}_2$ Invariant in 3D

Based on the definitions of the $\mathbb{Z}_2$ invariant in 2D from the last section, the next step is to generalize this to 3D.

Consider 2D planes in the 3D BZ along the principal axis. We denote these planes by xy, yz and xz. In all of these planes, we look at the effect of $\boldsymbol{\mathcal{T}}$ on the third momentum. As an example, we choose k_z for the xy plane. Only planes with $k_z = 0$ and $k_z = \pm\pi/a$ are $\boldsymbol{\mathcal{T}}$ symmetric, as they are projected onto themselves for $k_z \to -k_z$. Therefore two $\mathbb{Z}_2$ invariants per plane can be defined, which makes 6 in total for the 3 planes. Due to interdependencies, this can actually be reduced to 4. If an invariant changes between $k_z = 0$ and $k_z = \pi/a$, the system is a strong TI, otherwise a trivial insulator or weak TI. Again due to the interdependencies, the change in ν has to be the same for all three principal directions.

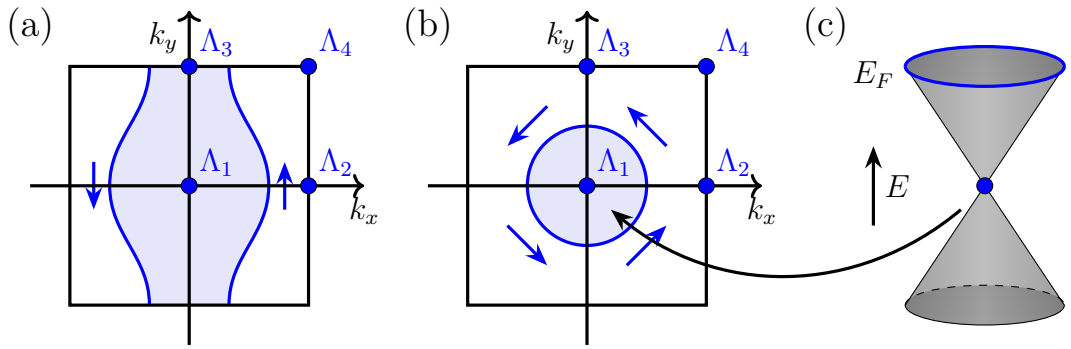


Figure 2.9. Surface BZ for a (a) weak TI and (b) strong TI. The blue lines denote the Fermi surface. (c) depicts the Dirac cone spectrum present at every TRIM Λ_a.

A compact notation for the 4 independent invariants is

$$(\nu_0; \nu_1\nu_2\nu_3)\,. \tag{2.26}$$

The first one is called the strong index. It states whether there is a change in $\mathbb{Z}_2$ invariants between the two special planes in each direction. Here $\nu_0 = 0$ means trivial and $\nu_0 = 1$ stands for topological. The other three are the weak indices. They are the $\mathbb{Z}_2$ invariants of the outer planes in each direction. The corresponding formulas are

$$(-1)^{\nu_0} = \prod_{n_j=0,\pi} \delta_{n_1,n_2,n_3}, \quad (-1)^{\nu_i} = \prod_{n_{j\neq i}=0,\pi;n_i=\pi} \delta_{n_1,n_2,n_3} \tag{2.27}$$

adapting the TRIM notation from Fig. 2.8 (b) together with Eq. (2.22).

These four invariants have direct physical consequences when we look at the surface BZ, see Fig. 2.9. At the 4 TRIMs $\Lambda_{1,2,3,4}$ the spectrum is Kramers degenerate. Otherwise, SOC lifts the degeneracy. Thus the TRIMs are Dirac points in the surface BZ. The question is how they connect to each other. Between any Λ_a and Λ_b the surface band structure looks like Fig. 2.10 (a) or (b). This plot illustrates the number of times the Fermi surface crosses the line connecting Λ_a and Λ_b. The number of crossings can be even or odd, where the odd surface states are topologically protected. One can distinguish weak and strong TIs, see Sec. 2.4, by the number of Kramers degenerate Dirac points the surface Fermi circle encloses. Odd for strong, even for weak TIs. Surface states of a weak TI are not protected by $\mathcal{T}$ symmetry.

2.2.4. Bulk-boundary Correspondence

The bulk topology, represented by the four $\mathbb{Z}_2$ invariants, has direct consequences for the surface physics as we already saw in the last section. Here we will deepen this bulk-boundary correspondence. This section is based on Ref. [FM13] Chap. 1, Secs. 2.3 & 3.3 & 4.3 & 5.2.2, Ref. [BH13] Sec. 8.8 and Ref. [She13] Sec. 2.2.

At interfaces of materials with different $\mathbb{Z}_2$ index, the gap has to close for the change in topology if the underlying symmetry is preserved throughout the system. The surface states associated with this gap closing at the boundary lie often in the bulk band gap, otherwise they are not stable due to coupling to the bulk bands. As a physical example, in systems where two bands have different symmetry, the topologically non-trivial (TnT) setup is related to a band inversion of these two bands, see. e.g. Sec. 2.4.1 and Sec. 2.4.2. This inversion has to be

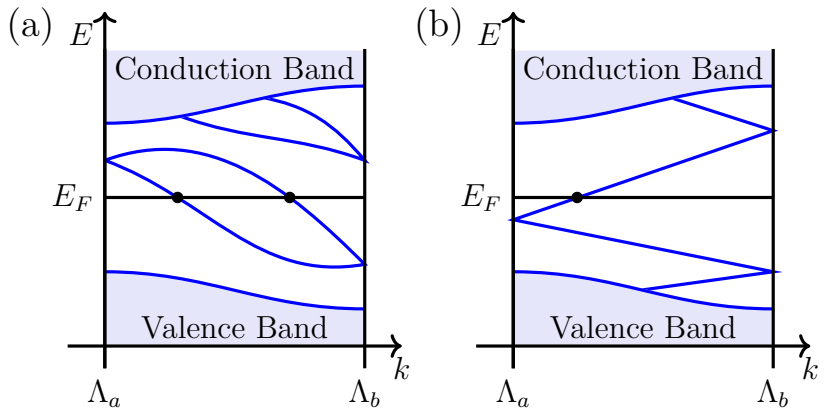


Figure 2.10. Dispersion relation between two TRIMs Λ_a and Λ_b. In (a) the surface state crosses the Fermi energy an even amount of times, in (b) an odd amount of times. The latter case is topologically protected, as the crossings have to occur due to the continuity of the bands.

undone at the boundary to change the topology to trivial. As the band structure is continuous, the gap has to close.

Let's assume boundary states in the bulk band gap as shown in Fig. 2.10. The bands have to be double degenerate at the TRIMs, but spin-orbit splits them everywhere else. If there is an odd number of intersections of the Fermi energy with the surface bands, these edge state can not be eliminated - they are topologically protected by the continuity of the band structure.

After this prelude, we give the bulk boundary correspondence. It relates the number of Kramers pairs of edge modes N_K to the change of $\mathbb{Z}_2$ at the boundary like

$$N_K = \Delta\nu \mod 2. \tag{2.28}$$

As these surface states originate from the bulk bands (they become bulk bands for high k), they can violate the fermion doubling theorem [WBZ06]. This makes them unique states in odd dimensions that are not possible without the higher dimensional bulk, see Ref. [QZ11] Sec. 2.C.3.

Let us have a look at a basic example of a boundary state in 1D. We take the Jackiw-Rebbi model, which in real space, with the substitution $k \to -i\partial_x$, is given by

$$H = -iv\sigma_x\partial_x + m(x)\,\sigma_y. \tag{2.29}$$

The spectrum is $E(k) = \pm\sqrt{(vk)^2 + m^2}$ and the model is in two different topological phases for $m(x \to +\infty) < 0$ and $m(x \to -\infty) > 0$. At the boundary of these two phases exists a zero energy solution, see Fig. 2.11. We construct it by multiplying Eq. (2.29) with $i\sigma_x$, the solution is then an eigenstate of σ_z denoted by $|z\pm\rangle$ with eigenvalue ± 1. The wave function is given by

$$\psi_0(x) = \mathrm{e}^{-\int_0^x \mathrm{d}x'(-m(x'))/v}|z+\rangle \tag{2.30}$$

which is localized around $m(x) = 0$. The state does not depend on the precise form of $m(x)$, but just stems from the sign change in m. This makes it topological.

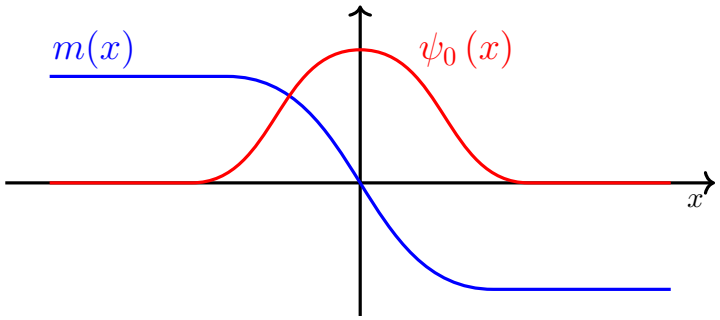


Figure 2.11. Sketch of the mass $m(x)$ and the zero energy solution wave function $\psi_0(x)$ as a function of x for the Jackiw-Rebbi model.

		TRS	PHS	SLS	1D	2D	3D
Standard	A (unitary)	0	0	0	-	$\mathbb{Z}$	-
(Wigner-Dyson)	AI (orthogonal)	+1	0	0	-	-	-
	AII (symplectic)	-1	0	0	-	$\mathbb{Z}_2$	$\mathbb{Z}_2$
Chiral	AIII (chiral unitary)	0	0	1	$\mathbb{Z}$	-	$\mathbb{Z}$
(sublattice)	BDI (chiral orthogonal)	+1	+1	1	$\mathbb{Z}$	-	-
	CII (chiral symplectic)	-1	-1	1	$\mathbb{Z}$	-	$\mathbb{Z}_2$
BdG	D	0	+1	0	$\mathbb{Z}_2$	$\mathbb{Z}$	-
	C	0	-1	0	-	$\mathbb{Z}$	-
	DIII	-1	+1	1	$\mathbb{Z}_2$	$\mathbb{Z}_2$	$\mathbb{Z}$
	CI	+1	-1	1	-	-	$\mathbb{Z}$

Table 2.1. Topological classification of matter with $\boldsymbol{\mathcal{T}}$ symmetry TRS, p-h symmetry PHS and chiral symmetry SLS. The symmetries can be absent, 0, or their operators can square to ± 1. Table adapted from Refs. [SRFL08, SRF+09].

2.2.5. Altland-Zirnbauer Classes

So far we looked at $\boldsymbol{\mathcal{T}}$ invariant systems and found that they can be classified by a $\mathbb{Z}_2$ invariant. This connection between preserved symmetries and topological invariants can be made more general. A topological classification of systems that can have $\boldsymbol{\mathcal{T}}$ symmetry TRS, particle-hole (p-h) PHS and chiral/sublattice symmetry, which is given by SLS=TRS$\times$PHS, is discussed in Ref. [She13] Sec. 12 and Refs. [SRFL08, SRF+09].

Here we just give the resulting classification table 2.1 and discuss some examples out of the ten possible classes. A representative of class A can be a 2DEG with magnetic field, having non of the three symmetries. Class AII has the $\boldsymbol{\mathcal{T}}$ symmetry operator squaring to -1, so an example would be an electron system with SOC.

Additional symmetries can provide an extended classification with additional edge states protected by $\mathbb{Z}_2$ invariants, see the comprehensive review in Ref. [CTSR16]. For spatial symmetries like $\boldsymbol{\mathcal{I}}$, see Ref. [vMOS16], these systems are called topological crystalline insulators (TCIs) and are discussed in Sec. 2.4.3.

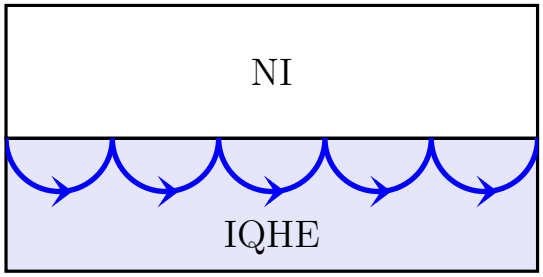


Figure 2.12. Sketch of electrons moving on skipping orbits along the edge of an IQH system.

2.3. Chern Insulator and Integer Quantum Hall Effect

The IQH effect was experimentally observed by von Klitzing, Ref. [KDP80], in 1980. It features a quantized Hall conductance in a 2DEG in a strong, perpendicular magnetic field. Before we discuss the connection to topology of this quantized response, we give a more straightforward explanation of the effect. In the magnetic field the electrons form quantized LLs with a flat dispersion. At the edge, the confining potential pushes the LLs through the Fermi energy. Thus for n filled LLs, one gets a Hall conductance of

$$\sigma_{x,y} = n\frac{e^2}{h}. \tag{2.31}$$

In a semi-classical picture, the electrons move on skipping orbits along the edge, see Fig. 2.12. As the magnetic field determines the direction of motion of the electrons, the edge states are chiral. No backscattering is allowed, as the edge state propagating in the opposite direction is located on the opposite side of the sample.

We will now make the connection between the IQH effect and the topological framework discussed to far. This section is based on Ref. [FM13] Chap. 1, Sec. 4, Ref. [BH13] Sec. 3, Ref. [Bud11] Sec. 1.3 and Ref. [She13] Sec. 4.3.

Integer Quantum Hall Effect and Chern Number

The quantized Hall conductance in a system with finite Chern number can directly be reasoned from the bulk-boundary correspondence discussed in Sec. 2.2.4. At a junction of a non-trivial material with finite Chern number and a trivial one with $n = 0$, the difference in number of right and left moving modes, $N_R - N_L$, is, similar to Eq. (2.28), given by the difference Δn in Chern numbers across the interface

$$N_R - N_L = \Delta n. \tag{2.32}$$

This analogy can be checked in linear response calculations, done e.g. by Thouless, Kohmoto, Nightingale, and den Nijs (TKNN) in Refs. [TKNdN82, Koh85]. They reveal the deep connection between Chern number n and Hall conductance σ_{xy} to be

$$\begin{aligned} \sigma_{xy} &= \frac{e^2}{h}n = \frac{e^2}{h}\frac{1}{2\pi}\int\int \mathrm{d}_{k_x}\mathrm{d}_{k_y}\mathcal{F}_{xy}\left(\boldsymbol{k}\right), \\ \mathcal{F}_{xy}\left(\boldsymbol{k}\right) &= \frac{\partial A_y\left(\boldsymbol{k}\right)}{\partial k_x} - \frac{\partial A_x\left(\boldsymbol{k}\right)}{\partial k_y}, \\ A_j &= i\sum_{n\ filled\ bands}\langle u_{n,\boldsymbol{k}}|\frac{\partial}{\partial k_j}|u_{n,\boldsymbol{k}}\rangle. \end{aligned} \tag{2.33}$$

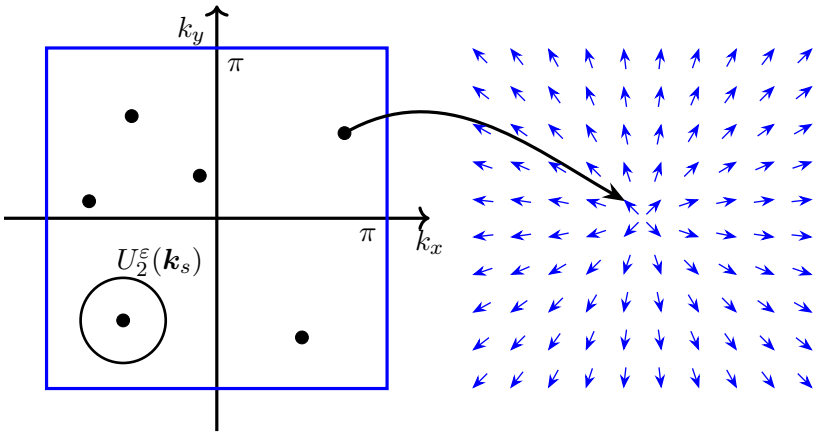


Figure 2.13. (left) BZ of an IQH system. The dots denote the zeros of the first component of the state $|\psi_1(\boldsymbol{k})\rangle$. (right) The phase around a simple zero.

Chern Number as an Obstruction to Stokes' Theorem over the whole Brillouin Zone

Apart from a direct transport calculation, the integer quantization of Chern number and Hall conductance can also be derived from more mathematical reasoning. The Berry curvature is defined as the curl of the Berry connection $\boldsymbol{A}$, see Sec. 2.2.2. Stokes' theorem would give the Chern number, Eq. (2.15), as the integral of $\boldsymbol{A}$ over the boundary of the BZ. As the BZ is a torus without a boundary, the Chern number is always 0 following this line of thought, which is not true. The way out is that $\boldsymbol{A}$ has to be well defined in the whole BZ for Stokes' theorem to apply. Thus for a finite Chern number, the Berry connection needs singularities. This is equivalent to the case that no global gauge can be defined that is continuous and single-valued over the entire BZ.

We show this now in detail. Let's define a smooth gauge: we choose a state $|\psi_1(\boldsymbol{k})\rangle = \mathrm{e}^{ih(\boldsymbol{k})}|u_{\boldsymbol{k}}\rangle$ by gauging the first component to be real. At some point $\boldsymbol{k}_s$ in the BZ the first component is zero, see Fig. 2.13, thus here the phase is ill defined. Yet the Bloch function can not be entirely zero. Therefore we define a different state $|\psi_2(\boldsymbol{k})\rangle = \mathrm{e}^{ig(\boldsymbol{k})}|u_{\boldsymbol{k}}\rangle$ in the epsilon neighborhood U_2^ε around $\boldsymbol{k}_s$. At the boundary, the states are related by the gauge transformation

$$\begin{aligned} |\psi_2(\boldsymbol{k})\rangle &= \mathrm{e}^{i\chi(\boldsymbol{k})}|\psi_1(\boldsymbol{k})\rangle \\ \boldsymbol{A}_2(\boldsymbol{k}) &= \langle\psi_2(\boldsymbol{k})|\partial_{\boldsymbol{k}}|\psi_2(\boldsymbol{k})\rangle = \boldsymbol{A}_1(\boldsymbol{k}) + i\nabla\chi(\boldsymbol{k}) \end{aligned} \tag{2.34}$$

with $\chi(\boldsymbol{k}) = g(\boldsymbol{k}) - h(\boldsymbol{k})$. Performing the integral over the BZ, one finds the winding number of the gauge transformation

$$\begin{aligned} n &= \frac{1}{2\pi}\int_{T^2}\mathcal{F}\mathrm{d}^2\boldsymbol{k} = \frac{1}{2\pi}\sum_j\oint_{\partial U_j}\boldsymbol{A}_j\cdot\mathrm{d}\boldsymbol{k} \\ &= \frac{1}{2\pi}\oint_{\partial U_2^\varepsilon}\nabla\chi(\boldsymbol{k})\cdot\mathrm{d}\boldsymbol{k}. \end{aligned} \tag{2.35}$$

The states $|\psi_2(\boldsymbol{k})\rangle = \mathrm{e}^{\chi(\boldsymbol{k})}|\psi_1(\boldsymbol{k})\rangle$ have to be single-valued, so after a rotation of 2π around the point $\boldsymbol{k}_s$, $\chi(\boldsymbol{k})$ can only have changed by a multiple of 2π in order to keep the state invariant. We conclude that the Chern number has integer values, and the winding number n counts the order of the zero at $\boldsymbol{k}_s$.

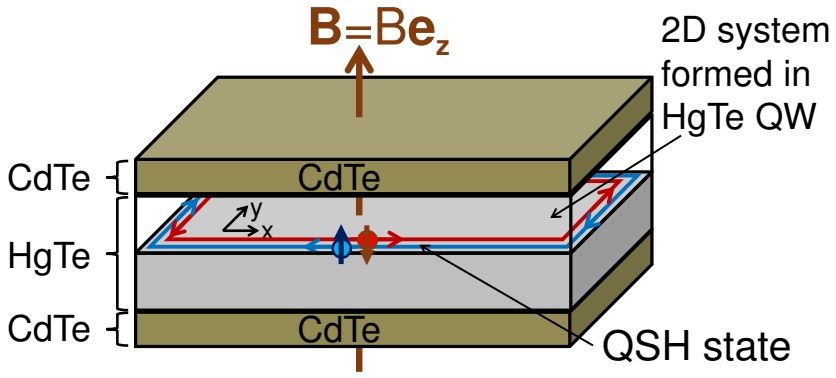


Figure 2.14. Sketch of a HgTe QW as a representative of the QSH insulators. The 1D helical edge states are indicated. The magnetic field $\boldsymbol{B}$ is zero when considering the QW as a 2D TI. A finite $\boldsymbol{B} \neq 0$ turns the system into a quantum Hall topological insulator (QHTI) as analyzed in Chap. III. Reprinted figure with permission from Ref. [SMAŽF15]. Copyright (2015) by the APS.

2.4. Topological Insulators

Topological insulators (TIs) differ from normal insulators (NIs) as they have a non-trivial topology, represented by a finite $\mathbb{Z}_2$ invariant, see Ref. [KM05b] and Sec. 2.2.3. This non-triviality usually stems from and is protected by $\boldsymbol{\mathcal{T}}$ symmetry. These systems have no Hall conductance as their Chern number is 0, but feature helical edge states, in which momentum and spin of the electrons are locked. Besides $\boldsymbol{\mathcal{T}}$, also spatial symmetries like $\boldsymbol{\mathcal{I}}$ can support topological states with finite $\mathbb{Z}_2$, see the discussion of the TCI in Sec. 2.4.3.

In the following, we will go through the details of TIs in 2D, Sec. 2.4.1, in 3D with a focus on strong TIs, Sec. 2.4.2, and finish with a short summary on TCI in Sec. 2.4.3. Good reviews of the different topics can be found in Refs. [HK10, QZ11, And13, BT13].

2.4.1. Quantum Spin Hall Insulator in 2D

The first proposed QSH insulator was graphene, see Ref. [KM05a]. Its topological properties stem from spin-orbit interaction, which acts in a way as a $\boldsymbol{\mathcal{T}}$ symmetric version of a magnetic field, as it switches its sign with the spin. A way of looking at this 2D system is thus an overlay of two IQH systems with differing spin. This gives a finite $\mathbb{Z}_2$ invariant according to Eq. (2.24). In graphene, one can write the spin-orbit term as $H_{SO} = \lambda_{SO}\sigma_z\tau_z s_z$ that is $\boldsymbol{\mathcal{T}}$ invariant. Here $s^z = \pm 1$ is for spin up / down and τ_z (σ_z) is the valley (sublattice) degree of freedom. But there is a practical problem: the spin-orbit gap in graphene is very small, on the order of 10^{-6} eV. So we take these ideas and look for a different material with stronger SOC. In this thesis, HgTe QWs will serve as the primary example of a 2D TI, see Fig. 2.14. The following section is based on Ref. [FM13] Chap. 1, Sec. 5.1 & Chap. 3, Sec. 2, Ref. [BH13] Secs. 9.3f, Ref. [She13] Secs. 2.5.2 & 6.1f as well as Ref. [Bud11] Sec. 1.4.

Let us consider the edge properties of a QSH insulator. Taking the quantized Hall conductance of each spin species, the charge part of the edge current cancels. The spin parts on the other hand add up, thus there is a finite spin current $\boldsymbol{J}_s = \frac{\hbar}{2e}(\boldsymbol{J}_\uparrow - \boldsymbol{J}_\downarrow)$ with spin Hall conductivity $\sigma^s_{xy} = \frac{e}{2\pi}$ at the edge, see Fig. 2.14. This simplified picture is usually not quite correct, as in a spin-orbit coupled system the S_z spin component is not conserved. Yet the quantized Hall conductivity is robust and no backscattering is allowed for preserved $\boldsymbol{\mathcal{T}}$ symmetry. Due

to the locking of spin and momentum, these edge states are called helical from the helicity

$$h = \pm \boldsymbol{p} \cdot \boldsymbol{\sigma}. \tag{2.36}$$

Left and right movers at the edge are Kramers partners. They can not scatter into one another which is straightforward to show. We define the states $|u_{k,\uparrow}\rangle = -\Theta|u_{-k,\downarrow}\rangle$, $|u_{-k,\downarrow}\rangle = \Theta|u_{k,\uparrow}\rangle$ and take the $\boldsymbol{\mathcal{T}}$ symmetry respecting potential V with $\Theta V \Theta^{-1} = V$. We find

$$\langle u_{k,\uparrow}|V|u_{-k,\downarrow}\rangle = 0 \tag{2.37}$$

thus there is no overlap between the Kramers partners for scattering potentials that respect $\boldsymbol{\mathcal{T}}$ symmetry. For additional information about the stability of edge states and the allowed scattering processes, see Refs. [XM06, WBZ06].

HgTe Quantum Wells and the Bernevig-Hughes-Zhang Model

The primary example for 2D QSH insulators in this thesis are HgTe/CdTe QWs, as depicted in Fig. 2.14, which we now introduce. We follow closely the original theory paper Ref. [BHZ06] and the experimental counterpart, Ref. [KWB+07].

HgTe and CdTe are II-VI group compound semiconductors with a zinc-blende crystal structure, forming via sp^3 hybridization. The bands near the Fermi energy are s-type (Γ_6) and p-type, where the latter split by SOC into a $J = 3/2$ band (Γ_8) and $J = 1/2$ (Γ_7), see Fig. 2.15 (a). CdTe has a band gap of ≈ 1.6 eV with normal ordering, such that the Γ_6 lie above the p-type bands. In this sense, HgTe has a negative band gap of ≈ -0.3 eV, as the order of the bands is inverted with Γ_6 below Γ_8. Actually, due to the degeneracy at $k = 0$ (Γ-point) between light- and heavy-hole bands, HgTe is a zero gap material, as shown in Fig. 2.15 (a).

In order to obtain a 2D TI, one builds QW heterostructures by putting a thin layer of HgTe of thickness d between CdTe layers, see Fig. 2.15 (b). The confinement splits the HgTe bands into subbands, where we take the stacking direction of the QW to be the z direction. The Γ_6 and light-hole Γ_8 bands form the new bands E_n. They hybridize because both have angular momentum $1/2$ along the z direction ([001] direction). The heavy-hole Γ_8 bands form subbands H_n. The energy of the subbands varies with QW thickness d, see Fig. 2.16. E_n decreases in energy with increasing width, while the heavy-hole H_n increase in energy. Thus for thin HgTe, the QW band structure has normal order, for a thick HgTe layer, it is inverted. At the sweet spot of $d_c \approx 6.3$ nm the 2D band structure is degenerate.

The goal is now to construct an effective low energy model for such a HgTe QW. As the basis one takes $\{|E_1+\rangle, |H_1+\rangle, |E_1-\rangle, |H_1-\rangle\}$ with $\pm$ for spin up / down. $|E_1\pm\rangle$ and $|H_1\pm\rangle$ are two sets of Kramers partners which are spin degenerate. The coupling between $|E_1\pm\rangle$ and $|H_1\pm\rangle$ has to be, in lowest order, linear in $\boldsymbol{k}$ due to the different parity of the bands. The rotational symmetry around the z-axis then dictates the form $k_\pm = k_x \pm i k_y$, such that the Hamiltonian is given by

$$H_{BHZ} = \begin{pmatrix} h(\boldsymbol{k}) & 0 \\ 0 & h^*(-\boldsymbol{k}) \end{pmatrix}, \quad h(\boldsymbol{k}) = E(\boldsymbol{k})\, \mathbb{1}_{2x2} + \boldsymbol{d}(\boldsymbol{k}) \cdot \boldsymbol{\sigma} \tag{2.38}$$

with

$$\begin{aligned} E(\boldsymbol{k}) &= \mathcal{C} - \mathcal{D}(k_x^2 + k_y^2), \\ \boldsymbol{d}(\boldsymbol{k}) &= \begin{pmatrix} \mathcal{A}k_x, & \mathcal{A}k_y, & \mathcal{M}(\boldsymbol{k}) \end{pmatrix}^T, \\ \mathcal{M}(\boldsymbol{k}) &= M - \mathcal{B}(k_x^2 + k_y^2). \end{aligned} \tag{2.39}$$

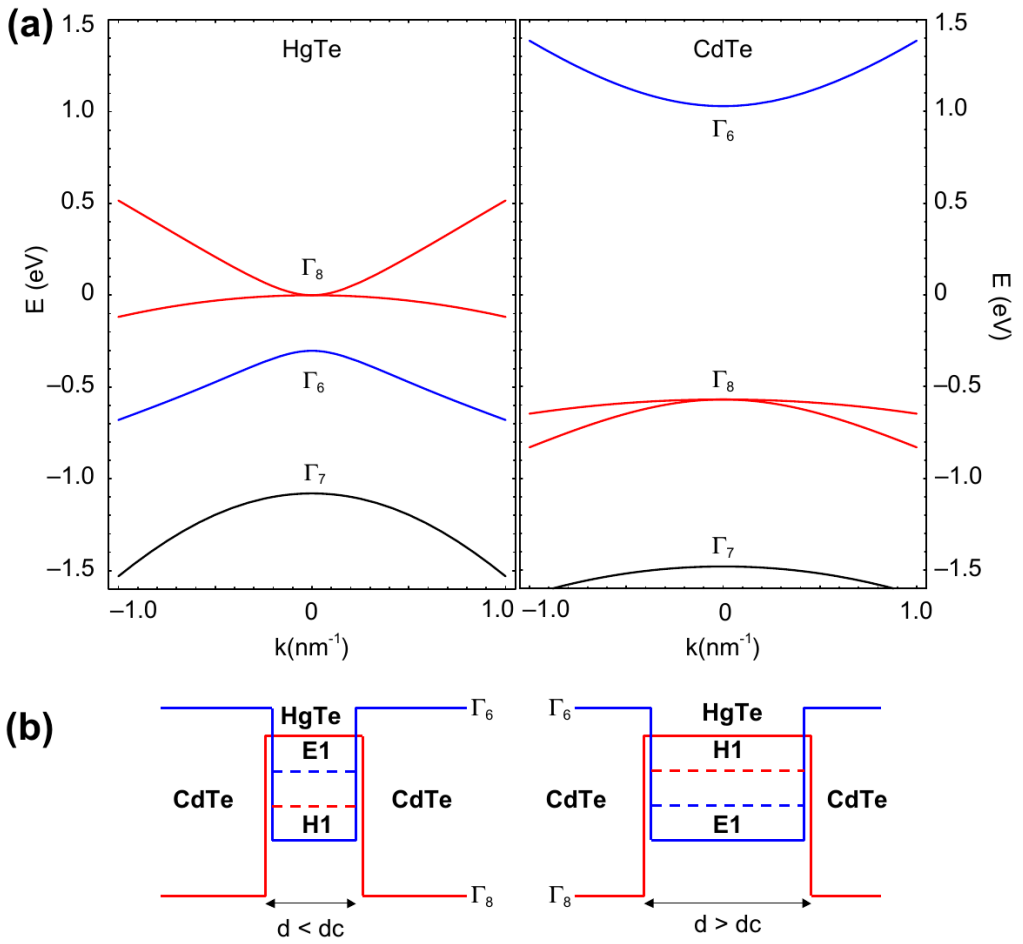


Figure 2.15. (a) Band structure of HgTe and CdTe. (b) Effective band structure of HgTe/CdTe QWs with different thickness d. From Ref. [BHZ06]. Reprinted with permission from AAAS.

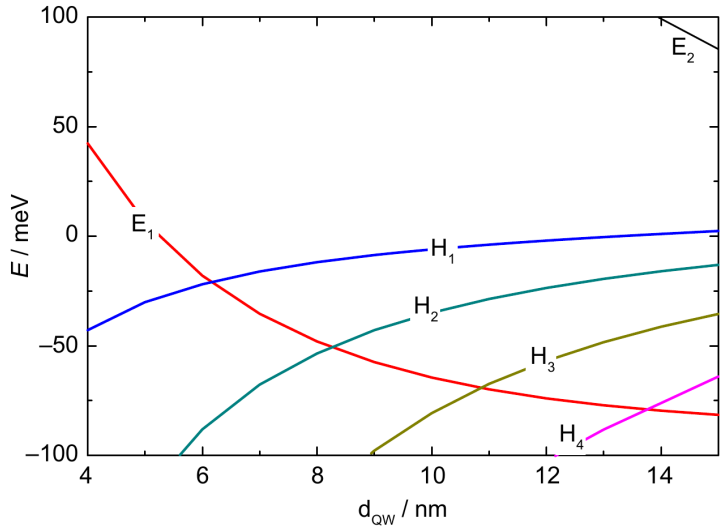


Figure 2.16. Energy of QW subbands E_n and H_n as a function of the QW thickness d_{QW}. ©[2008] The Physical Society of Japan, Ref. [KBW+08].

This Bernevig-Hughes-Zhang (BHZ) model has the parameters $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}, M$ which are determined by the material, depending on the QW geometry. Typical values for Hg(Cd)Te QWs are given in Sec. 4.4. Besides HgTe, this model can also be applied to QW structures based on different materials like InAs/GaSb/AlSb, see Refs. [KDS11, LWF+15a].

The BHZ model is block diagonal, such that one can restrict the analysis to one block. The other one follows from $\boldsymbol{\mathcal{T}}$ symmetry. The dispersion relation is given by

$$E_{k,\pm} = E(\boldsymbol{k}) \pm \sqrt{\mathcal{A}^2\left(k_x^2 + k_y^2\right) + \mathcal{M}^2(\boldsymbol{k})}. \tag{2.40}$$

It reproduces experimental band structures well for d close to d_c. In contrast to the Dirac model of graphene, the BHZ model is properly regularized, due to the finite quadratic part for $\mathcal{B} \neq 0$.

Considering the upper block of Eq. (2.38), one finds the Hall conductance to be

$$\begin{aligned} \sigma_h &= -\frac{e^2}{h}\frac{1}{4\pi}\int \mathrm{d}k_x \int \mathrm{d}k_y \hat{\boldsymbol{d}} \cdot \left(\frac{\partial \hat{\boldsymbol{d}}}{\partial k_x} \times \frac{\partial \hat{\boldsymbol{d}}}{\partial k_y}\right) \\ &= -\frac{1}{2}\left(\mathrm{sgn}(M) + \mathrm{sgn}(\mathcal{B})\right)\frac{e^2}{h} \end{aligned} \tag{2.41}$$

with $\hat{\boldsymbol{d}}(\boldsymbol{k}) = \boldsymbol{d}(\boldsymbol{k})/|\boldsymbol{d}(\boldsymbol{k})|$. σ_h is only finite if the inversion condition

$$M/\mathcal{B} > 0 \tag{2.42}$$

is fulfilled. Usually the mass term M is varied to change between the NI and TI phase, or explicitly tuned to the critical thickness in order to obtain a gapless Dirac system, as analyzed in Ref. [BLT+11]. The vector $\hat{\boldsymbol{d}}$ can be represented by a meron structure close to the Γ-point, as shown in Fig. 2.17. A meron is half a skyrmion, see Sec. 3.2.3, which are topological objects that form in (pseudo-)spin space. For larger momenta, the k^2 terms dominate and the vector $\hat{\boldsymbol{d}}$ points again out of plane. This completes the winding, making the spin structure a full skyrmion.

The total model consists then of two $\boldsymbol{\mathcal{T}}$ related copies of the Hamiltonian discussed above, thus that the finite $\mathbb{Z}_2$ invariant according to Eq. (2.24) is given by $\nu = 1$ for fulfilled inversion condition.

The Hamiltonian (2.38) was derived under the assumption of $\boldsymbol{\mathcal{I}}$ symmetry. This symmetry is slightly broken in HgTe QWs, which can be modeled by a bulk inversion asymmetry (BIA) term. Up to linear order in $\boldsymbol{k}$, it is given by

$$H_{BIA} = \begin{pmatrix} 0 & 0 & \Delta_e k_+ & -\Delta_z \\ 0 & 0 & \Delta_z & \Delta_h k_- \\ \Delta_e k_- & \Delta_z & 0 & 0 \\ -\Delta_z & \Delta_h k_+ & 0 & 0 \end{pmatrix}. \tag{2.43}$$

Importantly, it does not close the band gap such that the topological physics remains valid if one includes it in the analysis. Additional terms that are ignored are due to the crystal symmetry breaking by forming the QW. These effects can be considered in structural inversion asymmetry (SIA) terms, which we neglect for simplicity.

The QSH effect has been measured based on transport experiments in 4 and 6 terminal setups, see e.g. Refs. [KWB+07, KBW+08].

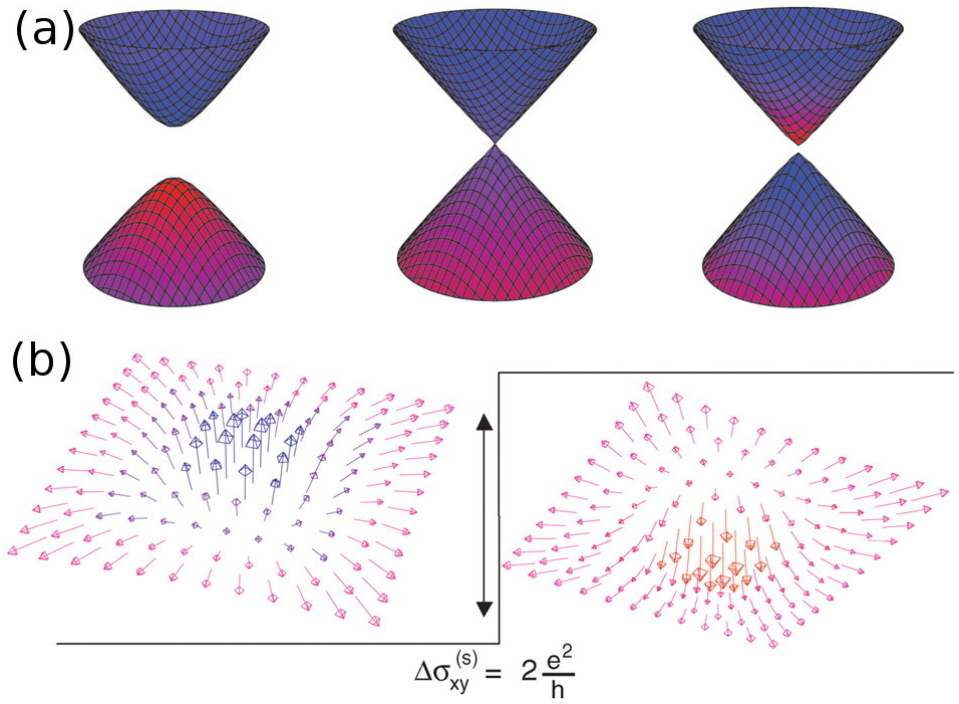


Figure 2.17. (a) Band structure and (b) meron configuration of the vector $\hat{\boldsymbol{d}}$ for the NI (left) and TI phase (right). From Ref. [BHZ06]. Reprinted with permission from AAAS.

2D Topological Insulator in a Magnetic Field

In Chap. III we will study HgTe QWs in a strong magnetic field, see Fig. 2.14, based on the BHZ Hamiltonian. In this section, we provide the underlying formulas, following Ref. [SMAF12].

The effect of a magnetic field $\boldsymbol{B} = B\boldsymbol{e}_z$ in the z direction of HgTe QWs can be taken into account in the BHZ Hamiltonian, Eq. (2.38), by two additions. The first one is a Zeeman term $\mu_B B g_{e/h}/2$, with μ_B the Bohr magneton and $g_{e/h}$ the effective g factor for an electron or a hole. The second addition is the minimal coupling, which changes the components of the wave vector to operators, $k_i \to \hat{\pi}_i/\hbar$, with $\hat{\pi}_i = \hat{p}_i + eA_i(\boldsymbol{r})$ the kinetic momentum operators. Here $\hat{p}$ are the momentum operators and $\boldsymbol{A}(\boldsymbol{r}) = -By\boldsymbol{e}_x$ is the magnetic vector potential from the magnetic field.

The LLs forming in the bulk are found analytically to be ($B > 0$ for magnetic field in z direction)

$$\begin{aligned} E_{\pm}^{\uparrow}(n) &= \mathcal{C} - \frac{2\mathcal{D}n + \mathcal{B}}{l_B^2} + \frac{g_e + g_h}{4}\mu_B B \pm \sqrt{\frac{2n\mathcal{A}^2}{l_B^2} + \left(M - \frac{2\mathcal{B}n + \mathcal{D}}{l_B^2} + \frac{g_e - g_h}{4}\mu_B B\right)^2} \\ E^{\uparrow}(0) &= \mathcal{C} + M - \frac{\mathcal{D} + \mathcal{B}}{l_B^2} + \frac{g_e}{2}\mu_B B \end{aligned} \tag{2.44}$$

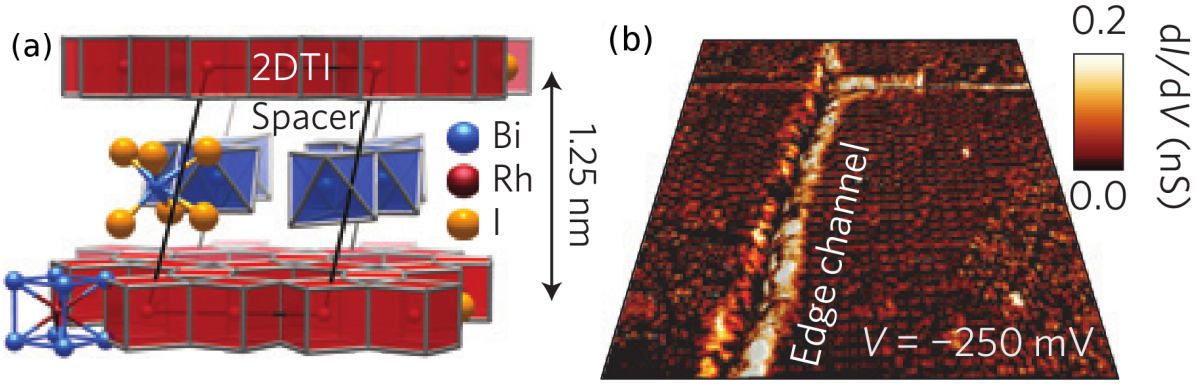


Figure 2.18. (a) Alternating stacking of a 2D TI material with a spacer layer. (b) STM image of the surface of the layered material. Strongest dI/dV signal at the edge. Adapted by permission from Macmillan Publishers Ltd: Nature Physics, Ref. [PRK+15], copyright (2015).

and

$$E_{\pm}^{\downarrow}(n) = \mathcal{C} - \frac{2\mathcal{D}n - \mathcal{B}}{l_B^2} - \frac{g_e + g_h}{4}\mu_B B \pm \sqrt{\frac{2n\mathcal{A}^2}{l_B^2} + \left(M - \frac{2\mathcal{B}n - \mathcal{D}}{l_B^2} - \frac{g_e - g_h}{4}\mu_B B\right)^2}$$
$$E^{\downarrow}(0) = \mathcal{C} - M - \frac{\mathcal{D} - \mathcal{B}}{l_B^2} - \frac{g_e}{2}\mu_B B \tag{2.45}$$

with $l_B = \sqrt{\frac{\hbar}{eB}}$ the magnetic length. These formulas are the basis for Fig. 7.1. Introducing edges to the system complicates the picture a lot. We refer for details to Ref. [SMAF12].

2.4.2. 3D Topological Insulator

TIs in 3D can be classified by the 4 $\mathbb{Z}_2$ invariants $(\nu_0; \nu_1\nu_2\nu_3)$ as described in Sec. 2.2.3. ν_0 is called the strong index, as it is 1 for strong TIs and 0 for weak TIs or trivial insulators. The three weak indices ν_i, $i = 1, 2, 3$, determine the surfaces which have topological surface states in the weak TI case. In the following, we will go into the details of the different types of TIs in 3D. These sections are based on Ref. [FM13] Chap. 1, Sec. 5.3, Chap. 2, Sec. 2.2, Chap. 3, Sec. 3 and Ref. [She13] Sec. 7.2.

Weak 3D Topological Insulator

A weak 3D TI can be thought of as a system of stacked 2D QSH insulators, where the layers are weakly coupled. The effective surface state will be anisotropic, and perpendicular to the direction of stacking there will be no surface state at all. This is the so called dark surface. Weak surface states are not protected by $\mathcal{T}$ symmetry. They exist for clean systems, but broken translational symmetry can destroy them without a gap closing in the bulk. This can be modeled by a dimerization of stacked layers.

An example of the stacking approach is the material $Bi_{14}Rh_3I_9$ examined in Refs. [RIR+13, PRK+15], consisting of alternately stacked layers of a 2D TI and a trivial insulator, as sketched in Fig. 2.18 (a). The surface of this material is dark, but at every step edge there is supposed

to form a topologically protected edge channel, see Fig. 2.18 (b). We will not go into further details here, as this thesis focuses on the strong 3D TI, which will be discussed next.

Strong 3D Topological Insulator

The first material experimentally considered as a strong TI was Bi_xSb_{1-x}, see Ref. [HQW+08]. This soon improved to Bi_2Se_3 with a large band gap of 0.3 eV, see Ref. [XQH+09], but still facing problems due to significant bulk conductivity. An alternative material is strained 3D HgTe. Instead of forming QWs out of it as for 2D TIs, the strain opens a band gap between the light- and heavy-hole Γ_8 bands. The strong 3D TI with $\nu_0 = 1$ is relevant for Chap. IV of this thesis, where we will couple it to a WSM.

We will stay with the BiSe family in this section. A low-energy model was derived in Refs. [ZLQ+09, LQZ+10], which we are now going to recap here. Bi_2Se_3 forms in a rhombohedral crystal structure with space group $D_{3d}^5\,(R\bar{3}m)$ with a quintuple layer structure as depicted in Fig. 2.19 (a). The energy level structure close to the Fermi energy is shown in Fig. 2.19 (b). The derivation includes steps (I) to (IV): (I) First the levels of Bi and Se repel each other, while they hybridize among themselves. (II) The resultant levels are sorted by parity and form bonding/antibonding states. (III) The crystal field is considered, which splits p_z from $p_{x,y}$. (IV) Finally SOC, $H_{SO} = \lambda \boldsymbol{L}\cdot\boldsymbol{S}$, is taken into account. It mixes spin and orbital angular momentum. Only the total angular momentum is conserved.

At last, one ends up with four states with definite angular momentum close to the Fermi energy: $|P1^+_-, \frac{1}{2}\rangle$, $|P2^-_+, \frac{1}{2}\rangle$, $|P1^+_-, -\frac{1}{2}\rangle$, $|P2^-_+, -\frac{1}{2}\rangle$. They are linear combinations of the original levels, where P1 stands for Bi and P2 for Se. SOC leads to an inversion of the states originating from Bi and Se, which makes the material topological. The low-energy Hamiltonian up to quadratic order in $\boldsymbol{k}$ is given by

$$H_{BiSe} = \varepsilon_{\boldsymbol{k}} + \begin{pmatrix} \mathcal{M}(\boldsymbol{k}) & -i\mathcal{B}(k_z)\,k_z & 0 & i\mathcal{A}(k_\parallel)k_- \\ i\mathcal{B}(k_z)\,k_z & -\mathcal{M}(\boldsymbol{k}) & i\mathcal{A}(k_\parallel)k_- & 0 \\ 0 & -i\mathcal{A}(k_\parallel)k_+ & \mathcal{M}(\boldsymbol{k}) & -i\mathcal{B}(k_z)\,k_z \\ -i\mathcal{A}(k_\parallel)k_+ & 0 & i\mathcal{B}(k_z)\,k_z & -\mathcal{M}(\boldsymbol{k}) \end{pmatrix} \tag{2.46}$$

with $\varepsilon_{\boldsymbol{k}} = C_0 + C_1k_z^2 + C_2k_\parallel^2$, $\mathcal{M}(\boldsymbol{k}) = M_0 + M_1k_z^2 + M_2k_\parallel^2$, $\mathcal{A}(k_\parallel) = A_0 + A_2k_\parallel^2$, $\mathcal{B}(k_z) = B_0 + B_2k_z^2$, $k_\parallel^2 = k_x^2 + k_y^2$ and $k_\pm = k_x \pm ik_y = k_\parallel e^{\pm i\phi_k}$. All free parameters need to be fitted to experimental or ab-initio band structures. The Hamiltonian is in the inverted, topological regime for $M_0M_j < 0$, $j = 1, 2$.

Surface Physics

The surface states are probably the most interesting features of TIs. We will have a closer look at them with the help of the Bi_2Se_3 Hamiltonian (2.46). Helpful sources are Ref. [FM13] Chap. 3, Secs. 5.1f and Ref. [She13] Secs. 7.4f.

The surface Hamiltonian is obtained from the bulk one, Eq. (2.46), following Ref. [LQZ+10]. The authors of Ref. [LQZ+10] find

$$H_{BiSe}^{sur} = C_0 + \alpha_2 M_0 + (C_2 + \alpha_3 M_2)\,k_\parallel^2 + A_0\alpha_1\,(\sigma_x k_y - \sigma_y k_x) \tag{2.47}$$

where the coefficients α_i can be determined, for instance, by fits to experimental data. The surface Hamiltonian shows spin-momentum locking, denoting the rotation of spin with momentum as depicted in Fig. 2.20. The term spin stands here as a synonym for the total angular

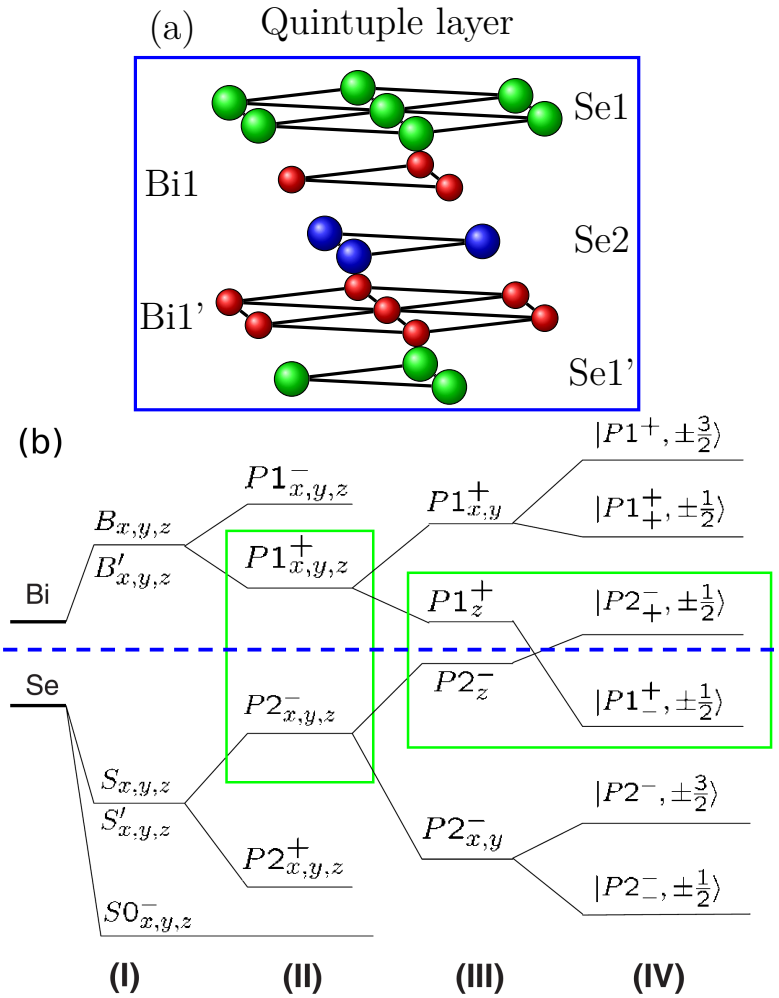


Figure 2.19. (a) Quintuple layer structure of Bi_2Se_3. (b) Hybridization of Bi and Se energy levels near the Fermi energy for different considered approximation steps. (I) Repulsion between Bi and Se levels, orbitals of the two materials hybridize among themselves. (II) Forming of bonding/antibonding orbitals, (III) including crystal field splitting, (IV) including SOC. Reprinted figure (b) with permission from Ref. [LQZ+10]. Copyright (2010) by the APS.

momentum.

Such a surface state is quite unique. Dropping all constant and quadratic in $\boldsymbol{k}$ terms, one obtains the simplest version, a single Dirac cone, as

$$H^{sur} = v_F \left(\sigma_x k_y - \sigma_y k_x\right). \tag{2.48}$$

This Hamiltonian describes half of an ordinary metal, as there is no spin degeneracy. The other half is located on the opposite surface, both obey Kramers degeneracy separately. The states are called helical because of the spin-momentum locking, resembling the helicity operator Eq. (2.36).

$\boldsymbol{\mathcal{T}}$ symmetry protects the gaplessness of the 2D surface state. One can not open a gap by adding a mass term to Eq. (2.48) in a $\boldsymbol{\mathcal{T}}$ invariant fashion, which is straightforward to show.

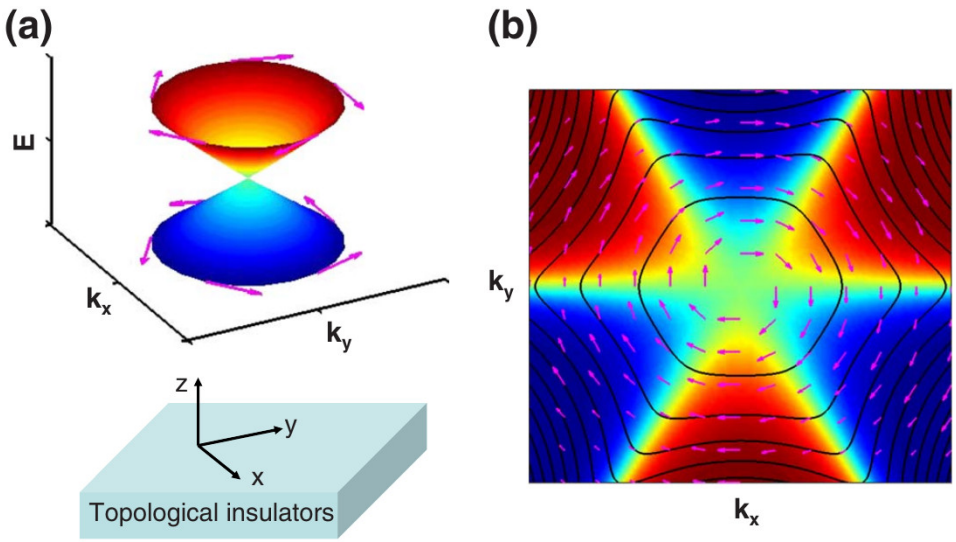


Figure 2.20. (a) Sketch of the Dirac cone dispersion relation of the surface state of a strong TI, including the spin texture. (b) Spin structure of the conduction band of the surface state showing spin-momentum locking. Arrows indicate $x-y$ planar spin polarization, while the color gives the z component. Red for spin up, blue for spin down. Reprinted figure with permission from Ref. [LQZ+10]. Copyright (2010) by the APS.

Take the $\boldsymbol{\mathcal{T}}$ operator

$$\Theta = i\sigma_y K \tag{2.49}$$

and let it act on the mass terms

$$H_{mass} = \sum_i m_i \sigma_i. \tag{2.50}$$

A finite $m_z \neq 0$ would open a gap in the spectrum of Eq. (2.48). One finds $\Theta H_{mass}\Theta^{-1} = -H_{mass}$, thus no mass term is allowed, the surface states are necessarily gapless.

We end this section by discussing two interesting properties of these helical edge states. The first one is that they show weak antilocalization instead of the usual weak localization, see Fig. 2.21 (b). This can be easily understood. Consider the interference of two time-reversed scattering paths of an electron in the surface state, as depicted in Fig. 2.21 (a). The eigenstate solution to Eq. (2.48) is the usual spinor of a spin 1/2 electron. Thus moving the electron in a circle, rotating it by 2π, gives a π Berry phase contribution to the wave function. This sign changes the interference of the two paths from constructive to destructive. The increased conductivity should be visible in experiments on sufficiently clean surfaces.

The second property is the absence of backscattering. Left and right movers in a surface state are Kramers partners. We define $|k,\uparrow\rangle$ and $|-k,\downarrow\rangle = \Theta|k,\uparrow\rangle$, take the $\boldsymbol{\mathcal{T}}$ symmetry respecting potential V with $\Theta V \Theta^{-1} = V$, and find

$$\langle -k,\downarrow|V|k,\uparrow\rangle = 0. \tag{2.51}$$

Thus electrons in the Dirac cone can not scatter into the state with opposite momentum. Yet the protection is not as strong as in the 1D edge state discussed in Sec. 2.4.1. Scattering elsewhere into the cone, where the momentum is not exactly opposite, is allowed in principle, although with a reduced overlap factor.

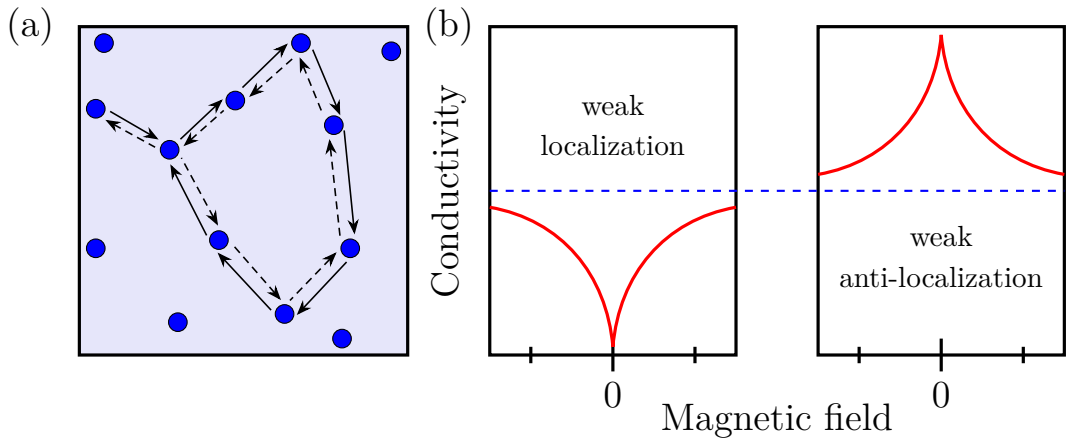


Figure 2.21. (a) Backscattering involving two time-reversed paths. (b) Weak localization and anti-localization in the magnetoconductivity. The dashed line denotes the classical conductivity.

2.4.3. Topological Crystalline Insulators

The usual symmetry that protects TIs is $\mathcal{T}$ symmetry. More recently, spatial symmetries moved into the focus of research which as well can support topological phases with finite $\mathbb{Z}_2$ invariant. We will give a short overview of the topic in the following. This section is based on Ref. [FM13] Chap. 1, Sec. 6.1, Ref. [And13] Sec. 4.5, as well as Refs. [SS14, CTSR16, vMOS16].

Spatial symmetries can modify the topological structure of a band theory. While the combination of $\mathcal{T}$ and p-h symmetry gives 10 different classes, see Sec. 2.2.5, a complete classification for all space groups is much more cumbersome. A comprehensive overview of the topic is given in Refs. [CTSR16, SS14, SSG16]. Insulators with non-trivial topology protected by point-group symmetries are called topological crystalline insulators (TCIs). Example symmetries include rotations and mirror symmetry as well as $\mathcal{I}$ symmetry, see Refs. [FGB12, LL14]. Also the combination of $\mathcal{T}$ with spatial symmetries can lead to new $\mathbb{Z}_2$ invariants. This was recently shown for the combination of $\mathcal{I}$ and $\mathcal{T}$ symmetry in 2D insulators in Ref. [vMOS16]. In the corresponding system that would be trivial with just $\mathcal{T}$ symmetry, the inclusion of $\mathcal{I}$ symmetry leads to a non-trivial topology.

Here we consider the example of mirror symmetry, where the corresponding invariant is the mirror Chern number. As a system we look at graphene, possessing a mirror symmetry under $z \to -z$. As the mirror operation is inversion times a π rotation, eigenvalues for the spin 1/2 particles are $\pm i$. Therefore the mirror operator is odd under $\mathcal{T}$. If one defines the Chern numbers $n_{\pm i}$ for the states with mirror eigenvalues $\pm i$, $n_{+i} + n_{-i} = 0$ is thus needed for $\mathcal{T}$ symmetry. The mirror Chern number, defined as $n_{\mathcal{M}} = (n_{+i} - n_{-i})/2$, is then a topological invariant, compare to Eq. (2.24), protected by the crystal symmetry.

In order to extend this to 3D TIs, one can define mirror invariant planes in the 3D BZ. The surfaces perpendicular to them, which retain the mirror symmetry, must have gapless surface modes. Breaking the mirror symmetry opens a gap in the surface spectrum.

An example material possessing a TCI phase is SnTe. Ab-initio calculations showed a band inversion at 4 TRIMs in the 3D BZ [HLL+12, TRS+12]. It is predicted to have a non-trivial mirror Chern number $n_{\mathcal{M}} = -2$, but a trivial $\mathbb{Z}_2$ with $(0;000)$. Due to the mirror Chern number, a double Dirac cone on the $\{001\}$ surface was predicted and later experimentally verified.

2.5. Weyl Semimetals

The last topological material that we introduce in this thesis is the Weyl semimetal (WSM). It comprises a gapless bulk band structure which discriminates it from the insulating systems considered so far. The gaplessness of the bulk as well as the surface states are protected by topology, as we will see in the following. This section is based on Ref. [FM13] Chap. 1, Sec. 6.2 & Chap. 11, Ref. [HQ13], Ref. [BH13] Sec. 11.1.2, as well as Refs. [YF17, JXH16]. We will use WSMs later in Chap. IV in conjunction with 3D TIs.

The defining property for a WSM is that the conduction and valence bands touch at so called Weyl points in the BZ, which are topologically protected. Locally one can expand around these points and obtain the Weyl Hamiltonian

$$H_{Weyl} = \sum_{i,j\in\{x,y,z\}} \hbar v_{ij}\sigma_i k_j, \tag{2.52}$$

where one defines the chirality of the node as

$$\chi = \mathrm{sgn}\left(\det\left(v_{ij}\right)\right). \tag{2.53}$$

The Chern number is finite and given by $n = \pm\chi$ for an electron (hole) Fermi surface. This corresponds to the Berry phase of $\pm 2\pi$ of the single Dirac cone already known from InfoBox 2.1 in Sec. 2.2.2.

The existence of Weyl points does not depend on any symmetry, but is a generic feature in 3 dimensions. The points act as sources and sinks of Berry curvature which makes them topologically stable. They can only be destroyed by the combination with Weyl points of opposite chirality. This can be done by either putting them on top of each other or via inter-cone scattering. Examples are induced superconductivity or the breaking of crystalline translational symmetry by disorder. The robustness of Weyl points can also be directly inferred from the Weyl Eq. (2.52). As all Pauli matrices are used, perturbations acting as mass terms can not open gaps in the spectrum. Therefore protection by $\boldsymbol{\mathcal{T}}$ symmetry is not needed as it is in the 2D case discussed in Sec. 2.4.2.

In a system with $\boldsymbol{\mathcal{T}}$ symmetry, Weyl points are required to come in pairs with momenta $\pm\boldsymbol{K}$, where the Berry curvature is related like $\mathcal{F}(\boldsymbol{K}) = -\mathcal{F}(-\boldsymbol{K})$. Due to this double reversal of signs, the chirality of both cones is the same, i.e. $\boldsymbol{\mathcal{T}}$ symmetry relates Weyl cones of the same chirality to one another. In order to keep the total Chern number of the system fixed as required from Refs. [NN81, NN83], there thus need to be a multiple of 4 Weyl points in a $\boldsymbol{\mathcal{T}}$ symmetric system.

Next we consider $\boldsymbol{\mathcal{I}}$ symmetric systems. The symmetry requires $\mathcal{F}(\boldsymbol{K}) = \mathcal{F}(-\boldsymbol{K})$, thus it relates pairs of Weyl nodes with opposite chirality. The minimum number of Weyl points in a $\boldsymbol{\mathcal{I}}$ symmetric system is thus only 2. At least one of the symmetries, $\boldsymbol{\mathcal{T}}$ or $\boldsymbol{\mathcal{I}}$, has to be broken in order to get a WSM. Otherwise the spectrum would be at least double degenerate and the system is then called a Dirac semimetal.

Similar to TIs, WSMs feature interesting surface physics. The surface states are called Fermi arcs and connect the bulk Weyl points on the surface, see Fig. 2.22. They are guaranteed to exist due to topological reasons. One way to see this is by slicing the BZ with planes which effectively describe Chern insulators, as done in Fig. 2.22 (left). Between the Weyl points, these Chern insulators would have a finite Chern number and thus edge states. The Fermi arcs are exactly these edge states. They are stable against hybridization with the bulk states

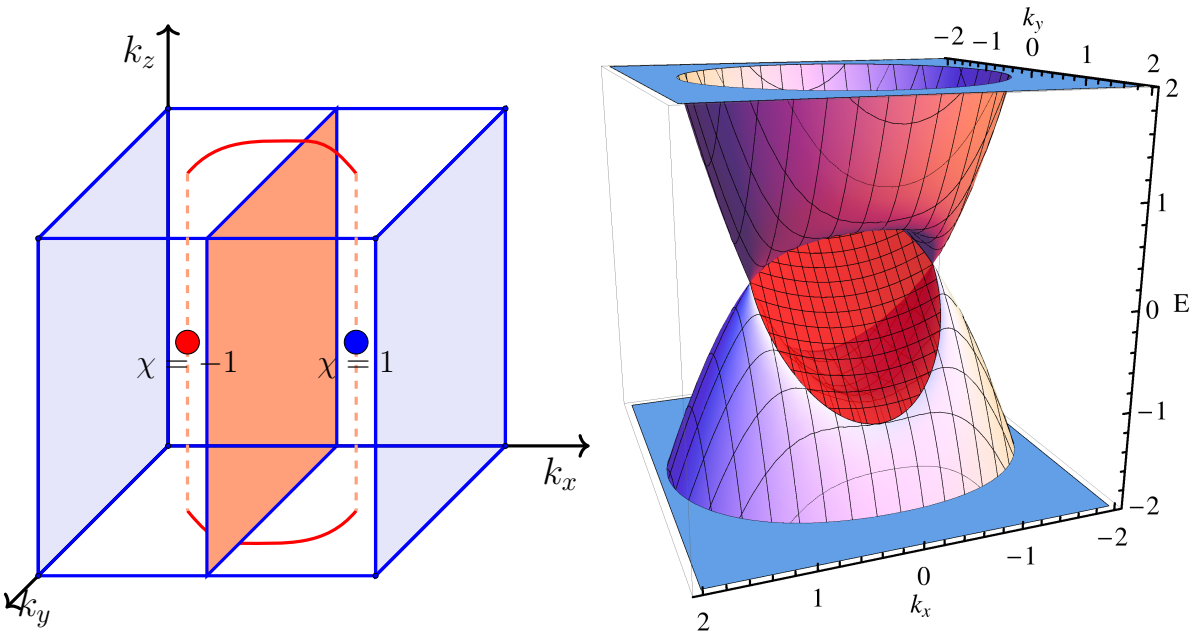


Figure 2.22. (left) Weyl points, depicted as red and blue dots, are connected on the surface via Fermi arcs, indicated as red lines. The red vertical plane is penetrated by a finite Berry curvature, while the blue planes are trivial. (right) Joint dispersion relation of bulk Weyl cones and surface Fermi arc (red surface), see Sec. 9.1.2 for details. Reprinted right figure with permission from Ref. [JT17]. Copyright (2017) by the APS.

due to translational symmetry, as long as surface and bulk states do not exist at the same energy and momentum. This implies that at the Weyl points, see Fig. 2.22 (right), the Fermi arcs are unstable and become delocalized.

Weyl Semimetal as Intermediate Phase between different Topological Insulator Phases in 3D

Before the WSM became popular as an interesting phase on its own, it was studied as an intermediate phase between a NI or weak TI and a strong TI, see Refs. [Mur07, MK08, OM14] and Fig. 2.23. The phase transition can be modeled as driven by a single parameter m, for example changing the system from a weak TI phase via a $\boldsymbol{\mathcal{T}}$ symmetric WSM phase to a strong TI phase. The Fermi arcs always connect two Weyl points of opposite chirality. When these points annihilate and the system becomes a strong TI, the Fermi arcs join into a Dirac cone.

The Chiral Anomaly

One of the most notable features of a WSM is the chiral anomaly. It results in a current originating from a node of chirality χ of the form

$$\nabla \cdot \boldsymbol{j}_\chi = -\chi \frac{e^3}{4\pi^2\hbar^2} \boldsymbol{E} \cdot \boldsymbol{B}. \tag{2.54}$$

The node with opposite chirality has a current with opposite sign. Thus effectively, the chiral anomaly leads to an extra current from one Weyl node to the other one for parallel applied

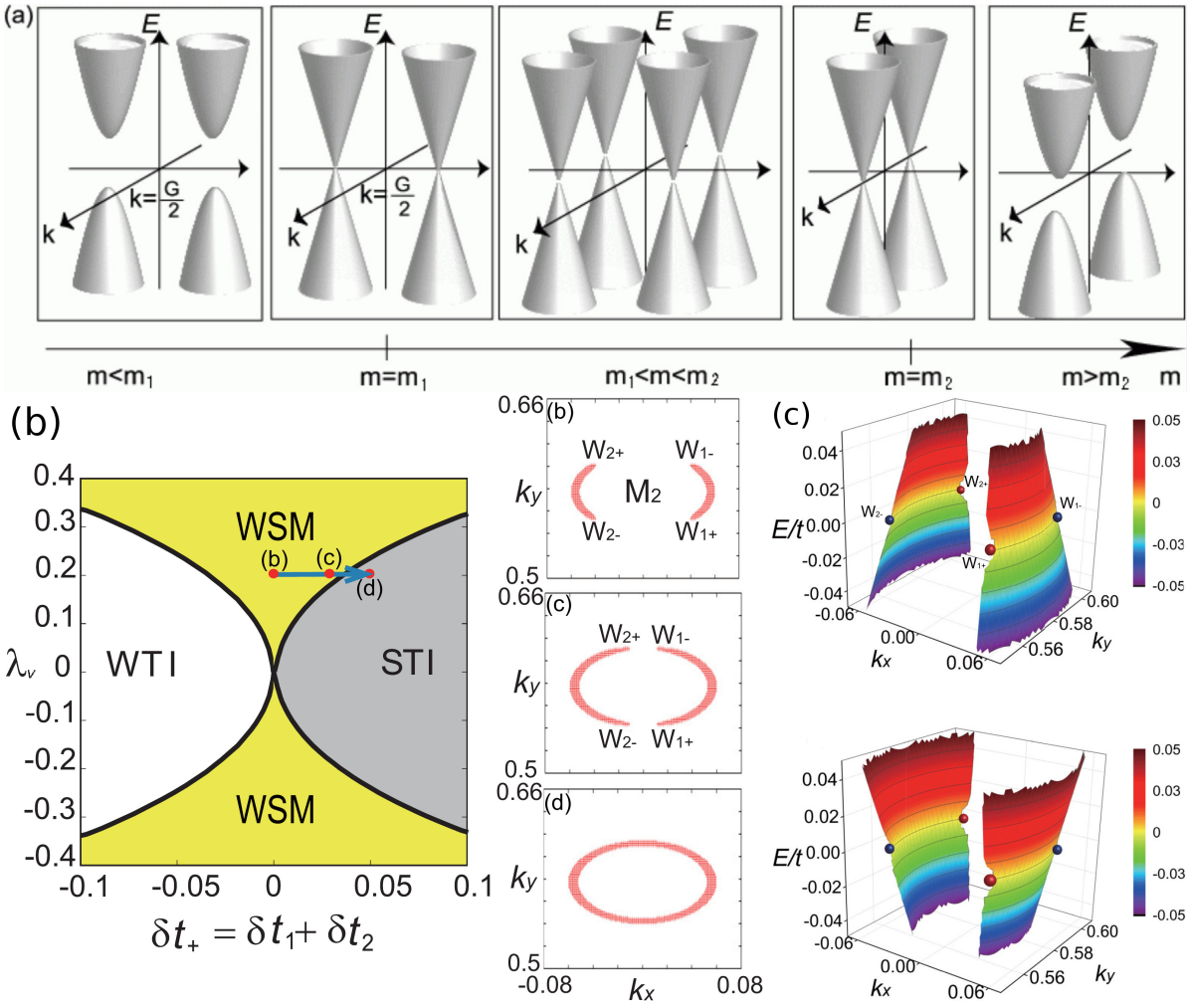


Figure 2.23. (a) Bulk spectrum for different topological phases, where the parameter m drives the transition. Transition from a gapped phase (left) via a $\mathcal{T}$ symmetric WSM phase (middle) to another gapped phase (right). (b) Phase diagram (left) and surface BZ (right) for a 3D topological material. The arrow indicates a phase transition between a WSM and a strong TI phase. The Fermi arcs join together at the phase transition to form the Dirac surface state. (c) 3D plot of the joint dispersion relation of bulk Weyl points and Fermi arcs. Upper and lower surface dispersion relation plotted. Figure (a) adapted from Ref. [Mur07]. Reprinted figures (b) and (c) with permission from Ref. [OM14]. Copyright (2014) by the APS.

magnetic and electric fields. The total current $\boldsymbol{j}_+ + \boldsymbol{j}_-$ is conserved as required from global current conservation.

One way to understand this is the following: in a magnetic field $\boldsymbol{B}$ electrons form LLs of degeneracy $g = \frac{BA_\perp}{h/e}$, with $A_\perp$ the cross section transverse to $\boldsymbol{B}$. The level spectrum of a single Weyl node becomes

$$\begin{aligned} \varepsilon_n &= v_F \mathrm{sgn}\,(n) \sqrt{2\hbar\,|n|\,eB + \left(\hbar \boldsymbol{k}\cdot\hat{\boldsymbol{B}}\right)^2}, \quad n = \pm 1, \pm 2, ..., \\ \varepsilon_0 &= -\chi \hbar v_F \boldsymbol{k}\cdot\hat{\boldsymbol{B}} \end{aligned} \tag{2.55}$$

with $B = |\boldsymbol{B}|$ and $\hat{\boldsymbol{B}} = \boldsymbol{B}/B$. The corresponding dispersion relation of two Weyl points with

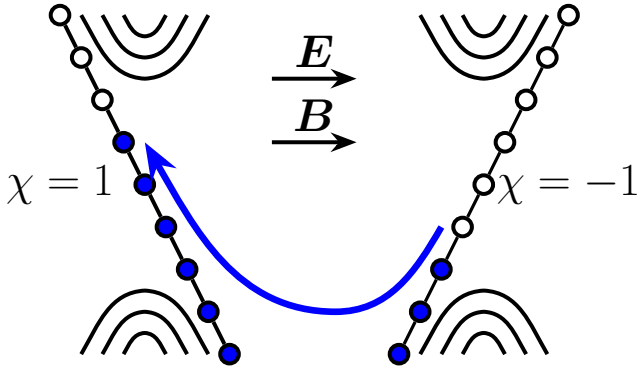


Figure 2.24. Plot of the LL spectrum around two Weyl points of different chirality. Lowest LLs are differently occupied due to applied electric field.

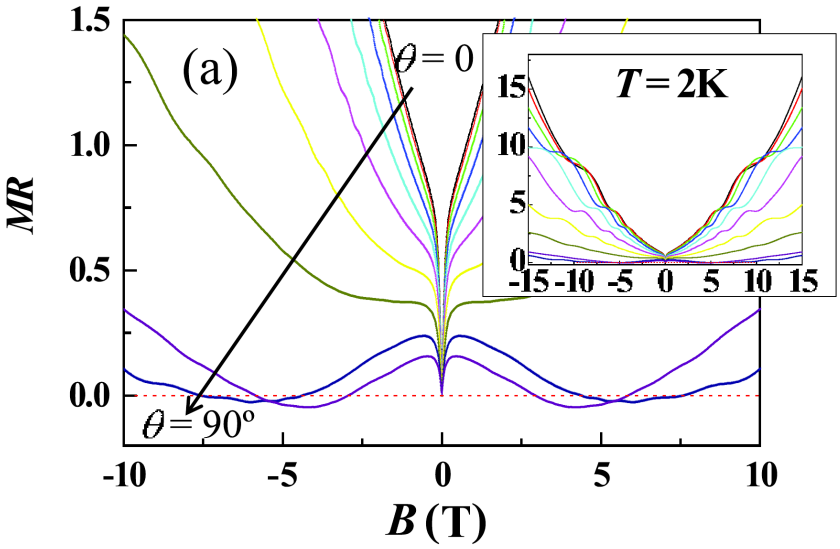


Figure 2.25. Measured magneto resistance (MR) as a function of the applied magnetic field. $\theta = 90°$ corresponds to parallel electric and magnetic field, where the negative MR is expected. For $\theta = 0$ the fields are perpendicular. Reprinted figure with permission from Ref. [KKW+13]. Copyright (2013) by the APS.

opposite chirality is depicted in Fig. 2.24. An applied electric field $\boldsymbol{E}$ shifts the momenta by $\hbar\dot{\boldsymbol{k}} = -e\boldsymbol{E}$. This produces an effective 1D current, if $\boldsymbol{E}$ is parallel to $\boldsymbol{B}$. Due to the chirality one has an effective transport of electrons from one Weyl node to the other. The 1D current is given by $\partial Q_\chi^{1D}/\partial t = e\chi L_B \left|\dot{\boldsymbol{k}}\right| /2\pi$ with L_B the system size in direction of $\boldsymbol{B}$. The full current due to the chiral anomaly is then

$$\frac{\partial Q_\chi^{3D}}{\partial t} = g\frac{\partial Q_\chi^{1D}}{\partial t} = -V\frac{e^3}{4\pi^2\hbar^2}\boldsymbol{E}\cdot\boldsymbol{B} \tag{2.56}$$

with the volume $V = A_\perp L_B$.

In experiments, this extra current should show up as a negative contribution to the magneto resistance, in dependence on the angle of the applied electric and magnetic field. An example measurement featuring such a negative magneto resistance is plotted in Fig. 2.25. The angle dependence is clearly visible and considered a smoking gun experimental signature.

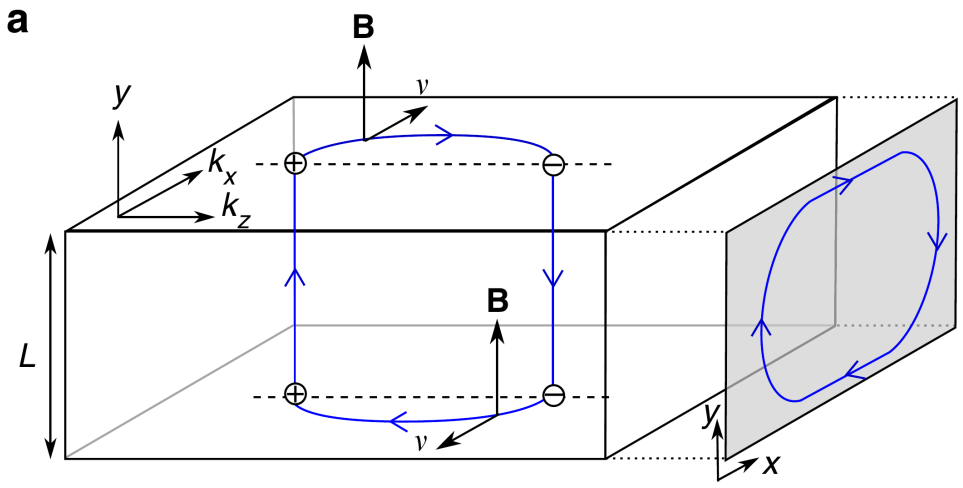


Figure 2.26. Sketch of closed orbital path in a WSM. The electrons on the orbit move along the Fermi arc at the surface, enter the bulk at the projections of the Weyl points denoted by circles with + and -, move through the bulk and reach the other surface of the system, moving again along the Fermi arc. Adapted by permission from Macmillan Publishers Ltd: Nature Communications, Ref. [PKV14], copyright (2014).

Weyl Orbits

In a WSM, both surface and bulk are gapless. Using this fact, one can define closed paths on which electrons move along the surface and through the bulk, as sketched in Fig. 2.26. These Weyl orbits lead to resonances in the density of states (DOS), which can be measured. For the original proposal and the experimental verification see Refs. [PKV14, MNH+16].

Minimal Model

The derivation and study of minimal models of the WSM phase is an active area of research, see, e.g., Refs. [OM14, DR16]. We close this section by giving a simple model for an $\mathcal{I}$ symmetric WSM with two Weyl points. It will be used in Chap. IV as a starting point for the study of combined TI and WSM surface states. It was first proposed in Ref. [YLR11], extended in Ref. [MKT17] and is given by

$$H_W = v_z k_z \tau_3 + v_y k_y \tau_2 + t(k)\tau_1 + \gamma_W t_x \left(k_x^2 - k_W^2\right) \tau_0 \tag{2.57}$$

with $t(k) = t_x \left(k_x^2 - k_W^2\right) + t_{yz} \left(k_y^2 + k_z^2\right)$ and the Pauli matrices $\boldsymbol{\tau}$. The two Weyl points are located at $k_x = \pm k_W$, and the type of the WSM can be switched from type I to type II [XZZ15, SGW+15, SGT16] by a finite $|\gamma_W| > 1$. For further details, see Chap. IV.

3. Correlations and Topological Matter

Taking into account Coulomb interactions beyond the mean-field level in topological matter, this gives rise to a fundamental problem: in most of Sec. 2 we used band theory to discuss and classify different topological phases and their properties. Yet band theory does not take into account correlation effects and is thus not valid in a correlated system. The invariants based on band theory loose their validity as well. Therefore we will introduce in Sec. 3.1 a complementary theory that allows for topological classification of systems without restrictions on the allowed correlations. Before we go into detail, let us consider what changes we could expect due to interactions, based on Ref. [FM13] Chap. 11, Sec. 2. The classification of phases could change by merging of prior distinct classes, or by refining an existing class into several new ones. One can think about it like adding a new axis in a phase diagram. Using this additional dimension, interactions could make it possible to define a path connecting two topological phases without going through a discontinuity, thus merging the topological phases. A classical example thereof are the liquid and vapor phases of water, which merge at high pressure.

Besides the topological classification of ground states of interacting systems, we are also interested in the study of their dynamics. For this we need effective theories describing the correlated electron gas. This is done in Sec. 3.2, where we introduce Fermi liquid theory, as well as perturbative theories for the weak and strong coupling limit, the Random Phase Approximation (RPA) and a formalism for quantum Hall ferromagnetic (QHFM) systems. The section ends with an overview of Luttinger liquid theory characterizing correlated 1D systems.

The last missing piece is then the connection between the introductory Chapter I and the main research part of this thesis in Chapters II-IV. This link is established in Sec. 3.3 where the underlying motivation for our research is presented. We consider the possible benefits of joining topological and interaction effects. This includes the use of interactions as a probe for topological phases, the idea of topological plasmonics as well as the study of coupled topological multilayer systems. The symmetry breaking due to interactions and the corresponding possible destruction or refinement of topological phases is considered as well. This overview should serve as motivation for the following research chapters.

3.1. Topological Field Theory

In Sec. 2 we classify topological matter with the help of a topological band theory. Is this classification still well-defined in system with relevant interactions? In order to answer this, we introduce a complementary theory that is not restricted to non-interacting systems. The central idea of this topological field theory (TFT) is the following: one classifies a system purely by its topological properties. These properties are for example the way the system responses to external electromagnetic fields. Thus TFT could also be called a topological response theory. All that is needed for the classification are the relevant symmetries of the system. This section is based on Ref. [FM13] Chap. 4, Secs. 1 & 3.2 and Chap. 2, Secs. 2.3f.

Throughout this section we stick to one example system. It is given by a 2D insulator with Hamiltonian $H\left[c_{i\alpha}, c_{i\alpha}^{\dagger}\right]$, with i for lattice sites and α for labeling spin and band indices. The details of the Hamiltonian are not important. We study its response to an external electromagnetic field via a finite A_μ, with $\mu = x, y, t$ for space and time indices. The coupling of system and field is determined by the gauge principle, e.g. $-tc_i^{\dagger}c_{i+\hat{x}}$ is replaced by $-tc_i^{\dagger}c_{i+\hat{x}}\mathrm{e}^{iA_{ix}}$ with $A_{ix} = \frac{e}{\hbar}\int_i^{i+\hat{x}} A_\mu \mathrm{d}r^\mu$. The Hamiltonian of the coupled system is then denoted as $H\left[c_{i\alpha}, c_{i\alpha}^{\dagger}, A_\mu\right]$. The action is, using imaginary time, given by

$$S\left[c_{i\alpha}\left(\tau\right), \bar{c}_{i\alpha}\left(\tau\right), A_\mu\left(\tau\right)\right] = \int_0^\beta \left[\sum_{i,\alpha} \bar{c}_{i\alpha}\partial_\tau c_{i\alpha} + H\right] \mathrm{d}\tau. \tag{3.1}$$

Here and in the following, we suppress arguments of functions and functionals for brevity. In order to study only the response of the system, we integrate out the internal fermionic degrees of freedom. This defines the effective action $S_{eff}\left[A_\mu\right]$ with

$$\mathrm{e}^{-S_{eff}[A_\mu]} = \int Dc_{i\alpha}\left(\tau\right) D\bar{c}_{i\alpha}\left(\tau\right) \mathrm{e}^{-S[c_{i\alpha}(\tau), \bar{c}_{i\alpha}(\tau), A_\mu(\tau)]}. \tag{3.2}$$

The action $S_{eff}\left[A_\mu\right]$ includes the response of the system to the external electromagnetic field. Observables can be calculated from it, for example, the average value of the charge current is given by

$$\langle J_\mu\left[A_\mu\right]\rangle = \langle\frac{\delta H}{\delta A_\mu}\rangle = \frac{\delta S_{eff}}{\delta A_\mu}. \tag{3.3}$$

Our goal is the derivation of a general theory not dependent on the specifics of the Hamiltonian. A way to achieve this is the expansion of the effective action in orders of A_μ. This results in

$$S_{eff} = \frac{1}{2}\int \mathrm{d}^3x\, \mathrm{d}^3x'\, \Pi_{\mu\nu}\left(x, x'\right) A_\mu\left(x\right) A_\nu\left(x'\right) + \mathcal{O}\left(A^2\right), \tag{3.4}$$

where $\Pi_{\mu\nu}$ are the linear response functions of the fermion system. Importantly, the action has to be consistent with gauge invariance

$$\mathrm{e}^{-S_{eff}[A_\mu]} = \mathrm{e}^{-S_{eff}[A_\mu + \partial_\mu\phi]} \tag{3.5}$$

for any gauge transformation $\phi\left(x\right)$ to every order in the response to A_μ.

We are now in a position to discuss general properties. For a gapped system, all connected correlation functions are short ranged in space-time. This makes the integration in Eq. (3.2) well defined. S_{eff} should be local in $A_\mu\left(x\right)$, thus $\Pi_{\mu\nu}\left(x, x'\right)$ needs to be short ranged in $|x - x'|$. The typical length scale is $l = \hbar v/E_g$ with the gap E_g and the average electron velocity v.

In the case that one considers the response of the system on a much larger length scale than l, one can perform a gradient expansion in momentum space

$$\begin{aligned}
\Pi_{\mu\nu}\left(x, x'\right) &= \int \frac{\mathrm{d}^3p}{\left(2\pi\right)^3}\mathrm{e}^{ip(x-x')}\Pi_{\mu\nu}\left(p\right) \\
&= \sum_{n=0}^{\infty} \Pi^{(n)}_{\mu\nu;\tau_1\tau_2\ldots\tau_n} \int \frac{\mathrm{d}^3p}{\left(2\pi\right)^3}\mathrm{e}^{ip(x-x')} p^{\tau_1}p^{\tau_2}\ldots p^{\tau_n} \\
&= \sum_{n=0}^{\infty} \Pi^{(n)}_{\mu\nu;\tau_1\tau_2\ldots\tau_n} \left(-i\right)^n \partial^{\tau_1}\partial^{\tau_2}\ldots\partial^{\tau_n}\delta\left(x - x'\right)
\end{aligned} \tag{3.6}$$

where we assumed translational symmetry. The expansion coefficients are here defined by $\Pi^{(n)}_{\mu\nu;\tau_1\tau_2\ldots\tau_n} = \frac{\partial^n}{\partial p^{\tau_1}\partial p^{\tau_2}\ldots\partial p^{\tau_n}} \Pi_{\mu\nu}|_{p=0}$. Each term has to respect gauge symmetry, which gives e.g. $\Pi^{(0)}_{\mu\nu} = 0$. The Maxwell term in 2nd order is usually the lowest contribution that respects gauge symmetry, translational symmetry and rotation symmetry. It is given by

$$\begin{aligned} \Pi^{(2)}_{\mu\nu;\tau_1\tau_2} &= \frac{1}{2}\left(g_{\mu\nu}g_{\tau_1\tau_2} - g_{\mu\tau_1}g_{\nu\tau_2}\right), \\ -\frac{1}{4}\int \mathrm{d}^3x\, F_{\mu\nu}F^{\mu\nu} &= \int \mathrm{d}^3x\, \mathrm{d}^3x'\, \Pi^{(2)}_{\mu\nu;\tau_1\tau_2}\partial^{\tau_1}\partial^{\tau_2}\delta\left(x-x'\right) A^{\mu}\left(x\right) A^{\nu}\left(x'\right). \end{aligned} \tag{3.7}$$

with the Lorentz metric $g_{\mu\nu}$ and the electromagnetic tensor $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$.

Chern-Simons Term

In 2+1 dimensions there is a special gauge invariant term of lower order than the Maxwell term. It is called Chern-Simons (CS) term and given by

$$\begin{aligned} \Pi^{(1)}_{\mu\nu;\tau} &= -\frac{\sigma_H}{2}\varepsilon_{\mu\nu\tau}, \\ S_{CS} &= -i\int \mathrm{d}^3x\, \mathrm{d}^3x'\, \Pi^{(1)}_{\mu\nu;\tau}\partial^{\tau}\delta\left(x-x'\right) A^{\mu}\left(x\right) A^{\nu}\left(x'\right) \\ &= i\frac{\sigma_H}{2}\int \mathrm{d}^3x\, \varepsilon_{\mu\nu\tau}A^{\mu}\partial^{\nu}A^{\tau}. \end{aligned} \tag{3.8}$$

It leads to the current, Eq. (3.3), of $J^\mu = \sigma_H \varepsilon^{\mu\nu\tau}\partial_\nu A_\tau$. The first component is thus $J_x = \sigma_H E_y$ and we can identity the CS term as being responsible for the Hall conductivity.

Let us check the gauge invariance of the CS term $\varepsilon_{\mu\nu\tau}A^\mu\partial^\nu A^\tau$. The gauge transformation $A^\mu \to A^\mu + \delta A^\mu$ leads to the variation $\delta\left(\varepsilon_{\mu\nu\tau}A^\mu\partial^\nu A^\tau\right) = 2\left(\delta A^\mu\right)\varepsilon_{\mu\nu\tau}\partial^\nu A^\tau - \partial_\mu\left(\varepsilon_{\mu\nu\tau}A^\nu\delta A^\tau\right)$, where the second term is a total derivative. If one takes the gauge transformation $\delta A^\mu = \partial^\mu\phi$ and ϕ is single-valued, one finds $\delta S_{CS} = 0$ and the gauge invariance is preserved.

We know from Sec. 2.3 that it is not always possible to define a single-valued ϕ across the BZ. Having a closer look at the gauge transformation, one actually only needs a single-valued $\mathrm{e}^{i\phi(x)}$ to fulfill gauge invariance. In 2+1 dimensions we assume a time-periodic system with $\phi\left(x,y,t+T\right) = \phi\left(x,y,t\right) + 2\pi n$. The change of the action due to the gauge transformation is then given by

$$\begin{aligned} \delta S_{CS} &= i\sigma_H \int \mathrm{d}^2x \int_0^T \mathrm{d}t\, \partial_\mu\left(\varepsilon^{\mu\nu\tau}\phi\partial_\nu A_\tau\right) \\ &= i\sigma_H \int_{S^2} \mathrm{d}^2x\, F_{xy}\left(x,y,0\right)\left[\phi\left(x,y,T\right) - \phi\left(x,y,0\right)\right] \\ &= i\sigma_H 4\pi^2 nm \end{aligned} \tag{3.9}$$

where we used that the total flux of F_{xy} on the spatial manifold is always quantized to $2\pi m$, $m \in \mathbb{Z}$.

The CS term is therefore not fully gauge invariant. Yet if one takes $\sigma_H = \frac{k}{2\pi}$ with k an integer, corresponding to $\sigma_H = k\frac{e^2}{h}$ in physical units, it is quantized to a multiple of 2π. That makes $\mathrm{e}^{-S_{CS}}$ gauge invariant and is thus a proof that the Hall conductivity σ_H needs to be quantized. This quantization makes the CS term robust as it can not be continuously tuned, but only changed abruptly if the description used above fails. The latter case corresponds then to a quantum phase transition.

Let's sum up the basic idea of topological response theory considered so far. One introduces an external field and writes down all topological terms of the external field that are consistent with gauge invariance and the symmetries of the system. Topological terms are defined by their quantization and serve as a way to topologically classify the state of the system. Two states with different discrete terms are adiabatically distinct and belong to different topological classes. Two states with the same quantized terms can be of the same topology, but this is not necessarily the case. The states could also differ in some other topological response terms not considered in the classification.

We close this section by considering some generalizations to the theory presented so far. The first one is a system that has more than one ground state. In this case, after integrating out the fermionic degrees of freedom, there will be one additional, low-energy degree of freedom left, describing the ground state degeneracy. Thus the action has more parameters than A_μ alone, and the theory is called dynamical topological field theory.

The second generalization is about a higher spatial dimension. The same principles presented above can be applied to a $\mathcal{T}$ invariant insulator in 3+1 dimensions. TFT can be used to define the $\mathbb{Z}_2$ invariant and study interesting topological effects like the magneto-electric effect or axion electrodynamics. We do not go into detail here.

Interaction Effect on Topological Classification

Finally we are in a position to answer the question about the changes in topological classification raised at the beginning of this section.

For finite correlation strength, the single-particle states are not well defined. In the response theory, additional terms appear in the expansion, due to the interaction, and may modify the response. But there is a way out: one can rewrite the theory in terms of single-particle Green functions which stay valid in the interacting case. One finds that the coefficients of the topological terms do not change for finite correlations, as long as no phase transition occurs. Thus we conclude: any interacting topological state that can be adiabatically connected to a non-interacting system has the same topology. Interactions need to be strong enough to close the band gap in order to drive a topological transition. This is a helpful conclusion for our studies in Chap. II, where we incorporate the effects of interactions perturbatively via RPA.

This TFT analysis is unfortunately not applicable to gapless systems like WSMs. In order to decide whether these systems are stable under the influence of interactions, one has to do an analysis on a case to case basis. For WSMs this has been done, with the result that electron-electron interactions are marginally irrelevant. Thus Weyl points are expected to be stable against weak disorder and the correlations just renormalize the system parameters, see e.q. Refs. [GC11, IN12, HPV12, IN13, YN14]. For the gapless BHZ model, we perform this analysis in Chap. II.

3.2. Methods for Incorporating Interactions

So far, we have considered the topological classification of matter, mostly on a non-interacting basis in Sec. 2 and in more general terms in Sec. 3.1. In the following section, we will introduce further concepts related to electron-electron interaction. We discuss methods to consider correlation effects in a controlled fashion. This serves as the second cornerstone in order to study correlations in topological matter in Chaps. II-IV.

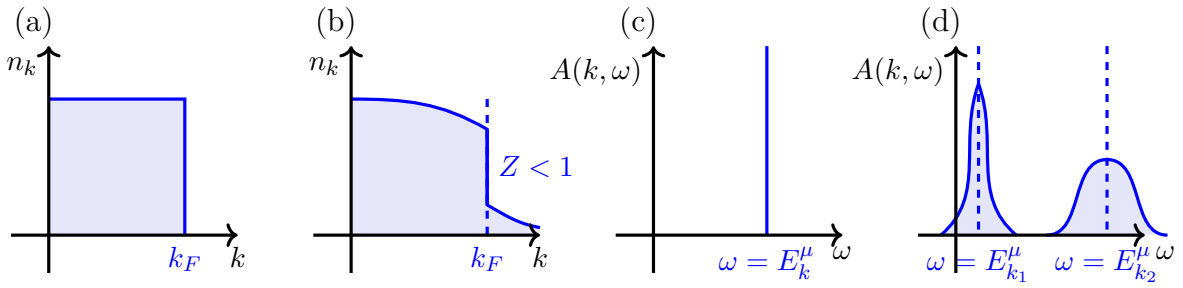


Figure 3.1. Occupation number n_k in the non-interacing case (a) and for finite interactions (b). The jump at k_F reduces due to interactions. (c) The spectral function $A(k,\omega)$ of an excitation is a δ peak in the non-interacting limt. Including interactions in (d), the peak broadens in dependence to its distance to the Fermi momentum.

This section is organized as follows. Fermi liquid theory is introduced in Sec. 3.2.1, offering a general way to consider interactions between electrons in 2 and 3 dimensions by modeling them as dressed quasiparticles. The dynamics of these quasiparticles is then further studied perturbatively in the weakly interacting limit in Sec. 3.2.2 using the Random Phase Approximation (RPA). The strongly interacting limit is treated in Sec. 3.2.3, where the system is assumed to be in a quantum Hall ferromagnetic (QHFM) state. Interactions in one dimension are special in the sense that they are always dominant. Thus we treat them separately in Sec. 3.2.4 in a Luttinger liquid description.

3.2.1. Fermi Liquid Theory

Coulomb interaction between the individual electrons in solids is often neither dominant nor negligible compared to their kinetic energy. This makes correlations inherently difficult to treat, as perturbation theory is not always justified. Fermi liquid theory can be understood as one of the fundamental theories tackling this problem in solid state physics, upon many of the more refined techniques build on. We use Ref. [Gia03] Sec. 1.1 as the source for this section.

Let us begin with a crystal system in D>1 dimensions, neglecting interactions as a start. The non-interacting electron bands are filled up to the Fermi level μ as discussed in Sec. 2.2.1. There at μ, or correspondingly at the Fermi wave vector k_F (assuming a Fermi circle or sphere), the occupation number n_k jumps from 1 to 0, see Fig. 3.1 (a). Excitations can be considered by looking at the effect of adding additional electrons. This can be done above the Fermi level, and the excitations have well defined momentum k and energy E_k. They are eigenstates of the Hamiltonian and thus have an infinite lifetime. The spectral function $A(k,\omega)$ gives the probability to find a state with frequency ω and momentum k. For these free electrons that we consider, it is given by

$$A(k,\omega) = \delta\left(\omega - E_k^\mu\right) \tag{3.10}$$

with $E_k^\mu = E_k - \mu$, E_k the energy of the particle and μ the chemical potential or Fermi level, used synonymously in the following. The δ peak is plotted in Fig. 3.1 (c).

As a next step we include interactions. The main idea and result of Fermi liquid theory is that the interacting case is essentially similar to the free fermion picture. The elementary particles are now electrons dressed by density fluctuations, forming quasiparticles. These quasiparticles are basically free spin 1/2 particles, with a definite momentum k and energy E_k. We elaborate on this in the following.

The occupation number n_k in the interacting case is still discontinuous at the Fermi level, but the jump is given by $Z < 1$, see Fig. 3.1 (b). Z gives the fraction of particles remaining in the quasiparticle state, while the rest is in a continuous background without well-defined structure. The quasiparticles have a well-defined relation between frequency and momentum given by $\omega = E_k^\mu$. The interaction renormalizes parameters of the system like the electron mass m, which becomes the quasiparticle mass m^*. The Fermi momentum on the other hand is unchanged due to the conservation law known as Luttinger theorem.

The low-energy physics of the excitations in this interacting system can be described by the functional

$$E\left[\{\delta n_{\boldsymbol{k}}\}\right] = \sum_{\boldsymbol{k},s} E^\mu_{\boldsymbol{k}} \delta n_{\boldsymbol{k},s} + \frac{1}{2V} \sum_{\boldsymbol{k},s,\boldsymbol{k}',s'} f^{ss'}_{\boldsymbol{k},\boldsymbol{k}'} \delta n_{\boldsymbol{k},s} \delta n_{\boldsymbol{k}',s'} \tag{3.11}$$

where $\delta n_{\boldsymbol{k},s}$ measures the deviations of the occupation number from the ground state. The Landau function $f^{ss'}_{\boldsymbol{k},\boldsymbol{k}'}$ contains the interaction or scattering effects between quasiparticles. These quasiparticles are thus not completely free. Their excitations have a finite lifetime τ, resulting in the time dependence

$$\mathrm{e}^{-iE^\mu_k t} \mathrm{e}^{-t/\tau}. \tag{3.12}$$

This finite lifetime translates into a finite width in the spectral function $A\left(k,\omega\right)$ as depicted in Fig. 3.1 (d). It is due to scattering between quasiparticles. As the scattering phase space goes to zero when approaching the Fermi surface, the lifetime diverges at the Fermi level. In 3D this happens as $\tau \propto 1/\left(E^\mu_k\right)^2$, making the excitations well-defined close to the Fermi surface.

In normal metals, the Fermi level is about 10000 K, so for practical purposes, with temperatures of $1 - 100$ K, one is often close to the Fermi energy. The quasiparticles behave as being essentially free excitations in this limit.

The validity of the theory summarized above can be established by perturbation theory. Yet Fermi liquid theory is valid beyond this regime and can also be based on simple and general phase space arguments. Therefore, this theory is expected to break down only due to exceedingly strong interactions or special instabilities occurring in the system. The latter happens e.g. in one dimension, see Sec. 3.2.4.

What we missed so far, these are the additional collective excitations that interactions bring along. They describe the response of the system to disturbances in the particle or spin density. Short range interactions, for example, lead to zero sound modes in the charge sector, while plasmon modes arise due to long range Coulomb interactions. We will study them in more detail in the next section.

3.2.2. Excitations and Screening in the Random Phase Approximation

We have argued in the last section 3.2.1 that in 2D and 3D the electrons in an interacting condensed matter system can be described by quasiparticles, which are electrons dressed by density fluctuations. In this section we are interested in the response of such an electron gas to perturbations, be it external potentials or the intrinsic electron-electron interaction between quasiparticles. We present a perturbation theory incorporating these effects, with a focus on the widely used Random Phase Approximation (RPA). This section is based on Ref. [GV05] Secs. 3.3, 4, 5 & 6 as well as Ref. [Mah00] Sec. 5.5 and Ref. [FW71] Chaps. 5, 8 & 9. The presented theory will be used in Chap. II.

(Proper) Density-density Response Function

Considering linear response theory, the coupling of a time-dependent external potential V_{ext} to the electron gas can be described by the term

$$\int V_{ext}(\boldsymbol{r},t)\, n(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r}. \tag{3.13}$$

Here $n(\boldsymbol{r}) = \sum_i \delta(\boldsymbol{r}-\boldsymbol{r}_i)$ is the density operator of the electrons. The electron gas reacts to this perturbation by screening it. This behavior induces a potential V_{ind} and the total, screened potential is then given by

$$V_{sc}(\boldsymbol{r},t) = V_{ext}(\boldsymbol{r},t) + V_{ind}(\boldsymbol{r},t), \quad V_{ind}(\boldsymbol{r},t) = \frac{e^2}{4\pi\varepsilon_0}\int \mathrm{d}\boldsymbol{r}'\, \frac{n_1(\boldsymbol{r}',t)}{|\boldsymbol{r}-\boldsymbol{r}'|}. \tag{3.14}$$

The induced density n_1 in linear response is defined by

$$n_1(\boldsymbol{r},\omega) = \int_{-\infty}^{\infty} \mathrm{d}t\, n_1(\boldsymbol{r},t)\, \mathrm{e}^{i\omega t} \tag{3.15}$$

$$= \int \mathrm{d}\boldsymbol{r}'\, \chi_{nn}(\boldsymbol{r},\boldsymbol{r}',\omega)\, V_{ext}(\boldsymbol{r}',\omega) \tag{3.16}$$

$$= \int \mathrm{d}\boldsymbol{r}'\, \tilde{\chi}_{nn}(\boldsymbol{r},\boldsymbol{r}',\omega)\, V_{sc}(\boldsymbol{r}',\omega). \tag{3.17}$$

Here, we introduced two different response functions. The first one is the density-density response function

$$\chi_{nn}(\boldsymbol{r},\boldsymbol{r}',\omega) = -\frac{i}{\hbar}\int_{-\infty}^{\infty} \mathrm{d}t\, \mathrm{e}^{(i\omega-0^+)t}\Theta(t)\left\langle [n(\boldsymbol{r},t), n(\boldsymbol{r}')]\right\rangle_0, \tag{3.18}$$

with the Heavyside step function Θ, and $\langle\ldots\rangle_0$ denotes the averaging over the thermal equilibrium ensemble. χ_{nn} gives the full response of the electron gas to the perturbation and is thus the quantity we are interested in. Full response means not only the response to the perturbation, but also the response to the response to the perturbation to all orders. Thus it includes all self-screening effects, which makes this quantity quiet involved and a bit cumbersome to compute.

The second response function $\tilde{\chi}_{nn}$, defined in Eq. (3.17), is the so called proper density-density response function. It describes the response of the electron gas as an answer to the already screened potential. One thus could expect that it is a simpler quantity, as the self-screening effects are already included in V_{sc}.

Comparing Eqs. (3.14)-(3.17), this provides the relation between the two response functions as

$$\tilde{\chi}_{nn}^{-1}(\boldsymbol{r},\boldsymbol{r}',\omega) = \chi_{nn}^{-1}(\boldsymbol{r},\boldsymbol{r}',\omega) + \frac{e^2}{4\pi\varepsilon_0}\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}. \tag{3.19}$$

For the homogeneous electron liquid, the physics only depends on the relative distance $|\boldsymbol{r}-\boldsymbol{r}'|$. In this case, a Fourier transform with respect to $\boldsymbol{r}-\boldsymbol{r}'$ can simplify the equations. We find

$$\chi_{nn}(\boldsymbol{q},\omega) = \frac{\tilde{\chi}_{nn}(\boldsymbol{q},\omega)}{1 - v_q\tilde{\chi}_{nn}(\boldsymbol{q},\omega)}, \tag{3.20}$$

$$\varepsilon(\boldsymbol{q},\omega) = 1 - v_q\tilde{\chi}_{nn}(\boldsymbol{q},\omega), \tag{3.21}$$

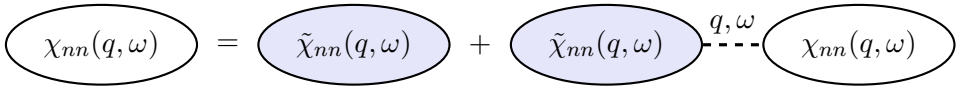


Figure 3.2. Summation over all orders of the proper response function $\tilde{\chi}_{nn}$ gives the full response χ_{nn}.

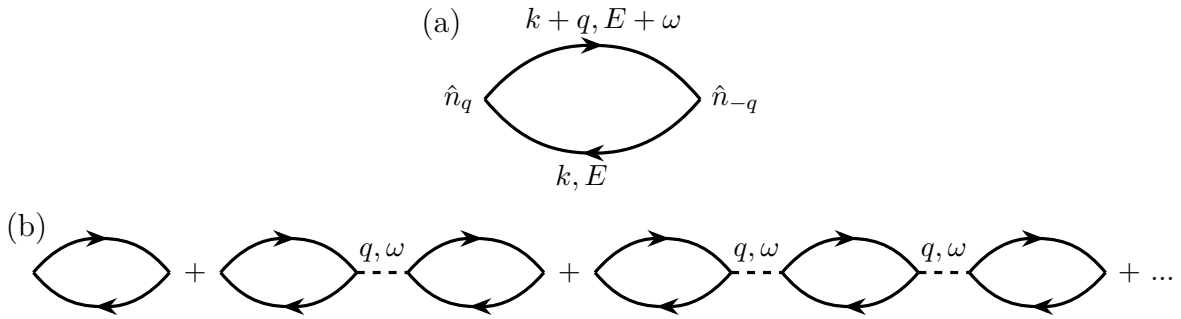


Figure 3.3. (a) Bubble diagram corresponding to the Lindhard response function (3.23). (b) The infinite sum of bubble diagrams gives the polarization function in RPA, Eq. (3.24).

where v_q is the Fourier transform of the Coulomb interaction. The dielectric screening function $\varepsilon(\boldsymbol{q},\omega)$ relates the screened and unscreened potential like

$$V_{sc}(\boldsymbol{q},\omega) = \frac{V_{ext}(\boldsymbol{q},\omega)}{\varepsilon(\boldsymbol{q},\omega)}. \tag{3.22}$$

With the help of the geometric series $\sum_{k=0}^{\infty} q^k = \frac{1}{1-q}$ for $|q| < 1$, one can understand from Eq. (3.20) that χ_{nn} is just the summation of all orders of $v_q\tilde{\chi}_{nn}$ as depicted in Fig. 3.2. This formalizes our idea that the full response should include not only the response to the perturbation, but also all orders of self-screening.

Lindhard Response Function and Random Phase Approximation

So far the definitions of the density-density response functions are very general. In order to calculate them we rely on approximations. A simple yet powerful one is the Random Phase Approximation (RPA). It is simply given by approximating the proper density-density response function $\tilde{\chi}_{nn}(\boldsymbol{q},\omega)$ by the so called Lindhard function

$$\chi^0_{nn}(\boldsymbol{q},\omega) = \chi_0(\boldsymbol{q},\omega) = \sum_{\sigma,\lambda,\lambda'} \int \frac{\mathrm{d}\boldsymbol{k}}{(2\pi)^d} \mathcal{F}^{\lambda,\lambda'}_{\boldsymbol{k},\boldsymbol{k}+\boldsymbol{q}} \frac{f(E_{\boldsymbol{k}\sigma}) - f(E_{\boldsymbol{k}+\boldsymbol{q}\sigma})}{\hbar\omega + E_{\boldsymbol{k}\sigma} - E_{\boldsymbol{k}+\boldsymbol{q}\sigma} + i0^+}. \tag{3.23}$$

The Fermi-Dirac function $f(E) = \frac{1}{\mathrm{e}^{\beta(E-\mu)}+1}$, with the inverse temperature $\beta = \frac{1}{k_B T}$ and k_B the Boltzmann constant, gives the occupation of the state at energy E. The overlap factor $\mathcal{F}^{\lambda,\lambda'}_{\boldsymbol{k},\boldsymbol{k}+\boldsymbol{q}}$ takes into account that scattering between different bands λ and λ' can be suppressed. $\sigma = \pm$ considers the spin of the electrons.

The full density-density response function χ_{nn} can be calculated in diagrammatic perturbation theory based on Green functions. One finds that the lowest order contributing diagram is the bubble diagram in Fig. 3.3 (a). It exactly corresponds to the Lindhard function (3.23). The density-density response function in RPA, Π^{RPA}, is thus given by the infinite sum of

bubble diagrams as shown in Fig. 3.3 (b), and the formula

$$\Pi^{RPA}(q,\omega) = \frac{\chi_0(q,\omega)}{1 - v_q\chi_0(q,\omega)}, \tag{3.24}$$

$$v_q^{RPA} = \frac{v_q}{1 - v_q\chi_0(q,\omega)}, \tag{3.25}$$

$$\varepsilon^{RPA}(q,\omega) = 1 - v_q\chi_0(q,\omega). \tag{3.26}$$

We close this subsection be giving some useful properties and formulas that are needed in later sections. For χ_0 one finds

$$\chi_0(\boldsymbol{q},0) \leq 0, \quad \chi_0(-\boldsymbol{q},\omega) = \chi_0(\boldsymbol{q},\omega), \quad -\Im[\chi_0(\boldsymbol{q},-\omega)] = \Im[\chi_0(\boldsymbol{q},\omega)], \tag{3.27}$$

and from $\chi_0(\boldsymbol{q},0) < 0$ we directly conclude $1/\varepsilon(\boldsymbol{q},0) < 1$. The real and imaginary parts are connected via the Kramers-Kronig relation

$$\Re[\chi_0(\omega)] = \frac{2}{\pi}\mathcal{P}\int_0^\infty \frac{\nu\Im[\chi_0(\nu)]}{\nu^2 - \omega^2}\mathrm{d}\nu \tag{3.28}$$

where P denotes the principal value part. The response function fulfills a set of sum rules, of which we only consider one. It is called the f-sum rule which we will denote as $f\sum$ in the following. It states that for all valid approximations to χ_{nn} the integral

$$f\sum(\boldsymbol{q}) = -\frac{2}{\pi}\int_0^\infty \omega\Im[\chi_{nn}(\boldsymbol{q},\omega)]\,\mathrm{d}\omega \tag{3.29}$$

is always the same. The function is different for each band structure considered, e.g. for a 2DEG it is $f\sum(\boldsymbol{q}) = \frac{nq^2}{m}$ with n the density and m the mass.

The polarization function can be connected to a quantity directly measurable in experiments, the static structure factor $A(\boldsymbol{q}) = \int_0^\infty \frac{\mathrm{d}\omega}{\pi} A(\boldsymbol{q},\omega)$. $A(\boldsymbol{q},\omega)$ is the spectral function of the density operator, where peaks in $A(\boldsymbol{q},\omega)$ correspond to excitations in the density, see also Sec. 3.2.1. It is connected to χ_0 and v_q via

$$A(\boldsymbol{q},\omega) = \frac{1}{nv_q}\Im\left[-\frac{1}{\varepsilon(\boldsymbol{q},\omega)}\right] = \frac{1}{n}\frac{\Im[\chi_0(\boldsymbol{q},0)]}{\Im[\varepsilon(\boldsymbol{q},\omega)]^2 + \Re[\varepsilon(\boldsymbol{q},\omega)]^2} \tag{3.30}$$

with n the particle density and the loss function $\Im\left[-\frac{1}{\varepsilon(\boldsymbol{q},\omega)}\right]$.

Single-Particle Excitations and Plasmons

The Lindhard function describes the ability to excite the electron gas. Depending on the dimensionality and dispersion of the system, this is only possible for certain combinations of momenta and energies. Fig. 3.4 gives an idea about the possible intraband electron-hole excitation continuum in an electron gas.

The polarization function in RPA goes beyond this. It does not only describe the single-particle excitations (SPEs) known from the Lindhard function, but also collective excitations. For long range Coulomb interactions, these are collective density-density oscillations called plasmons, sketched in Fig. 3.5. They occur as a divergency of the effective Coulomb interaction (3.25). Their dispersion can be obtained from

$$\varepsilon(q,\omega_p - i\gamma) = 1 - v_q\chi_0(q,\omega_p - i\gamma) = 0 \tag{3.31}$$

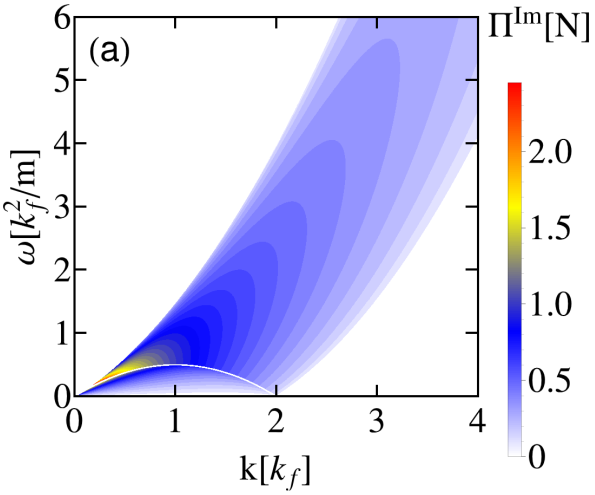


Figure 3.4. The electron-hole continuum of a non-interacting electron gas in 2 dimensions with quadratic dispersion. The boundaries $\omega_{\pm}(q) = \frac{\hbar q^2}{2m} \pm v_F q$ confine the continuum. In one spatial dimension, the lower region between 0 and $2k_f$ is absent. The plot is further discussed in App. A. Reprinted figure with permission from Ref. [JMT14b]. Copyright (2014) by the APS.

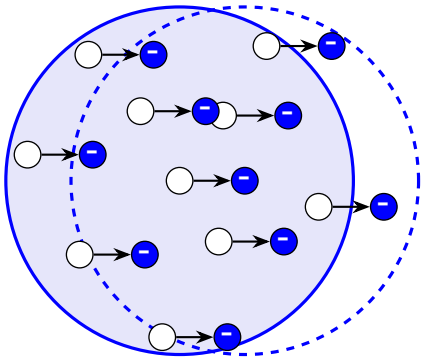


Figure 3.5. Plasmons are the quasiparticles of the collective oscillation modes of the electron gas.

with the plasmon frequency ω_p. Often single-particle continuum and plasmons occur at the same energy and momentum. In this case, the plasmons are damped which is modeled by a finite damping γ in Eq. (3.31). In the limit of weak damping, one can perform an expansion of Eq. (3.31) in γ. To first order one finds

$$\begin{aligned} 1/v_q &= \Re\left[\chi_0(q,\omega)\right] + \gamma\partial_\omega \Im\left[\chi_0(q,\omega)\right]|_{\omega\to\omega_p}, \\ 0 &= \Im\left[\chi_0(q,\omega)\right] - \gamma\partial_\omega \Re\left[\chi_0(q,\omega)\right]|_{\omega\to\omega_p} \end{aligned} \tag{3.32}$$

which can be solved for ω_p and γ.

3.2.3. Quantum Hall Ferromagnetism and the nonlinear σ-model

The Fermi liquid theory, Sec. 3.2.1, provides the possibility to incorporate interactions by considering interacting electrons in 2 and 3 dimensions as quasiparticles, dressed by electron-hole excitations. The RPA, Sec. 3.2.2, describes the dynamics of these quasiparticle based on a

perturbation theory in the interaction strength. It is thus only valid for weak and intermediate interaction strength.

Complementary to these quite general theories, in this section, we introduce a theory that is tailored for a specific class of systems in the strongly interacting limit. The system is assumed to order in a specific way and one can develop a theory that describes distortions, thus excitations, to this perfect order. In a way, it is again a perturbative theory like the RPA, but not around the non-interacting limit, but around the strongly interacting one. This section is based on Refs. [SKKR93, MMY+95] as well as Ref. [Gir99] Sec. 1.10. It provides the theory basis for Chap. III. Additional details can be found in Refs. [YMZ+94, YMB+96, AKL99, YDM06].

Quantum Hall Ferromagnetism Basics

Let us consider a system in a strong magnetic field. The electrons form LLs, with a completely flat dispersion relation in the bulk. For the effective theory we would like to present, we consider only two of these levels, which are the highest filled one and the lowest unfilled. We assume that this two level system is at half filling $\nu = 1$, meaning that there is one electron per two orbitals and the low lying level is completely filled, while the higher level is empty.

It is convenient to describe the occupation of this two level system by a pseudospin degree of freedom, where e.g. spin up (down) means level a (b) is occupied. The pseudospin can take any value on the Bloch sphere, so intrinsically in this description all superpositions of partially occupied states are included. The nature of the two levels a and b is not important. They can have different spin, then pseudospin and spin coincide, but they could also correspond to states in the two layers of a bilayer system.

In the limit of strong interactions between the electrons in different orbitals, one can invoke Hund's rule which states that the total system can lower its energy if it (pseudo)spin polarizes. The reason is that the state with maximum global spin is symmetric under spin exchange, meaning the spatial wave function is fully antisymmetric. This increases the interparticle distance and decreases Coulomb repulsion. The developing order is usually counteracted by the increase in kinetic energy due to the Pauli principle. Yet in a LL system, the kinetic energy is quenched by the magnetic field such that the LLs have a flat dispersion relation and are massively degenerate. Therefore there is no increase in kinetic energy and one could expect a 100% spin polarization. The system is supposed to be in the so called quantum Hall ferromagnetic (QHFM) ground state.

Due to this full order, the ground state of N electrons is expected to be of simple Slater determinant form. It can be explicitly written down as

$$\Psi = \Psi_V |\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow \ldots \uparrow\rangle \tag{3.33}$$

with the Vandermonde determinant wave function

$$\Psi_V = \Pi_{i<j} \left(z_i - z_j\right) \Pi_k \exp\left(-\left|z_k\right|^2 / 4l_B^2\right) \tag{3.34}$$

with the magnetic length $l_B = \sqrt{\frac{\hbar}{eB}}$. The two-particle distribution function is

$$g\left(\left|\boldsymbol{r} - \boldsymbol{r}'\right|\right) = 1 - \mathrm{e}^{-|\boldsymbol{r}-\boldsymbol{r}'|^2/2l_B^2}. \tag{3.35}$$

One can use it to directly calculate the exchange hole that surrounds each particle and lowers its Coulomb energy by

$$\begin{aligned} E_x &= \frac{1}{2\pi l_B^2} \frac{e^2}{4\pi\varepsilon_0\varepsilon_r} \int \mathrm{d}^2 r \frac{1}{r} \left[g\left(r\right) - 1 \right] \\ &= \sqrt{\frac{\pi}{2}} \frac{e^2}{4\pi\varepsilon_0\varepsilon_r} \frac{1}{l_B}. \end{aligned} \tag{3.36}$$

This puts our argument about using Hund's rule on a concrete basis.

The explicit wave function (3.33) offers the chance to study the ground state physics analytically. This is done by projecting all relevant operators of the Hamiltonian onto the two level system, see Ref. [MMY+95] for details.

In this thesis, we follow a different approach. We focus on deriving an effective model for the two level system, where the free parameter is the pseudospin degree of freedom. As all spins are aligned in a QHFM system, the ground state can be represented by the isospin order parameter

$$\boldsymbol{n}\left(\boldsymbol{r}\right) = \Big(\sin\left(\theta\left(\boldsymbol{r}\right)\right)\cos\left(\phi\left(\boldsymbol{r}\right)\right), \quad \sin\left(\theta\left(\boldsymbol{r}\right)\right)\sin\left(\phi\left(\boldsymbol{r}\right)\right), \quad \cos\left(\theta\left(\boldsymbol{r}\right)\right) \Big)^T \tag{3.37}$$

on the Bloch sphere. The physics can then be cast into a model for this order parameter, called nonlinear σ-model. This is done in detail for a concrete example in Sec. 7.

Here we conclude this description by shortly discussing the form of excitations of such a QHFM state, compared to the systems discussed beforehand. Projecting the operators on the two lowest LLs, one finds a physical interesting property: the spin and charge density operators do not commute anymore. Spin and charge are entangled in the QHFM state, which becomes physically observable for the low-energy charge excitations.

In the Fermi liquid, the low-energy excitations are electronic quasiparticles excited above the Fermi sea, see Sec. 3.2.1. Their dynamics can effectively be described perturbatively, e.g. in RPA (Sec. 3.2.2) for weak interactions. In the QHFM state, we are in the opposite limit of strong interactions, where the electrons are strongly correlated. An excitation of a single electron would have to overcome the exchange energy, Eq. (3.36), which is costly. It turns out that due to this exchange energy, not electron-hole pairs are the low-energy charge excitations of the QHFM state, but skyrmions, as calculated in Ref. [MMY+95]. A skyrmion anti-skyrmion pair costs only half the energy of an electron-hole pair. The energy per skyrmion is

$$E_s = \frac{1}{4}\sqrt{\frac{\pi}{2}} \frac{e^2}{4\pi\varepsilon_0\varepsilon_r} \frac{1}{l_B} \tag{3.38}$$

for the state defined in Eq. (3.33).

Skyrmions

Skyrmions are smoothly, on the length scale of l_B, varying deformations of the isospin order parameter $\boldsymbol{n}$. These spin structures deform in a way that includes an integer amount of windings n_{sk} around the Bloch sphere. For an overview, see Ref. [NT13]. A simple example for a skyrmion with a single winding $n_{sk} = 1$ is shown in Fig. 3.6 (a). The winding marks skyrmions as topological excitations, with the winding number

$$n_{sk} = \int \frac{\mathrm{d}^2 r}{4\pi} \boldsymbol{n} \cdot \left(\partial_x \boldsymbol{n} \times \partial_y \boldsymbol{n}\right) \tag{3.39}$$

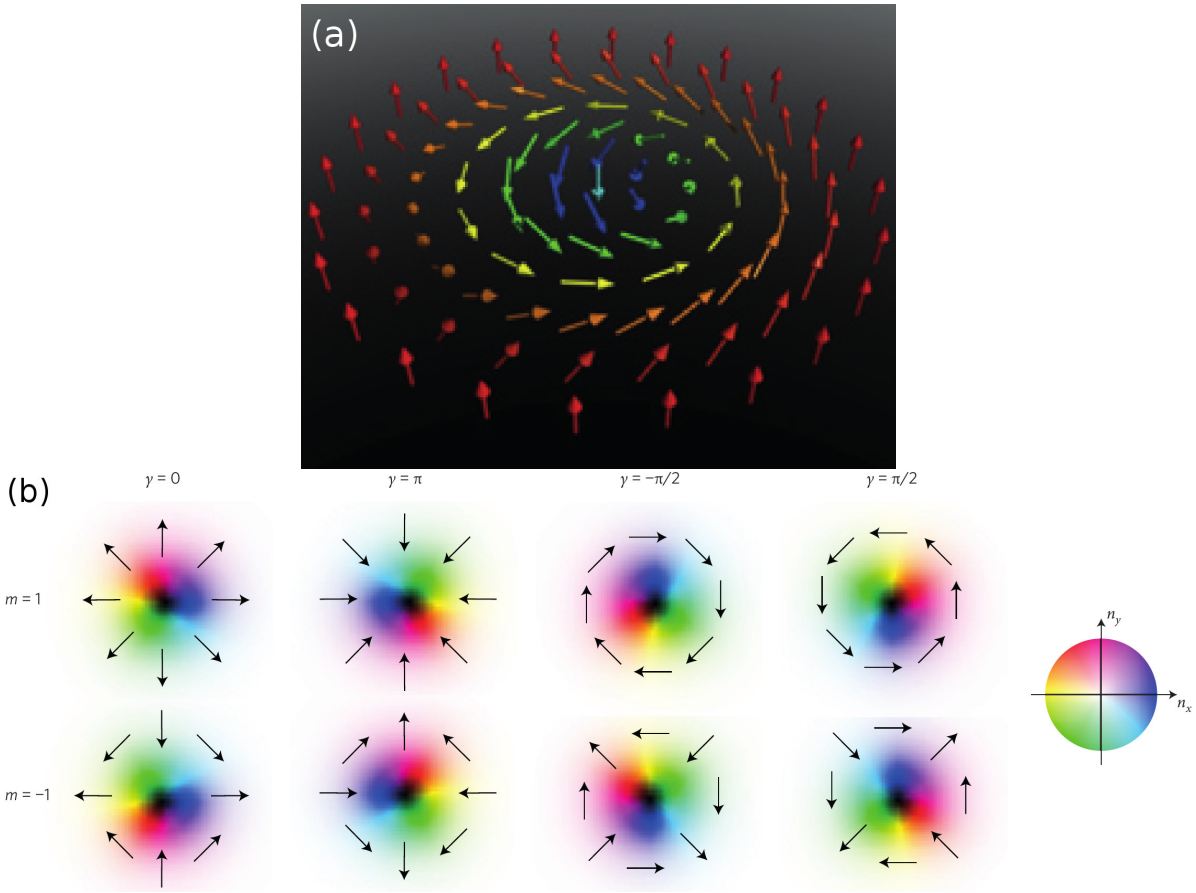


Figure 3.6. (a) 3D sketch of a skyrmion spin structure. (b) Different skyrmion types for parameters γ and m defined in Eq. (3.41). The black arrows denote the in-plane spin orientation. Adapted by permission from Macmillan Publishers Ltd: Nature Nanotechnology, Ref. [NT13], copyright (2013).

being quantized to integer values for skyrmions and half-integer values for merons. Due to the entanglement of charge and spin in the QHFM state, the topological charge n_{sk} describes also the electrical charge

$$Q_{sk} = \nu n_{sk} e \tag{3.40}$$

of the excitation with the filling factor ν.

Skyrmions come in different configurations, of which a few are shown in Fig. 3.6 (b). There, the parametrization $\boldsymbol{r} = \left(\begin{array}{cc} r\cos(\varphi), & r\sin(\varphi) \end{array} \right)^T$ is used with

$$\phi(\varphi) = m\varphi + \gamma \tag{3.41}$$

and $\boldsymbol{n}(\boldsymbol{r})$ given by Eq. (3.37). We conclude from these sketches that the spin-momentum locking for Dirac cones, described in Sec. 2.4.2, can be nicely related to skyrmion physics, compare Fig. 3.6 (b) to Fig. 2.20. Instead of the isospin $\boldsymbol{n}$ it is the vector $\hat{\boldsymbol{d}}$ from the Hamiltonian $H = \boldsymbol{d} \cdot \boldsymbol{\sigma}$ that has a non-trivial winding in this case. The difference is that the winding is only in the x-y plane in 2D, such that one gets a winding number of $n_{dirac} = \frac{1}{2}$ in accordance to the Berry phase of π found for the 2D Dirac states in Sec. 2.4.2. These objects with half integer winding are called merons. Another important difference is that for the Dirac cone the winding happens in momentum space, while in the QHFM systems it is located in real space.

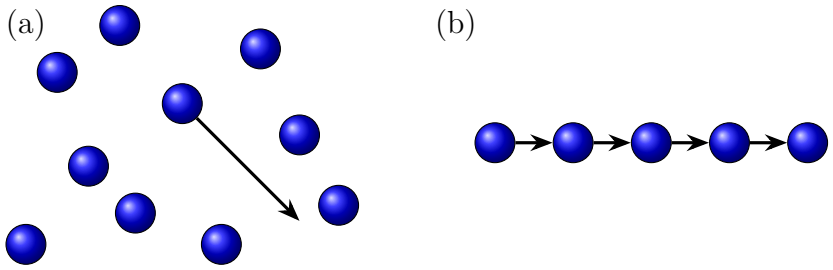


Figure 3.7. (a) Electron motion in $D \geq 2$ dimensions. (b) Electrons in 1D are always strongly correlated due to spatial restrictions.

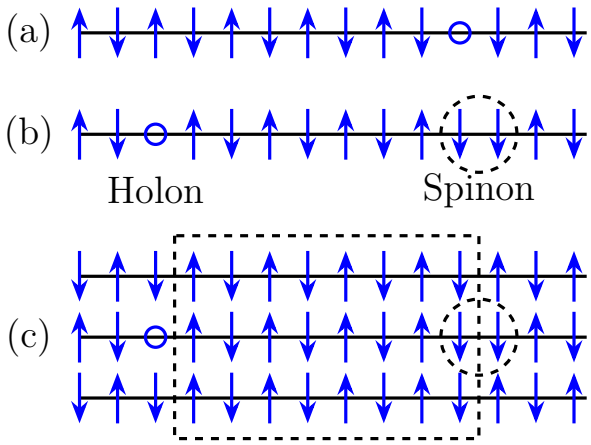


Figure 3.8. (a) Antiferromagnetic spin chain with a vacancy. (b) The vacancy moves, which effectively looks like two propagating excitations, a Holon and a Spinon. (c) In 2D, Holon and Spinon are always bound together by the exchange coupling to the surrounding spin chains.

3.2.4. Interaction in One Dimension: Luttinger Description

We conclude the method section with a brief description of 1D systems. Interaction effects in 1D are special, and the Fermi liquid picture recapped in Sec. 3.2.1 does not apply. Its place will be taken by the Luttinger liquid, which we introduce in the following. This section is based on Ref. [Gia03] and Ref. [QZ11] Sec. 2.C.2. The described physics will be relevant for the strongly interacting edge states considered in Chap. III.

Why interactions in 1D are different

In higher spatial dimensions ($D \geq 2$), electrons exist as nearly free quasiparticles according to Fermi liquid theory. In the following, we show, based on Ref. [Gia03] Secs. 1.2 and 3.2, that this picture fails in 1D.

Electrons in $D \geq 2$ can move around each other in the electron gas, which naturally mitigates correlations between them. In 1D instead, every individual electron pushes its neighbor due to electron-electron interaction, see Fig. 3.7. Therefore every excitation of the electron system is collective and there can be no Fermi liquid theory of weakly interacting quasiparticles.

Let us extend our analysis to particles with spin, depicted as an antiferromagnetic spin chain in Fig. 3.8 (a). One electron is missing, and this vacancy starts propagating along the chain. It looks like the single fermionic excitation splits into a charge and a spin collective

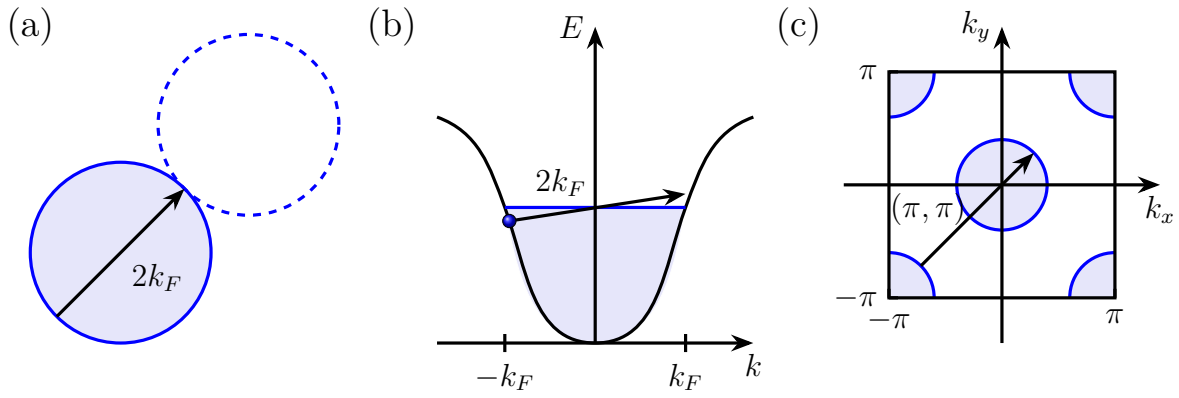


Figure 3.9. (a) Weak nesting condition $q = 2k_F$ for circular Fermi surface in 2D. (b) Fermi surface in 1D consists of two points resulting in perfect nesting. (c) More involved Fermi surfaces in 2D also lead to nesting conditions for special wave vectors.

excitation moving into opposite directions, see Fig. 3.8 (b). Thus the electron breaks apart into two elementary excitations, the holon (charge) and the spinon (spin). These spin and charge density waves have in general different velocities. In 2D systems, these two particles would be bound together by the exchange coupling to the other chains, see Fig. 3.8 (c). Thus particle fractionalization (in particular spin-charge separation) is a special feature in 1D systems.

Another way to see that one needs a special theory for 1D systems is looking directly at the failure of the perturbation theory summed up in Sec. 3.2.2. Assuming linear response theory works, we look at the density-density correlation function as response to an external perturbation

$$H_{ext} = \int \mathrm{d}^d r V_{ext}\left(\boldsymbol{r}, t\right) n\left(\boldsymbol{r}\right) \tag{3.42}$$

with n the density of the system. The susceptibility measuring the response is the Lindhard function, given by

$$\chi_0\left(\boldsymbol{q}, \omega\right) = \frac{1}{V} \sum_{\boldsymbol{k}} \frac{f\left(E_{\boldsymbol{k}}\right) - f\left(E_{\boldsymbol{k}+\boldsymbol{q}}\right)}{\omega + E_{\boldsymbol{k}} - E_{\boldsymbol{k}+\boldsymbol{q}} + i0^+} \tag{3.43}$$

with volume $V = L^d$ for systems with linear dimension L and f the Fermi-Dirac function. In the $\omega = 0$ limit, one expects a divergency if $E_{\boldsymbol{k}} = E_{\boldsymbol{k}+\boldsymbol{q}}$ is fulfilled, but due to the integration over $\boldsymbol{k}$ in Eq. (3.43) it is smoothed out. Especially in higher dimensions only the derivatives of $\chi_0\left(\boldsymbol{q}, 0\right)$ are discontinuous, resulting in phenomena like Friedel oscillations.

There is a way for stronger divergencies to occur, which is nesting. It describes the case of a finite domain of values of $\boldsymbol{k}$ for which $E_{\boldsymbol{k}} = E_{\boldsymbol{k}+\boldsymbol{q}}$ is fulfilled. In higher dimensions this is rarely realized. A circular Fermi surface in 2D has only weak nesting for $q = 2k_f$, see Fig. 3.9 (a), although there are exceptions for more involved Fermi surfaces, e.g. involving electron and hole puddles as sketched in Fig. 3.9 (c). The 1D Fermi surface on the other hand consists only of two points, such that $E_k = E_{k+2k_F}$ is fulfilled at the Fermi momentum for inversion symmetric systems. This results in perfect nesting in 1D. Therefore there is a divergency in χ_0 at $q = 2k_F$ and we expect the failure of perturbation theory. The ground state of the interacting system is thus expected to differ substantially from the non-interacting one.

Is there a simple picture to better describe 1D systems in order to avoid divergencies such as

we just discussed? Let's consider, in 1D, a p-h excitation around the Fermi level with energy

$$E_k^{p,h}(q) = E_{k+q} - E_k \tag{3.44}$$

where the state for E_k is supposed to be occupied and the state at E_{k+q} originally empty. Expanding around the Fermi momentum for small q, one gets the average energy $E^{p,h}(q)$ and dispersion $\delta E(q)$ of a p-h pair as

$$\begin{aligned} E^{p,h}(q) &= v_F q, \\ \delta E(q) &= \frac{E^{p,h}(q)^2}{m v_F^2} \end{aligned} \tag{3.45}$$

with Fermi velocity v_F and electron mass m. Similar to the quasiparticles in Fermi liquid theory in $D \geq 2$, p-h excitations are well-defined quasiparticles close to the Fermi energy in 1D systems. Their dispersion goes to zero like energy squared as for the quasiparticles in the Fermi liquid description. As p-h excitations are bosonic in nature, we can understand this as a hint that the bosonization of our system will solve the problems occurring in 1D.

Luttinger Liquids

Following the conclusion from the last section, we want to derive an effective theory for the low-energy excitations of a 1D system. In contrast to the fermionic quasiparticles of Fermi liquid in $D \geq 2$, this Luttinger liquid description is based on bosonic p-h excitations. We closely follow Ref. [Gia03] Sec. 3.1f & App. D in the derivation. The major results needed in Chap. III will be the Hamiltonian (3.60) and the action (3.61).

The effective theory will contain two degrees of freedom, the fields ϕ, where $\nabla\phi$ is the density variation and thus linked to Coulomb energy from interactions, and θ. The latter is related to the particle creation operator and thus linked to the kinetic energy. We rederive the particle Hamiltonian in this language, which corresponds to bosonization. The described excitations are bosons as argued before, and ϕ and θ can be represented by bosonic operators.

We start by writing down the 1D density

$$\rho(x) = \sum_i \delta(x - x_i) \tag{3.46}$$

with x_i the position operator of the ith particle. A labeling field ϕ_l can be introduced, which is continuous and takes the value $\phi_l(x_i) = 2\pi i$ at the position of the ith particle. It constitutes a unique way of numbering particles, because in 1D particles can be distinctively ordered. So the field is well defined, and we use it to rewrite the density into

$$\rho(x) = \sum_i \delta(x - x_i) \tag{3.47}$$

$$= \sum_n |\nabla\phi_l(x)| \, \delta(\phi_l(x) - 2\pi n). \tag{3.48}$$

Imagine $\phi_l(x)$ as an increasing function of x as sketched in Fig. 3.10. Thus we drop the absolute value, choose a representation of the δ function based on the Poisson summation formula, and find

$$\rho(x) = \frac{\nabla\phi_l(x)}{2\pi} \sum_p \mathrm{e}^{ip\phi_l(x)} \tag{3.49}$$

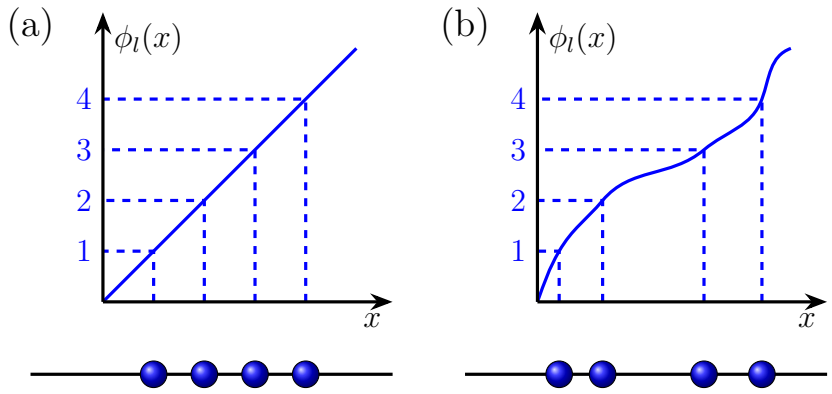


Figure 3.10. Examples for the labeling field $\phi_l(x)$ for (a) a homogeneous system and (b) one with spatial deviations.

with p an integer. The actual field of interest ϕ describes the mean-field fluctuations around the perfect crystalline lattice solution. The latter is given by $\phi_l^0(x) = 2\pi\rho_0 x$ with $\rho_0 = 1/d$ the equilibrium density of particles and d the average distance between particles. $\phi(x)$ is then defined via

$$\phi_l(x) = 2\pi\rho_0 x - 2\phi(x). \tag{3.50}$$

Using this notation, the density becomes

$$\rho(x) = \left[\rho_0 - \frac{1}{\pi}\nabla\phi(x)\right] \sum_p \mathrm{e}^{i2p(\pi\rho_0 x - \phi(x))}. \tag{3.51}$$

We look now at the long wavelength limit, at which the terms in Eq. (3.51) oscillating on the interparticle distance scale d vanish. The averaged or smeared density is thus

$$\rho_{q\approx 0}(x) = \rho_0 - \frac{1}{\pi}\nabla\phi(x). \tag{3.52}$$

The second field $\theta(x)$ needed for our rewriting of the Hamiltonian is introduced by expressing the particle creation operator like

$$\psi^\dagger(x) = [\rho(x)]^{1/2}\,\mathrm{e}^{-i\theta(x)}. \tag{3.53}$$

The (anti-)commutation relations between the operators ψ impose commutation relations between the density operator and $\theta(x)$. For bosons they are given by

$$\left[\psi_B(x), \psi_B^\dagger(x')\right] = \delta(x - x'). \tag{3.54}$$

Under the assumption that the field θ commutes with itself, $[\theta(x), \theta(x')] = 0$, a sufficient condition satisfying Eq. (3.54) is

$$\left[\rho(x), \mathrm{e}^{-i\theta(x')}\right] = \delta(x - x')\,\mathrm{e}^{-i\theta(x')}. \tag{3.55}$$

Again focusing on the long wavelength limit, we assume that the fields ϕ and θ vary slowly on the interparticle distance ρ_0^{-1}. Hence we replace the exact density in Eq. (3.55) by the smeared one, Eq. (3.52). The commutator is then fulfilled by

$$\left[\frac{1}{\pi}\nabla\phi(x), \theta(x')\right] = -i\delta(x - x') \tag{3.56}$$

where higher oscillating terms, not fulfilling the relation, are neglected in the continuum limit. θ and $\frac{1}{\pi}\nabla\phi(x)$ are canonically conjugated, hence an integration by parts leads to

$$\pi\Pi(x) = \nabla\theta(x) \tag{3.57}$$

with $\Pi(x)$ the canonically conjugate momentum to $\phi(x)$. The bosonic creation operator

$$\psi_B^\dagger(x) = \left[\rho_0 - \frac{1}{\pi}\nabla\phi(x)\right]^{1/2} \sum_p \mathrm{e}^{i2p(\pi\rho_0 x - \phi(x))}\mathrm{e}^{-i\theta(x)} \tag{3.58}$$

follows from Eqs. (3.51) and (3.53).

For fermions one has to fulfill an anticommutation relation instead of Eq. (3.54). The extra minus sign can be accounted for by the definition of the fermion creation operator

$$\psi_F^\dagger(x) = \psi_B^\dagger(x)\,\mathrm{e}^{i\frac{1}{2}\phi_l(x)}. \tag{3.59}$$

As ϕ and Π are canonically conjugated, they can be expressed in terms of bosonic operator b_p and $b_p^\dagger$. These bosons represent the small oscillations of the density as p-h excitations.

Applying this newly introduced notation, we consider the dominant contributions to the many-body Hamiltonian. The interactions $\int \mathrm{d}x\rho(x)^2$ will give the term $(\nabla\phi)^2$ to leading order, following from Eq. (3.51). The kinetic energy $\int \mathrm{d}x\frac{1}{2m}(\nabla\psi^\dagger(x))(\nabla\psi(x))$ adds the part $(\nabla\theta)^2$. Here, all the terms in the particle operators involving derivatives like $\nabla\phi$ are considered as higher order contributions. Due to inversion symmetry which we assume for 1D systems, there are no cross terms. The energy should be invariant under the substitution $x \to -x$, as should be the operators ρ and ψ. From the latter condition it follows $\nabla\phi(x) = \nabla\phi(-x)$ and $\psi(x) = \psi(-x)$, so cross terms like $\nabla\psi\nabla\phi$ would change sign.

In conclusion, the Luttinger liquid Hamiltonian containing the dominant parts of the energy of a 1D system is given by

$$H = \frac{\hbar}{2\pi}\int \mathrm{d}x\left[\frac{v\mathcal{K}}{\hbar^2}(\pi\Pi(x))^2 + \frac{v}{\mathcal{K}}(\nabla\phi(x))^2\right] \tag{3.60}$$

with $\pi\Pi(x) = \nabla\theta(x)$ and the Luttinger liquid action is

$$S/\hbar = \frac{1}{2\pi\mathcal{K}}\int \mathrm{d}x\mathrm{d}\tau\left[\frac{1}{v}(\partial_\tau\phi)^2 + v(\partial_x\phi(x))^2\right]. \tag{3.61}$$

The two parameters v and $\mathcal{K}$ completely characterize the low-energy properties of a massless 1D system. v is considered a velocity, while $\mathcal{K}$ is the interaction parameter. $\mathcal{K} = 1$ for non-interacting, $0 < \mathcal{K} < 1$ for repulsive interactions.

Currents in the system can be directly obtained from the continuity equation $\frac{\partial\rho}{\partial t} + \nabla j = 0$. One finds the formula for the current

$$j = \partial_t\phi/\pi. \tag{3.62}$$

With our topological materials in mind, we define, in a system with left (L) and right (R) moving electrons, also left and right moving densities, via

$$\nabla\phi(x) = -\pi\left[\rho_R(x) + \rho_L(x)\right], \tag{3.63}$$

$$\nabla\theta(x) = \pi\left[\rho_R(x) - \rho_L(x)\right]. \tag{3.64}$$

This closes the section about ways of treating correlation effects in condensed matter systems. Next, we give a few examples of connecting interaction effects with topological properties.

3.3. Interactions and Topology - a Plethora of Possibilities

So far, we have looked separately at the effects of topology and the consequences of interactions in condensed matter systems. As a prelude for the following chapters II-IV, here, we will motivate why the combination of the two can be a fruitful endeavor. As this is a boundless task, we restrict ourselves to the cases relevant for the subsequent chapters. Our examples include interactions as a bulk probing tool for topology in Sec. 3.3.1, the combination of plasmonics and topology in Sec. 3.3.2, the study of symmetry breaking by interactions in Sec. 3.3.3 as well as the consideration of topological multilayer systems in Sec. 3.3.4.

3.3.1. Interactions as a Probing Tool for Topology

The most obvious and measurable difference between topological trivial and non-trivial phases is the occurence of surface states at the boundary. On the other hand, also trivial surface states can form, e.g. due to band bending effects. Additionally, pure surface effects could be hard to detect for experimental tools that are also bulk sensitive.

In contrast to this, one could look for definite topological distinctions in the bulk properties of a system. This is a much more subtle task. The topology is not encoded in the energy spectrum, but in the structure of the Berry phase, see Sec. 2.2.2. This is a quantity quite hard to access experimentally. On the plus side, bulk quantities are generally more accessible in experiments.

The bulk property we would like to consider is the excitation spectrum of a material. It depends on the structure of the Berry phase of the involved bands. An example is the spin-momentum locking in graphene, giving severe limitations to the scattering of electrons in the bulk bands. A concrete physical quantity where this enters is the Lindhard response function $\chi_0(\boldsymbol{q},\omega)$ defined in Eq. (3.23). Here, the overlap factor $\mathcal{F}^{\lambda,\lambda'}_{\boldsymbol{k},\boldsymbol{k}'}$ directly depends on the Berry phase structure of the involved bands.

In an interacting systems, the RPA response function Π^{RPA}, Eq. (3.24), provides insight into the excitation spectrum. It is directly connected to measurable quantities like the dynamic structure factor $A(\boldsymbol{q},\omega)$ given in Eq. (3.30). The study of the influence of topology on the bulk excitation spectrum has been one of the goals of the project presented in Chap. II.

3.3.2. Applications: Plasmonics and Topology

Π^{RPA} includes both single-particle as well as collective excitations, see Sec. 3.2.2. These collective charge excitations of the electron gas are called plasmons.

An interesting application of plasmons is in the connection of electronics with optics. Plasmons can be used to enhance the coupling between light and electronic degrees of freedom of the electron gas. This goal results in the name plasmonics for this field of research, from joining plasmon and photonics. The hope is to integrate photonic systems, which are usually on the micrometer scale, into electronic chips, where nanometers are the relevant length scale. See Ref. [Ozb06] for an introduction.

The whole field gained a boost with the advent of graphene, which seems to be an ideal platform in order to tailor the interaction between light and electrons at will. We consider two examples from Refs. [KCG11, GPN12] in Fig. 3.11. The first one is a dipole molecule on the surface of graphene. The coupling of the two is enhanced by a plasmonic resonance, leading to the possibility of strong light-matter interaction. Practical applications as nanoantennas,

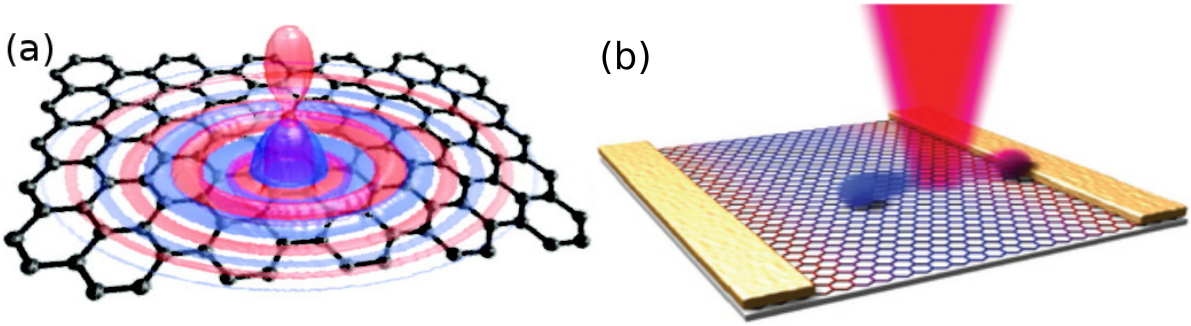


Figure 3.11. (a) Dipole molecule on the surface of graphene leads to a strong coupling between the two systems. (b) Incident light can create electron-hole pairs in graphene. (a) reprinted with permission from Ref. [KCG11]. Copyright (2011) American Chemical Society. (b) adapted by permission from Macmillan Publishers Ltd: Nature Photonics, Ref. [GPN12], copyright (2012).

gas detectors or for Raman scattering on single molecules come to mind. The second example is less profound. Due to a plasmonic resonance, incident light can easier create electron-hole pairs in graphene. This would boost the efficiency of solar cells.

A natural question is now to think about going from graphene to topological Dirac systems found in TIs, see e.g. Ref. [DOL+13]. The search for topological benefits in the field of plasmonics is just at the beginning. A comprehensive overview over the results so far is provided in Ref. [Sta14]. We contributed to it with the project presented in Chap. II.

3.3.3. Stability: Symmetry Breaking due to Interactions

Usually topological properties, like gapless edge states, are guaranteed by the presence of certain symmetries in the system. Examples for TIs are $\mathcal{T}$ symmetry, see Sec. 2.2.3, or some spatial symmetries like inversion, discussed in Sec. 2.4.3.

The presence of interaction can break these symmetries. This can happen explicitly, i.e. the interaction does not have the symmetries of the underlying non-interacting system. Another way is spontaneous symmetry breaking. Whether and how topological properties are preserved or destroyed by interactions was studied by us for a 2D TI, with HgTe QWs in mind, in Chap. III. The interaction strength in the system can be tuned by changing an applied magnetic field. The topological properties directly depend both on the assumed spatial symmetry of the interaction terms as well as their respective strength.

The effect of $\mathcal{T}$ symmetry breaking is studied in Chap. IV. There we consider the coupling of a $\mathcal{T}$ broken WSM to a TI. The surface Dirac cones can still be gapless, even if $\mathcal{T}$ symmetry is broken in the combined system.

3.3.4. Multilayer Systems: Interactions as Mediator between Topological Phases

Correlations can both destroy and enrich topological physics as we have argued so far. But one can also think about Coulomb interaction as a tool to create something new that could not exist without it. In this section, we talk about topological multilayer systems: (different) 2D topological phases combined together by an interlayer Coulomb interaction. These can be layers of 2D TIs or surface states of 3D topological systems. Their joint topological properties will crucially depend on the properties and symmetries of the interlayer interaction.

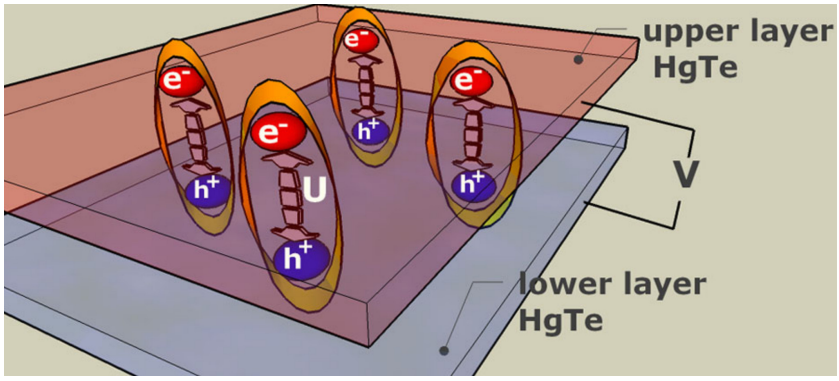


Figure 3.12. Double layer system of HgTe QWs. For sufficient strong interlayer Coulomb interaction, electrons and holes of different layers bind together to form excitons. Reprinted figure with permission from Ref. [BTM14]. Copyright (2014) by the APS.

The study of multilayer system can be performed on various levels. A simple approach is the assumptions of δ-layers of 2D systems, coupled only weakly by Coulomb interaction. In this limit, the extension of a perturbative calculation from one to several layers is manageable, e.g. done for graphene multilayers in Ref. [PPA+10]. System properties concerning screening effects and electronic excitations can be easily obtained.

One experimental relevant application of such multilayers systems are drag setups. There one drives a current through one layer and measures the Coulomb induced current or voltage in the adjacent layers. For a basic introduction see Ref. [JS93], more detailed linear-response calculations of Coulomb drag effects are found in Ref. [FHJK95]. As with plasmonics, graphene is an interesting system in this context, see e.g. Ref. [GGK+12]. Remarkably, plasmons can also play a role in this context. They enhance the interlayer Coulomb interaction in certain regimes, strengthening the drag effect. An example calculation for graphene is given in Ref. [SJ14]. A comprehensive review on drag effects is found in Ref. [NL16].

More recently, drag experiments on TI film surfaces have been studied theoretically, see Ref. [LLC16]. The newest systems considered are WSMs. In Ref. [BS17], the intrinsically present Weyl orbits, see Sec. 2.5, are used to create non-local Coulomb drag between two opposite surfaces of the system.

In order to address the question of topological classification of such a combined system, instead of transport calculations it is more interesting to find out whether the ground state of the coupled systems changes. An example for a new ground state of the combined system is the formation of an exciton condensate. In the case that the interlayer Coulomb interaction is dominant, a pairing of electrons and holes of different layers occurs, as depicted in Fig. 3.12. The coupled system is then in a ground state that is inherently different from the one in the former separated layers. As the interlayer Coulomb interaction is usually weak compared to other energy scales of the problem, the pairing can be enhanced by quenching the kinetic energy of the system with a strong magnetic field. Then, the electrons form flat LLs as already discussed for QHFM systems in Sec. 3.2.3, and Coulomb energy becomes dominant. Electrons of one layer then bound to vacant states, represented as holes, of the second layer. For low temperature, the resulting excitons can condensate. For an overview about exciton condensation in 2D quantum Hall (QH) systems see Refs. [EM04, Eis13].

Whether exciton condensates can also form without the help of LLs in realistic setups is

still under debate. For graphene it was answered positively in Ref. [MBSM08] and declined in Ref. [KE08]. Even more interesting for us is the formation of a topological exciton condensate based on TIs systems, see e.g. Refs. [SMF09, WHFZ12]. For HgTe QWs this is sketched in Fig. 3.12, based on Ref. [BTM14]. There, the authors find that the topology of the combined system is explicitly depending on the Coulomb interaction, and can even differ from the topology of the former decoupled subsystems. Very recently, also condensates in Weyl bilayers, see Ref. [MT17], have been discussed.

Our own work in Chap. IV considers the combination of a 3D TI and a WSM system. Special attention is paid to the emerging hybrid surface states combining properties of Fermi arcs and Dirac surface states.

Part II.

Correlation Effects within Random Phase Approximation

Topological insulators (TIs) are among the most actively investigated systems in condensed matter physics today. In reality, there is experimental evidence for their existence in two and three spatial dimensions, see Secs. 2.4.1 and 2.4.2. Due to the bulk-boundary correspondence discussed in Sec. 2.2.4, non-trivial topological states of matter have edge states at their boundaries with peculiar transport and optical properties. For instance, the two-dimensional (2D) time-reversal ($\boldsymbol{\mathcal{T}}$) symmetric quantum spin Hall (QSH) state, realized in Hg(Cd)Te quantum wells (QWs), is known to come along with helical edge states that are protected against elastic backscattering of non-magnetic impurities, see Sec. 2.4.1 for details. The helicity (2.36), i.e. coupling of momentum and spin degrees of freedom, introduces new interesting phenomena in this context.

However, not only the edge state physics of these systems is interesting, but also the 2D bulk physics bears exciting novelties. The low-energy excitations of Hg(Cd)Te QWs are described by the Bernevig-Hughes-Zhang (BHZ) model, see Sec. 2.4.1, that interpolates between the limiting cases of Schrödinger and Dirac fermions. This interplay between Schrödinger and Dirac physics constitutes an opportunity for new, rich phenomena to emerge.

An important aspect of condensed matter physics is also the influence of Coulomb interaction on observables. In the Dirac fermion system graphene, see Refs. [GN07, CGP+09], this research has been intensified in recent years, Refs. [Shu86, And06, BPBP+07, HD07, WSSG06, KUP+12], resulting even in the development of plasmon technology, summed up in Sec. 3.3.2. Plasmons, as described in Sec. 3.2.2, are collective density oscillations commonly occurring at finite doping in an electronic system. Plasmons in Dirac fermion systems have been experimentally observed in graphene, Refs. [JGH+11, CWS12, FRA+12], as well as on the surface of three-dimensional (3D) topological insulators (TIs) in Ref. [DOL+13].

Usually absent in the undoped limit, intrinsic plasmons have been predicted in graphene in Refs. [Vaf06, DL13], when the electron and hole gas have a finite density due to thermal excitations. Additionally, with the inclusion of excitonic effects through ladder-type vertex corrections in the calculation of the dielectric response function, intrinsic plasmons have been predicted in Ref. [GFM08]. Yet, the latter result is still under debate due to the neglect of diagrams of the same order, see, e.g., Ref. [SPM12]. Interestingly, a reduction of the dimension to 1D gives rise to intrinsic plasmons in metallic armchair nanoribbons, as discovered in Ref. [BF07]. These situations are physically distinct from our predictions below, where we show that plasmons, in the intrinsic limit, can appear due to an interplay of Dirac and Schrödinger physics.

In the following sections, we present a comprehensive analysis of the polarization function of the BHZ model in the static and full dynamic limit, at zero and finite doping. We study the screening properties and the collective charge excitations on the basis of the Random Phase Approximation (RPA), with the application to Hg(Cd)Te QWs in mind.

Continuously tuning the parameters of the BHZ model, we reproduce the limits of pure Dirac and pure Schrödinger fermions and explore intermediate regimes, in order to understand how analogies and differences emerge. We support our numerical calculations of the polarization functions with analytical expressions derived by f-sum rules. In the static limit, we calculate the screening properties due to the intrinsic system and at finite doping, analyzing the induced charge density (with Friedel oscillations) in response to a charged impurity. Different to the Dirac fermion system graphene, where static screening in the intrinsic limit is momentum independent and can therefore be absorbed into an effective dielectric constant as discussed in Refs. [WSSG06, HD07], the BHZ model shows a significant momentum dependence that translates into a finite extent of the induced charge density.

In the dynamic limit, we pay particular attention to the emergence of a new, intrinsic interband plasmon in the undoped limit. It results from the interplay between Schrödinger and Dirac fermion physics, and is absent in both limiting cases of pure Dirac and pure two dimensional electron gas (2DEG)[1] system. We discuss the presence of the plasmons and their damping rate for varies parameters, including the TI and the normal insulator (NI) phase as well as experimentally realistic scenarios. Furthermore, we calculate the bulk optical conductivity, which offers a way to quantitatively resolve between NI and TI phase, following the ideas presented in Sec. 3.3.1.

At finite doping, under certain conditions specified below, e.g. broken particle-hole (p-h) symmetry applicable to Hg(Cd)Te QWs, we find a coexistence between the novel interband plasmon and an ordinary intraband plasmon. Both plasmons can be rather weakly damped by single-particle excitations (SPEs) and should therefore be observable in experiments. Interestingly, the two plasmons respond to the topology of the band structure with a distinctive behavior. They merge together in a normal insulating phase, while they remain clearly resolved when the system realizes a TI. Thus besides the optical conductivity, also the plasmon spectrum offers ways to distinguish between topological trivial and non-trivial bulk physics.

Generally, RPA is known to provide reliable predictions at large densities and in systems with a large number of fermionic degrees of freedom. While its validity was indeed questioned for the intrinsic Dirac limit, where the system is unable to screen the Coulomb interaction and strong renormalization effects are expected as discussed in Ref. [KUP+12], the RPA has been shown to yield a quantitative description of many-body effects in graphene in Refs. [BHTD14, HBD14]. It has been widely used for the study of plasmons in the Dirac model, including various forms of (multilayer) graphene and TI surface states, see Ref. [Sta14] for a comprehensive review.

Closely related to our work, the intraband plasmons of black phosphorous have been studied on the basis of the RPA and an extended version of the BHZ model including anisotropy in Ref. [LA14]. A similar study has been done for MoS_2 in Ref. [SSS13]. Yet, in both works the authors restrict themselves to fixed parameters obtained from experiments and simulations. We on the other hand want to provide a systematic exploration of the parameter space.

This chapter is based on the two publications Refs. [JMT14a, JMT14b], which we reorganized into the following three sections and App. A. In Sec. 4, we introduce the BHZ model and present the general formalism we employ to calculate the static and dynamical dielectric function and the induced charge density. The nature of the nontrivial pseudospin, the origin of possible interband plasmons, experimentally relevant parameters and the different contributions to the f-sum rule are also discussed here. Subsequently, in Sec. 5, we present the static screening properties, the dynamical excitation spectrum and the f-sum rule in the undoped regime. Special attention is paid to the new interband plasmon and how to make it visible in experiments based on Raman or electron loss spectroscopy. In Sec. 6, this analysis is extended to the case of finite doping where inter- and intraband excitations equally matter. Besides the ability of the BHZ model to interpolate between Dirac and 2DEG physics, we focus on the parameter regimes relevant for Hg(Cd)Te QWs as well as on the analysis of the topological trivial and non-trivial phases. Interestingly, we find that broken p-h symmetry as well as a large, negative mass make it possible to observe both kinds of plasmons in the same system, just separated in energy in the interacting excitation spectrum. Finally a conclusion and a brief outlook are given.

[1] We use 2DEG synonymously for Schrödinger fermions in the following sections.

4. Model and Formalism

In this section, we introduce the necessary model and mathematical formalism used in the following sections 5 and 6. We do this by making the connection to the introductory Chap. I. While there the focus lay on deriving model and RPA formalism, here, we focus more on the physics contained in them.

This section is organized as follows. In Sec. 4.1 we introduce the BHZ model. It contains an intrinsic energy and momentum scale discussed in Sec. 4.1.1, where we compare it to the Fermi scale. The Lindhard polarization function in these intrinsic coordinates is given in Sec. 4.2. Special emphasis is put on the examination of the overlap factor. The Coulomb interaction on the RPA level is introduced in Sec. 4.3. In the subsections we focus on the conditions for antiscreening to arise, as well as on the static limit important for screening and calculating induced charge distributions. Experimental parameters for the case of Hg(Cd)Te QWs are given in Sec. 4.4. We end in Sec. 4.5 with a discussion of the f-sum rule.

4.1. The BHZ Model

The BHZ model is derived for Hg(Cd)Te QWs as a low-energy model around the Γ-point in Sec. 2.4.1. It is given by

$$H_{BHZ} = \begin{pmatrix} h(\boldsymbol{k}) & 0 \\ 0 & h^*(-\boldsymbol{k}) \end{pmatrix}, \quad h(\boldsymbol{k}) = E(\boldsymbol{k})\,\mathbb{1}_{2x2} + \boldsymbol{d}(\boldsymbol{k})\cdot\boldsymbol{\sigma} \tag{4.1}$$

with

$$\begin{aligned} E(\boldsymbol{k}) &= \mathcal{C} - \mathcal{D}(k_x^2 + k_y^2), \\ \boldsymbol{d}(\boldsymbol{k}) &= \begin{pmatrix} \mathcal{A}k_x, & \mathcal{A}k_y, & \mathcal{M}(\boldsymbol{k}) \end{pmatrix}^T, \\ \mathcal{M}(\boldsymbol{k}) &= M - \mathcal{B}(k_x^2 + k_y^2). \end{aligned} \tag{4.2}$$

Here, $\boldsymbol{\sigma}$ are the Pauli matrices associated with the band-pseudospin degree of freedom (band E_1 and H_1 in Hg(Cd)Te QWs), and the quadratic in $\boldsymbol{k}$ terms are taken to be overall positive, such that $\mathcal{B}, \mathcal{D} < 0$. The described system possesses $\boldsymbol{\mathcal{T}}$ symmetry and H_{BHZ} is block diagonal in the Kramers partners or spin degree of freedom. We thus restrict ourselves to the block $h(\boldsymbol{k})$, the results can then be extended to the other one by applying the $\boldsymbol{\mathcal{T}}$ operator. The model describes a 2D TI for the inversion condition $M/\mathcal{B} > 0$. The block structure is related to an assumed inversion ($\boldsymbol{\mathcal{I}}$) symmetry. If one considers the bulk inversion asymmetry (BIA) terms, introduced in Eq. (2.43), the two blocks get coupled.

$h(\boldsymbol{k})$ describes fermions with intermediate properties between a Dirac and a conventional 2DEG system. The off-diagonal term ($\mathcal{A}$ parameter) is typical for a Dirac system, e.g. $\mathcal{A} \hat{=} \hbar v_f$ in graphene. The Dirac mass M gives rise to a gap of $2|M|$. We consider positive and negative masses, where the latter one corresponds to an inversion of the band structure. The system is then topologically non-trivial (TnT) with a finite $\mathbb{Z}_2$ topological invariant of $\nu = 1$. This

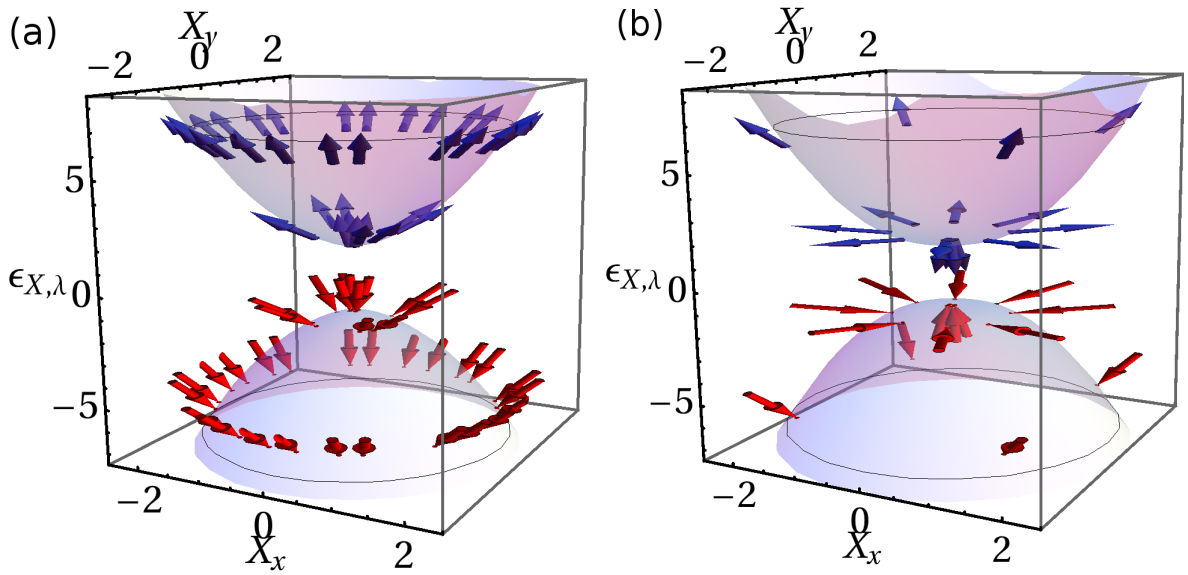


Figure 4.1. Dispersion relation and pseudospin of (a) a NI phase with $\xi_M = \frac{1}{2}$, and (b) a TI phase with $\xi_M = -\frac{1}{2}$. The bands are separated by an additional $2\epsilon_{X,\lambda}$ for better illustration of the pseudospin. Reprinted figures with permission from Ref. [JMT14a]. Copyright (2014) by the APS.

follows directly from the finite Hall conductance of the single block, given in Eq. (2.41). For simplicity, we restrict ourselves to band structures with extrema at the Γ-point, which limits the mass to $M > -\frac{1}{2}\frac{\mathcal{A}^2}{|\mathcal{B}|}$. This excludes Mexican-hat-shaped band structures.

In analogy to a 2DEG, the diagonal elements in Eq. (4.1) bear kinetic energy elements which preserve ($\mathcal{B}$ parameter) and break ($\mathcal{D}$ parameter) p-h symmetry. For Schrödinger fermions with the quasi-particle mass m this corresponds to $-\mathcal{B} \mp \mathcal{D} \mathrel{\widehat{=}} \frac{1}{2m}$.

The eigenstates of Eq. (4.1) are described by the following dispersion relation and pseudospin

$$E_{k,\lambda} = E(k) + \lambda \left| \boldsymbol{d}\left(\boldsymbol{k}\right) \right|, \quad \lambda \hat{\boldsymbol{d}}\left(\boldsymbol{k}\right) = \langle \boldsymbol{k}, \lambda | \boldsymbol{\sigma} | \boldsymbol{k}, \lambda \rangle \tag{4.3}$$

with $\lambda = \pm$ for conduction and valence band. Here $\hat{\boldsymbol{d}}$ is the direction of the $\boldsymbol{d}$ vector, illustrated in Fig. 4.1. Note that we consider electrons to be perfectly localized in the 2D x-y plane and therefore we neglect the real shapes of the envelope functions due to the quantum confinement along the z direction [MPER12].

4.1.1. Energy and Momentum Scales

Examining Eq. (4.1), one finds that the BHZ model is characterized by the intrinsic scales for momentum and energy

$$q_0 = \frac{\mathcal{A}}{|\mathcal{B}|}, \quad E_0 = \mathcal{A} q_0. \tag{4.4}$$

They reflect the interpolating character of the model between Dirac ($\mathcal{A}$ parameter) and Schrödinger ($\mathcal{B}$ parameter) system. The Fermi momentum k_f and chemical potential μ provide additional, externally tunable momentum and energy scales. They are called Fermi scales in the following.

We expect the ratio between Fermi and intrinsic scales to govern the physics of this system. Thus we define the dimensionless quantities

$$\boldsymbol{X} = \frac{\boldsymbol{q}}{q_0}, \quad \Omega = \frac{\omega}{E_0}, \quad X_f = \frac{k_f}{q_0}, \quad \Omega_f = \frac{\mu}{E_0}, \quad \xi_M = \frac{M}{E_0}, \quad \xi_{\mathcal{D}} = \frac{\mathcal{D}}{|\mathcal{B}|}, \tag{4.5}$$

where we set $\hbar = 1$ in the following. Ω_f is defined to be the energy to the wave vector X_f, such that $\Omega_f = \pm |\xi_M|$ if $X_f = 0$. The Hamiltonian in Eq. (4.1) can then be brought into the form

$$h(\boldsymbol{k}) = E_0 h(\boldsymbol{X}) = E_0 \begin{pmatrix} \xi_M + (1 - \xi_{\mathcal{D}}) \boldsymbol{X}^2 & X_- \\ X_+ & -(\xi_M + (1 + \xi_{\mathcal{D}}) \boldsymbol{X}^2) \end{pmatrix} \tag{4.6}$$

with $X_\pm = X_x \pm i X_y$ and $X = |\boldsymbol{X}|$. For $X \sim 1$, we therefore expect intermediate physics, while in the limit $X, \Omega \to 0$ ($X, \Omega \to \infty$) the Dirac (2DEG) physics should be recovered.

4.2. Polarization Function and Overlap Factor

The linear response of a homogeneous system to an externally applied potential is described by the density-density susceptibility or retarded polarization function $\Pi^R(\boldsymbol{q}, \omega)$. It is exactly the Lindhard function, Eq. (3.23), introduced in Sec. 3.2.2. This response comprises two main phenomena. The first one is screening, described by the real part $\Re\left[\Pi^R(\boldsymbol{q}, \omega)\right]$. The second one is dissipation by SPEs, contained in the imaginary part $\Im\left[\Pi^R(\boldsymbol{q}, \omega)\right]$.

The retarded polarization function (Lindhard function), written in the intrinsic scales, yields the expression

$$\Pi^R(X, \Omega) = \frac{g_s}{|\mathcal{B}|} \sum_{\lambda, \lambda'} \int \frac{\mathrm{d}^2 \tilde{X}}{4\pi^2} \mathcal{F}^{\lambda,\lambda'}_{\tilde{\boldsymbol{X}}, \tilde{\boldsymbol{X}}'} \frac{f(\epsilon_{\tilde{X},\lambda}) - f(\epsilon_{\tilde{X}',\lambda'})}{\Omega + i0^+ + \epsilon_{\tilde{X},\lambda} - \epsilon_{\tilde{X}',\lambda'}}. \tag{4.7}$$

Here we define $\tilde{\boldsymbol{X}}' = \tilde{\boldsymbol{X}} + \boldsymbol{X}$, 0^+ as a positive infinitesimal, $g_s = 2$ for spin degeneracy and the dispersion relation

$$\epsilon_{X,\lambda} = E_{q_0 X,\lambda}/E_0 = -\xi_{\mathcal{D}} X^2 + \lambda \sqrt{(\xi_M + X^2)^2 + X^2}. \tag{4.8}$$

The square root in Eq. (4.8) makes it necessary to calculate $\Pi^R(X, \Omega)$ numerically in most of the cases considered in the following. The Fermi-Dirac function $f(\epsilon) = \frac{1}{\mathrm{e}^{\bar{\beta}(\epsilon - \Omega_f)} + 1}$, with $\bar{\beta} = \frac{E_0}{k_B T}$ and k_B the Boltzmann constant, gives the occupation of the states. In the following we assume zero temperature, $T = 0$. Eq. (4.7) implies that $|\mathcal{B}|\,\Pi^R(X, \Omega)$ is only a function of the reduced dimensionless variables X and Ω. It further parametrically depends on ξ_M, $\xi_{\mathcal{D}}$ and X_f.

The overlap factor is given by

$$\mathcal{F}^{\lambda,\lambda'}_{\boldsymbol{X},\boldsymbol{X}'} = \left|\langle \boldsymbol{k}, \lambda | \boldsymbol{k}', \lambda' \rangle\right|^2 = \frac{1}{2}\left[1 + \lambda\lambda' \hat{\boldsymbol{d}}(q_0 \boldsymbol{X}) \cdot \hat{\boldsymbol{d}}(q_0 \boldsymbol{X}')\right]. \tag{4.9}$$

In the massless Dirac limit ($\mathcal{B} = \mathcal{D} = M = 0$), the eigenspinors of the Hamiltonian (4.1) are characterized by their helicity, see Eq. (2.36). Consequently the overlap factor $\mathcal{F}^{\lambda,\lambda'}_{\boldsymbol{k},\boldsymbol{k}'} = \frac{1}{2}(1 + \lambda\lambda' \cos\theta)$ only depends on the angle θ between $\boldsymbol{k}$ and $\boldsymbol{k}'$. It is strictly one (zero) for states with the same (opposite) helicity.

Next we consider the influence of the quadratic terms $\mathcal{B} < 0$. In the BHZ model, they have the effect of turning the pseudospin of the eigenstates out of plane at large X, in opposite directions for conduction and valence bands as depicted in Fig. 4.1. This results in a decay of the overlap factor (4.9) down to zero in the limit of a conventional 2DEG system ($\mathcal{A} \to 0$ or $X \to \infty$).

A finite mass $\xi_M \neq 0$ has a similar effect, but in the limit of small momenta $X \leq |\xi_M|$. Here, the pseudospin turns in the same direction as for the quadratic term for positive mass, and in the opposite direction for negative mass, see Fig. 4.1. In the latter case, this directly reveals the non-trivial winding of the pseudospin, resembling a skyrmion, as discussed in Sec. 3.2.3, in momentum space.

This turning of the pseudospin has direct consequences for the overlap factor. In the interband case, it is reduced for a NI phase, while for a TI phase it is increased. The winding in the TI phase ensures that valence and conduction band states with unitary overlap can always be found for finite X. Thus, the interband overlap factor is enhanced in the region $X \gtrsim |\xi_M|$ with respect to the NI phase. The intraband overlap factor shows the contrary effect. Here, a positive (negative) mass enhances (diminishes) the overlap. This picture is also confirmed in Sec. 4.5 by calculating the f-sum rule.

4.3. Coulomb Interaction

The bare 2D Coulomb interaction $v_q = \frac{e^2}{2\varepsilon_0 q}$ is modified in an electron gas by screening effects. This results in the effective interaction $v_q^{RPA}(q,\omega) = \frac{v(q)}{\varepsilon(q,\omega)}$ from Eq. (3.25). There, screening is described by the dynamical dielectric function in RPA approximation, where we omit the superscript RPA in the following. Employing dimensionless units, it acquires the form

$$\frac{\varepsilon(X,\Omega)}{\varepsilon_r} = 1 - \alpha g(X,\Omega), \tag{4.10}$$

where we have introduced the interaction strength parameter α as an effective Dirac fine-structure constant [KUP+12]. Together with the dimensionless function $g(X,\Omega)$ it is given by

$$\alpha = \frac{1}{\mathcal{A}} \frac{e^2}{4\pi\varepsilon_0\varepsilon_r}, \quad g(X,\Omega) = 2\pi \frac{|\mathcal{B}|}{X} \Pi^R(X,\Omega). \tag{4.11}$$

In graphene one finds [GPN12] $\alpha = 2.2/\varepsilon_r$, while in Hg(Cd)Te QWs it is of the order $\alpha \approx 4/\varepsilon_r$ [BLT+11, SNKT09]. Here, ε_r is the background dielectric constant, accounting for screening of internal electronic shells, while $-\alpha g(X,\Omega)$ gives the dynamic screening due to the electrons in the bands near the Fermi level. The background screening can be seen just as an additional screening source, and we added it to Eq. (3.26) in order to arrive at Eq. (4.10).

Zeros of $\varepsilon(X,\Omega)$ describe a density-density (longitudinal) perturbation of the system that is able to sustain itself. It forms a collective mode called plasmon, see Sec. 3.2.2, and is defined by

$$\varepsilon(X, \Omega_p - i\Gamma) = 0 \tag{4.12}$$

with the plasma frequency Ω_p. The finite imaginary part $\Gamma = \frac{\gamma}{E_0}$ accounts for possible damping due to SPEs.

The total dissipation in the interacting system, including both SPEs and the plasmon mode, is then described by the imaginary part of the interacting polarization function $\Pi^{RPA}(X,\Omega) =$

$\frac{\Pi^R(X,\Omega)}{\varepsilon(X,\Omega)}$. In order to compare to the non-interacting one, we will plot the normalized functions

$$\Pi^{\mathrm{Im}}_{\mathrm{rpa}} \equiv \varepsilon_r \Im\left[\Pi^{RPA}\right], \quad \Pi^{\mathrm{Im}} \equiv \Im\left[\Pi^{R}\right], \quad \Pi^{\mathrm{Re}} \equiv \Re\left[\Pi^{R}\right] \tag{4.13}$$

in the following, with $\varepsilon_r \Pi^{RPA} \underset{\alpha\to 0}{=} \Pi^R$. A directly related quantity is the loss function

$$\Im\left[-\frac{1}{\varepsilon}\right] = -2\pi \frac{\alpha\,|B|}{\varepsilon_r}\frac{1}{X}\Pi^{\mathrm{Im}}_{\mathrm{rpa}} = \frac{1}{\varepsilon_r \alpha}\frac{\Im\left[g\left(X,\Omega\right)\right]}{|\frac{1}{\alpha} - g\left(X,\Omega\right)|^2}, \tag{4.14}$$

which is linked to the dynamic structure factor $A\left(q,\omega\right)$ given in Eq. (3.30). As loss function and polarization function $\Pi^{\mathrm{Im}}_{\mathrm{rpa}}$ are directly proportional, they contain the same physics. The additional factor $1/X$ puts an additional emphasis on the long wavelength limit, which we will see in Sec. 5.3. While the polarization function has units of $1/\mathcal{B}$, the loss function is a dimensionless quantity.

4.3.1. (Anti-)Screening and Intrinsic Plasmons

In the RPA, the dielectric function, Eq. (4.10), characterizes the screening of the interaction between two electrons exchanging momentum X and energy Ω. This screening happens through the creation of electron-hole (e-h) pairs in the electron gas with the energy Ω_{eh} and same momentum X.

If these pairs are resonant in energy with $\Omega_{eh} = \Omega$, they correspond to a physical process leading to dissipation and a lowering of the Coulomb interaction. This is described by the imaginary part of the polarization function (4.7). If they are off-resonant, $\Omega_{eh} \neq \Omega$, we have only virtual e-h pairs. They either still screen the interaction, if $\Pi^{\mathrm{Re}} < 0$, or they enhance it (antiscreening effect), if $\Pi^{\mathrm{Re}} > 0$.

Thus these effects depend on the energy of the created e-h pair. For $\Omega_{eh} < \Omega$ one finds antiscreening, while $\Omega_{eh} > \Omega$ leads to screening of the bare Coulomb interaction. This can directly be seen from the definition of the polarization function. For every allowed (by the Fermi-Dirac functions) p-h excitation in the spectrum, the real part of the integrand in Eq. (4.7) can be rewritten into

$$\mathcal{F}^{1,\lambda}_{\tilde{\boldsymbol{X}},\tilde{\boldsymbol{X}}'}\frac{2\Omega_{eh}[\tilde{\boldsymbol{X}},\tilde{\boldsymbol{X}}']}{\Omega^2 - \Omega_{eh}[\tilde{\boldsymbol{X}},\tilde{\boldsymbol{X}}']^2}. \tag{4.15}$$

Here, $\lambda = 1$ ($\lambda = -1$) labels intraband (interband) excitations. Therefore every e-h process with energy less than Ω increases Π^{Re}. This lowers ε (remember the - sign in Eq. (4.10)) and thus increases the interaction. This is the antiscreening effect.

In the intrinsic, undoped Dirac system one finds $\Pi^{\mathrm{Re}} = 0$ for all energies Ω where e-h excitations are allowed, see Ref. [Sta14]. Thus the screening effect of virtual excitations with $\Omega_{eh} > \Omega$ cancels exactly with the one from excitations with $\Omega_{eh} < \Omega$. The only screening then comes, at least within RPA, from the resonant process $\Omega_{eh} = \Omega$.

In the BHZ model with finite $\mathcal{B} < 0$, the high-energy excitations become less likely compared to the graphene limit. On the one hand, this is due to the decoupling of the electron and the hole band for large Ω. Additionally, the excitation energy of the pairs is higher as in the Dirac case for the same momentum X. This provides an additional reduction of their influence on Π^{Re}. Furthermore, low-energy excitations become more important, as processes are allowed that where forbidden in Dirac systems by helicity. This is discussed in detail in Sec. 6.3.

Combining these effects, one finds the virtual excitations which increase the Coulomb interaction, $\Omega_{eh} < \Omega$, dominating for larger frequency Ω. As a result, the effective interaction is increased and there is the possibility of intrinsic plasmons in the BHZ model, see Sec. 5.3.

More mathematically speaking, the described effects alter the high-energy behavior of Π^{Im} from a decay like Ω^{-1} in the Dirac case to a Ω^{-2} decay in the BHZ model, as is shown in Sec. 5.2. Rewriting the Kramers-Kronig relation, Eq. (3.28), in intrinsic scales gives

$$\Pi^{\mathrm{Re}}(X,\Omega) = \frac{1}{\pi}\int_0^\infty \mathrm{d}\Omega' \frac{2\Omega'}{\Omega'^2-\Omega^2}\Pi^{\mathrm{Im}}(X,\Omega')\,. \tag{4.16}$$

Inserting the different high-energy behaviors, one finds that the real part of the polarization changes sign for $\Pi^{\mathrm{Im}} \propto \Omega^{-2}$, but not for $\Pi^{\mathrm{Im}} \propto \Omega^{-1}$.

In more general terms, one can expect intrinsic interband plasmons to appear in all models for which Π^{Im} decays faster as Ω^{-1} for high energies.

4.3.2. Static limit and Screening

The static limit of the polarization function is obtained by sending $\Omega \to 0$ at finite momentum X in Eq. (4.7). This way we can easily analyze the response of the system to the application of a static (or sufficiently slowly varying) external potential. An important physical problem of this kind is the screening of a charged impurity by the electronic system.

The static polarization is a strictly real function, that we define as

$$\Pi(X) \equiv \Pi^R(X,0) = \Pi_0(X) + \Pi_\mu(X). \tag{4.17}$$

In a multiband system, like the BHZ model, it is useful to separate the contributions to the static polarization coming from the intrinsic neutral system, $\Pi_0(X)$ obtained for $\mu = 0$, and the contribution due to a finite charge density, $\Pi_\mu(X)$ with finite μ. Correspondingly, the dielectric function (4.10) can be rearranged into

$$\varepsilon(X) \equiv \varepsilon(X,0) = \varepsilon_r\left[1 - \alpha g_0(X) - \alpha g_\mu(X)\right]. \tag{4.18}$$

We will use it in the following to calculate the induced charge density in response to a test charge Ze placed at the origin. The variation of the electronic charge density in momentum space is given by $Zen(X)$, where $n(X)$ is defined as

$$\begin{aligned} n(X) &= \frac{1}{\varepsilon(X)} - 1 = \frac{1}{\varepsilon_r[1-\alpha g(X)]} - 1 \\ &= n_r(X) + n_0(X) + n_\mu(X). \end{aligned} \tag{4.19}$$

This directly follows from Eq. (3.16) or can be taken from Ref. [FW71] Sec. 14.

The induced charge density can be seen as a sum of three contributions of different physical origin. The first one, $n_r(X)$, is due to the background polarization and also called the high-energy polarization of the system. It stems from the internal electronic shells. The second contribution is the intrinsic, undoped polarization $n_0(X)$ and the third, $n_\mu(X)$, comes from

the polarization of the finite charge density in the system. In total, they can be written as

$$n_r(X) = \frac{1}{\varepsilon_r} - 1, \tag{4.20}$$

$$n_0(X) = \frac{1}{\varepsilon_r} \frac{\alpha g_0(X)}{1 - \alpha g_0(X)}, \tag{4.21}$$

$$n_\mu(X) = \frac{1}{\varepsilon_r} \frac{1}{1 - \alpha g_0(X)} \frac{\alpha g_\mu(X)}{1 - \alpha g(X)}. \tag{4.22}$$

In real space, the density fluctuation, using physical dimensional units, is given by

$$n(r) = \frac{1}{2\pi} \int \mathrm{d}q\, q\, J_0(qr) n(q), \tag{4.23}$$

with J_0 the zero-th order Bessel function. It will be closer examined in the undoped limit in Sec. 5.1 and the doped one in Sec. 6.1. In the latter case, we will find Friedel oscillations with interesting properties in the intermediate regime between Dirac and Schrödinger.

4.4. Experimental Parameters

Including Coulomb interaction, we now have a 4-dimensional parameter space consisting of ξ_M, $\xi_\mathcal{D}$, X_f and α. This parameter space will be explored systematically in the following sections 5 and 6.

While the exploration of the different physical behaviors featured by the BHZ model in different regions of this parameter space has a clear theoretical significance, we want to stress that our discussion is also relevant for experiments. In particular, realistic parameters for Hg(Cd)Te QW structures, based on Refs. [BLT+11, SNKT09], are roughly

$$\xi_\mathcal{D} \leq -0.5, \quad q_0 \approx 0.4\ {}^1\!/_{\mathrm{nm}}, \quad E_0 \approx 140\ \mathrm{meV} \tag{4.24}$$

and masses M with absolute values up to several meV. The interaction strength is around $\alpha \approx 4/\varepsilon_r \approx 0.3$ with an average $\varepsilon_r = 15$ from the CdTe substrate ($\varepsilon_r = 10$) and HgTe ($\varepsilon_r = 20$).

Considering the experimental acceptable damping rate for plasmons, we refer to experiments on the surface states of a 3D TI in Ref. [DOL+13]. There, plasmons with a ratio of $\frac{\Gamma}{\Omega_p} = 0.5$ are perfectly resolvable.

4.5. The f-sum Rule

The f-sum rule for the polarization function, Eq. (3.29), provides the total spectral weight of all excitations in the system. It is identical for the interacting and noninteracting system, as the interaction conserves the number of particles, and therefore commutes with the electron density operator, see Eq. (4.25). Thus the sum rule is a powerful tool to check our numerics. Additionally, it offers a deeper insight concerning the shift of spectral weight between the inter and intra SPEs as well as the different plasmons in the system.

Definition and Calculation

Based on Ref. [NP66], the f-sum rule is defined by

$$f\sum(\boldsymbol{q}) = -\frac{2}{\pi}\int_0^\infty \mathrm{d}\omega\,\omega\Im\left[\Pi(\boldsymbol{q},\omega)\right] = g_s\left\langle 0\left|\left[\left[n_{\boldsymbol{q}},H^0\right],n_{\boldsymbol{q}}^\dagger\right]\right|0\right\rangle \tag{4.25}$$

with the density operator $n_{\boldsymbol{q}}^\dagger = \sum_{\boldsymbol{k}}\Psi_{\boldsymbol{k}+\boldsymbol{q}}^\dagger\Psi_{\boldsymbol{k}}$ and the Hamiltonian $H^0 = \sum_{\boldsymbol{k}}\Psi_{\boldsymbol{k}}^\dagger h(\boldsymbol{k})\Psi_{\boldsymbol{k}}$, with $h(\boldsymbol{k})$ as defined in Eq. (4.1). $\Psi_{\boldsymbol{k}}$ is a spinor associated with the band-pseudospin degree of freedom, corresponding to band E_1 and H_1 in Hg(Cd)Te QWs. The real spin degree of freedom enters via the degeneracy factor $g_s = 2$.

For the calculation we follow the steps outlined in the appendix of Ref. [SNC08], where the f-sum rule for the Dirac model is obtained. For the BHZ model, the computational steps are the same, therefore we only present important intermediate results, stressing the differences to the Dirac limit. The commutator in Eq. (4.25) is given by

$$\begin{aligned}\left[\left[n_{\boldsymbol{q}},H^0\right],n_{\boldsymbol{q}}^\dagger\right] = &\sum_{\boldsymbol{k}}\left(\Psi_{\boldsymbol{k}}^\dagger H_{\boldsymbol{k},\boldsymbol{q}}^0\Psi_{\boldsymbol{k}} - \Psi_{\boldsymbol{k}+\boldsymbol{q}}^\dagger H_{\boldsymbol{k}+\boldsymbol{q},\boldsymbol{q}}^0\Psi_{\boldsymbol{k}+\boldsymbol{q}}\right)\\ &- 2q^2\sum_{\boldsymbol{k}}\Psi_{\boldsymbol{k}+\boldsymbol{q}}^\dagger\left(\mathcal{D}\sigma_0+\mathcal{B}\sigma_z\right)\Psi_{\boldsymbol{k}+\boldsymbol{q}},\end{aligned} \tag{4.26}$$

where $H_{\boldsymbol{k},\boldsymbol{q}}^0 = \mathcal{A}\boldsymbol{q}\cdot\boldsymbol{\sigma} - \mathcal{D}\boldsymbol{q}\left(2\boldsymbol{k}+\boldsymbol{q}\right)\sigma_0 - \mathcal{B}\boldsymbol{q}\left(2\boldsymbol{k}+\boldsymbol{q}\right)\sigma_z$. A simple shift of the momentum sums in Eq. (4.26) would put the first line to zero, but this is not allowed. In the same way as in the Dirac system, the operators are unbounded and one has to work with a large momentum cutoff κ. The momentum shift then results in different cutoffs for the sums. While in the Dirac limit, one simply finds $H_{\boldsymbol{k},\boldsymbol{q}}^0 = \mathcal{A}\boldsymbol{q}\cdot\boldsymbol{\sigma}$ and the second line of Eq. (4.26) would be zero, now the latter gives rise to a contribution depending on the chemical potential, as expected for a 2DEG.

The sums in Eq. (4.26) are then converted into integrals and explicitly calculated in the limit of large κ. Care has to been taken when converting the momentum cutoff κ into the frequency cutoff λ, such that both integrals cover the same phase space.

Formulas

For a pure Dirac system, the authors of Ref. [SNC08] find the f-sum rule

$$\int_0^\lambda \mathrm{d}\omega\,\omega\Im\left[\Pi(q,\omega)\right] = -\frac{g_s q^2\lambda}{16}, \tag{4.27}$$

where the cutoff λ is needed as the Dirac spectrum is unbounded. In a 2DEG system, the f-sum rule is given by

$$\int_0^\infty \mathrm{d}\omega\,\omega\Im\left[\Pi(q,\omega)\right] = \frac{g_s}{4}\left(\mathcal{B}\pm\mathcal{D}\right)k_f^2q^2 = -\frac{\pi N q^2}{2m} \tag{4.28}$$

with $N = \frac{g_s}{4\pi}k_f^2$ the electron density. As $\Im\left[\Pi(q,\omega)\right] \neq 0$ only over a finite range of ω, the integral has a natural cutoff.

Similar to a Dirac system, the BHZ spectrum is unbounded which complicates the evaluation of the sum rule and makes it necessary to introduce a high-energy cutoff $\Lambda = \frac{\lambda}{E_0}$. We find

approximately for large cutoff $\Lambda \gg 1$

$$\begin{aligned} f\sum &\equiv -\int_0^\Lambda \mathrm{d}\Omega\, \Omega\, |\mathcal{B}|\, \Im\left[\Pi\left(X,\Omega\right)\right] \\ &= \frac{g_s}{8} X^2 \Bigg[-1 - 2\xi_M + 2\left|\Omega_f\right| + \ln\left(\frac{2\Lambda}{1 + 2X_f^2\left(1+\eta\xi_{\mathcal{D}}\right) + 2\xi_M + 2\left|\Omega_f\right|}\right) \\ &+ \frac{1 - X^2 + 4\xi_M}{\Lambda} - \frac{2X^4 + \left(1+4\xi_M\right)^2 - 4X^2\left(2+7\xi_M\right)}{4\Lambda^2}\Bigg] + \mathcal{O}\left(\frac{\xi_{\mathcal{D}}}{\Lambda^2}\right) + \mathcal{O}\left(\frac{1}{\Lambda^3}\right) \end{aligned} \tag{4.29}$$

with $\eta = \mathrm{sgn}\left[\Omega_f\right]$. Thus the leading order term diverges logarithmically with the cutoff Λ. This is due to the fact that $\Im\left[\Pi\left(X,\Omega\right)\right]$ decays like Ω^{-2} for $\Omega \gg 1$, and not as Ω^{-1} as for a Dirac system. The sum rule is exact up to order Λ^{-1} (Λ^{-2}) for a finite (zero) p-h symmetry breaking term $\xi_{\mathcal{D}}$.

The f-sum rules for BHZ, Dirac and 2DEG models are always proportional to $q^2 \propto X^2$ in leading order, but otherwise distinct from one another. Taking the limit $\mathcal{A} \to 0$ in the BHZ result, Eq. (4.29), gives the 2DEG case, Eq. (4.28). The same is not possible for the Dirac limit $\mathcal{B} \to 0$, due to the details of the derivation of the analytical expansion in Eq. (4.29). There, the chosen cutoff $\Lambda = \lambda\frac{|\mathcal{B}|}{\mathcal{A}^2}$ would go to zero.

Weight of the Contributions from different Orders $\mathcal{O}$ in Cutoff Λ

We begin our discussion of the f-sum rule (4.29) by comparing the contributions from the different orders $\mathcal{O}\left(\ln\left(\Lambda\right)\right)$ (where we also include the Λ independent terms), $\mathcal{O}\left(\Lambda^{-1}\right)$ and $\mathcal{O}\left(\Lambda^{-2}\right)$. Our aim is to find for the cutoff the right size and functional dependence on X, such that higher order terms in Eq. (4.29) can be neglected. Of course one could just take the limit $\Lambda \to \infty$, but that would not work out well with our goal of comparing the analytical result Eq. (4.29) to numerical data of $\Pi\left(X,\Omega\right)$. The data will be given only over a finite range of Ω, for cutoffs exceeding this frequency an extrapolation of the data is necessary. This provides a source for errors.

In the limit of $\xi_M = \xi_{\mathcal{D}} = X_f = 0$ we find

$$\frac{f\sum_{\mathcal{O}(\Lambda^{-1})}}{f\sum_{\mathcal{O}(\ln(\Lambda))}} = -\frac{X^2 - 1}{\Lambda\left(\ln\left(2\Lambda\right) - 1\right)}, \tag{4.30}$$

$$\frac{f\sum_{\mathcal{O}(\Lambda^{-2})}}{f\sum_{\mathcal{O}(\ln(\Lambda))}} = -\frac{2X^4 - 8X^2 + 1}{4\Lambda^2\left(\ln\left(2\Lambda\right) - 1\right)}. \tag{4.31}$$

Therefore, the ratio $\frac{X^2}{\Lambda}$ determines the importance of higher order corrections for $X \gg 1$. This advises us to take

$$\Lambda = 2\left(\beta X_{max}\right)^2 \tag{4.32}$$

for the cutoff in the following. Already for $\beta = 2$ and a maximal momentum $X_{max} = 6$, the ratio (4.30) is about 2%, and the ratio (4.31) drops down to 0.1%.

We conclude that a modest cutoff $\Lambda = 2\left(\beta X_{max}\right)^2$ with $2 \leq \beta \leq 5$ works best for comparing Eq. (4.29) to numerical data. It is the best compromise between analytical and numerical accuracy.

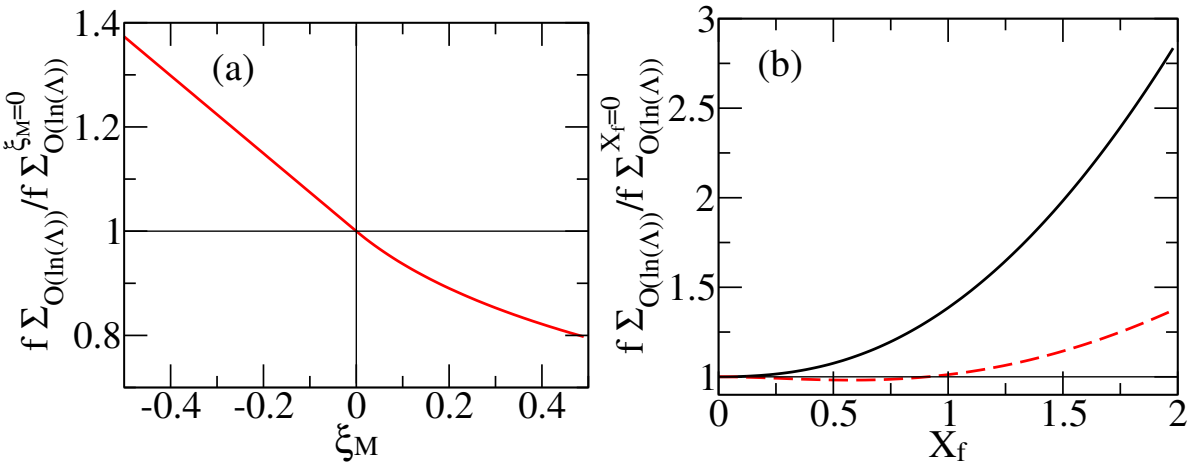


Figure 4.2. (a) Ratio $\frac{f\sum_{\mathcal{O}(\ln(\Lambda))}}{f\sum_{\mathcal{O}(\ln(\Lambda))}^{\xi_M=0}}$ of the lowest order in Λ contributions to the f-sum rule including mass, over the ones without mass. $\xi_{\mathcal{D}} = 0$ and $X_f = 0$. (b) Ratio $\frac{f\sum_{\mathcal{O}(\ln(\Lambda))}}{f\sum_{\mathcal{O}(\ln(\Lambda))}^{X_f=0}}$ of the lowest order in Λ contributions to the f-sum rule including finite doping, over the ones without doping. Electron doped, $\Omega_f > 0$, as a black, solid line. Hole doped, $\Omega_f < 0$, as a red, dashed line. $\xi_M = 0$ and finite p-h asymmetry $\xi_{\mathcal{D}} = -0.5$. In both plots we set $\beta = 2$ and $X_{max} = 6$, see Eq. (4.32). Reprinted figures with permission from Ref. [JMT14b]. Copyright (2014) by the APS.

Influence of finite Mass ξ_M, particle-hole Asymmetry $\xi_{\mathcal{D}}$ and Doping X_f

Next, we investigate changes to the f-sum rule and therefore to the total spectral weight by varying the mass. In contrast to graphene, see Ref. [SNC08], a mass term directly enters the f-sum rule (4.29) in leading order. Thus the influence of a finite gap in the system should be significant.

We study this in Fig. 4.2 (a) for the undoped, $X_f = 0$, p-h symmetric, $\xi_{\mathcal{D}} = 0$, case. Evidently, a positive mass lowers the f-sum rule, while a negative mass increases it linearly. This is a direct consequence from the change of the overlap factor discussed in Sec. 4.2. A negative mass enhances the coupling between the two bands, while a positive mass diminishes it, as in the latter case the pseudospins do not match. This is also consistent with the increase in the optical conductivity observed in the undoped limit with negative mass, see Sec. 5.4 below.

Finally, we consider the effects of finite doping. Again like for a finite mass, doping enters the f-sum rule (4.29) directly in the leading order term, while for graphene doping has no influence on the f-sum rule, see Ref. [SNC08]. Physically there are two effects competing here. On the one hand, a finite doping level blocks interband transitions close to the Dirac point. Yet due to the small density of states, these transitions carry only a small spectral weight. On the other hand, doping enables intraband transitions, which carry a large spectral weight due the combined effects of larger overlap factor, density of states and smaller excitation energies compared to interband transitions.

Therefore, a finite doping usually increases the f-sum rule, as seen in Fig. 4.2 (b), where we plot the ratio $\frac{f\sum_{\mathcal{O}(\ln(\Lambda))}}{f\sum_{\mathcal{O}(\ln(\Lambda))}^{X_f=0}}$ for positive (black, solid line) and negative (red, dashed line) doping with $\xi_{\mathcal{D}} = -0.5$ and $\xi_M = 0$. A finite p-h asymmetry $\xi_{\mathcal{D}} < 0$ adds a term $\mp\frac{g_s}{4}\xi_{\mathcal{D}} X_f^2 X^2$ to

the leading order of the f-sum rule, + (−) for positive (negative) doping. It can be seen as an increased (decreased) contribution from the 2DEG part of the spectrum, Eq. (4.28), and leads to the slight decrease of the f-sum rule for negative doping in panel (b) of Fig. 4.2.

4.5.1. Judging and Comparing the Spectral Weight of Excitations

In order to compare the importance of different excitations in the system, one should compare their spectral weight and thus their contribution to the f-sum rule. The latter has the benefit of being independent of the Coulomb interaction strength and the position of the excitation peaks, in contrast to the polarization function $\Pi^{\mathrm{Im}}_{\mathrm{rpa}}$.

As an example, we assume that the excitation spectrum, $\Pi^{\mathrm{Im}}_{\mathrm{rpa}}$, is governed by a single plasmonic peak, following a Lorentzian shape with width Γ and peak height $\frac{1}{c\Gamma}$. Here c is a normalization constant that we want to determine below. The f-sum rule is then proportional to

$$\int_0^\infty \mathrm{d}\Omega\, \Omega \frac{1}{c} \frac{\Gamma}{\Gamma^2 + (\Omega - \Omega_p)^2} = \frac{\Omega_p}{c} \int_0^\infty \frac{\mathrm{d}\Omega}{\Omega_p} \frac{\Omega}{\Omega_p} \frac{\frac{\Gamma}{\Omega_p}}{(\frac{\Gamma}{\Omega_p})^2 + (\frac{\Omega}{\Omega_p} - 1)^2}. \tag{4.33}$$

The value of this integral should be independent of α and thus of the peak position Ω_p. This is fulfilled by setting $c \propto \Omega_p$, such that the peak height of a resonance in $\Pi^{\mathrm{Im}}_{\mathrm{rpa}}$ naturally has to scale with $1/\Omega_p$ to keep the f-sum rule constant. Therefore peaks at higher energy will appear less pronounced in $\Pi^{\mathrm{Im}}_{\mathrm{rpa}}$ as the ones at lower energies, although they carry the same spectral weight.

We conclude that the importance of a resonance in $\Pi^{\mathrm{Im}}_{\mathrm{rpa}}$ should be judged by its spectral weight. The latter can be estimated by multiplying the peak height with its position Ω_p. The relevant width of the peak is given by $\frac{\Gamma}{\Omega_p}$, with Γ being the width of the resonance in $\Pi^{\mathrm{Im}}_{\mathrm{rpa}}$:

$$\begin{aligned} \text{Peak height} &: \frac{1}{\Gamma} \to \frac{\Omega_p}{\Gamma}, \\ \text{Peak width} &: \Gamma \to \frac{\Gamma}{\Omega_p}. \end{aligned} \tag{4.34}$$

5. Undoped System

Our focus in this section is on the intrinsic, undoped limit ($\mu = 0$) of the BHZ model. Therefore the Fermi level is at the Dirac point or in the middle of the gap, and only interband transitions can occur in the limit of zero temperature. The most interesting result in this section will be a new kind of interband plasmons, which are not present in Dirac or 2DEG systems. Only the merging of the two makes these new collective excitations possible. Another interesting aspect will be the optical conductivity. We propose it as a bulk quantity that differs strongly in the topological trivial and non-trivial phase. In the latter, it can even diverge close to the band gap.

This section is structured as follows. First, we analyze the static polarization function and the screening properties in Sec. 5.1. Special emphasis is put on the BHZ specific features of the screening compared to the Dirac and 2DEG limit. Afterwards, we consider the long wavelength limit of the dynamical polarization function in Sec. 5.2, providing an analytical expansion. The interacting excitation spectrum, and especially the new interband plasmons, are explored in Sec. 5.3. The optical conductivity as a bulk quantity that is topology dependent is analyzed in Sec. 5.4. We close with a short check of our numerics based on the f-sum rule in Sec. 5.5.

5.1. Static Limit and Screening

We begin our analysis of the undoped BHZ system with a focus on the static limit, $\Omega \to 0$. The physics described by the RPA formalism in this case is the (static) screening of the Coulomb interaction by the electron gas. We study the polarization function and induced charge density and find a resemblance to the (massive) Dirac system for small momenta. Yet on scales $X \gtrsim 1$ the physics changes qualitatively, crossing over into the 2DEG limit.

In order to set a reference with a closely related and analytically solvable model, we discuss the static intrinsic polarization of a massive Dirac system. It is given by [KPU08, Pya09]

$$\Pi_0(q) = -\frac{gq}{8\pi\mathcal{A}}\left[\frac{2M}{\mathcal{A}q} + \left(1 - \frac{4M^2}{\mathcal{A}^2q^2}\right)\arctan\frac{\mathcal{A}q}{2M}\right] \overset{M\to 0}{\longrightarrow} -\frac{gq}{16\mathcal{A}}, \tag{5.1}$$

where the index 0 stands for the intrinsic limit $\mu = 0$ and $g = g_s g_v$ accounts for possible spin and band degeneracy. In graphene one has $g = 4$. In the limit of a massless Dirac system with $M = 0$, $\Pi_0(q)$ is a linear function of the momentum q. A finite Dirac mass suppresses the polarization for $q \lesssim M/\mathcal{A}$, where $\Pi_0(q)$ shows a super-linear behavior. In the opposite limit of $q \gg M/\mathcal{A}$, the mass is negligible instead and the result of the massless limit is reproduced.

The static polarization function of the BHZ model is simply obtained by direct numerical evaluation of Eq. (4.7) at zero frequency. In Fig. 5.1, we show $\Pi_0(X)$ calculated for a p-h symmetric BHZ system with $\xi_{\mathcal{D}} = 0$. Note that we obtain the massless Dirac case in the limit $\mathcal{B} \to 0$ and therefore $X \to 0$, where $\lim_{X\to 0}|\mathcal{B}|\Pi_0(X)/X = -\frac{g_s}{16}$. A finite $\mathcal{B}$ parameter determines a fundamental qualitative change with respect to the Dirac system. It is responsible for $\Pi_0(X)$

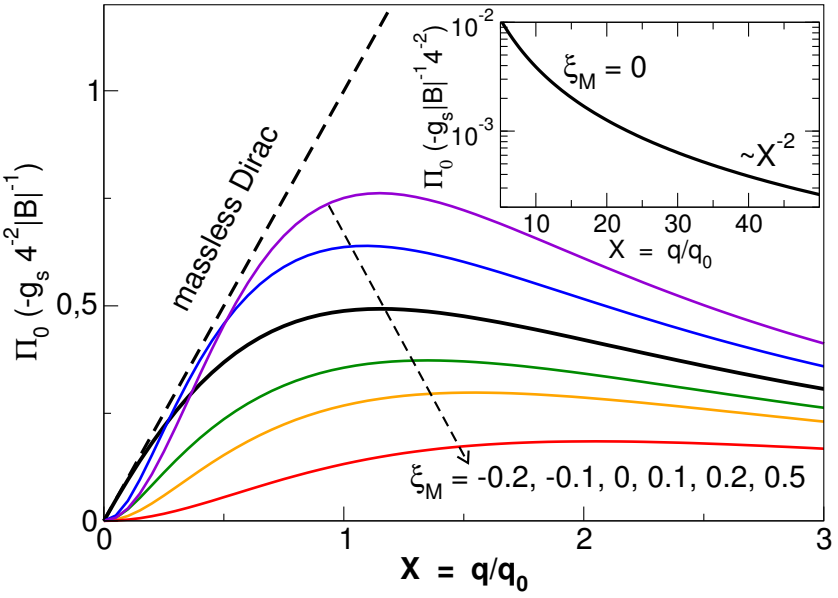


Figure 5.1. Static intrinsic polarization function of the BHZ model for different ξ_M and $\xi_{\mathcal{D}} = 0$. Reprinted figure with permission from Ref. [JMT14b]. Copyright (2014) by the APS.

reaching a maximum at $X \approx 1$ and then decaying as $1/X^2$ for $X \gg 1$, as shown in the inset of Fig. 5.1.

A finite and positive Dirac mass M leads to a general suppression of the polarization function with respect to the massless case. In the region $X < 1$, where quadratic terms are less important, $\Pi_0(X)$ resembles the massive Dirac case, with a super-linear increase in the region $X \lesssim \xi_M$. This is due to the suppression of the interband overlap factor discussed in Sec. 4.2, determining a reduction of the polarization at small momentum. For intermediate values $\xi_M \lesssim X \lesssim 1$, analogously to the massive Dirac limit, $\Pi_0(X)$ is approximatively linear in X. Considering larger momenta $X \gtrsim 1$, the behavior is dominated by the quadratic terms and the polarization eventually vanishes for $X \to \infty$. In general, the interplay of quadratic terms and a finite Dirac mass shifts the maximum of $\Pi_0(X)$ as shown in Fig. 5.1.

Taking into account a negative Dirac mass $M < 0$, corresponding to the TI phase, we observe a less pronounced suppression of the polarization for $X < \xi_M$ with respect to a massive Dirac system in the NI phase with equal modulus of M. Moreover, $\Pi_0(X)$ is enhanced at large X with respect to the massless, p-h symmetric limit, or the $M > 0$ case. This behavior is due to the enhanced overlap factor between electron and hole bands in the TI phase.

In Fig. 5.2 (a)-(c), we analyze the effects of a finite value of the parameter $\xi_{\mathcal{D}}$ in the BHZ model, in combination with different masses $\xi_M = -0.2$, 0 and 0.2. A finite $\xi_{\mathcal{D}}$ breaks p-h symmetry by changing the effective masses of conduction and valence bands. We only find minor quantitative changes to $\Pi_0(X)$, which is progressively increased for increasing $|\xi_{\mathcal{D}}|$.

Based on this analysis of $\Pi_0(q)$, we proceed to study screening in the BHZ model. Again, we discuss the massless Dirac limit first. There, the static polarization, Eq. (5.1), is linear in q, resulting in a constant dielectric function

$$\varepsilon(q) = \varepsilon_r \left(1 + \frac{g_s g_v \pi}{8} \alpha\right) \equiv \varepsilon. \tag{5.2}$$

This is an interesting result, telling us that the contribution coming from the intrinsic polar-

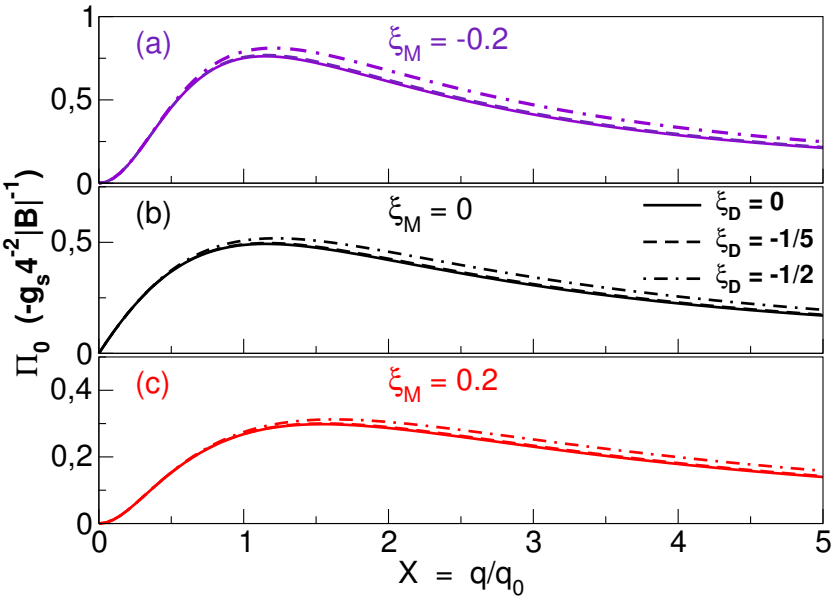


Figure 5.2. Static intrinsic polarization function of the BHZ model for a finite p-h asymmetry term $\xi_{\mathcal{D}}$. Reprinted figure with permission from Ref. [JMT14b]. Copyright (2014) by the APS.

ization can be absorbed into an effective background dielectric constant ε. As a consequence, a test charge Ze placed at the origin, see Sec. 4.3.2, induces a screening electronic density of

$$Ze\left[n_0(\boldsymbol{q}) + n_r\right] = Ze\left(\frac{1-\varepsilon}{\varepsilon}\right). \tag{5.3}$$

In real space, this corresponds to a screening image charge

$$n(\boldsymbol{r}) = \left(\frac{1-\varepsilon}{\varepsilon}\right)\delta(\boldsymbol{r}) \tag{5.4}$$

placed exactly at the same position as the external one. Note that the screening charge only due to the low-energy electronic system, without background contribution, is a fraction $-(\varepsilon - \varepsilon_r)/\varepsilon\varepsilon_r$ of the external one.

We move on to finite mass. In a massive Dirac system, the large q behavior of $\Pi_0(q)$ reproduces the massless limit. Therefore a screening charge given by Eq. (5.4) is also developed at vanishing distances r ($q \to \infty$) in response to an external test charge. However, in the long wavelength limit $q < M/\mathcal{A}$, $\Pi_0(q)$ has a superlinear behavior and thus $n_0(0) \propto \lim_{q\to 0} \Pi_0(q)/q = 0$. Thus far away from the test charge, one does not feel any induced screening charge arising from the polarization function, just the part from the background screening. This is due to an induced charge density of the same sign as the external charge, developed at finite distances and canceling the one at the origin. See Ref. [KPU08] for details. The canceling charge exactly sums up to $Ze(\varepsilon - \varepsilon_r)/\varepsilon\varepsilon_r$, so that the test charge feels only the background screening over long distances, as expected in an insulator.

For the BHZ model, we find similar to Eq. (5.2)

$$\lim_{X\to 0} \varepsilon(X) = \varepsilon_r\left(1 + \frac{g_s\pi}{8}\alpha\right) \tag{5.5}$$

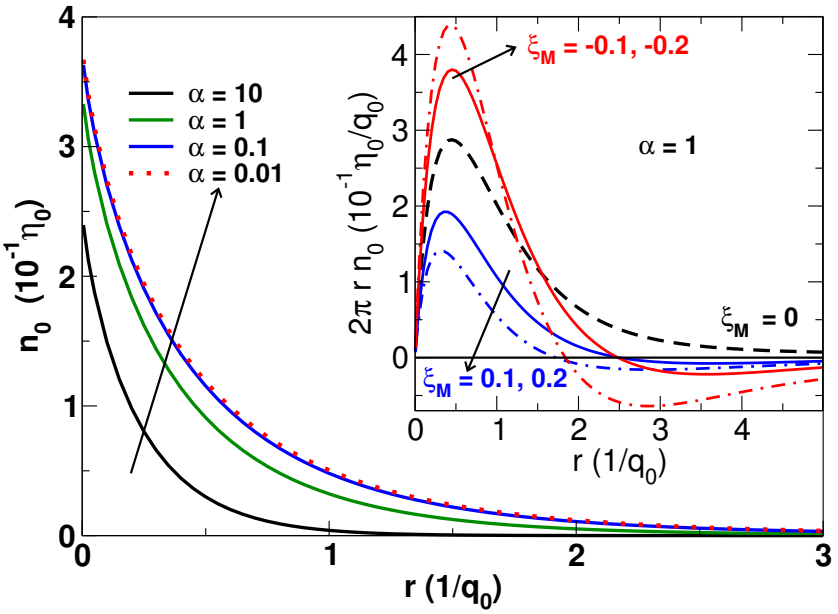


Figure 5.3. Induced charge $n_0(r)$ due to a test charge in the intrinsic limit of the BHZ model for $\xi_{\mathcal{D}} = \xi_M = 0$. The plot is invariant under a change of the $\mathcal{B}$ parameter and only depend on the effective fine structure constant α. In the inset, $r\, n_0(r)$ is calculated for $\alpha = 1$ and finite Dirac mass. Reprinted figure with permission from Ref. [JMT14b]. Copyright (2014) by the APS.

in the long wavelength limit, but $\lim_{X\to\infty} \varepsilon(X) = \varepsilon_r$. We can understand this by discussing the induced charge density in real space, Eq. (4.23), for the BHZ model. It is given by

$$n_0(r) = \eta_0 \int \mathrm{d}X\, J_0(Xrq_0)\, \frac{|\mathcal{B}|\, \Pi_0(X)}{1 - \alpha g_0(X)} \tag{5.6}$$

with $\eta_0 = \frac{\alpha}{\varepsilon_r} q_0^2$ a natural charge density constant of the model. We note that $n_0(r)$ is proportional to q_0^2 and α, but $n_0(r)$ has an additional dependence on α through its integrand. It also parametrically depends on ξ_M and $\xi_{\mathcal{D}}$ through $\Pi_0(X)$ and $g_0(X)$. In Fig. 5.3, we plot the induced charge density $n_0(r)$ in real space for $\xi_{\mathcal{D}} = \xi_M = 0$ with different interaction strengths α. Opposite to a Dirac system, the induced charge density has a finite extent over a distance of the order of $1/q_0$. This is clearly related to the $1/X^2$ decay of Π_0 at large wave vectors, caused by the presence of quadratic $\mathcal{B}$ terms. $n_0(r)$ decays at large distances as r^{-2}. Thus an electron far away from this induced charge, $r \gg 1/q_0$, does not see the finite extent of it and therefore experiences the same screening as in the Dirac system. This explains the similarity of Eqs. (5.2) and (5.5). In the opposite limit, where the electron sits on top of the induced charge, $r \ll 1/q_0$, it does not feel the screening cloud at all, resulting in no screening besides ε_r.

In the inset of Fig. 5.3, we study the effect of a mass term. With a finite ξ_M, the induced density, as in the case of pure Dirac systems, shows a qualitatively different behavior. $n_0(r)$ changes sign for sufficiently large r, ensuring a vanishing total induced charge. From a quantitative point of view, a finite negative (positive) M enhances (suppresses) the features of $n_0(r)$, due to its effect on the interband overlap factor.

5.2. Long Wavelength Expansion of $\Pi^R(X,\Omega)$ and Plasmon Dispersion

We proceed to finite frequencies $\Omega > 0$. In general, the polarization function (4.7), and therefore all related quantities in RPA approximation, can only be calculated numerically. This is done in the following section 5.3. Here, we look at the limit $X \ll 1$, where an analytic discussion of the polarization function is possible. As we want to extract an analytical formula of the plasmon dispersion, we focus on the limit of vanishing mass $\xi_M \to 0$.

An expansion of $\Pi^R(X,\Omega)$ in X for $\Omega > X$ gives

$$\begin{aligned}\Pi^R &= -\frac{g_s}{8|\mathcal{B}|}X^2\left[\frac{2}{\pi}\frac{\Omega-\sqrt{1+\Omega^2}\,\mathrm{ArcSinh}(\Omega)}{\Omega^3}+i\frac{1}{\Omega(1+\sqrt{1+\Omega^2})}\right]+\mathcal{O}\left(X^4\right)\\ &\underset{\Omega\gg1}{=} -\frac{g_s}{8|\mathcal{B}|}\frac{X^2}{\Omega^2}\left[\frac{2}{\pi}\left(1-\ln(2)-\ln(\Omega)\right)+i\right]+\mathcal{O}\left(\Omega^{-3}\right)+\mathcal{O}\left(X^4\right).\end{aligned} \tag{5.7}$$

One finds in the high frequency limit an Ω^{-2} behavior with an additional logarithmic correction for the real part. Calculating the plasmon dispersion is possible by performing an expansion of the plasmon equation (4.12) as described in Sec. 3.2.2 after Eq. (3.31). Including contributions up to second order in Γ/Ω, we find in the small frequency regime the linear dispersion

$$\Omega_p = \frac{1}{8}\pi\left(g_s\alpha X+\frac{1}{6}g_s^2\alpha^2X^2\right)+\mathcal{O}\left(X^3\right). \tag{5.8}$$

It is only valid for sufficiently large α, such that the conditions $\Re\left[\Pi^R(X,\Omega_p)\right] > 0$ (anti-screening) and $\Omega > X$ are fulfilled.

The linearity of the dispersion follows from Eq. (5.7) only by inclusion of the damping via a finite Γ. Without the substitution $\Omega \to \Omega - i\Gamma$, $\Re\left[\varepsilon(X,\Omega_p)\right] = 0$ has no sensible solution for Ω_p. The damping ratio in the small frequency limit is given by

$$\frac{\Gamma}{\Omega}\underset{\Omega\ll1}{=}1-\frac{\Omega^2}{8}+\mathcal{O}\left(\Omega^3\right) \tag{5.9}$$

underlining the importance of damping in this case. Thus the plasmon is considerably modified by interacting with the SPE background.

The plasmon is only well defined for a finite $\Omega_p > \Omega_c$, such that the damping ratio (5.9) has an upper bound c with $\frac{\Gamma}{\Omega}\Big|_{\Omega=\Omega_c} \lesssim c$. Here, $0 < c < 1$ sets the limit for the detectability of the plasmons. In recent experiments on the surface of the TI Bi_2Se_3, reported in Ref. [DOL+13], c was shown to be of the order 0.5. We insert Eq. (5.8) into Eq. (5.9) and thus translate this damping restriction into a finite momentum scale $q > \frac{q_0}{g_s\alpha}\frac{8\Omega_c}{\pi}$. We identify the intrinsic plasmon length scale

$$l_0 = g_s\frac{\alpha}{q_0} \tag{5.10}$$

given by the Coulomb interaction strength times the charge decay length $\frac{1}{q_0}$, see Fig. 5.3. We interpret l_0 as the length scale up to which charge separation due to Coulomb interaction can occur and give rise to the interband plasmons, in an undoped and therefore overall neutral system.

In the opposite limit of high frequencies, the term $\ln(\Omega)$ in Eq. (5.7) spoils a simple $\sqrt{X}$ behavior of the plasmon dispersion. We can extract the analytical form of the damping rate

$$\begin{aligned}\frac{\Gamma}{\Omega} &\underset{\Omega\gg 1}{=} \frac{1}{3\pi}\left[3-\ln(4\Omega^2)+\sqrt{3\pi^2+\left(\ln\left(4\Omega^2\right)-3\right)^2}\right]+\mathcal{O}\left(\frac{1}{\Omega}\right)\\ &\underset{\Omega\to\infty}{=} \frac{\pi}{2\ln\left(4\Omega^2\right)}+\frac{3\pi}{2\ln\left(4\Omega^2\right)^2}+\mathcal{O}\left(\frac{1}{\ln\left(4\Omega^2\right)^3}\right)+\mathcal{O}\left(\frac{1}{\Omega}\right),\end{aligned} \tag{5.11}$$

yet the plasmon dispersion can only be calculated numerically.

In the following discussion of the different excitation spectra, we will use these analytic results to check our numerics in the limits of small momenta and low and high frequencies.

5.3. Excitation Spectrum

Now we have a closer look at the excitation spectrum of the undoped BHZ model. We begin with the non-interacting one, given by Π^{Im}, and proceed from there by looking for plasmons in the system. We consider different parameter regimes, like finite mass and broken p-h symmetry. We conclude that the new interband plasmons we find should be experimentally observable.

The imaginary part of the polarization function, Π^{Im}, is calculated from Eq. (4.7) by employing the relation $\Im\left[\frac{1}{\omega+i\mathbf{0}^+}\right]=-\pi\delta(\omega)$. The real part is then obtained by the Kramers-Kronig relation (4.16). In Fig. 5.4 (a) and (b) we plot $\Pi^R(X,\Omega)$ for $\xi_M=\xi_{\mathcal{D}}=0$. The imaginary part in panel (a) is strictly zero below the cut-off frequency Ω_{min}, where electronic excitations are forbidden by energy-momentum conservation. In the pure Dirac limit $X,\Omega\to 0$, the cut-off frequency is $\Omega_{min}=X$ where a divergent behavior is observed, as discussed in Ref. [GFM08]. The reason behind this is the linear spectrum that allows for the existence of a divergent number of p-h excitations satisfying the energy-momentum conditions. In the opposite limit of $X,\Omega\to\infty$, the polarization function goes to zero due to the vanishing overlap factor, as expected in the 2DEG limit. In this high frequency limit, the imaginary part decays as Ω^{-2}, while in a pure Dirac system it shows a Ω^{-1} decay.

In Fig. 5.4 (b) we consider the real part of Π^R. For graphene within RPA, it is always negative. This is fundamentally different in the BHZ model, where Π^{Re} changes sign and becomes positive in the region above Ω_{min} indicated by the green dashed line. In this antiscreening region, we can search for solutions to Eq. (4.12), describing plasmonic resonances. Yet, Landau damping of the plasmon mode by SPE processes can be expected due to the finite value of Π^{Im} in the same region.

Next we take into account the effects of broken p-h symmetry. In Fig. 5.4 (c), we plot line cuts of Π^{Im} for fixed $X=0.5$ with and without a finite $\xi_{\mathcal{D}}$. In the latter case, the polarization strongly changes for small X, as Π^{Im} goes to zero at Ω_{min}, instead of exhibiting the divergency known from the pure Dirac system. Due to the breaking of p-h symmetry, low-energy electrons from the conduction band are excited close to the Dirac point, where the density of states goes to zero. Importantly, this reduction of Π^{Im} greatly diminishes the Landau damping of plasmons. Therefore we will pay it special attention in Sec. 5.3.2.

Lastly, in Fig. 5.4 (d) line cuts of Π^{Re} are shown for fixed $X=2$ and different Dirac masses. One nicely sees that a negative mass enhances and a positive mass diminishes the features of the polarization function with respect to the $M=0$ case. Again, this is caused by the changes in the overlap factor as elaborated in Sec. 4.2. Therefore, in a TnT phase the antiscreening effect, $\Pi^{\mathrm{Re}}>0$, gets enhanced, which increases the chance of observing plasmons.

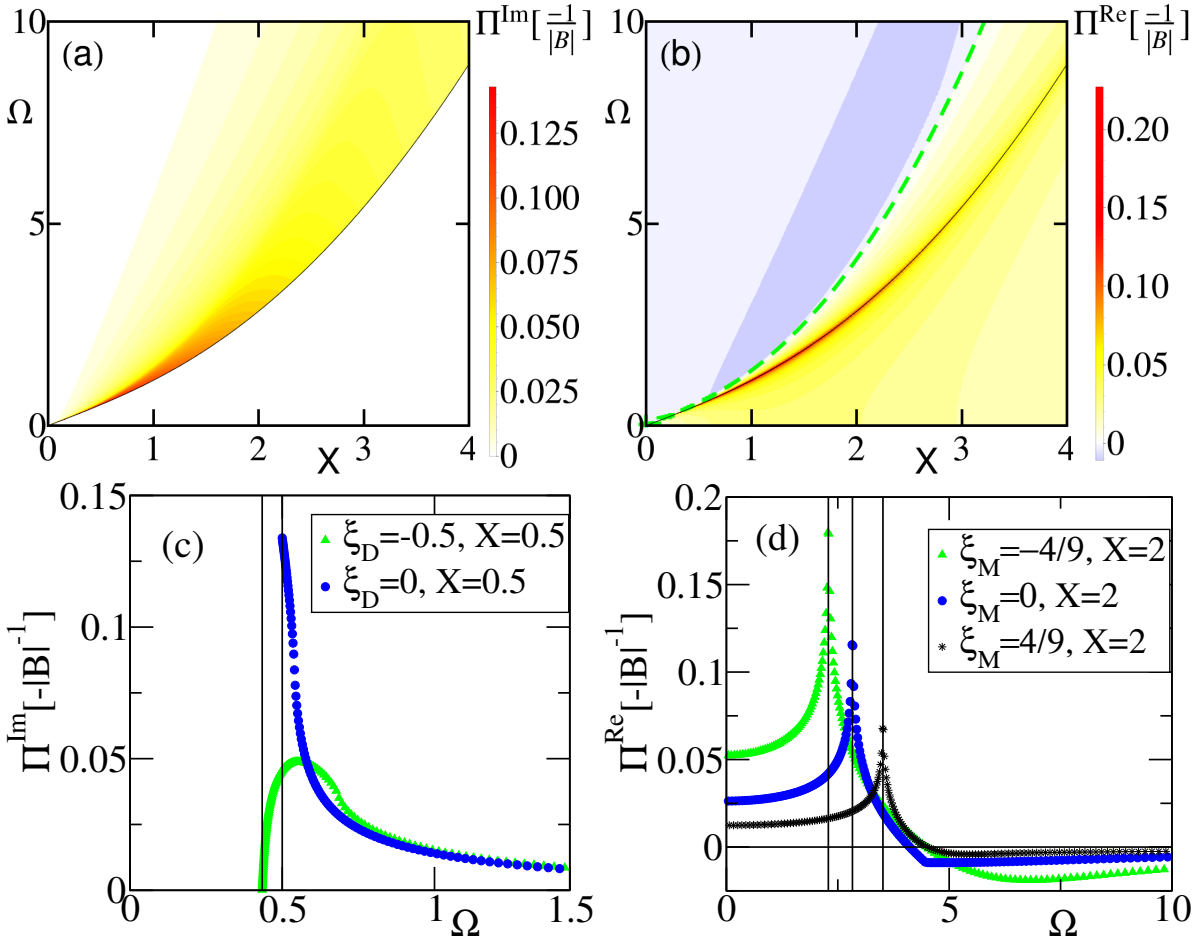


Figure 5.4. (a) Imaginary and (b) real part of $\Pi^R(X,\Omega)$ of the BHZ model for $\xi_M = \xi_\mathcal{D} = 0$. In (c) line cuts are drawn of Π^{Im} for $X = 0.5$ with $\xi_\mathcal{D} \in \{-0.5, 0\}$. The line cuts in (d) show Π^{Re} for $X = 2$ with $\xi_M = -\frac{4}{9}$, 0 and $\frac{4}{9}$ with green triangles, blue dots and black stars, respectively. Reprinted figures with permission from Ref. [JMT14a]. Copyright (2014) by the APS.

5.3.1. Interband Plasmons

Now, we look for solutions of Eq. (4.12) corresponding to plasmon quasiparticles of definite energy and momentum, i.e. with limited damping of $\Gamma/\Omega < 1$. For this purpose, Eq. (4.12) is expanded up to order $\mathcal{O}(\Gamma/\Omega)^2$ as was already done in Sec. 5.2. The results are two equations for real and imaginary part, from which we obtain the plasmon dispersion and its damping factor, respectively. We find that the damping ratio $\frac{\Gamma}{\Omega}(X,\Omega)$ parametrically depends on ξ_M and $\xi_\mathcal{D}$ only, but not on the interaction strength α. This behavior implies that Landau damping is due to the non-interacting SPEs.

Different from the expansion in Sec. 5.2, in the following we solve numerically for the damping ratio and plasmon frequency. In Fig. 5.5 (a), $\frac{\Gamma}{\Omega}$ is depicted for a TI phase with $\xi_M = -\frac{4}{9}$. For frequencies $\Omega \gg \Omega_{min}$, the system is sufficiently undamped and the expansion in Γ/Ω is justified. In order to study the plasmon dispersion curve, we look at isolines $\frac{1}{\Re[g(X,\Omega-i\Gamma)]} = \alpha$ for different interaction strengths α. They are shown in Fig. 5.5 (b) as a contour plot of $\frac{1}{\Re[g(X,\Omega-i\Gamma)]}$ for $\xi_M = -\frac{4}{9}$. The plasmon dispersion relation has a square root dependence on

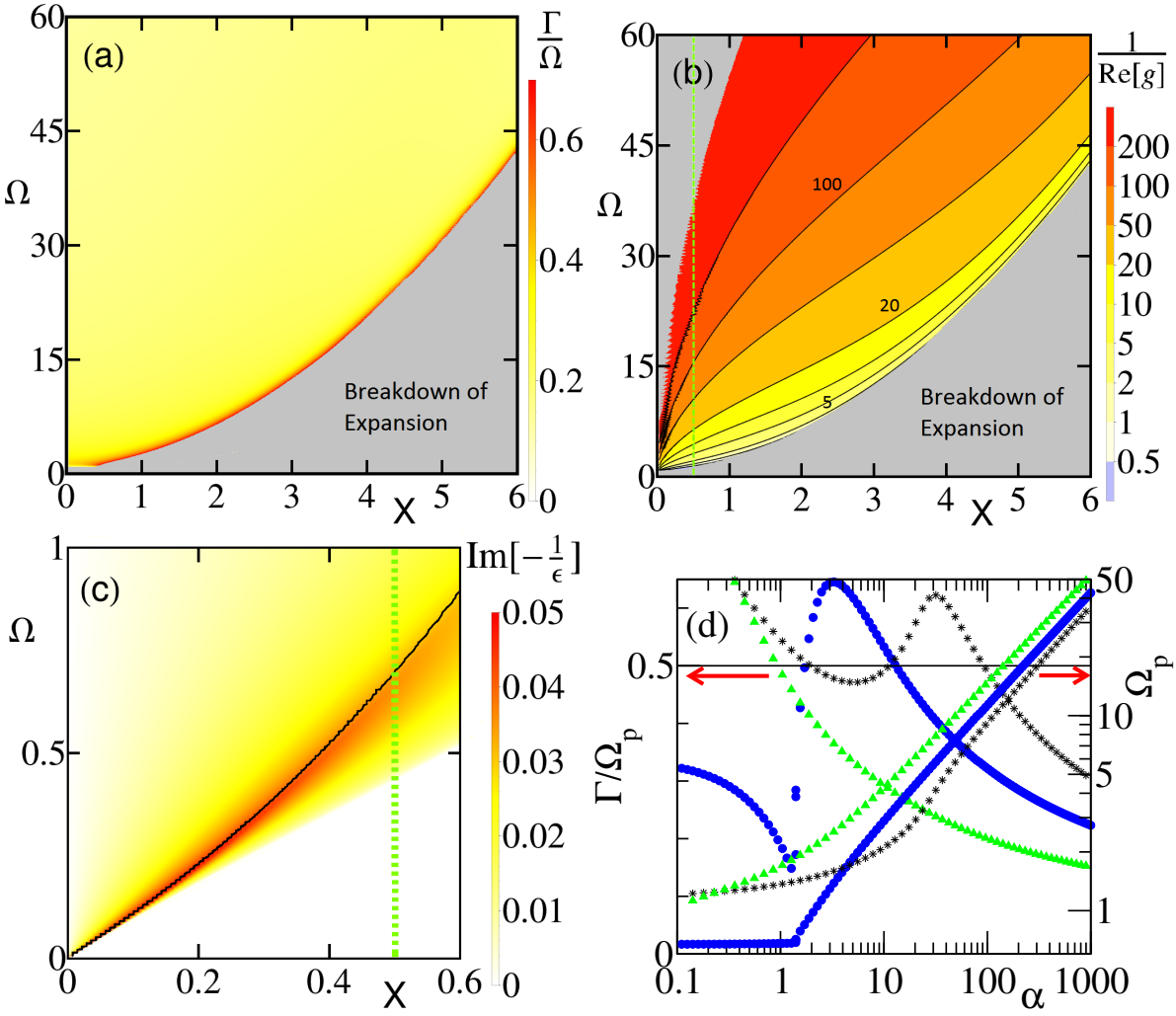


Figure 5.5. (a) Illustration of the damping ratio $\frac{\Gamma}{\Omega}$ and (b) the function $\frac{1}{\Re[g(X,\Omega-i\Gamma)]}$, for $\xi_M = -\frac{4}{9}$. (c) Plasmon dispersion together with the loss function $\Im\left[-\frac{1}{\varepsilon}\right]$, Eq. (4.14), for $\xi_{\mathcal{D}} = -0.5$ and $\alpha = 0.4$. (d) Plasmon frequency and damping at $X = 0.5$ (dashed line in panels (b) and (c)) for $\xi_M \in \left\{-\frac{4}{9}, \frac{4}{9}\right\}$ and $\xi_{\mathcal{D}} = -0.5$ in green triangles, black stars and blue dots, respectively. Reprinted figures with permission from Ref. [JMT14a]. Copyright (2014) by the APS.

X, known from doped graphene studied in Ref. [HD07] and ordinary 2DEGs, for $\alpha \to \infty$. In the opposite limit of $\alpha \to 0$, it shows instead the almost linear dependence from Eq. (5.8) for small X. A similar flattening has been predicted for intrinsic plasmons in graphene at finite temperature in Ref. [DL13].

Plasmons are also revealed as peaks in the loss function $\Im\left[-\frac{1}{\varepsilon}\right]$ defined in Eq. (4.14). In Fig. 5.5 (c) we compare the plasmon dispersion calculated numerically from Eq. (4.12) with the loss function for an interaction strength of $\alpha = 0.4$ and broken p-h symmetry term $\xi_{\mathcal{D}} = -0.5$. Evidently, plasmons are easily resolved and in perfect agreement with our analytical calculation. This is one of the key results of our research: the interplay between Dirac and Schrödinger fermions leads to a plasmonic excitation, which is absent in the limiting cases of a pure Dirac or Schrödinger system.

Indeed, plasmons have been proposed to exist also in intrinsic graphene, yet only under special circumstances. One is a finite temperature, examined in Refs. [Vaf06, DL13]. The other are a special kind of higher order diagrams going beyond the RPA, taken into account in Ref. [GFM08]. The latter result is still under debate, e.g. in Ref. [SPM12]. Hence, the physical reason for the appearance of plasmons in our case is rather different from the previously reported ones: The plasmonic solutions found here are due to the mixed Dirac/Schrödinger nature of the electronic system.

We compare the effect of different masses and the p-h breaking to one another in Fig. 5.5 (d). There, the plasmon frequency and damping as a function of α is plotted at fixed $X = 0.5$ for a NI and a TI phase with $\xi_{\mathcal{D}} = 0$ as well as for the case $\xi_M = 0$ and $\xi_{\mathcal{D}} = -0.5$. As the plasmon frequency increases with α, for $\alpha \to \infty$ the damping ratio $\frac{\Gamma}{\Omega_p}$ decreases to values below $\frac{\Gamma}{\Omega_p} \lesssim 0.2$. Notably, at large α the TI phase yields a larger plasmon frequency and is considerably less damped than the NI phase. This behavior directly stems from the strengthening of the antiscreening region of the polarization function due to the overlap factor enhancement in the TI phase, see Fig. 5.4 (d). In the opposite limit, $\alpha \to 0$, the excessive damping leads to a breakdown of our expansion. For finite $\xi_{\mathcal{D}}$ on the other hand, the plasmon damping has a minimum around $\alpha \approx 1$, which prevails to even smaller interaction strengths. This behavior is a direct consequence of the smaller Π^{Im} depicted in Fig. 5.4 (c), leading to reduced Landau damping for small plasmon frequencies.

5.3.2. Broken particle-hole Symmetry

The broken p-h symmetry with $\xi_{\mathcal{D}} < 0$ ensures that the lowest-energy interband excitations correspond to processes exciting particles from the valence band to the proximity of the Dirac point. There, however, the density of states is zero and these excitations are prohibited. This makes the p-h broken regime the most suitable one to observe plasmons, as their damping by SPEs is suppressed. Therefore we choose this regime to study the polarization function in further details. The SPE spectrum is given by Π^{Im}, which we plot in Fig. 5.6 (a) for $\xi_M = 0$ and $\xi_{\mathcal{D}} = -0.5$. In contrast to graphene, here Π^{Im} increases continuously from 0. The excitation spectrum shows a maximum for small momenta $X < 1$ which lies beneath the plasmon dispersion given by the black line, perturbatively calculated from Eq. (4.12) up to order $\mathcal{O}\left(\Gamma/\Omega\right)^2$ for $\alpha = 0.4$.

Considering a finite Coulomb interaction strength, the excitation spectrum is given by $\Pi^{\mathrm{Im}}_{\mathrm{rpa}}$ plotted in Fig. 5.6 (b) for $\alpha = 0.4$ and (c) for $\alpha = 10$. The maximum of the spectrum shifts to higher energies compared to the non-interacting one, indicating the formation of a collective excitation in the system - the plasmon. This is proven by the perturbatively calculated dispersion plotted as a black line on top of the spectrum. Additionally, the plasmon dispersion relations based on the long wavelength expansion of Π^R, Sec. 5.2, for $\Omega \ll 1$ and $\Omega \gg 1$ are plotted as gray lines in Fig. 5.6 (c).

The plasmon dispersion relation starts linearly for small q, as one would expect for a neutral system without doping. At high energies on the other hand, a free-particle behavior could be expected, leading to the usual $\sqrt{q}$ dispersion known from doped systems. Although Eq. (5.7) shows that this picture is only partly true due to the logarithmic correction of Π^{Re}, Fig. 5.6 (c) indicates a qualitative agreement. Comparing $\Pi^{\mathrm{Im}}_{\mathrm{rpa}}$ to the loss function in Fig. 5.5 (c), both are clearly in accord with one another. Due to the additional factor $1/X$, the loss function puts a stronger emphasis on the long wavelength part of the excitation spectrum.

The line cuts of $\Pi^{\mathrm{Im}}_{\mathrm{rpa}}$ shown in Fig. 5.6 (d) are for fixed $X = 0.7$ with $\alpha \in \{0, 0.2, 0.4, 10\}$.

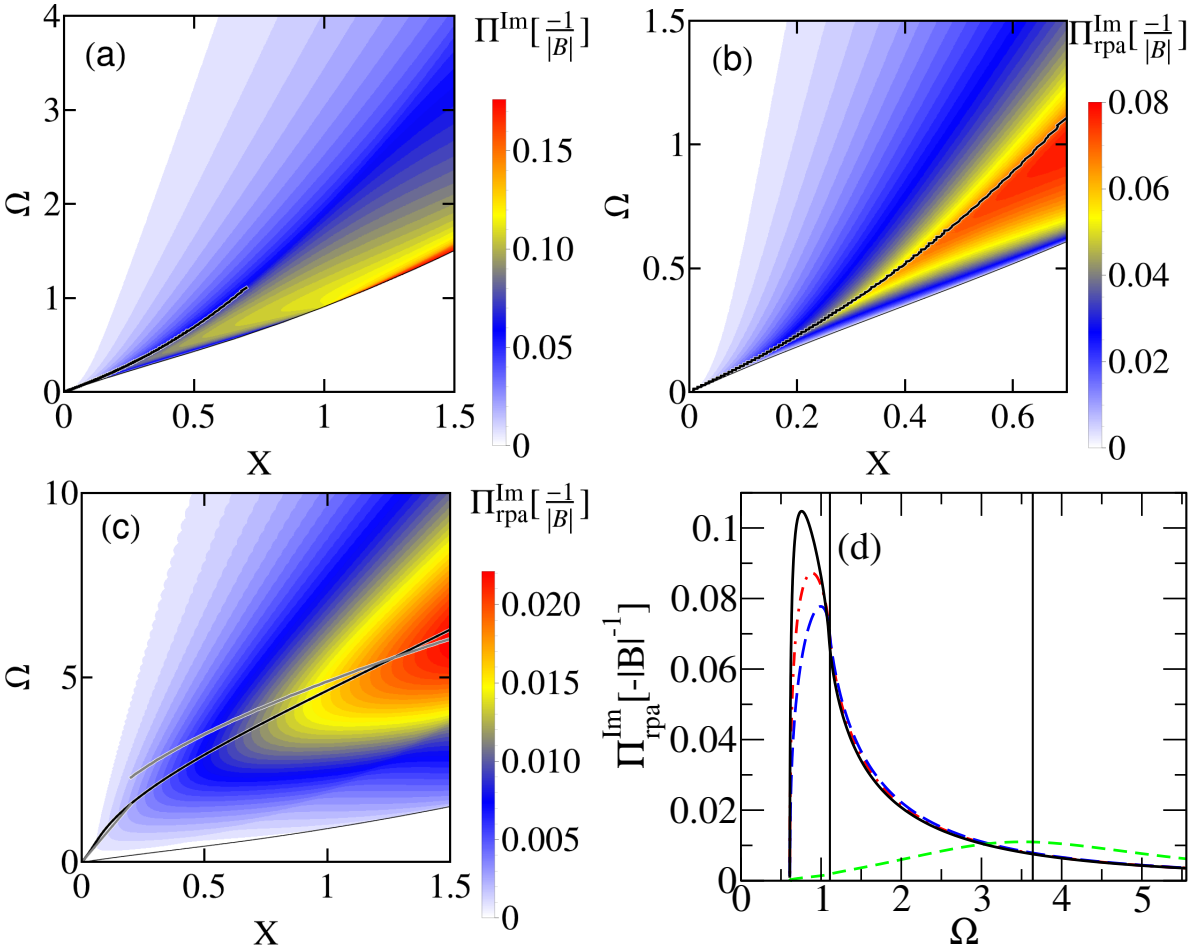


Figure 5.6. Plots of (a) Π^{Im} and $\Pi^{\mathrm{Im}}_{\mathrm{rpa}}$ for (b) $\alpha = 0.4$ and (c) $\alpha = 10$. (d) shows line cuts for fixed $X = 0.7$ with $\alpha \in \{0, 0.2, 0.4, 10\}$ in black solid, red dot-dashed, blue long dashed and green short dashed lines, respectively. $\xi_M = 0$ and $\xi_{\mathcal{D}} = -0.5$. Reprinted figures with permission from Ref. [JMT14b]. Copyright (2014) by the APS.

Additionally, the black vertical lines indicate the plasmon frequency for $\alpha \leq 0.4$ (left line) and $\alpha = 10$ (right line). For $\alpha = 0.2$, the maximum of the interacting spectrum lies between the maximum of the non-interacting spectrum and the plasmon frequency, signifying that single-particle and collective excitations are equally strong. Increasing the interaction to $\alpha = 0.4$, the maximum of the interacting spectrum and the plasmon frequency almost coincide. We conclude therefore that the plasmon dominates over the SPEs. Finally, at very large interactions $\alpha = 10$, the plasmon is the only relevant excitation in the system.

Increasing the Coulomb interaction broadens the plasmon peak and reduces its height as shown in Fig. 5.6 (d). This seems contrary to the picture of a plasmon as a sharp interaction-induced charge resonance, suggesting that these interband plasmons may not be well-defined for high energies. Yet this is a false conclusion. In Sec. 4.5.1 we discussed that the contribution of the resonance to the f-sum rule is the actual measure of importance of a resonance. Following Eq. (4.34), the latter can be estimated by multiplying the peak height in $\Pi^{\mathrm{Im}}_{\mathrm{rpa}}$ by Ω_p, while the relevant peak width is given by $\frac{\Gamma}{\Omega_p}$. The latter is decreasing with Ω_p according to Eq. (5.11).

From this normalization of the peak we conclude that the discussed interband plasmons

fulfill the interpretation as sharp interaction-induced charge resonances. The width $\frac{\Gamma}{\Omega_p}$ of the collective excitation decreases with increasing plasmon frequency, above the critical frequency Ω_c as defined in Sec. 5.2.

5.3.3. Experimental Realization

In Sec. 4.4 we considered the experimentally relevant parameters for Hg(Cd)Te QWs, and concluded that plasmons with a damping ratio of $\frac{\Gamma}{\Omega_p} = 0.5$ are still observable in experiments. From Fig. 5.5 (d) we find $\frac{\Gamma}{\Omega_p} \approx 0.3$ for the interaction strength $\alpha \approx 0.27$ and $\xi_{\mathcal{D}} = -0.5$. The wave vector and frequency of the plasmon is extracted from Fig. 5.5 (c) and Fig. 5.6 (b) to be

$$\begin{aligned} q &\in [0.1, 0.6]\, q_0 = [0.04, 0.24]\, {}^1\!/_{\mathrm{nm}}, \\ \omega &\in [0.1, 0.8]\, \frac{E_0}{\hbar} = [14, 112]\, \frac{\mathrm{meV}}{\hbar} = [21, 170]\, \mathrm{THz}, \end{aligned} \tag{5.12}$$

where the lower bound stems from the merging of plasmon and single-particle background for $X, \Omega \to 0$. This momentum and frequency range is of the right order for experimental techniques like Raman spectroscopy.

A finite temperature in experiments can lead to doping by thermal excitations, which in turn can result in the formation of intraband plasmons, studied in Refs. [Vaf06, DL13]. At the temperature of liquid helium, one finds $k_B T_{He} \approx 0.35$ meV with k_B the Boltzmann constant. Thus, the plasmons resulting from thermal excitations occur on an energy and momentum scale at least two orders of magnitude smaller than the plasmons discussed in this section. This makes it possible to fully separate them or to suppress them with a small gap $k_B T < |M|$.

We conclude that the plasmonic resonances discussed above are measurable, e.g. with Raman spectroscopy on Hg(Cd)Te QWs.

5.4. Optical Conductivity

We have seen that plasmons differ for the NI and TI phases, but the difference is of a quantitative fashion and probably hard to distinctively detect in experiments. Therefore we still look for a bulk quantity where topology leads to a clearly measurable signal. The bulk optical conductivity could be such a quantity.

From the knowledge of the polarization function, we can calculate the bulk optical conductivity of the system, defined by Ref. [Ste67], as

$$\sigma_0(\Omega) = \lim_{X \to 0} \frac{\Omega}{X^2} |\mathcal{B}|\, \Pi^R(X, \Omega)\,. \tag{5.13}$$

An analytical calculation yields for the BHZ model

$$\Im\left[\sigma_0(\Omega)\right] = -\left[\frac{1}{W} + \frac{1 + 4\xi_M}{\Omega^2}\left(\frac{1 + 2\xi_M}{W} - \frac{1}{2}\right)\right] \Theta\left(\Omega - 2\,|\xi_M|\right), \tag{5.14}$$

where $W = 2\sqrt{1 + 4\xi_M + \Omega^2}$. Notably, σ_0 is a universal function depending only on Ω and ξ_M. In Fig. 5.7, we plot $\Im\left[\sigma_0(\Omega)\right]$ for the different masses $\xi_M = -\frac{4}{9}$, 0 and $\frac{4}{9}$.

Compared to the massless case, both positive and negative Dirac mass lead to a peak just above $\Omega_{min} = |2\xi_M|$. The signal from the TI phase is much stronger, even diverging for $|\xi_M| \to \frac{1}{2}$. The latter is the threshold of turning the band structure into a Mexican hat shape.

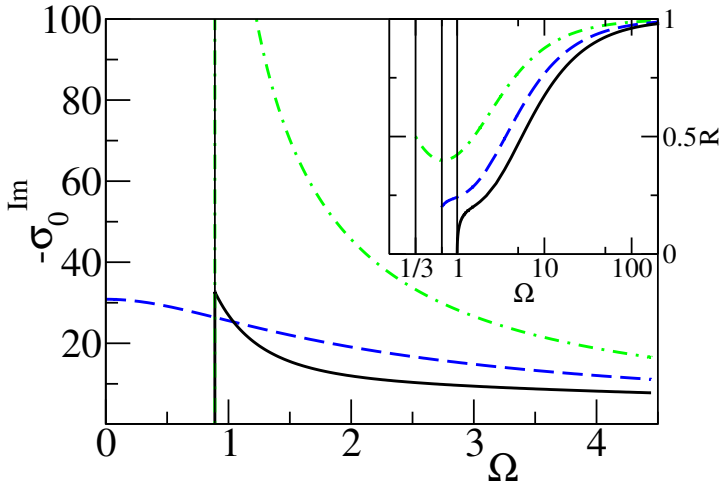


Figure 5.7. Illustration of $\Im[\sigma_0(\Omega)]$ for $\xi_M = -\frac{4}{9}$, 0 and $\frac{4}{9}$ with dot-dashes, long dashes and a solid line, respectively. The inset shows the ratio $R = \frac{\Im[\sigma_0(\Omega)]_{M>0}}{\Im[\sigma_0(\Omega)]_{M<0}}$ for $|\xi_M| \in \left\{\frac{1}{6}, \frac{1}{3}, \frac{1}{2}\right\}$ in dot-dashes, long dashes and a solid line, respectively. Reprinted figure with permission from Ref. [JMT14a]. Copyright (2014) by the APS.

Such difference between the trivial and topological phases is emphasized in the inset of Fig. 5.7, where we plot the ratio $R = \frac{\Im[\sigma_0(\Omega)]_{M>0}}{\Im[\sigma_0(\Omega)]_{M<0}}$.

The divergent behavior can be explained by considering the combined effects of the overlap factor, Sec. 4.2, and the phase space for the excitation processes. In the TI phase of the BHZ model, the conduction and valence bands flatten for small momenta $X < |\xi_M|$, with respect to the NI phase. This fact enormously increases the number of finite momentum states, available for an excitation just above $\Omega_{min} = |2\xi_M|$. The increased overlap factor leads to an additional increase of $\Im[\sigma_0]$ for small and intermediate momenta. This is contrasted with a slight general decrease of the optical conductivity for positive mass.

We conclude: the increase and divergence of the optical conductivity in the TI phase could be strong enough as to serve as a distinction mark in experiments. This can be combined with the increase in the f-sum rule for negative masses, explored in Sec. 4.5.

5.5. F-sum Rule - a Numerics Check

The f-sum rule provides a check for our numerics, as dicussed in Sec. 4.5. In Fig. 5.8 we plot the ratio $\frac{\Delta_{f\sum}^{num}}{f\sum}$, with $\Delta_{f\sum}^{num} = f\sum^{num} - f\sum$ being the difference between the numerical calculated f-sum rule $f\sum^{num}$ and the analytic $f\sum$ from Eq. (4.29). The deviation is of the order 10^{-3}, comparable to the analytical uncertainty, see Sec. 4.5, and thus negligible.

The f-sum has to be the same for interacting and non-interacting systems. We find a slight dependence on the interaction strength α, which could be a numerical artifact, depending on the cutoff Λ, or a real α dependence like in graphene. In the latter Dirac case, spectral weight is missing for small frequencies, cf. Eq. (14) in Ref. [SNC08], such that $\Pi^{RPA} < \Pi^R \ \forall q, \omega$ for the undoped Dirac model. As the observed dependence declines with increasing cutoff Λ, we conclude that the RPA approximation in the BHZ model misses no spectral weight compared to the full Coulomb interaction, even in the undoped limit.

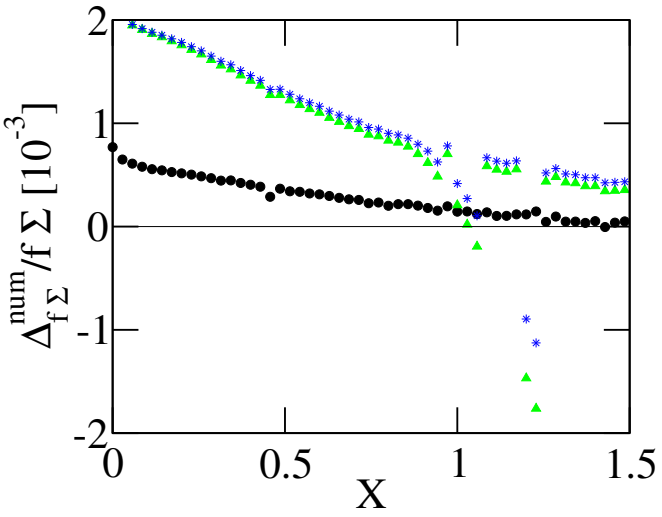


Figure 5.8. The ratio $\frac{\Delta^{num}_{f\sum}}{f\sum}$, with $\Delta^{num}_{f\sum}$ being the difference between the numerical and analytical f-sum rule, Eq. (4.29). Black dots are for the non-interacting spectrum, while blue stars stand for $\alpha = 0.4$ and green triangles for $\alpha = 10$. $\beta = 3$, $X_{max} = 6$, $\xi_M = 0$ and $\xi_{\mathcal{D}} = -0.5$. The deviations around $X \gtrsim 1$ stem from numerical instabilities, which are however negligibly small. Reprinted figure with permission from Ref. [JMT14b]. Copyright (2014) by the APS.

6. Doped System

In this section, we extend our analysis of the polarization function of the BHZ model to finite doping $\mu > 0$, where a net charge density is present in the system. Doping the system has two effects: one is the Fermi blocking of interband excitations for small X and Ω, represented by the red arrow in Fig. 6.3 (a). The other one is the appearance of intraband excitations, indicated by the green arrow in Fig. 6.3 (a), which are absent in the intrinsic limit.

Again, we first study the polarization and screening properties of the system in the static limit in Sec. 6.1. The discussed Friedel oscillations are relevant, e.g., in the case of scattering on a charged impurity. Then, we take a look at the dynamic polarization function in the long-wavelength limit in Sec. 6.2. Here we obtain an analytical expression for the dispersion and damping of the collective plasmonic modes of the system. Finally we numerically compute the dynamical polarization function in the full range of momenta and frequencies, considering the full parameter space of the BHZ model. Its ability to interpolate between the Dirac and 2DEG limit is explored in Sec. 6.3. The experimental relevant case for Hg(Cd)Te QWs of broken p-h symmetry and small to negligible mass terms is studied in Sec. 6.4. As in the undoped limit, Sec. 5.3.2, the interband plasmon is especially strong in this case. Due to a blocking effect, it is possible to separate inter- and intraband plasmons in the interacting excitation spectrum, for experimentally relevant interaction strengths α, momenta and energies. We conclude in Sec. 6.5 by investigating the effect of large mass terms. The topological trivial and non-trivial excitation spectrum are qualitatively distinct from one another.

6.1. Static Limit and Screening

We begin our discussion again with the static limit of the polarization function. An interesting feature will be the Friedel oscillations, which change their decay behavior depending on the doping level of the system.

In Fig. 6.1, we present the static polarization function $\tilde{\Pi}(X) = \Pi(X)/G(\mu)$ at finite doping, normalized by the density of states (DOS) at the Fermi level $G(\mu)$. This normalization stands out naturally from the long wavelength property of the polarization function

$$\lim_{q\to 0} \Pi(q) = \Pi_\mu(0) = G(\mu). \tag{6.1}$$

For the BHZ model at finite doping, $\tilde{\Pi}(X)$ has a pronounced dependence on the extrinsic parameter $X_F = k_F/q_0$. For $X_F \ll 1$ ($X_F \gg 1$) the Fermi level falls in a region where locally the dispersion curve has predominant Dirac (2DEG) character. In a 2DEG system, the static polarization assumes the known analytic form, cf. Ref. [GV05], of

$$\tilde{\Pi}(q) = 1 - \Theta(q - 2k_F)\frac{\sqrt{q^2 - 4k_F^2}}{q}. \tag{6.2}$$

It is constant up to $q = 2k_F$, where a jump in the first derivative occurs. In the limit of large momenta, it decays like $1/q^2$. The Dirac limit is analyzed in Refs. [GGMS02, And06], one

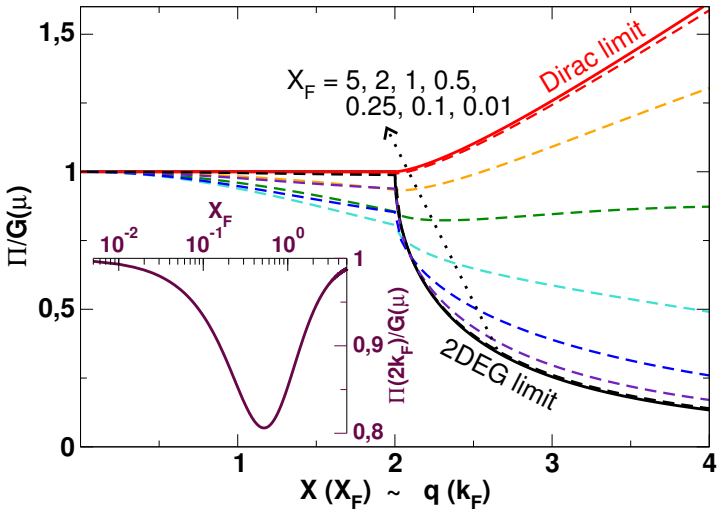


Figure 6.1. Static polarization function $\tilde{\Pi}(X)$ of the BHZ model, normalized by the DOS $G(\mu)$. We choose the massless, p-h symmetric case $\xi_M = \xi_{\mathcal{D}} = 0$ and vary the doping with different values of X_F. In the inset, details on the value of $\tilde{\Pi}(2X_F)$ as a function of X_F are given. Reprinted figure with permission from Ref. [JMT14b]. Copyright (2014) by the APS.

finds for the normalized polarization

$$\tilde{\Pi}(q) = 1 - \Theta(q - 2k_F)\left[\frac{\sqrt{q^2 - 4k_F^2}}{2q} - \frac{q}{4k_f}\arctan\frac{\sqrt{q^2 - 4k_F^2}}{2k_F}\right]. \tag{6.3}$$

It is constant up to $q = 2k_F$ as well, but diverges linearly for large momenta as depicted in Fig. 6.1.

Our own calculations for the BHZ model with $\xi_M = \xi_{\mathcal{D}} = 0$ correctly reproduce the Dirac and 2DEG limits for $X_F \ll 1$ and $X_F \gg 1$, respectively. We note that with a finite $\mathcal{B}$ term and nonzero doping X_F, the polarization will always have a decay behavior for $q > q_0$, in contrast to the divergence of the pure Dirac case. In the 2DEG and Dirac limit one finds $\tilde{\Pi}(q) = 1$ for $q < 2k_F$, which seems like a coincidence due to the balancing effect of dispersion curve and overlap factor. Interestingly, in the BHZ model we observe instead a deviation from unity with a maximum for $X_F \approx 0.5$, shown in details in the inset of Fig. 6.1. At the Fermi wave vector $X = 2X_F$, a discontinuity occurs in the polarization function. In the 2DEG limit, $X_F \gg 1$, the first derivative of $\tilde{\Pi}(X)$ has a large jump. Decreasing X_F, this discontinuity decreases as well and finally vanishes in the Dirac limit, where the discontinuity affects only the second derivative of $\tilde{\Pi}(X)$.

Let us consider the consequences of the described behavior for the screening in the system. We already analyzed in Sec. 5.1 the intrinsic response of a BHZ system to a test charge, when no net charge density is present in the system. While the intrinsic response is realized on the intrinsic scale $1/q_0$ of the model, the 'metallic' response at finite electronic density is characterized by the Fermi wave length π/k_F. Therefore it is convenient to express $n_\mu(X)$ as a function of dimensionless units $\tilde{X} = X/X_F = k/k_F$, due to the presence of the discontinuity at $\Pi_\mu(2X_F)$. The induced charge density in real space follows from Eqs. (4.22) and (4.23) and

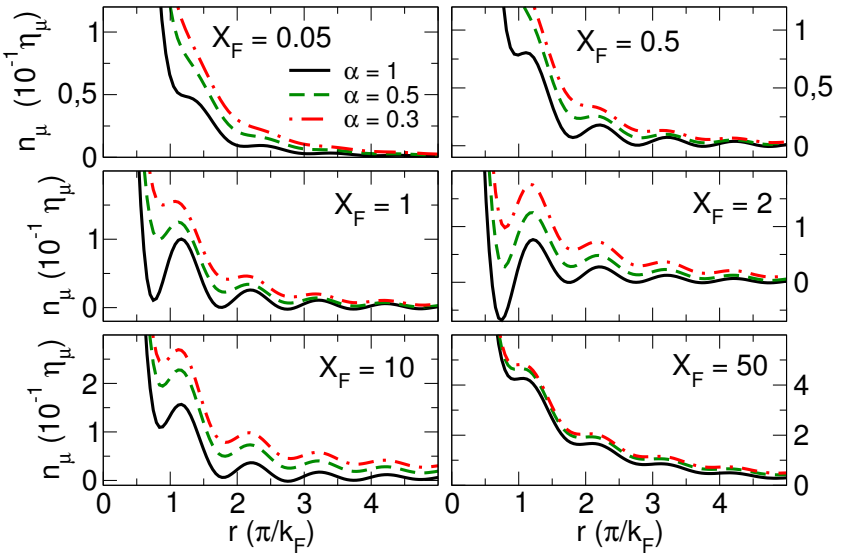


Figure 6.2. Induced charge density in real space for the BHZ model for $\xi_M = \xi_{\mathcal{D}} = 0$ and $\alpha = 0.3,\ 0.5$ and 1. Data in different panels belong to systems with different doping levels $X_F \in \{0.05, 0.5, 1, 2, 10, 50\}$. All calculation are obtained by keeping $\varepsilon_r = 10$ and $k_F = 0.1$ $^1/_{\text{nm}}$, and varying $\mathcal{A} \in \{0.143, 0.29, 0.47\}$ eVnm in order to change $\alpha \in \{0.3, 0.5, 1\}$, respectively. The parameter $\mathcal{B}$ is varied accordingly to X_F. In the panel $X_F = 50$ (2DEG limit), the three curves with $\alpha = 0.3$, 0.5 and 1 are quite close and correspond to similar β parameters, $\beta = 0.01$, 0.008 and 0.007, respectively. Reprinted figure with permission from Ref. [JMT14b]. Copyright (2014) by the APS.

is given by

$$n_\mu(r) = \eta_\mu \int \mathrm{d}\tilde{X} \frac{J_0(r k_F \tilde{X})}{1 - \alpha g_0(\tilde{X} X_F)} \frac{\tilde{\Pi}_\mu(\tilde{X} X_F)}{1 - \alpha g(\tilde{X} X_F)}. \tag{6.4}$$

The scaling factor $\eta_\mu = \eta_0 |B| G(\mu) X_F = \frac{e^2 k_F}{4\pi\varepsilon_0 \varepsilon_r^2} G(\mu)$ is directly proportional to the DOS and the Fermi momentum. We note that the integral explicitly depends on the parameters α and X_F, and naturally on ξ_M and $\xi_{\mathcal{D}}$, when finite.

In Fig. 6.2, we present the radial density, induced by a point-like test charge, for the massless, p-h symmetric BHZ model. Each panel corresponds to a different value of the ratio $X_F = k_F/q_0$, and within each panel curves differing by the Dirac fine-structure constant α are presented. Friedel oscillations appear of period π/k_F, which become more pronounced for larger α. We also note that the density oscillations are more prominent in the intermediate regime of $k_F \approx q_0$ than in the Dirac ($X_F \ll 1$) and 2DEG limits ($X_F \gg 1$). In the 2DEG limit, the $\alpha = \frac{1}{\mathcal{A}} \frac{e^2}{4\pi\varepsilon_0\varepsilon_r}$ parameter is ill defined and should be replaced by the more general parameter

$$\beta = \frac{e^2 G(\mu)}{2\varepsilon_0 \varepsilon_r k_F}, \tag{6.5}$$

characterizing the dielectric response of the system.

We checked that Friedel oscillations reproduce the expected behavior in the Dirac and 2DEG limits, where oscillations decay as r^{-3} and r^{-2}, respectively, while for intermediate regimes they decay as r^{-p} with $p \in (2, 3)$. The presence of Friedel oscillations and their asymptotic behavior are related through the Lighthill theorem, presented in Ref. [Lig58], to discontinuities in the static polarization function and its derivatives. A detailed discussion can be found, for

example, in Ref. [BMAVF10]. A discontinuity like $|q-2k_F|^\nu \Theta(q-2k_f)$ in $\Pi(q)$, with Θ the Heaviside step function and $\nu \in \mathbb{R}$, translates into a decay of the oscillations in $n(r)$ with leading order $r^{-\nu-3/2}$. One finds $\nu = 1/2$ ($\nu = 3/2$) for the leading order discontinuity of a 2DEG (Dirac) system, such that the first (second) and all higher derivatives of the static polarization function are discontinuous at $q = 2k_F$. Analyzing the Friedel oscillations for the BHZ model, one finds a composition of two different contributions with an asymptotic decay at large distances as r^{-2} (2DEG contribution) and r^{-3} (Dirac contribution), respectively.

As a consequence, the discontinuity in the RPA polarization function of the BHZ model at $q = 2k_F$ can be very well approximated by a combination of 2DEG ($\nu = 1/2$) and Dirac ($\nu = 3/2$) contribution. In the Dirac (2DEG) limit, achieved by low (high) doping, the effect of the discontinuity in the second (first) derivative becomes predominant and oscillations purely decay in leading order as r^{-3} (r^{-2}).

6.2. Long Wavelength Expansion of $\Pi^R(X,\Omega)$ and Plasmon Dispersion

At finite doping, for small momenta X, the polarization function is governed by intraband excitations, as the interband excitations are Fermi-blocked. We perform an expansion in this limit, for $\Omega > X$, in order to gain an analytical insight into the physics at finite doping and derive an analytical formula for the plasmon dispersion. In particular, intraband plasmons are expected to be the dominant excitation for small momenta, similarly to the 2DEG and Dirac case. The obtained solutions will be compared to the full numerics in the following Sec. 6.3.

We expand the polarization function up to order X^4

$$|\mathcal{B}|\,\Re\left[\Pi^R(X,\Omega)\right] = \Pi_{44}\frac{X^4}{\Omega^4} + \Pi_{42}\frac{X^4}{\Omega^2} + \Pi_{40}X^4 + \Pi_{22}\frac{X^2}{\Omega^2} + \Pi_{20}X^2 + \mathcal{O}\left(\Omega^2\right) \tag{6.6}$$

and use it to solve the plasmon equation (4.12) for the damping coefficient and the plasmon dispersion. The latter is given by

$$\Omega = \sqrt{2\pi\alpha\Pi_{22}}\sqrt{X} + \left(\frac{\Pi_{44}}{\Pi_{22}^{\frac{3}{2}}}\frac{1}{\sqrt{8\pi\alpha}} + \sqrt{2\pi^3\alpha^3\Pi_{22}}\Pi_{20}\right)X^{\frac{3}{2}} \tag{6.7}$$

with the leading coefficient

$$\begin{aligned}
\Pi_{22} =& g_s\left(\frac{X_f^2(1+2X_f^2+2\xi_M)}{4\pi\sqrt{X_f^2+(\xi_M+X_f^2)^2}} - \eta\xi_{\mathcal{D}}\frac{X_f^2}{2\pi}\right) = g_s\frac{1}{4\pi}(\,|\Omega_f| + \underbrace{\Pi_{inter}(X_f)}_{-|\xi_M|<\ldots<|\Omega_f|}\,) \\
\underset{\substack{\xi_M=0\\ X_f\to 0}}{=}& g_s\frac{X_f}{4\pi} + \mathcal{O}\left(X_f^2\right) = g_s\frac{|\Omega_f|}{4\pi} + \mathcal{O}\left(X_f^2\right) \qquad (6.8) \\
\underset{\substack{\xi_M=0\\ X_f\to\infty}}{=}& g_s\frac{X_f^2}{2\pi}(1-\eta\xi_{\mathcal{D}}) + \mathcal{O}\left(\frac{1}{X_f^2}\right) = g_s\frac{|\Omega_f|}{2\pi} + \mathcal{O}\left(\frac{1}{X_f^2}\right) \qquad (6.9)
\end{aligned}$$

with $\eta = \operatorname{sgn}[\Omega_f]$ and $\Pi_{inter}(X_f) = \frac{X_f^4-\xi_M^2}{\sqrt{X_f^2+\left(\xi_M+X_f^2\right)^2}} - \eta\xi_{\mathcal{D}}X_f^2$. In the limit of zero mass, $\Pi_{inter}(X_f)$ interpolates smoothly between 0 for $X_f \to 0$ and $|\Omega_f|$ for $X_f \to \infty$. The former

case corresponds to the Dirac limit, where one finds the plasmon frequency

$$\omega = A\sqrt{\frac{g\alpha k_f}{2}}\sqrt{q} = \sqrt{\frac{ge^2\mu}{8\pi\varepsilon_0\varepsilon_r}}\sqrt{q} \tag{6.10}$$

in the literature, cf. Ref. [WSSG06], being identical to Eqs. (6.7) and (6.8). The latter case is the 2DEG limit, with the plasmon dispersion

$$\omega = \sqrt{\frac{e^2N}{2m\varepsilon_0\varepsilon_r}}\sqrt{q} = \sqrt{\frac{ge^2\mu}{4\pi\varepsilon_0\varepsilon_r}}\sqrt{q} \tag{6.11}$$

found in the literature, e.g. in Ref. [GV05]. Here, $N = \frac{g}{4\pi}k_f^2$ is the carrier density and $m = \frac{1}{2|\mathcal{B}+\mathcal{D}|}$. This is in agreement with Eqs. (6.7) and (6.9). Thus the BHZ model as a function of its parameters reproduces the plasmon dispersion in the Dirac and 2DEG limits and smoothly interpolates between them.

We note that for $k_f \to 0$ the term Π_{22} goes to zero and the intraband plasmon disappears. In this limit, the leading order contribution $\mathcal{O}\left(\frac{X^2}{\Omega^2}\right)$ of the intrinsic polarization, Eq. (5.7), takes the place of Π_{22}. A crucial difference between the extrinsic and the intrinsic polarization is that the latter has a finite imaginary part of order $\mathcal{O}\left(\frac{X^2}{\Omega^2}\right)$, leading to the linear dispersion (5.8) of the interband plasmons. Yet for finite $k_f > 0$, these interband plasmons are suppressed due to the Fermi blockade of the interband excitations. Then they only exist if their plasmon frequency exceeds both the chemical potential Ω_f and the critical frequency Ω_c as defined in Sec. 5.2. This is nicely presented in Fig. 6.5.

Besides the different scaling with momenta in the limit $X \to 0$, also the scaling with the interaction stength α is different for the inter- and intraband plasmon dispersions, Eqs. (5.8) and (6.7): linear vs. square root. This will have important consequences in the following when we discuss how to separate the two different collective excitations in the spectrum.

6.3. Excitation Spectrum of the massless, p-h symmetric BHZ Model: Interpolation between Dirac and 2DEG Regime

We proceed, with the analytical results at hand, to the full numerical discussion of the polarization function and its related quantities. We begin with the analysis of the doped excitation spectrum by looking at the limiting results of 2DEG and Dirac system first. This is done in App. A. From there, we find that we can interpolate between these two limits by changing the Fermi momentum. Interestingly, by considering the cases of broken p-h symmetry and large masses, we also discover regimes which are distinct from the Dirac and 2DEG limits in Secs. 6.4 and 6.5. As an example, these regimes support both inter- and intraband plasmons, at parameters which are realistic for HgTe QWs.

In all the following plots, the boundaries of the single-particle spectrum will be indicated by faint black lines, the isolines $\Pi^{\mathrm{Re}} = 0$ by red lines. The plasmon dispersions are plotted as black curves (full result from perturbation theory) and gray curves (expanded result in the limit $X \to 0$).

We consider first the different possible excitations in the system. For this, the band structure of the BHZ model without mass and p-h symmetry breaking terms is shown in Fig. 6.3 (a). The interband SPEs lying lowest in energy are symmetric in momentum as shown by the red

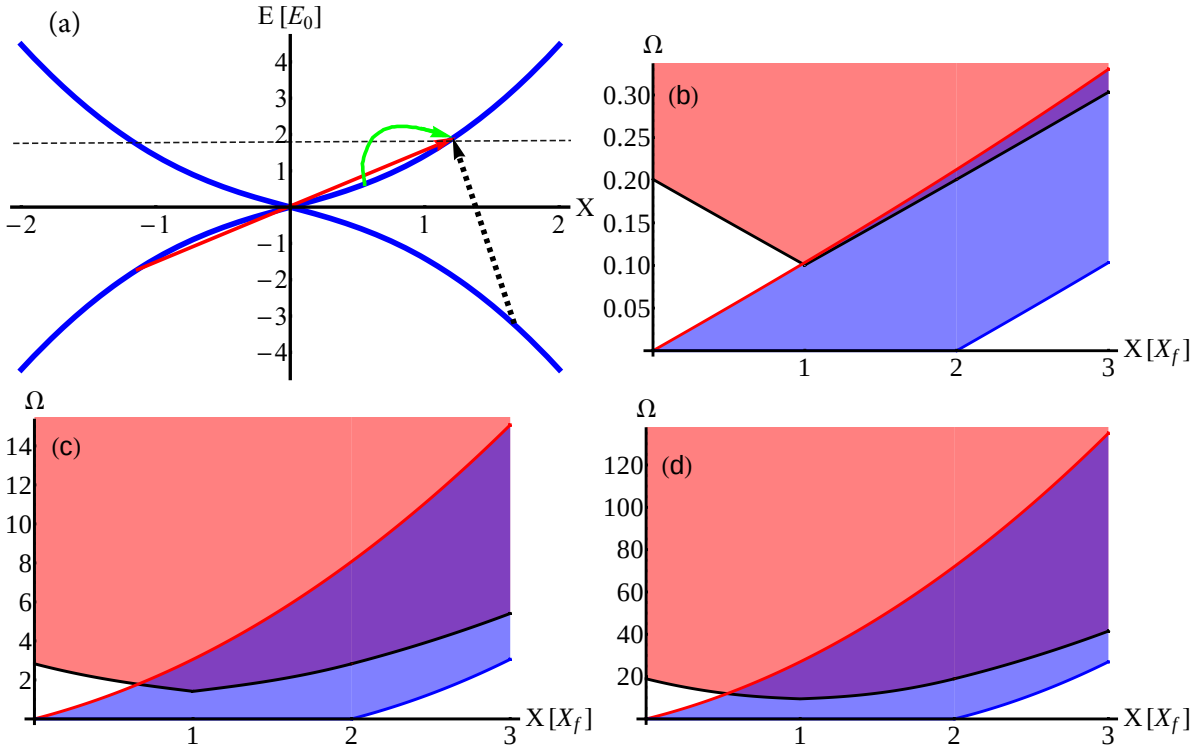


Figure 6.3. (a) Band structure of the BHZ model with finite chemical potential. The intraband (interband) transitions are indicated by green (red and black, dashed) arrows. (b)-(d) Boundaries of the excitation spectrum for Fermi momenta $X_f \in \{0.1, 1, 3\}$. The interband spectrum in red, intraband spectrum in blue and mixed area in purple. Reprinted figures with permission from Ref. [JMT14b]. Copyright (2014) by the APS.

arrow in Fig. 6.3 (a), going from $-\boldsymbol{X}$ to $\boldsymbol{X}$ closely above the Fermi level. Due to p-h symmetry, this leads to nesting and thus we expect these excitations to dominate the interband spectrum. Another kind of interband excitations are indicated by the dashed, black arrow. They go from momentum $\boldsymbol{X} + \boldsymbol{X}_f$ to $\boldsymbol{X}_f$ with $\boldsymbol{X} \| \boldsymbol{X}_f$, and are suppressed due to imperfect nesting of the different sized electron and hole cones, as well as by a small overlap factor resulting from opposing pseudospins. The latter can be cured by introducing a large negative mass, as will be shown in Sec. 6.5. There, these excitations have a considerable influence on the polarization function for higher energies. They help with the formation of interband plasmons, following the ideas presented in Sec. 4.3.1. In the pure Dirac system on the other hand, these processes are completely forbidden by helicity.

The excitation spectrum of the system can be modified by varying the doping level as presented in Fig. 6.3 (b)-(d). It resembles the spectrum of a Dirac system ($X_f \ll 1$), of a 2DEG ($X_f \gg 1$), or we can obtain an intermediate behavior for $X_f \sim 1$. In the figures we highlight the boundaries of the excitation spectra, with the red (blue) area corresponding to the interband (intraband) spectrum. The overlap between the two is indicated by the purple area. The boundaries of the spectra vary from the linear Dirac behavior to the q^2 dependence of the 2DEG. In general, the mixing of linear ($\mathcal{A}$) and quadratic ($\mathcal{B}$) dispersion relations leads to an overlap of the inter- and intraband spectrum. This affects the visibility of the interband plasmons, which can be hidden in this area of parameter space due to the strong single-

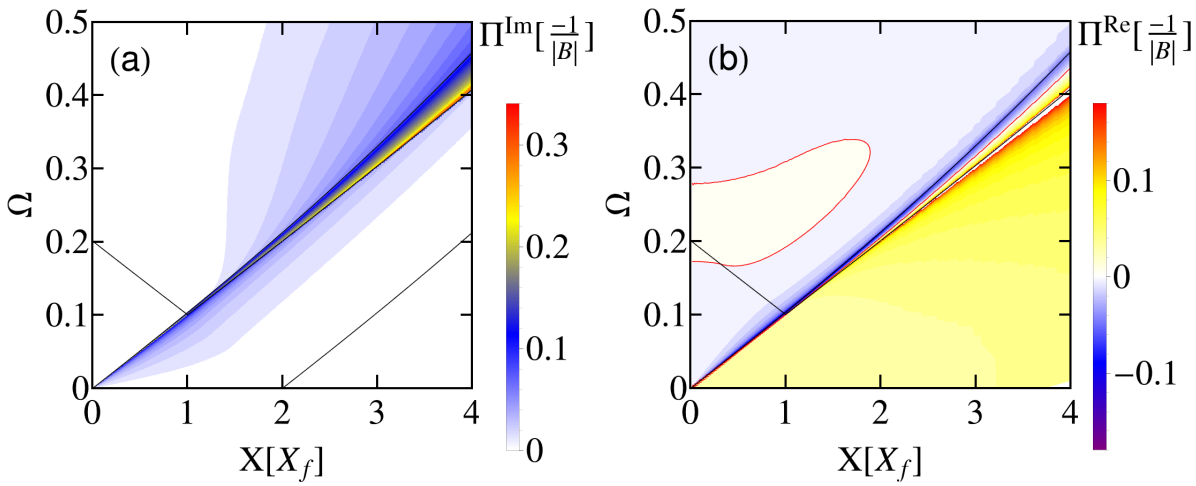


Figure 6.4. (a) Imaginary and (b) real part of the polarization function for $X_f = 0.1$. The red line indicates $\Pi^{\mathrm{Re}} = 0$. Reprinted figures with permission from Ref. [JMT14b]. Copyright (2014) by the APS.

particle damping. A clear separation of the two regions is possible for broken p-h symmetry, as discussed in Sec. 6.4.

6.3.1. Weak doping of $X_f = 0.1$

In the limit of weak doping, the extrinsic (k_f) and intrinsic (q_0) momentum scales of the system are separated by one order of magnitude. As the Fermi surface lies in the almost linear part of the spectrum, we expect that on the k_f scale the physics resembles the one of graphene. On the q_0 scale on the other hand, the excitation spectrum should be more or less untouched by the doping, and the system will behave as in the intrinsic limit.

We plot Π^R in Fig. 6.4. Comparing panel (a) to Fig. A.2 in App. A, one finds a good agreement with the Dirac case. The biggest deviation is given by the peak of Π^{Im} at Ω^{inter}_{min}, which is not symmetric as for a Dirac system due to the overlap of inter- and intraband spectrum, see Fig. 6.3 (b). The finite quadratic part in the spectrum cures the divergency formerly occurring in the Dirac limit.

The real part of $-\Pi^R$ presented in Fig. 6.4 (b) is strongly negative only at the upper boundary of the intraband spectrum. This indicates that for small interaction strengths, only one plasmon will dominate the excitation spectrum on the Fermi scale. As we are interested in the regime where both inter- and intraband plasmons are visible, we look at the interacting spectrum for interaction strength $\alpha = 10$.

This is done in Fig. 6.5 by plotting $\Pi^{\mathrm{Im}}_{\mathrm{rpa}}$. On the Fermi scale k_f, panel (a), the intraband plasmon absorbs all spectral weight from the intraband spectrum. The plasmon dispersion agrees well with the perturbative one from the expansion in Eq. (6.7) in the limit $X \to 0$, plotted as a gray curve. The green, dashed line shows the linear dispersion of the interband plasmon in the undoped limit, based on Eq. (5.8). On the Fermi scale, it is not obvious that there is an interband plasmon, although the interacting polarization function develops a smeared resonance around the perturbative interband plasmon dispersion for high momenta.

Switching to the intrinsic scale, presented in Fig. 6.5 (b), one finds the interband plasmon as the dominant excitation. It corresponds to the single peak in $-\Pi^{\mathrm{Im}}_{\mathrm{rpa}}$, unperturbed by doping

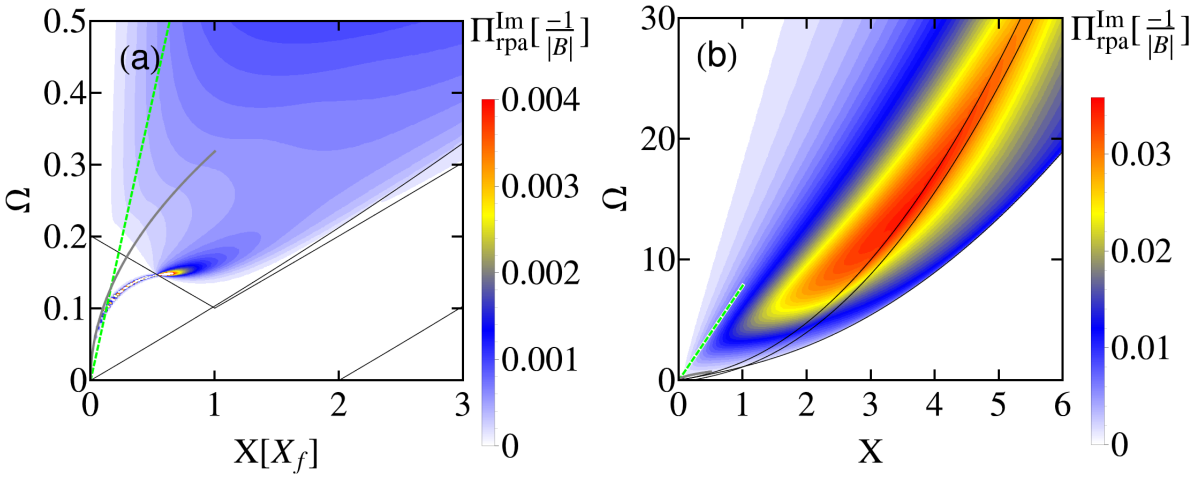


Figure 6.5. $\Pi^{\mathrm{Im}}_{\mathrm{rpa}}$ for $X_f = 0.1$ with $\alpha = 10$ on (a) the k_f scale and (b) the q_0 scale. We add an artificial damping in the regions of $\Pi^{\mathrm{Im}} = 0$ to make the plasmons visible. Reprinted figures with permission from Ref. [JMT14b]. Copyright (2014) by the APS.

for momenta much larger than k_f. Its dispersion is the same as for a plasmon in the undoped limit, discussed in Sec. 5.3.1. The two black lines near the peak are just the boundaries of the intraband excitation spectrum, which does not play a role here.

In the limit of $X \to 0$ the interband plasmon dispersion scales linearly with the interaction strength, $\Omega_p \propto \alpha$, see Eq. (5.8), while the intraband plasmon frequency is proportional to $\sqrt{\alpha}$. Therefore by lowering the interaction strength we can induce a merging of the two resonances below some critical α.

6.3.2. Strong doping of $X_f = 3$

A strong doping of the system significantly increases the total spectral weight, as shown in Fig. 4.2 (b). This increase results from strengthened intraband excitations, while most of the interband excitations get Fermi-blocked, leading therefore to an effective decoupling of the two bands in the BHZ model. Thus we expect the overall spectrum for $X_f = 3$ to be governed by intraband excitations and to resemble the excitation spectrum of a 2DEG, as the Fermi surface lies in the almost quadratic part of the band structure.

The corresponding Π^R is plotted in Fig. 6.6. The single-particle spectrum in panel (a) is peaked at small momenta and at energies close to the upper bound of the intraband spectrum. The interband transitions lead only to minor deviations from the 2DEG case, compare with Fig. A.1 (a).

The real part $-\Pi^R$ in panel (b) is strongly negative at the upper boundary of the intraband spectrum, indicating that only a single intraband plasmon will dominate the interacting spectrum. Interestingly, the static limit property of the polarization being a constant $\tilde{\Pi}(X) = 1$ for $X < 2X_F$, discussed in Sec. 6.1, extends also to an area of finite Ω.

The interacting spectrum is shown in Fig. 6.7 by plotting $\Pi^{\mathrm{Im}}_{\mathrm{rpa}}$ for the interaction strength $\alpha = 10$. Even for this strong Coulomb interaction, we only find the intraband plasmon. This is as expected due to the combined effects of Fermi blocking of interband excitations and increased spectral weight for intraband transitions. The interband plasmon lies in the large overlap of inter- and intraband spectrum, cf. Fig. 6.5 (b), and it is therefore damped beyond

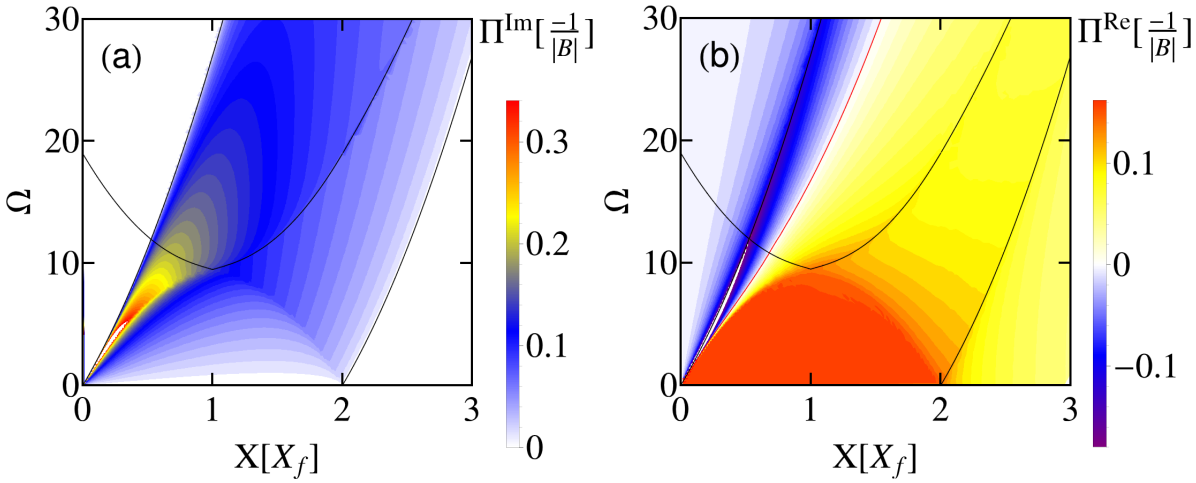


Figure 6.6. The (a) imaginary and (b) real part of the polarization function for $X_f = 3$. The red line indicates $\Pi^{\mathrm{Re}} = 0$. Reprinted figures with permission from Ref. [JMT14b]. Copyright (2014) by the APS.

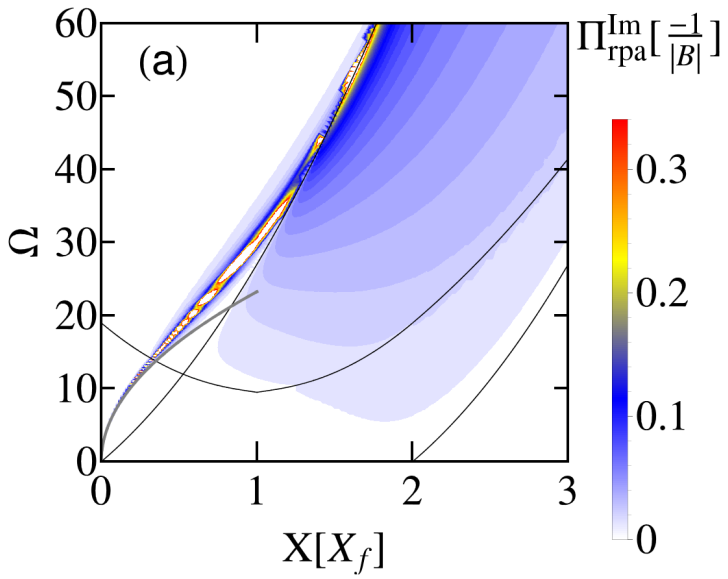


Figure 6.7. Plot of $\Pi^{\mathrm{Im}}_{\mathrm{rpa}}$ for $X_f = 3$ and $\alpha = 10$. We add an artificial damping in the regions of $\Pi^{\mathrm{Im}} = 0$ to make the plasmons visible. Reprinted figure with permission from Ref. [JMT14b]. Copyright (2014) by the APS.

recognition and not visible in the overall spectrum.

6.3.3. Intermediate Doping of $X_f = 1$

For intermediate doping levels like $X_f = 1$, a mixture of Dirac and 2DEG behavior is expected, due to the similar importance of inter- and intraband excitations. The polarization function Π^R is presented in Fig. 6.8. Indeed, the single-particle spectrum in panel (a) looks like a combination of Figs. 6.4 (a) and 6.6 (a). While the shape of the polarization resembles the one of the 2DEG, the interband spectrum is now more pronounced and even dominating for

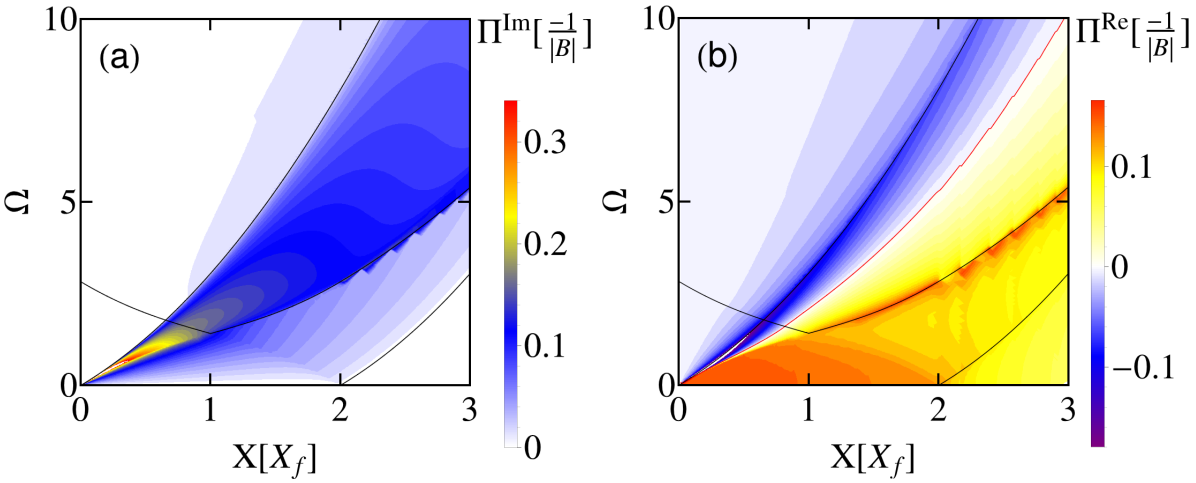


Figure 6.8. (a) Imaginary and (b) real part of the polarization function for $X_f = 1$. The red line indicates $\Pi^{\mathrm{Re}} = 0$. Reprinted figures with permission from Ref. [JMT14b]. Copyright (2014) by the APS.

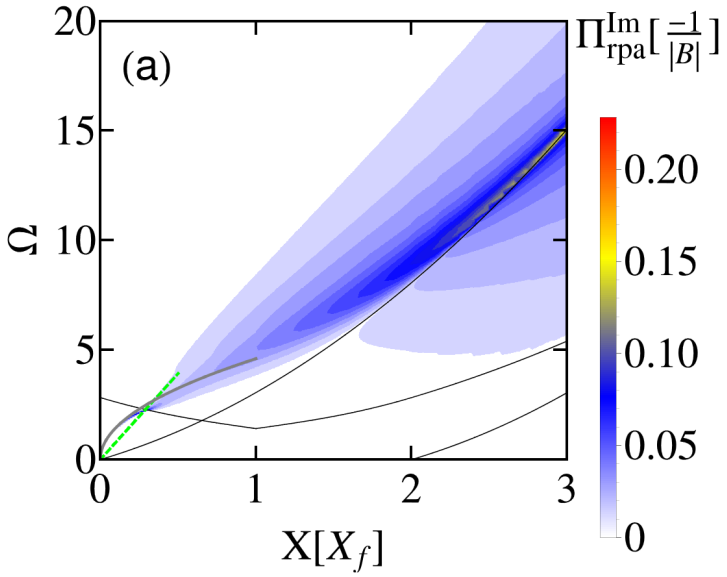


Figure 6.9. Plot of $\Pi^{\mathrm{Im}}_{\mathrm{rpa}}$ for $X_f = 1$ with $\alpha = 10$. We add an artificial damping in the regions of $\Pi^{\mathrm{Im}} = 0$ to make the plasmons visible. Reprinted figure with permission from Ref. [JMT14b]. Copyright (2014) by the APS.

larger momenta of $X > 2X_f$. Therefore we could expect both kinds of excitations giving rise to a plasmon mode. Yet the real part $-\Pi^R$ in panel (b) shows again just a single minimum, following the upper boundary of the intraband spectrum. The deviations from the constant behavior $\tilde{\Pi}(X) = 1$ for $X < 2X_F$ in the case of intermediate doping, analyzed in Sec. 6.1, are also found for finite Ω.

In Fig. 6.9 we show the interacting spectrum by plotting $\Pi^{\mathrm{Im}}_{\mathrm{rpa}}$ for an interaction strength of $\alpha = 10$. It is dominated by a single resonance, lying above the intraband part of the single-particle spectrum. For small momenta, this resonance corresponds to the intraband plasmon

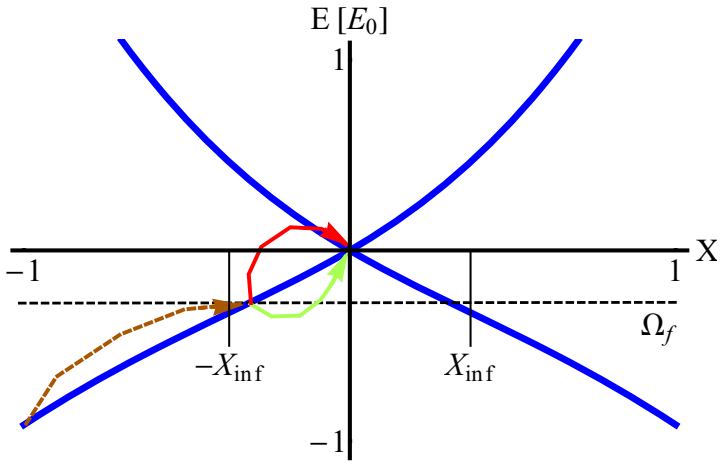


Figure 6.10. Band structure for $\xi_{\mathcal{D}} = -0.5$ and doping to the inflection point. Both low energy interband excitations (red arrow) and high energy intraband excitations (green arrow) involve the Dirac point. Reprinted figure with permission from Ref. [JMT14b]. Copyright (2014) by the APS.

as verified by the perturbative plasmon dispersion (gray line). Yet for intermediate momenta, a comparison with the interband plasmon dispersion in Fig. 6.5 (b) indicates that also the interband plasmon contributes to the resonance. A clear distinction between the two is thus not possible anymore.

In summary, doping the system offers the possibility to change the excitation spectrum on the Fermi scale from a Dirac to a 2DEG type. The interacting spectrum is usually governed by a single intraband plasmon, while the interband plasmon is hidden in the single-particle background. Only large interaction strengths offer a possibility to distinguish both plasmons in the spectrum. In the following, we will now analyze the influence of both broken p-h symmetry, Sec. 6.4, and finite masses, Sec. 6.5. Both options offer a way to separate the two plasmons and make them visible in the total excitation spectrum.

6.4. Hg(Cd)Te Quantum Wells: BHZ Model with finite $\xi_{\mathcal{D}}$

The BHZ model with a broken p-h symmetry and a small or vanishing mass describes the experimental relevant case for Hg(Cd)Te QWs. The finite $\xi_{\mathcal{D}} \neq 0$ offers the possibility of blocking the interband SPE spectrum close to the minimal excitation energy Ω_{min}^{inter}, resulting in less damped interband plasmons as discovered in Sec. 5.3.2.

6.4.1. Excitation Spectrum

Here, we want to use a similar blocking effect for the intraband excitations in order to separate the inter- and intraband spectrum as well as the two plasmon modes. The broken p-h symmetry introduces an inflection point into the band structure, $\frac{\partial^2}{\partial X^2}\varepsilon_{X,\lambda} = 0|_{X=X_{inf}}$, with momentum X_{inf} and energy Ω_{inf}. For $\xi_{\mathcal{D}} < 0$, it lies in the hole part, $\lambda = -1$, of the dispersion relation. With a sufficiently small Fermi momentum, $X_f \lesssim X_{int}$, the highest-energy intraband excitations involve the Dirac point for momenta on the order of the Fermi momentum, see Fig. 6.10. The same is true for the lowest-energy interband excitations. Due to the vanishing DOS at the Dirac point, both kinds of excitations are suppressed, and therefore the inter- and

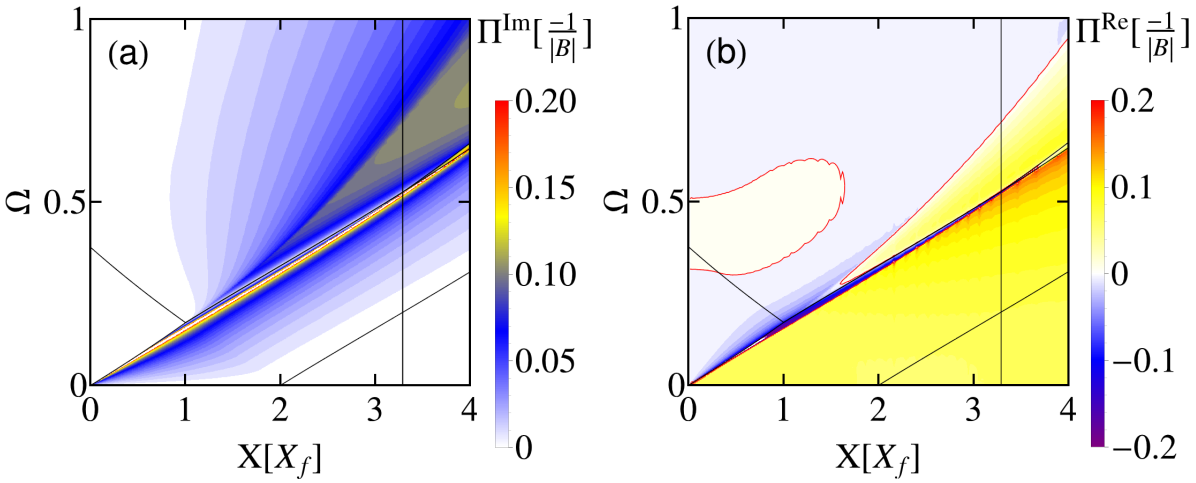


Figure 6.11. Plot of (a) Π^{Im} and (b) Π^{Re} for $X_f = \frac{1}{2}X_{inf} = 0.185$ and $\xi_{\mathcal{D}} = -0.5$. The red line indicates $\Pi^{\mathrm{Re}} = 0$. The separation of inter- and intraband excitations and the existence of two possible plasmon branches is clearly visible. Reprinted figures with permission from Ref. [JMT14b]. Copyright (2014) by the APS.

intraband SPE spectra are effectively separated in energy and momentum up to the scale of X_{inf}.

This effect is demonstrated in Fig. 6.11 (a) and Fig. 6.12 (c)-(e), where the imaginary part of the polarization goes to zero between inter- and intraband parts of the spectrum, fully separating them. There, we choose the Fermi momentum $X_f = \frac{1}{2}X_{inf}$ and parameters $\xi_{\mathcal{D}} = -0.5$ and $\xi_M = 0$. This blocking effect holds for small momenta up to roughly $2X_{inf}$, indicated by the black vertical line in Fig. 6.11 at $X \approx 3.2$. For larger momenta, the high energy intraband excitations go from deep in the valence band directly to the Fermi surface and the blocking effect involving the Dirac point is gone, as depicted by the dashed arrow in Fig. 6.10.

In Fig. 6.11 (b), Π^{Re} shows one major and fundamental difference in comparison to the p-h symmetric case of weak doping in Fig. 6.4 (b). At the border of intra- and interband spectrum, a strong antiscreening region is formed. A plasmon should exist there for sufficiently small α, clearly separated from the second antiscreening region at higher Ω. This gives rise to the possibility of observing both intra- and interband plasmons at the same time.

In order to check this, we plot $\Pi^{\mathrm{Im}}_{\mathrm{rpa}}$ in Fig. 6.12 for $\alpha = 2$ (a) and $\alpha = 0.4$ (b). Panels (c)-(e) show line cuts for fixed momenta $X \in \{1, 1.4, 3\}X_f$ and different $\alpha \in \{0, 0.2, 0.4, 2\}$. For large interaction strength $\alpha = 2$, the intraband plasmon decays into the interband SPE spectrum, see panel (a) for $X \approx X_f$ and the green short dashed line in panel (c). Most of the spectral weight stays there also for larger momenta, as $\Pi^{\mathrm{Im}}_{\mathrm{rpa}}$ is close to 0 in the intraband SPE region and the resonance between inter- and intraband SPE spectrum is weak. The latter can be best seen in the insets of panels (d) and (e), where the green short dashed line is peaked slightly above, panel (d), or below, panel (e), the black vertical line separating intra- and interband SPE regions. Even with the peak being small, it indicates the formation of a slightly damped plasmon, yet with small spectral weight.

The missing spectral weight is transferred to higher energies into the interband SPE region. For intermediate momenta, a second plasmon branch forms as can be seen in panel (a) for $X_f < X < 2X_f$ and from the second peak of the green short dashed line in the inset of panel

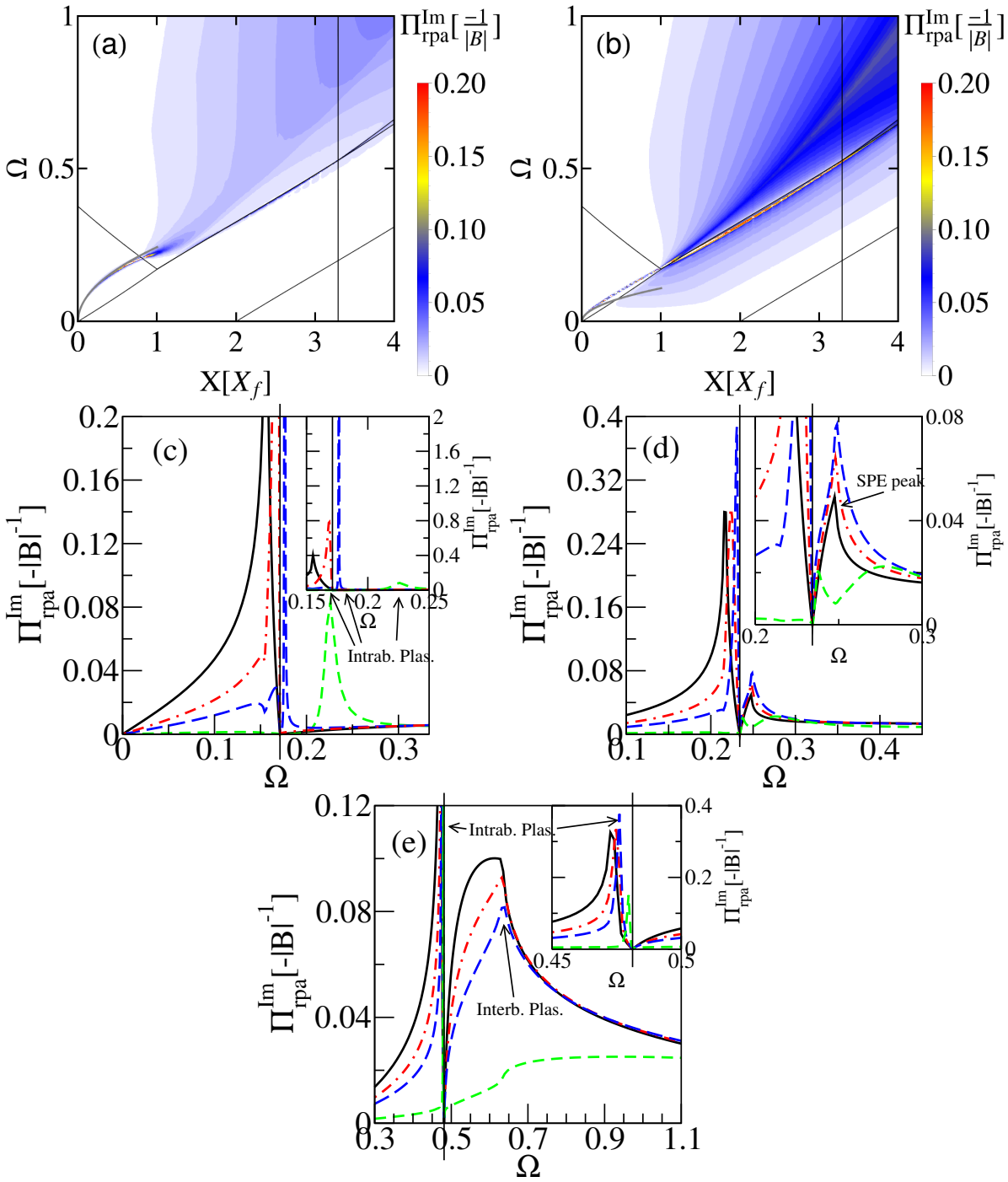


Figure 6.12. Interacting polarization function $\Pi^{\mathrm{Im}}_{\mathrm{rpa}}$ for $X_f = \frac{1}{2}X_{inf} = 0.185$, $\xi_{\mathcal{D}} = -0.5$, $\xi_M = 0$ and $\alpha = 2$ (a) and $\alpha = 0.4$ (b). (c)-(e) show linecuts for fixed $X = X_f$, $X = 1.4X_f$ and $X = 3X_f$, respectively, with $\alpha \in \{0, 0.2, 0.4, 2\}$ in black solid line, red dot-dashed line, blue long dashed line and green short dashed line, respectively. The black, vertical line separates the inter- and intraband SPE region. Reprinted figures with permission from Ref. [JMT14b]. Copyright (2014) by the APS.

(d). At even higher momenta, $X > 2X_f$, the branch overlaps with the forming interband plasmon leading to a broad charge resonance without clear peak, as represented by the green short dashed line in panel (e) for $\Omega > 0.7$.

The picture changes for smaller interaction strength. For $\alpha = 0.4$, the intraband plasmon decays in the region between inter- and intraband SPE spectrum, indicated by the strong peak of the blue long dashed line in panel (c). As the SPEs in this region are suppressed due to the Dirac point, the plasmon leads to a high and narrow peak of $\Pi^{\mathrm{Im}}_{\mathrm{rpa}}$. This resonance splits for higher momenta $X > X_f$: one part forms an intraband plasmon in the intraband SPE region, represented by the blue long dashed line peaked slightly below the black vertical line in panels (d) and (e). The second part stays in the interband SPE region, where it enhances the SPE peak (black line in the inset of panel (d)) for intermediate momenta $X_f < X < 2X_f$. The interband plasmon fully forms for even larger momenta $X \gtrsim 2X_f$, as shown in panel (e). There, the broad single-particle peak (black line) around $\Omega = 0.6$ gets reshaped into a clear peaked resonance (blue long dashed line) - the interband plasmon.

6.4.2. Experimental Parameters

We take the experimental parameters for HgTe QWs from Sec. 4.4, $q_0 \approx 0.4\ {}^1/_{\mathrm{nm}}$ and $E_0 \approx 140$ meV. For the plots in Fig. 6.12 this gives the Fermi momentum $k_f \approx 0.07\ {}^1/_{\mathrm{nm}}$ and the chemical potential $\mu \approx -24\ \mathrm{meV} = -\hbar \cdot 36\ \mathrm{THz}$. The plot range is therefore

$$\begin{aligned} q &\in [0, 0.74]\, q_0 = [0, 0.3]\ {}^1/_{\mathrm{nm}} \\ \omega &\in [0, 1]\, \frac{E_0}{\hbar} = [0, 210]\ \mathrm{THz} \end{aligned} \tag{6.12}$$

and thus of the right order of magnitude for experimental techniques like Raman spectroscopy or electron loss spectroscopy. The interaction strength of $\alpha = 0.4$ is semi-realistic as well.

We compare to the undoped case discussed in Sec. 5.3.3 and find that the scales are the same. Thus in experiments one could tune the Fermi level through the Dirac point and study the interband plasmon alone or its interplay with the intraband plasmon at the same momentum and energy range.

6.4.3. Spectral Weight and the f-sum Rule

The two plasmonic resonances in Fig. 6.12 (b) overlap for $X \approx X_f$, before they separate for higher momenta. Therefore the question arises whether one can really speak of a clear distinction between inter- and intraband plasmons for larger momenta. In order to answer this, we want to study the f-sum rule and thus the spectral weight of the different excitations. As a numerics check, we find the relative deviations of numerical to analytical f-sum rule to be again of the order 10^{-3}, comparable to the undoped case presented in Sec. 5.5, and thus negligible.

We begin our investigation by splitting the f-sum rule over different regions in parameter space. Fig. 6.13 (a) shows the ratio of spectral weight in the interband SPE region for the interacting over the non-interacting case, $\frac{f\sum^{\mathrm{RPA}}_{\mathrm{inter}}}{f\sum^{0}_{\mathrm{inter}}}$, and panel (b) the same for the intraband SPE region. This division excludes the intraband plasmon lying outside of these two regions for small momenta $X < X_f$. For cutoffs $\Lambda = 2\,(\beta X_{max})^2$ with $\beta > 1$ one usually has $f\sum^{0}_{\mathrm{inter}} \gg f\sum^{0}_{\mathrm{intra}}$. Therefore an absolute transfer of spectral weight from one region to the other one can lead to quantitatively different relative changes of spectral weight in panels (a) and (b).

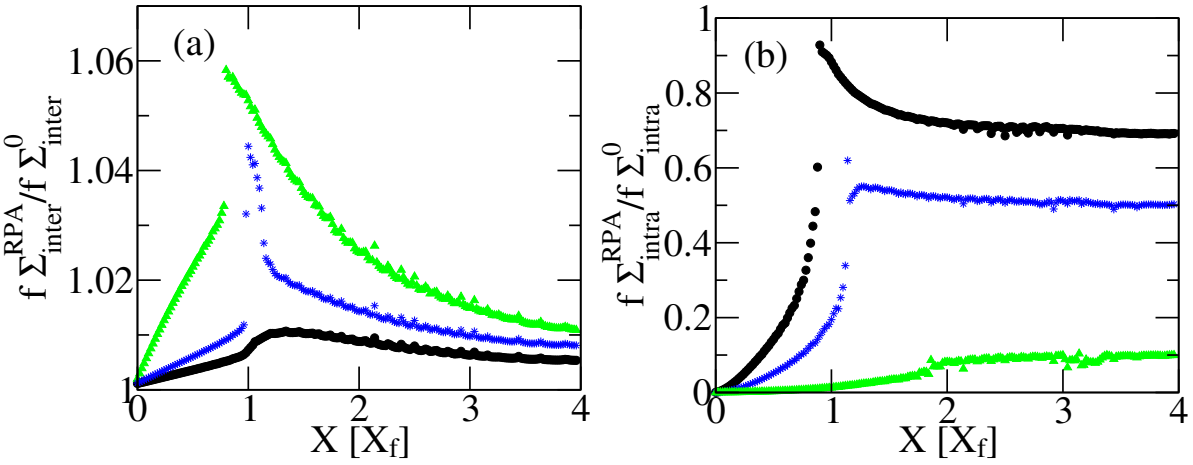


Figure 6.13. (a) The ratio $\frac{f\sum_{\text{inter}}^{\text{RPA}}}{f\sum_{\text{inter}}^{0}}$ of interacting over non-interacting f-sum rule of the interband excitations. Black dots are for $\alpha = 0.2$, blue stars stand for $\alpha = 0.4$ and green triangles for $\alpha = 2$. (b) The same for the ratio $\frac{f\sum_{\text{intra}}^{\text{RPA}}}{f\sum_{\text{intra}}^{0}}$ of the intraband excitations. For all plots: $X_{max} = 4X_f$ and $\beta = 4$. See Sec. 4.5 for definitions. Reprinted figures with permission from Ref. [JMT14b]. Copyright (2014) by the APS.

As a key result of Fig. 6.13, we find that there is always spectral weight missing in the intraband SPE region. For small momenta, $X < X_f$, the weight goes into the undamped intraband plasmon. This follows directly from the conservation of the f-sum rule for interacting and non-interacting systems. At larger momenta, the weight is transferred to higher energies into the interband SPE region. Yet, the increase there is only about 2% at $X \gtrsim 2X_f$. We therefore conclude that the plasmon forming between the inter- and intraband SPE region is a pure intraband plasmon with a reduced spectral weight. The plasmon in the interband region is the interband plasmon we already know from the undoped system, see Fig. 5.6. Here, it gains additional spectral weight from the intraband SPE region.

6.4.4. A small Gap $\xi_M \neq 0$

Deviations in the thickness of Hg(Cd)Te QWs lead to the opening of a small gap in the band structure, resulting in a topological trivial, $\xi_M > 0$, or non-trivial, $\xi_M < 0$, system. In the latter case, one-dimensional (1D) edge states may appear, which we do not want to consider here. The interplay of edge states with the bulk in a 2D TI is a main topic in Sec. 8. Here, we focus on the finite density of states at $X = 0$ generated by a small mass, which works in opposition to the blocking effect of broken p-h symmetry.

In Fig. 6.14 we plot the non-interacting and interacting spectrum for $\xi_{\mathcal{D}} = -0.5$ and a small mass $\xi_M = 0.01 \approx 1.4\ {}^{\text{meV}}/_{E_0}$. A comparison with Figs. 6.11 and 6.12 shows that the small mass has just the effect of an additional separation of the inter- and intraband SPE region.

Thus we conclude that our idea of observing both plasmons simultaneously in experiments is robust against slight deviations in the mass and therefore the thickness of the Hg(Cd)Te QWs.

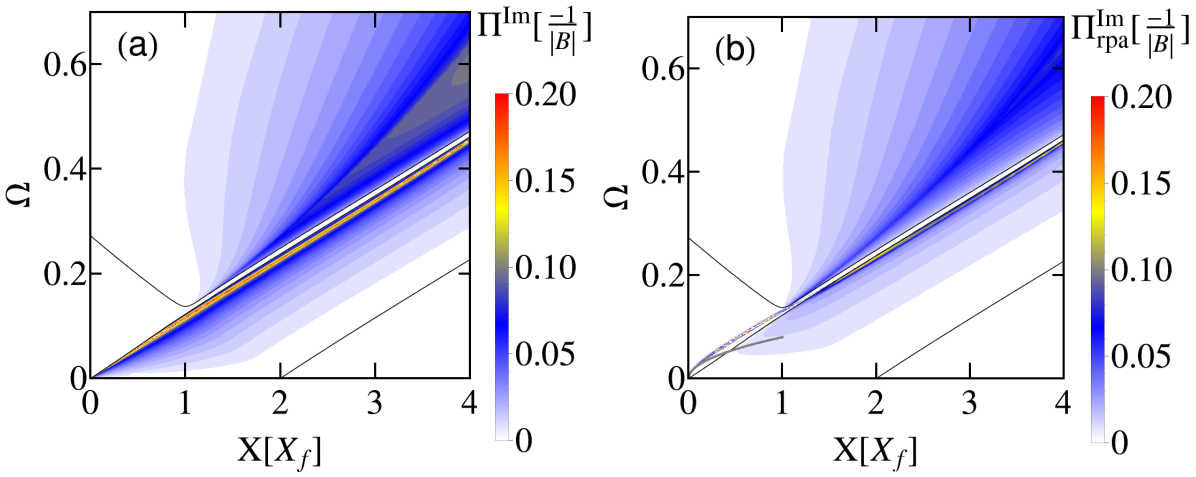


Figure 6.14. Plot of (a) Π^{Im} and (b) $\Pi^{\mathrm{Im}}_{\mathrm{rpa}}$ with interaction strength $\alpha = 0.4$. In both cases, we choose $X_f = 0.133 < \frac{1}{2}X_{inf}$, $\xi_M = 0.01$ and $\xi_{\mathcal{D}} = -0.5$. Reprinted figures with permission from Ref. [JMT14b]. Copyright (2014) by the APS.

6.5. Topology: BHZ Model with large Masses

A finite Dirac mass opens a gap in the band structure and changes the pseudospin, and therefore the overlap factor (4.9), in a non-trivial fashion. Thus we can expect in general a quite different behavior for positive and negative masses. Yet, for these differences to occur on the intrinsic scale and thus influence the interband plasmons, $|\xi_M|$ needs to be of the order of 1.

In the following, we study such large masses, both negative and positive, with p-h symmetry. While not experimentally relevant for HgTe QWs, they offer the possibility to study the effects of topologically distinct band structures on the electronic excitations, including plasmons. We also note here that the dispersion of the BHZ model becomes purely parabolic for the mass $\xi_M = -\frac{1}{4}$,

$$\epsilon_{X,\lambda} \underset{\xi_M=-\frac{1}{4}}{=} \frac{\lambda}{4} + X^2\left(\lambda - \xi_{\mathcal{D}}\right). \tag{6.13}$$

In this limit, the polarization function (4.7) can be calculated analytically.

6.5.1. Topologically non-trivial: Large, negative Mass

We begin with the TnT regime. For the parameters $X_f = 0.33$ and $\xi_M = -\frac{4}{9}$, Fig. 6.15 (a) and (b) show the polarization Π^R. The large mass separates intra- and interband SPE regions for momenta $X \lesssim 2X_f$. Compared to the massless cases of $X_f = 0.1$, Fig. 6.4, and $X_f = 1$, Fig. 6.8, the interband SPE spectrum is enhanced due to the combination of larger overlap factor and low doping, thus small Fermi blockade. As a result of the flat band structure, the chemical potential is just barely above the gap even for $X_f = 0.33$. An interesting consequence of these strong interband transitions can be seen in panel (b), where we find two distinct areas where $-\Pi^{\mathrm{Re}}$ becomes negative. As a consequence, inter- and intraband plasmons will always be separated, with the intraband plasmon being confined to low energies. This stems from the fact that the electrons in the conduction band are pseudospin polarized. Intraband excitations

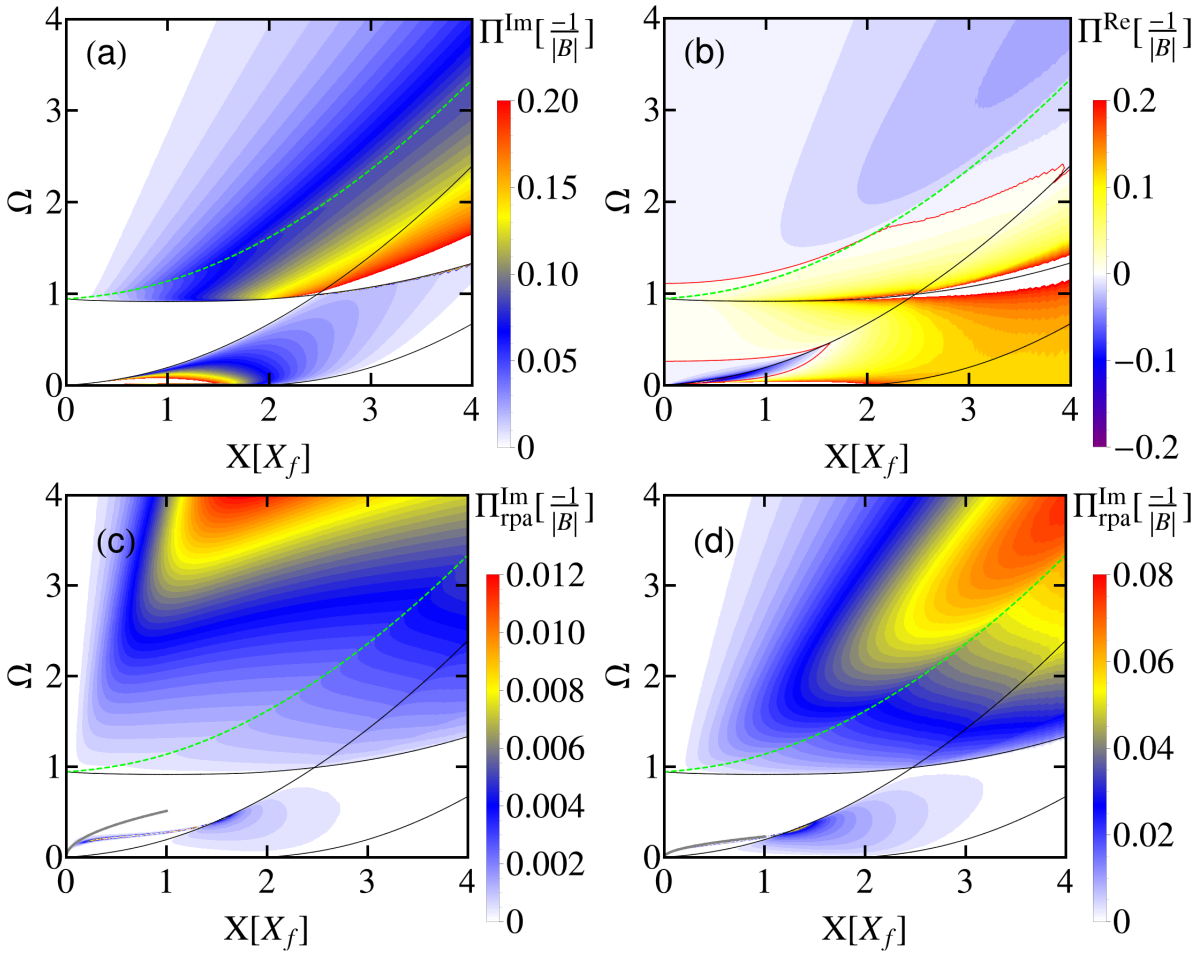


Figure 6.15. (a) Imaginary and (b) real part of the polarization Π^R. The two antiscreening areas are clearly distinct. (c) and (d) show $\Pi^{\mathrm{Im}}_{\mathrm{rpa}}$ with $\alpha = 10$ and $\alpha = 2$, respectively. The two plasmonic resonances can be uniquely identified. $\xi_M = -\frac{4}{9}$ and $X_f = 0.33$ in all plots. Reprinted figures with permission from Ref. [JMT14b]. Copyright (2014) by the APS.

to much higher momenta and energies, where the pseudospin shows in the opposite direction, are not possible.

This is confirmed in panels (c) and (d) of Fig. 6.15, where we plot $\Pi^{\mathrm{Im}}_{\mathrm{rpa}}$ with $\alpha = 10$ and $\alpha = 2$, respectively. All the spectral weight of the intraband SPE region goes into the plasmon, which, at least for $\alpha = 2$, follows very well the $\sqrt{X}$ law. The interband spectrum is dominated by the interband plasmon, having of course a much broader peak due to damping (finite Π^{Im}).

The dashed, green line in the interband spectrum in Fig. 6.15 indicates the energy, at which excitation processes going from momentum $\boldsymbol{X} + \boldsymbol{X}_f$ in the valence band to $\boldsymbol{X}_f$ in the conduction band, with $\boldsymbol{X} \| \boldsymbol{X}_f$, are possible. They are represented by the black, dashed arrow in Fig. 6.3 (a). Usually suppressed by opposing pseudospins, a large negative mass enhances the overlap factor of these excitations to near unity for small Fermi momenta. Fig. 6.15 (b) and (d) show that the interband plasmons mainly occur above this line, indicating that the described excitation process is important for the collective excitation. As the process is forbidden by helicity in the pure Dirac system, it is one reason why the BHZ model supports intrinsic

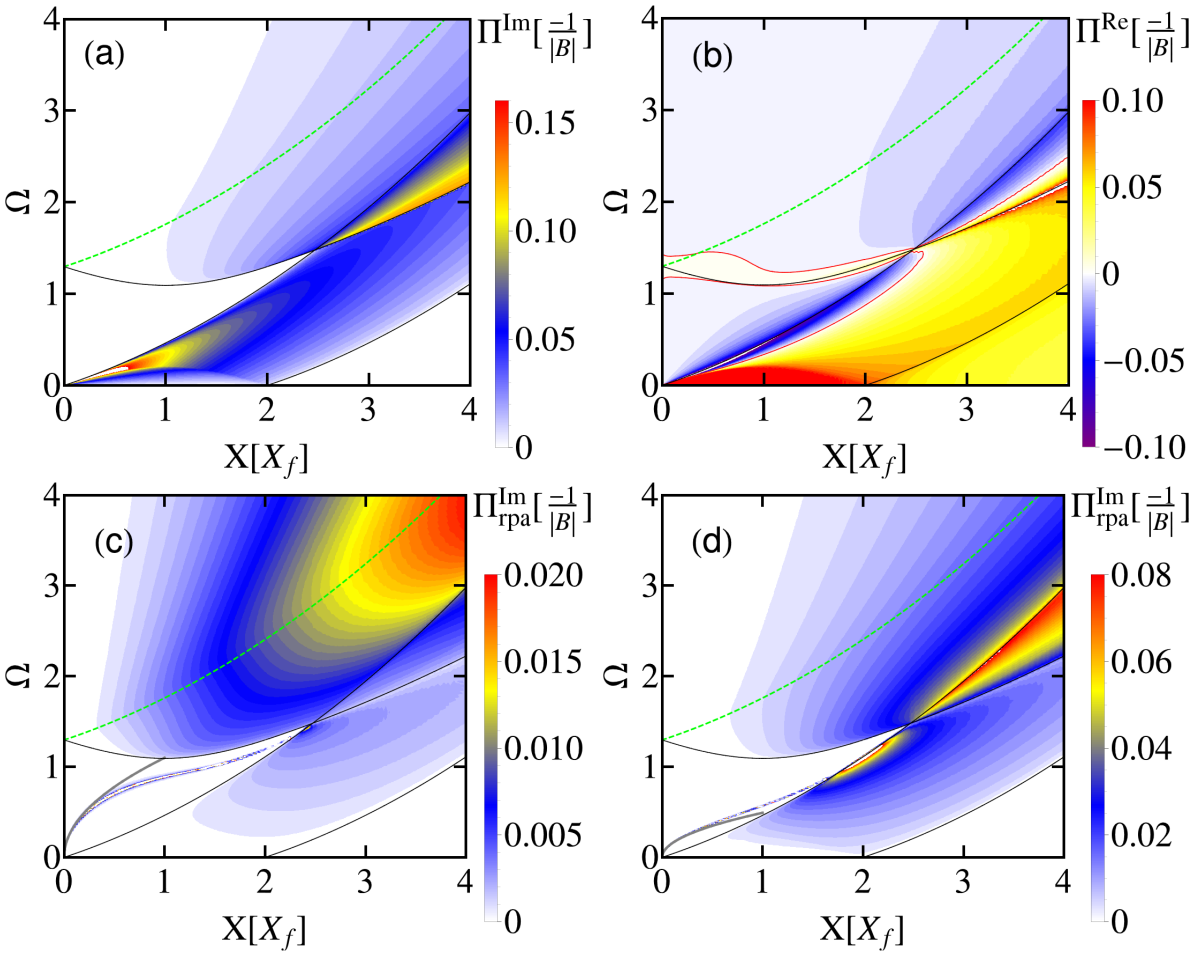


Figure 6.16. (a) Imaginary and (b) real part of the polarization Π^R. (c) and (d) show $\Pi^{\mathrm{Im}}_{\mathrm{rpa}}$ with $\alpha = 10$ and $\alpha = 2$, respectively. $\xi_M = \frac{4}{9}$ and $X_f = 0.33$ in all plots. Reprinted figures with permission from Ref. [JMT14b]. Copyright (2014) by the APS.

plasmons while the Dirac model does not.

6.5.2. Topologically trivial: Large, positive Mass

We switch now to the case of a large, trivial mass term. For the parameters $X_f = 0.33$ and $\xi_M = \frac{4}{9}$ we plot the polarization function in Fig. 6.16 (a) and (b). Compared to the case of negative mass, the interband spectrum is much weaker. This is a result of the lower overlap factor, partly decoupling the two bands, as well as of the higher chemical potential. The band structure is not as flat as in the TI phase, thus for the same Fermi wave vector $X_f = 0.33$ one gets a stronger Fermi blockade. For the real part of the polarization, this has the effect that the two former distinct areas of sign reversal, see Fig. 6.15, now almost merge. The interband excitations are so weak that the minimum $-\Pi^{\mathrm{Re}}$ always lies closely above the intraband SPE region - indicating that the latter is the main source for plasmons.

Taking a look at the interacting excitation spectrum, in Fig. 6.16 panel (c) for $\alpha = 10$ one can identify both inter- and intraband plasmon. Interestingly, the polarization is clearly higher in the pure interband SPE region than in the mixed inter- and intraband SPE spectrum,

suggesting that the latter one serves as an additional damping for the interband plasmon. Going to the smaller interaction strength $\alpha = 2$ in panel (d), one finds just a single resonance following the upper boundary of the intraband SPE spectrum.

Thus we conclude that the interacting spectrum for moderate interaction strengths is governed by just the intraband plasmons. The interband excitations are too weak to support an additional plasmon but for very high interactions - a consequence of the effective decoupling of the bands by the overlap factor.

Conclusion

We have analyzed the dynamical and static polarization properties of the BHZ model based on RPA, with the experimental realization of Hg(Cd)Te QWs in mind. In the static undoped limit, due to the presence of quadratic terms in the model and hence to the natural charge decay length $1/q_0 = |\mathcal{B}|/\mathcal{A}$, the induced charge density in response to a test charge has a finite spatial extent. This is in contrast to the point-like screening charge obtained with the continuous Dirac model of graphene. In the doped regime, we have observed Friedel oscillations with an intermediate decay behavior between the Dirac (r^{-3}) and the 2DEG (r^{-2}) limits.

The discussion of the full dynamical polarization function has been focused on the appearance of new interband plasmons due to the interplay of Dirac and Schrödinger physics. In principle, we expect these plasmons to appear in all multiband systems where the imaginary part of the polarization function decays faster with energy than the one in the Dirac case (ω^{-1}). For the BHZ model, with a decay as ω^{-2}, this is fulfilled. These plasmons appear already in the undoped system at experimentally relevant parameters, but it is also possible to observe them in the doped regime, where they coexists with the usual intraband plasmons. This is favored by broken p-h symmetry in the BHZ model, which allows for the presence of both a Dirac point and an inflection point in the band structure, giving rise to a crucial blocking of the SPEs. These new plasmons should appear for momenta and energies on the right order of magnitude for experimental techniques like Raman spectroscopy or electron loss spectroscopy on Hg(Cd)Te QWs.

The behavior of these two collective modes is also influenced by the topology of the band structure. The two plasmons tend to merge into one another in a trivially gapped insulator, while they remain distinct resonances, separated in energy, in the TI phase. Thus the plasmons are a bulk quantity that directly depends on the topological state of the system. Furthermore, we have predicted that a measurement of the optical conductivity at finite frequency, above the bandgap of $\Omega = |2\xi_M|$, yields an additional and more direct way to distinguish between the two topological phases of Hg(Cd)Te QWs.

The wide range of parameters considered in this chapter, including the regime of topological trivial and non-trivial insulators, should make our results applicable to all kinds of materials described by phenomenological models interpolating between Dirac and Schrödinger fermion physics.

Throughout this chapter, we have only discussed bulk excitations of this peculiar 2D system. Hence, we have totally ignored the influence of edge states in the TnT regime of the model in the presence of physical boundaries. This will change in the next chapter III, where in the quantum Hall ferromagnetic (QHFM) regime both 2D bulk and 1D edge will be treated together on the same footing.

Part III.

Quantum Hall Ferromagnetism in 2D Topological Insulators

Most of the materials to date that have been experimentally established as topological insulators (TIs), such as HgTe/CdTe, BiSb, or BiSe compounds as discussed in Sec. 2.4, are rather weakly interacting due to efficient screening of the Coulomb interactions. These correlation effects can thus be treated perturbatively as we did based on the Random Phase Approximation (RPA) in the last Chapter II. Nevertheless, strongly interacting topological systems are currently an active area of research, see Ref. [HA13] for a comprehensive review. Of particular interest are the situations, when interactions change the single-particle picture in a qualitative way and lead to effects not present in the non-interacting system. Theoretical proposals of such a non-trivial behavior include a "topological Mott insulator" presented in Ref. [PB10], a "topological Kondo insulator" discussed in Refs. [DSGC10, DXGC16] and different interaction-induced topological phases in graphene, analyzed in Refs. [LAF12, CTV12]. First-order topological phase transitions driven by interactions are investigated by the authors of Ref. [ABC+15], and the stability of one-dimensional (1D) (helical) edge channels with respect to different scattering mechanisms in the presence of disorder is the topic of a number of publications [XM06, WBZ06, SJJ10, SRvOG12, BDRT12, CBD+12, LOB12, GCT14, KGCM14].

Usually these predictions require sufficiently strong electron interactions and were made for strongly correlated materials. With respect to the naturally weakly interacting established TIs, it is thus desirable to expand the range of possibilities to attain the regime of strong effective interactions in topological systems. Even better, we would like to be able to tune the strength of interactions by experimentally feasible means. In the following sections, we identify a class of topological systems in which these conditions can be realized even for weak bare interactions, by applying an orbital magnetic field.

In these quantum Hall topological insulators (QHTIs), the interactions are tunable by the magnetic field and their strength is controlled by the proximity to the topological phase transition. The vicinity of the transition is automatically the regime of strong effective interactions, in which Coulomb interactions are crucial for both bulk and edge properties and lead to a non-trivial interplay of topological and interacting phenomena.

An important theoretical advantage of such a system is that it can be analyzed in a well-controlled way. In particular, this allows us to determine the symmetry requirements for topological protection in this system, which is one of the key questions raised in the studies of interacting topological systems. Based on the nonlinear σ-model formalism, we analyze the bulk and edge properties and establish that an effective continuous U(1) symmetry with respect to uniaxial isospin rotations must be preserved. In this case, the topologically non-trivial (TnT) phase persists and the edge is a helical Luttinger liquid with highly tunable effective interactions. However, U(1) symmetry may be broken, either spontaneously or by U(1)-asymmetric interactions. In either case, interaction-induced transitions occur to the respective topologically trivial (TT) phases with gapped edge charge excitations.

The work presented in the following two sections has been published by us in Ref. [KJT16]. In Sec. 7, a general introduction to the considered system class of QHTIs is given and the effective low-energy nonlinear σ-model is derived. Sec. 8 focuses then on the physics contained in the model. Starting from a discussion of the bulk, the edge is introduced and the excitations of the combined system are analyzed. For preserved U(1) symmetry, the σ-model is mapped to a helical Luttinger liquid Hamiltonian, in which the interaction parameter $\mathcal{K}$ is tunable by the magnetic field applied to the system. Special attention is paid to the role of the U(1) symmetry offering topological protection in this highly interacting phase with broken time-reversal ($\boldsymbol{\mathcal{T}}$) symmetry.

7. Setup and Mapping to nonlinear σ-model

This section introduces the system class of QHTIs in Sec. 7.1, which are TnT states in the presence of a strong magnetic field. In the forming Landau levels (LLs), the kinetic energy is quenched and the system is in a highly ordered quantum Hall ferromagnetic (QHFM) ground state, see Sec. 3.2.3 for an introduction. The Hamiltonian projected onto these LLs is discussed in Sec. 7.2, where we identify the important symmetries responsible for the topological protection of the system. We close in Sec. 7.3 with the derivation of the low-energy theory for this QHFM state, on which the physics discussion in Sec. 8 is built on.

7.1. Quantum Hall Topological Insulators

The two-dimensional (2D) electron systems we consider in the following have zero Chern number $n = 0$ at finite magnetic field B, but can still be TnT and exhibit helical edge states. We call this class of systems QHTIs. Since the $\boldsymbol{\mathcal{T}}$ symmetry is broken by the magnetic field, the TnT phase must be protected by some other symmetry. A few possible examples were already discussed when we looked at topological crystalline insulators (TCIs) in Sec. 2.4.3. In a system with considerable spin-orbit coupling (SOC), such a symmetry is some spatial symmetry like inversion ($\boldsymbol{\mathcal{I}}$), reflection, or rotation. In the case that SOC is negligible, an axial spin rotation symmetry can also play the role of such a symmetry. We will refer to this symmetry, responsible for the topological protection of a *non-interacting* QHTI, as the *physical symmetry*. In contrast to it, we will demonstrate that for an *interacting* system the *effective*, emergent U(1) symmetry is of central importance.

A possible type of LL structure of a QHTI is plotted in Fig. 7.1. Due to the TnT phase, a crossing of two LLs at some value B^* of the magnetic field occurs. The band inversion ensures that one LL, labeled a, originates from the valence band and moves upwards with increasing B, while the other LL, labeled b, originates from the conduction band and moves downwards. The crossing of the two levels marks the point of the topological phase transition of a QHTI. It separates the TnT phase with counter-propagating edge states at lower $B < B^*$ from the TT phase with gapped edge states at higher $B > B^*$, see Fig. 7.2. Other variants of the LL structure in QHTIs are also possible.

A number of previously studied theoretical models and real physical systems are relevant to the class of QHTIs. The single-particle behavior depicted in Figs. 7.1 and 7.2 has been identified in the BHZ model for an applied magnetic field perpendicular to the 2D structure, as indicated in Fig. 2.14. This behavior is likely to have a topological origin and we expect it to be protected by a spatial symmetry. The BHZ model is directly relevant to HgTe/CdTe and InAs/GaSb heterostructures, see Sec. 2.4.1, which are established 2D TIs at zero magnetic field, protected by $\boldsymbol{\mathcal{T}}$ symmetry. Other likely QHTI systems are single- and multi-layer structures of graphene. Non-interacting graphene exhibits counter-propagating edge states at finite magnetic field due to the spin splitting by the Zeeman field and its semimetallic character, see Ref. [ALL06]. It can be seen as a QHTI protected by the continuous axial spin rotation symmetry, as proposed in Ref. [YSYH+13]. Although directly relevant, graphene also has a few peculiarities and its

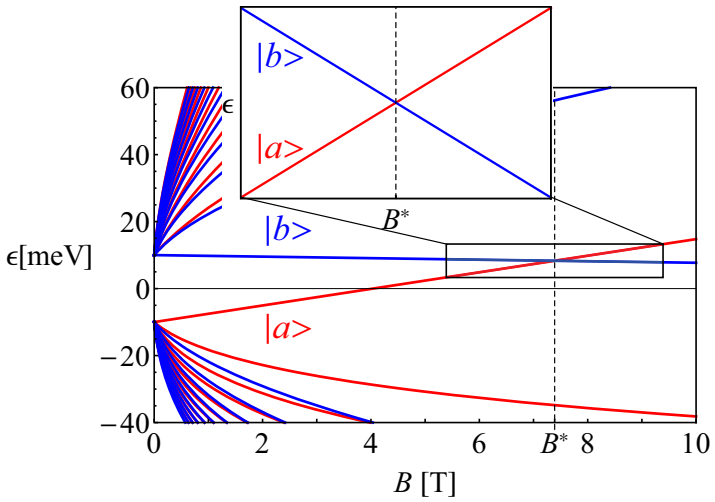


Figure 7.1. One possible type of LL structure of a QHTI. At the point of the single-particle topological phase transition, LLs of different symmetry, labeled a (red) and b (blue), cross. The spectrum shown is calculated for the BHZ model with a perpendicularly applied magnetic field, see Sec. 2.4.1 for formulas and references. In this case, LLs a and b are distinguished by the spatial $\boldsymbol{\mathcal{I}}$ parity, even and odd. Reprinted figure with permission from Ref. [KJT16]. Copyright (2016) by the APS.

LL structure differs from that in Fig. 7.1. In the following, we focus on QHTIs with a spectrum of the type drawn in Fig. 7.1.

We emphasize that according to the above definition, QHTIs are not necessarily new topological systems symmetry-wise, in regard to the existing classifications discussed in Secs. 2.2.5 and 2.4.3. The key requirement here is the explicitly present orbital magnetic field, with the system being in the quantum Hall (QH) regime. This leads to physical phenomena stemming mainly from the flat-band property of the LL spectrum in the bulk (Fig. 7.2), that are specific to the QH regime and may hardly be realized otherwise.

7.1.1. Quantum Hall Ferromagnet at the Topological Phase Transition

The QHTIs with the LL structure as in Fig. 7.1 are particularly appealing for the study of the interplay of interactions and topology. Due to the near degeneracy of the two LLs in vicinity of the single-particle topological phase transition point B^*, electron interactions become the dominant effect driving the physics there. Thus, the effective interactions in QHTIs are tunable through the magnetic field, making the regime of *strong effective* interactions experimentally accessible even in a system with *weak bare* interactions. We thus investigate in the following sections the effect of electron interactions on the topological properties of a QHTI with the LL structure as shown in Fig. 7.1. Particular attention is paid to the regime of strong effective interactions in the vicinity of the single-particle topological phase transition.

We follow the quantum Hall ferromagnetism (QHFMism) reasoning laid out in Sec. 3.2.3. The zero Chern number $n = 0$ corresponds to half-filling of the two crossing LLs a and b, with an average of one electron per two states. At such a commensurate filling factor, interactions make the electron system particularly prone to polarization in the 2D ab space, analogous to the Hund exchange mechanism in atoms. This results in the formation of a ferromagnetic ground state as sketched in Fig. 7.3. The electrons in each orbital occupy exactly the same

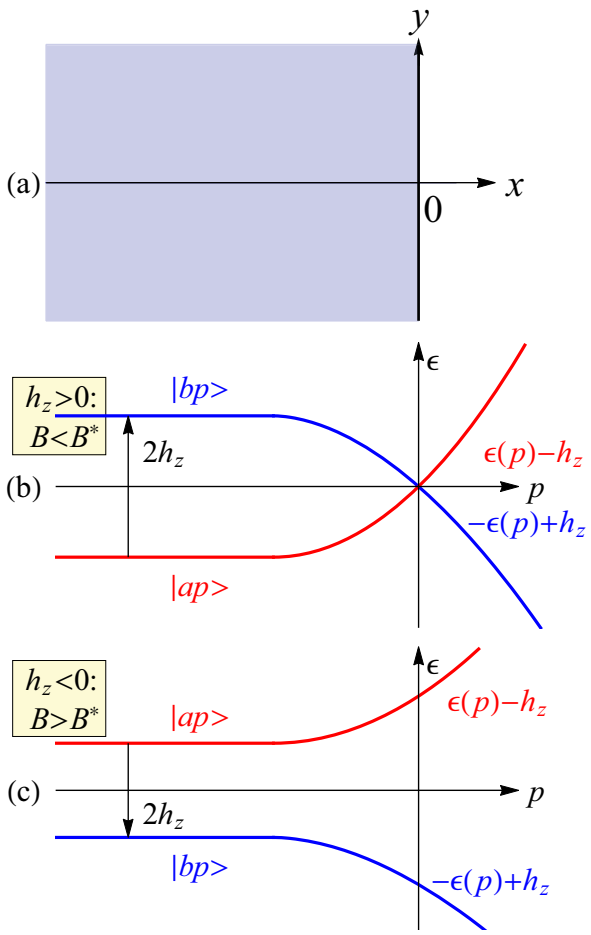


Figure 7.2. (a) Half-infinite sample occupying the region $x \leq 0$. (b) and (c) Schematics of the edge spectrum of a QHTI in the Landau gauge, including only the two intersecting LLs a and b of interest, see Fig. 7.1. The single-particle states are labeled by the conserved 1D momentum p in the y direction. The states located in the bulk and at the edge correspond to the values $p \lesssim 0$ and $p \gtrsim 0$, respectively. In the TnT phase at lower fields $B < B^*$ (b), the edge states cross and are gapless. In the TT phase at higher fields $B^* < B$ (c), the edge states do not cross and are gapped. Reprinted figure with permission from Ref. [KJT16]. Copyright (2016) by the APS.

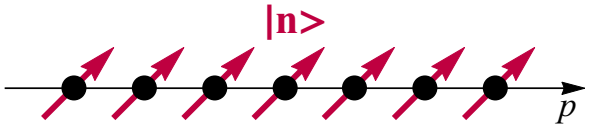


Figure 7.3. QHFM state realized at the crossing of LLs at half-filling. For each momentum p, one electron occupies the state $|\boldsymbol{n}\rangle$ with isospin $\boldsymbol{n}$, see Eqs. (7.1) and (7.2) and Fig. 7.4. Reprinted figure with permission from Ref. [KJT16]. Copyright (2016) by the APS.

state

$$
\begin{aligned}
|\boldsymbol{n}\rangle &= \chi_a(\boldsymbol{n})|a\rangle + \chi_b(\boldsymbol{n})|b\rangle, \\
\chi(\boldsymbol{n}) &= \begin{pmatrix} \chi_a(\boldsymbol{n}) \\ \chi_b(\boldsymbol{n}) \end{pmatrix} = \begin{pmatrix} \mathrm{e}^{-\frac{\mathrm{i}}{2}\varphi} \cos\frac{\theta}{2} \\ \mathrm{e}^{\frac{\mathrm{i}}{2}\varphi} \sin\frac{\theta}{2} \end{pmatrix}.
\end{aligned}
\tag{7.1}
$$

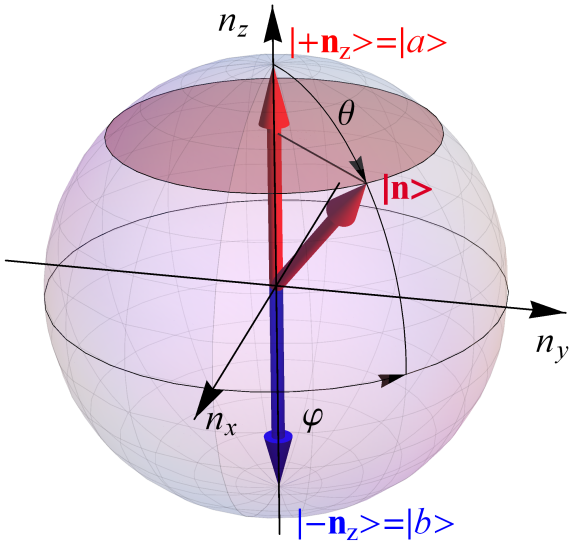


Figure 7.4. The isospin $\boldsymbol{n}$ in the 2D space spanned by the LLs a and b, Eqs. (7.1) and (7.2), on the Bloch sphere. The north and south poles, $|+\boldsymbol{n}_z\rangle = |a\rangle$ and $|-\boldsymbol{n}_z\rangle = |b\rangle$, correspond to the occupation of the a and b LLs, respectively. Any other state is a coherent superposition of the two. Reprinted figure with permission from Ref. [KJT16]. Copyright (2016) by the APS.

The latter is characterized by the unit-vector isospin order parameter (OP)

$$\boldsymbol{n} = (n_x, n_y, n_z) = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta),\ \ \boldsymbol{n}^2 = 1, \tag{7.2}$$

$$|\boldsymbol{n}\rangle\langle\boldsymbol{n}| = \frac{1}{2}(\tau_0 + \boldsymbol{\tau}\cdot\boldsymbol{n}), \tag{7.3}$$

depicted on the Bloch sphere in Fig. 7.4. Throughout this chapter, τ_0 and $\boldsymbol{\tau} = (\tau_x, \tau_y, \tau_z)$ denote the unity and Pauli matrices in the ab space.

Exactly at the crossing point of the two LLs, if we approximate all interactions in the ab space to be SU(2)-symmetric, the isospin $\boldsymbol{n}$ can be completely arbitrary. Thus fixing the state $|\boldsymbol{n}\rangle$ describes the spontaneous breaking of this SU(2) symmetry. This phenomenon is referred to as QHFMism, as introduced in Sec. 3.2.3. All the bulk single-particle and interaction effects responsible for the deviation from this fully degenerate SU(2)-symmetric situation, as well as the effect of the edge, can be taken into account within a low-energy field theory. Such a nonlinear σ-model for the isospin OP $\boldsymbol{n}(\boldsymbol{r};t)$, here developed for configurations slowly varying in time and space, will be derived in the following. Crucially, the effect of the edge is incorporated as a boundary condition for the OP. This approach allows us to study most properties of interest analytically.

7.2. Projected Hamiltonian

We begin the derivation by considering all terms in the Hamiltonian relevant for our two LLs a and b. The σ-model for the OP $\boldsymbol{n}$ is then deduced from this reduced Hamiltonian in Sec. 7.3.

7.2.1. Restricted Hilbert Space of two intersecting Landau Levels

In the following, we take into account only the two intersecting LLs a and b in Fig. 7.1, neglecting all other LLs. This is a standard approximation for QH systems, justified for weak Coulomb interactions. In this limit, the energy separation between the LLs of interest and other LLs is much larger than the interaction energy scale

$$\frac{e_*^2}{l_{B_z}} = \frac{e^2}{4\pi\varepsilon_0\varepsilon_r}\frac{1}{l_{B_z}} \tag{7.4}$$

set by the Coulomb energy. Here, e_* is the electron charge screened by the dielectric environment with the constant ε_r, and

$$l_{B_z} = \sqrt{\frac{\hbar}{eB_z}} \tag{7.5}$$

is the magnetic length with B_z the component of the magnetic field perpendicular to the sample plane, $B^2 = B_z^2 + B_\parallel^2$. The magnetic field can be arbitrarily orientated relative to the quasi 2D sample. We work in the Landau gauge, in which the single-particle states are characterized by the 1D momentum p along the edge y direction. The two single-particle states of the LLs of interest are

$$|ap\rangle \quad \text{and} \quad |bp\rangle, \tag{7.6}$$

where we assume no discrete degeneracies, such as valleys, of these LLs. The 2D sample under consideration is half-infinite with $x < 0$, see Fig. 7.2. The states with $p \lesssim 0$ are then the bulk states, for which the usual coordinate-momentum correspondence holds, while the states with $p \gtrsim 0$ correspond to the edge states, localized over l_{B_z} near the edge. The corresponding electron annihilation operators are denoted as c_{ap} and c_{bp}, which we join into the two-component spinor

$$\hat{c}_p = \begin{pmatrix} c_{ap} \\ c_{bp} \end{pmatrix} \tag{7.7}$$

for compactness.

7.2.2. U(1)-symmetric projected Hamiltonian

We first consider the following many-body projected Hamiltonian, operating within the states (7.6) of the intersecting LLs:

$$\hat{H} = \hat{H}_{1\circ} + \hat{H}_{1\circ}^{\text{edge}} + \hat{H}_{2\odot} + \hat{H}_{2\circ}, \tag{7.8}$$

$$\hat{H}_{1\circ} = -h_z \sum_p \hat{c}_p^\dagger \tau_z \hat{c}_p, \tag{7.9}$$

$$\hat{H}_{1\circ}^{\text{edge}} = \sum_p \epsilon(p) \hat{c}_p^\dagger \tau_z \hat{c}_p, \tag{7.10}$$

$$\hat{H}_{2\odot} = \frac{1}{2} \sum_{p_1+p_2=p_1'+p_2'} V\binom{0}{0}^{p_1 p_1'}_{p_2 p_2'} : [\hat{c}_{p_1}^\dagger \hat{c}_{p_1'}][\hat{c}_{p_2}^\dagger \hat{c}_{p_2'}] :, \tag{7.11}$$

$$\hat{H}_{2\circ} = \frac{1}{2} \sum_{p_1+p_2=p_1'+p_2'} \sum_{\alpha=x,y,z} V\binom{\alpha}{\alpha}^{p_1 p_1'}_{p_2 p_2'} : [\hat{c}_{p_1}^\dagger \tau_\alpha \hat{c}_{p_1'}][\hat{c}_{p_2}^\dagger \tau_\alpha \hat{c}_{p_2'}] :, \quad V\binom{x}{x}^{p_1 p_1'}_{p_2 p_2'} = V\binom{y}{y}^{p_1 p_1'}_{p_2 p_2'}. \tag{7.12}$$

The labels 1 and 2 designate single-particle and two-particle interaction terms, while $\circ$ and $\odot$ denote U(1)-symmetric and SU(2)-symmetric terms, respectively. The part (7.9) describes

the energy spacing of $2h_z$ between the two LLs of interest. The energy $h_z(B)$ is a function of the magnetic field B, as can be expected from Figs. 7.1 and 7.2. It decreases monotonically with increasing magnetic field, starting from positive values in the TnT regime and changing to negative values at the crossing point $B = B^*$. Close to the transition, one may expand it to linear order as

$$h_z(B) \approx -|\partial_B h_z(B^*)|(B - B^*). \tag{7.13}$$

Next, the single-particle effect of the edge is described by the term (7.10). The dispersion function $\epsilon(p) > 0$ is shown schematically in Fig. 7.2. It has a plateau $\epsilon(p) \approx 0$ in the bulk ($p \lesssim 0$), but grows monotonically at the edge $p \gtrsim 0$. This forces the bands at the edge into the normal, TT order. The two branches at the edge do not have to be exactly particle-hole (p-h) symmetric, so an additional energy term $\epsilon_0(p)\tau_0$ could be added. Yet it would produce only a trivial, $\boldsymbol{n}$-independent term in the σ-model derived below, thus we neglect it.

The single-particle spectrum $\pm[-h_z + \epsilon(p)]$ of $\hat{H}_{1\circ} + \hat{H}^{\text{edge}}_{1\circ}$ describes two LLs with counter-propagating edge states for $0 < h_z$ (TnT phase) and a fully gapped spectrum, both in the bulk and at the edge, for $h_z < 0$ (TT phase). Crucially, due to the assumed topological protection by the physical symmetry, the Hamiltonians $\hat{H}_{1\circ} + \hat{H}^{\text{edge}}_{1\circ}$ do not couple the $|ap\rangle$ and $|bp\rangle$ states. As a result, the single-particle Hamiltonian possesses U(1) symmetry with respect to continuous rotations about the isospin z axis. This can be described by the matrix

$$\hat{D}(\phi) = \begin{pmatrix} \mathrm{e}^{-\mathrm{i}\frac{\phi}{2}} & 0 \\ 0 & \mathrm{e}^{\mathrm{i}\frac{\phi}{2}} \end{pmatrix} \tag{7.14}$$

acting on the spinor (7.1) in the ab space as

$$\hat{D}(\phi)\chi(\theta, \varphi) = \chi(\theta, \varphi + \phi). \tag{7.15}$$

Here, θ and φ are the angles of the spherical parametrization of the isospin given in Eq. (7.1). As we will find in Sec. 8, this effective continuous U(1) symmetry is central to the properties of the edge charge excitations of the interacting system and the associated topological properties.

Following this reasoning, we consider in Eqs. (7.11) and (7.12) the two-particle interactions that preserve this U(1) symmetry. They can be split into a SU(2)-symmetric part $\hat{H}_{2\odot}$ and the SU(2)-asymmetric U(1)-symmetric part $\hat{H}_{2\circ}$. Terms with the structure $\hat{1} \otimes \tau_z$ are also U(1)-symmetric, but get discarded in the following. They would only lead to an inconsequential shift of the position of the single-particle transition point in the σ-model. The exact form of the matrix elements $V(^{\alpha}_{\alpha}|^{p_1 p_1'}_{p_2 p_2'})$ with $\alpha = 0, x, y, z$ will not matter for our considerations, only the presented structure of the terms in the isospin space is important. The central condition we assume is that the U(1)-asymmetric terms are much smaller than the SU(2)-symmetric ones, in order to make the low-energy field theory a controlled expansion. The SU(2)-symmetric interactions have the typical scale of the Coulomb energy, given by

$$\sum_{p_2} V(^{0}_{0}|^{p_1 p_2}_{p_2 p_1}) \sim \frac{e_*^2}{l_{B_z}}. \tag{7.16}$$

The total Hamiltonian $\hat{H}$, Eq. (7.8), thus possesses U(1) symmetry. We now consider possible U(1)-asymmetric terms.

7.2.3. U(1)-asymmetric Terms

The mechanisms of (non-spontaneous) U(1) symmetry breaking can be grouped into two categories according to an important symmetry distinction between them.

1) In the first category, already the *physical symmetry* responsible for the topological protection of the non-interacting system is *violated*. Consequently, the U(1) symmetry is then broken directly at the single-particle level. The corresponding terms have the form of an isospin Zeeman field acting in the xy plane, thus coupling the a and b LLs, with the Hamiltonian

$$\hat{H}_{1\varnothing} = -h_\perp \sum_p \hat{c}_p^\dagger (\tau_x \cos\varphi_{1\varnothing} + \tau_y \sin\varphi_{1\varnothing}) \hat{c}_p. \tag{7.17}$$

The orientation of this field in the xy plane, set by the angle $\varphi_{1\varnothing}$, depends on the choice of the relative phase factor between the $|ap\rangle$ and $|bp\rangle$ states and is largely arbitrary.

2) In the second category, the *physical symmetry* is *preserved*. Here, no single-particle terms breaking U(1) symmetry are allowed. However, interactions that preserve the physical symmetry but break U(1) symmetry can be present, described by

$$\hat{H}_{2\varnothing} = \frac{1}{2} \sum_{p_1+p_2=p_1'+p_2'} \sum_{\alpha_1\alpha_2}^{\varnothing} V\left(\begin{smallmatrix}\alpha_1 & p_1 p_1' \\ \alpha_2 & p_2 p_2'\end{smallmatrix}\right) : [\hat{c}_{p_1}^\dagger \tau_{\alpha_1} \hat{c}_{p_1'}][\hat{c}_{p_2}^\dagger \tau_{\alpha_2} \hat{c}_{p_2'}] : \tag{7.18}$$

where the sum $\sum_{\alpha_1\alpha_2}^{\varnothing}$ contains only U(1)-asymmetric terms. The structure of such interactions depends on the specific physical symmetry, providing some constraints on the matrix elements $V\left(\begin{smallmatrix}\alpha_1 & p_1 p_1' \\ \alpha_2 & p_2 p_2'\end{smallmatrix}\right)$. Yet for most physical symmetries such U(1)-asymmetric interactions are allowed. For the σ-model approach we employ, however, the detailed knowledge of their structure is not necessary. Only the corresponding anisotropy function is required, which is derived in Sec. 7.3.3 using symmetry considerations.

7.3. Low-energy nonlinear σ-model

Now, we derive from the Hamiltonian (7.8) an effective model for the isospin OP $\boldsymbol{n}$, that will form the basis of our analysis in Sec. 8.

7.3.1. Quantum Hall Ferromagnet

Close to the crossing point $h_z = 0$ of the LLs, the SU(2)-symmetric part $\hat{H}_{2\odot}$ of the electron interactions is the dominant term in the total Hamiltonian (7.8). Its typical scale e_*^2/l_{B_z} exceeds the energies of all other terms. Following Sec. 3.2.3 and the references therein, it is then straightforward to show that, at half-filling of the two LLs $|ap\rangle$ and $|bp\rangle$, the Slater determinant state

$$\Psi(\boldsymbol{n}) = \prod_{\text{bulk } p} c_{\boldsymbol{n}p}^\dagger |0\rangle, \quad c_{\boldsymbol{n}p}^\dagger = \chi_a(\boldsymbol{n}) c_{ap}^\dagger + \chi_b(\boldsymbol{n}) c_{bp}^\dagger \tag{7.19}$$

is an exact eigenstate of $\hat{H}_{2\odot}$. Here, $|0\rangle$ is the vacuum state with both LLs empty and $c_{\boldsymbol{n}p}^\dagger$ is the operator creating an electron in the state $|\boldsymbol{n}\rangle$, defined in Eq. (7.1), with the isospin $\boldsymbol{n}$. The latter is visualized by means of the Bloch sphere in Fig. 7.4. The isospin at the poles $\boldsymbol{n} = \pm\boldsymbol{n}_z$ ($\theta = 0, \pi$) corresponds to pure $|\boldsymbol{n}_z\rangle = |a\rangle$ or $|-\boldsymbol{n}_z\rangle = |b\rangle$ states, while any other

state with $-1 < n_z < 1$ ($0 < \theta < \pi$) is a coherent superposition of $|a\rangle$ and $|b\rangle$. For a wide class of repulsive interactions, one can expect this eigenstate to be an exact ground state by the Hund-rule argument. This is the main assumption of the QHFM theory, introduced in Sec. 3.2.3. Importantly, the many-body wave-function $\Psi(\boldsymbol{n})$ is an eigenstate of $\hat{H}_{2\odot}$ for any choice of the isospin $\boldsymbol{n}$. It thus describes the state with spontaneously broken SU(2) symmetry. The unit vector $\boldsymbol{n}$ represents the OP of a family of degenerate ground states.

7.3.2. U(1)-symmetric nonlinear σ-model

The effects of all other, SU(2) symmetry breaking, terms on the QHFM ground state and its excitations can be taken into account within a low-energy field theory, the σ-model. As long as the energy scales of these terms are much smaller than the Coulomb scale (7.16) of the SU(2)-symmetric interactions, the σ-model presents a controlled, systematic low-energy expansion about the exact ground state (7.19) of $\hat{H}_{2\odot}$. For the Hamiltonian given by Eqs. (7.8), (7.17), and (7.18), the derivation of the σ-model is rather standard and follows the general recipe, see references in Sec. 3.2.3. The new aspect is incorporating the effect of the edge with counter-propagating states into the real-space σ-model, which we perform at the end of this section. For this, one generalizes the homogeneous and static isospin OP $\boldsymbol{n}$ of the state (7.19) to configurations $\boldsymbol{n}(\boldsymbol{r};t)$ that vary slowly in time and space. The constraint $\boldsymbol{n}^2(\boldsymbol{r};t) = 1$ is then satisfied locally.

The low-energy dynamics and energetics of the system is described by the Lagrangian functional. For the bulk part, i.e. all terms except $\hat{H}_{1\circ}^{\text{edge}}$, of the U(1)-symmetric Hamiltonian (7.8), it has the form

$$\mathbb{L}[\boldsymbol{n}] = \mathbb{K}[\boldsymbol{n}] - \mathbb{E}[\boldsymbol{n}] = \int \frac{\mathrm{d}^2\boldsymbol{r}}{s} L[\boldsymbol{n}],\ L[\boldsymbol{n}] = K[\boldsymbol{n}] - E[\boldsymbol{n}], \tag{7.20}$$

$$\mathbb{K}[\boldsymbol{n}] = \int \frac{\mathrm{d}^2\boldsymbol{r}}{s} K[\boldsymbol{n}],\ K[\boldsymbol{n}] = \frac{\dot{\varphi}}{2}\cos\theta, \tag{7.21}$$

$$\mathbb{E}[\boldsymbol{n}] = \int \frac{\mathrm{d}^2\boldsymbol{r}}{s} E[\boldsymbol{n}],\ E[\boldsymbol{n}] = \frac{\rho}{2}(\nabla\boldsymbol{n})^2 + \mathcal{E}(n_z), \tag{7.22}$$

$$\mathcal{E}(n_z) = \frac{u}{2}n_z^2 - h_z n_z = \frac{u}{2}\cos^2\theta - h_z\cos\theta. \tag{7.23}$$

As usual, the Lagrangian $\mathbb{L}[\boldsymbol{n}]$ is given by the difference of the kinetic $\mathbb{K}[\boldsymbol{n}]$ and energy $\mathbb{E}[\boldsymbol{n}]$ terms. The spatial integration $\int \mathrm{d}^2\boldsymbol{r}\ldots$ is performed over the half-space $x < 0$, where we introduce the normalization factor

$$s = 2\pi l_{B_z}^2 \tag{7.24}$$

equal to the area threaded by one magnetic flux quantum. $1/s$ is the electron density per one LL. This way, the respective densities $L[\boldsymbol{n}]$, $K[\boldsymbol{n}]$, $E[\boldsymbol{n}]$ are defined per area s and have the dimension of energy.

The kinetic term $\mathbb{K}[\boldsymbol{n}]$ contains a time derivative and can be written explicitly in terms of the spherical angles θ and φ parameterizing the isospin (7.2). The form of $K[\boldsymbol{n}]$ is not unique, but defined up to a full time derivative, resulting in an inconsequential constant contribution to the action $S = \int \mathrm{d}t\,\mathbb{L}[\boldsymbol{n}]$. The energy functional $\mathbb{E}[\boldsymbol{n}]$ (7.22) consists of the gradient term $\frac{\rho}{2}(\nabla\boldsymbol{n})^2$ and the energy function $\mathcal{E}(n_z)$ (7.23). The former describes the energy cost of a spatially inhomogeneous configuration. In leading order, the stiffness

$$\rho = \frac{l_{B_z}^4}{4}\sum_{p_2} V\left(\begin{smallmatrix}0 & p_1 p_2\\ 0 & p_2 p_1\end{smallmatrix}\right)(p_1 - p_2)^2 \tag{7.25}$$

can be expressed in terms of the SU(2)-symmetric $\hat{H}_{2\odot}$ interactions (7.11). The effect of all bulk terms $\hat{H}_{1\circ} + \hat{H}_{2\circ}$ with a symmetry lower than SU(2) is taken into account by the energy function $\mathcal{E}(n_z)$. It is derived by taking the expectation value

$$\mathcal{E}(n_z) = \frac{1}{\mathcal{N}} \langle \Psi(\boldsymbol{n}) | \hat{H}_{1\circ} + \hat{H}_{2\circ} | \Psi(\boldsymbol{n}) \rangle \tag{7.26}$$

of the corresponding terms with respect to the state $\Psi(\boldsymbol{n})$ (7.19). The normalization factor $\mathcal{N} = \sum_p 1 = \int \frac{\mathrm{d}^2 \boldsymbol{r}}{s} 1$ is the number of orbital states, equal to the number of flux quanta threading the sample. The most relevant anisotropy term we consider, arising from the SU(2)-asymmetric two-particle interactions $\hat{H}_{2\circ}$, is the quadratic-in-$\boldsymbol{n}$ contribution $\frac{u}{2} n_z^2$. The corresponding energy equals

$$\begin{aligned} u &= u_z - u_\perp, \quad u_\perp \equiv u_x = u_y, \\ u_\alpha &= \sum_{p_2} [V(^{\alpha}_{\alpha}|^{p_1 p_1}_{p_2 p_2}) - V(^{\alpha}_{\alpha}|^{p_1 p_2}_{p_2 p_1})], \quad \alpha = x, y, z. \end{aligned} \tag{7.27}$$

We consider in the following the more interesting case of positive, *easy-plane* anisotropy energy

$$u > 0, \tag{7.28}$$

where the energy $\frac{u}{2} n_z^2$ alone is minimized by the isospin $n_z = 0$ lying in the plane.

The last missing term from the U(1)-symmetric Hamiltonian (7.8) is $\hat{H}_{1\circ}^{\mathrm{edge}}$, describing the edge. We take it into account by an effective boundary condition for the order parameter $\boldsymbol{n}(\boldsymbol{r}; t)$. As the edge states are half-filled, one electron per two states, the filling factor remains the same for both bulk ($p \lesssim 0$) and edge ($p \gtrsim 0$) states. Thus the occupation of the edge can be described by the same isospin OP as the bulk. At such momenta p that the edge potential $\epsilon(p)$ is much greater than the energies u and h_z of the SU(2)-asymmetric terms, electrons will always occupy the hole branch of the edge spectrum. This corresponds to the normal, TT order with the states $|bp\rangle$ of negative energy $-\epsilon(p)$ filled, leading to the OP of $\boldsymbol{n} = -\boldsymbol{n}_z$. While the edge states with $p \gtrsim 0$ are localized at spatial scales $\sim l_{B_z}$ near the edge $x = 0$ of the sample, $\boldsymbol{n}(\boldsymbol{r}; t)$, by assumption, varies at much larger scales. Hence, the effect of the edge may be described in real space by the boundary condition

$$\boldsymbol{n}(x = 0, y; t) = -\boldsymbol{n}_z, \quad \boldsymbol{n}_z = (0, 0, 1). \tag{7.29}$$

It pins the OP in the state that corresponds to the occupation of the hole branch of the edge spectrum.

Equations (7.20)-(7.23) for the Lagrangian and Eq. (7.29) for the boundary condition constitute the closed-form low-energy σ-model in the coordinate space for the considered QHFM system with an edge. Naturally, the model inherits the U(1) symmetry of the Hamiltonian (7.8) and is thus invariant under the rotations of the isospin

$$\varphi(\boldsymbol{r}; t) \to \varphi(\boldsymbol{r}; t) + \phi \tag{7.30}$$

about the z axis. The additional terms in the σ-model originating from the U(1)-asymmetric terms (7.17) and (7.18) in the full Hamiltonian, and thus breaking this invariance, are considered next.

7.3.3. U(1)-asymmetric Terms

We begin with the single-particle term $\hat{H}_{1\varnothing}$ (7.17) of the first type, that breaks U(1) symmetry by violating the physical symmetry directly. It produces an additional contribution

$$\begin{aligned}\mathcal{E}_{1\varnothing}(\boldsymbol{n}) &= \frac{1}{\mathcal{N}}\langle\Psi(\boldsymbol{n})|\hat{H}_{1\varnothing}|\Psi(\boldsymbol{n})\rangle = -h_\perp(n_x\cos\varphi_{1\varnothing} + n_y\sin\varphi_{1\varnothing}) \\ &= -h_\perp\sin\theta\cos(\varphi-\varphi_{1\varnothing})\end{aligned} \tag{7.31}$$

to the energy function $\mathcal{E}(n_z)$ (7.23).

The structure of the anisotropy energy function arising from the U(1)-asymmetric two-particle interactions $\hat{H}_{2\varnothing}$ (7.18) depends on the specific physical symmetry. It can be derived via group-theoretical considerations without using any information about the interaction matrix elements of $\hat{H}_{2\varnothing}$. As an example, we consider the $\boldsymbol{\mathcal{I}}$ symmetry, where the LL states a and b are characterized by opposite $\boldsymbol{\mathcal{I}}$ parities + and −. Therefore, the isospin components transforming according to $|\boldsymbol{n}\rangle\langle\boldsymbol{n}|$ in Eq. (7.3) as $n_{x,y} \sim |a\rangle\langle b|$ and $n_z \sim |a\rangle\langle a| - |b\rangle\langle b|$ have − and + parities, respectively. The anisotropy function arising from two-particle interactions is a quadratic function of $\boldsymbol{n}$. It must be invariant under $\boldsymbol{\mathcal{I}}$, i.e., have + parity. All quadratic functions with + parity are

$$\{n_z^2, n_x^2, 2n_xn_y, n_y^2\}. \tag{7.32}$$

The most general form of the anisotropy function is then an arbitrary linear combination of these terms. It is convenient to choose the basis functions as

$$\{n_x^2+n_y^2+n_z^2, n_z^2-n_x^2-n_y^2, n_x^2-n_y^2, 2n_xn_y\}. \tag{7.33}$$

Here, the combination $n_x^2+n_y^2+n_z^2 = 1$ preserves SU(2) symmetry and is $\boldsymbol{n}$-independent due to the constraint $\boldsymbol{n}^2 = 1$ in Eq. (7.2). The next combination $n_z^2 - n_x^2 - n_y^2 = 2n_z^2 - 1$ preserves U(1) symmetry and depends only on n_z. These two functions produce, up to a constant, the U(1)-symmetric anisotropy function $\frac{u}{2}n_z^2$ in $\mathcal{E}(n_z)$. An arbitrary linear combination

$$\begin{aligned}\mathcal{E}_{2\varnothing}(\boldsymbol{n}) &= \frac{1}{2}[u_+(n_x^2-n_y^2) + u_\times 2n_xn_y] \\ &= \frac{1}{2}u_{2\varnothing}\sin^2\theta\cos 2(\varphi-\varphi_{2\varnothing})\end{aligned} \tag{7.34}$$

of the two remaining functions represents the U(1)-asymmetric contribution to the anisotropy function in case of $\boldsymbol{\mathcal{I}}$ as the physical symmetry.

8. Quantum Hall Ferromagnet and Helical Luttinger Liquid

We proceed with the analysis of the obtained σ-model, with the focus on the properties of the edge excitations. The reduction of the effect of the edge to the boundary condition (7.29) in the coordinate space will prove advantageous in the analysis of the problem. It enables us to obtain explicit analytical expressions for many quantities of interest in the following.

The approach we use to study the edge excitations follows that developed in Refs. [FI99, FB06, MSF14, MSF16, TSFM16], where the authors mainly focus on the $n = 0$ QH state in graphene with armchair-type boundaries. There, the transition between the ferromagnetic (F) and canted antiferromagnetic (CAF) phases, see Refs. [Her07, Kha12c, Kha12a, Kha12b] for details, can be seen in analogy to our TnT to TT phase transition. Thus although graphene has a few additional peculiarities, like the presence of valley degrees of freedom which makes the QHFM physics richer, there are mathematical and physical similarities to our model. Another related system is a QH bilayer with an inverted band structure, studied theoretically in Ref. [PSH16]. While the Hamiltonian considered there essentially coincides with the U(1)-symmetric part of our model, the focus and methods of this analysis differs from ours in several aspects. We point out the analogies between our and these two systems as we move along.

As originally recognized in Ref. [FB06] for the $\nu = 0$ state in graphene, the physics of the edge in the QHFM state is governed by the fact that the order favored at the edge due to the propagating edge states may be different from that favored in the bulk. This leads to a spatially inhomogeneous OP texture at the edge, called *domain wall* in the following, which connects the bulk and edge orders and determines the properties of the edge excitations. Here, we also mention Refs. [FB97, LKS99, HKLN99, KLS99] on QHFM systems with chiral edge states and the recent Ref. [BZY+17] on bilayer graphene with double counter-propagating edge states.

The central result we find in this section is that the topological properties of our interacting many-body system, the QHTI, are directly tied to its effective symmetry. The U(1) symmetry with respect to rotations about the isospin z axis is responsible for the topological protection. We demonstrate that if U(1) symmetry is preserved, the single-particle TnT phase with fully filled $|a\rangle$ LL, corresponding to $\boldsymbol{n} = \boldsymbol{n}_z = (0, 0, 1)$ isospin in the QHFM formalism, remains TnT in the presence of interactions in most of the range $0 \leq B < B^*$. The edge excitations stay gapless, but take the form of collective excitations described by the helical Luttinger liquid. We obtain explicit analytical expressions for its parameters, and find that the effective interactions in this Luttinger liquid are highly tunable. They are weak at small magnetic fields $B \ll B^*$, with the interaction parameter $\mathcal{K} \approx 1$, see Sec. 3.2.4, but strong ($\mathcal{K} \ll 1$) in the QHFM regime in the vicinity of the single-particle phase transition at B^*.

Here, preserved U(1) symmetry means that *both* the bulk ground state and the many-body Hamiltonian are U(1)-symmetric. Violating one of these conditions, we identify two mechanisms of U(1) symmetry breaking, which lead to the loss of topological protection and the eventual transition to the TT phases. First, U(1) symmetry can be broken spontaneously.

This is done by increasing B until a second-order phase transition from the TnT phase $\boldsymbol{n} = \boldsymbol{n}_z$ to the phase with broken symmetry occurs. The gap in the edge excitation spectrum of this broken U(1) symmetry phase grows monotonically upon further increasing B, starting from the zero value at the phase transition. Interestingly, upon approaching this phase transition from the TnT phase, the edge Luttinger liquid becomes infinitely strongly interacting with $\mathcal{K} \to 0$.

The second possibility is a U(1)-asymmetric many-body Hamiltonian. Here, the interaction terms may fully respect the physical symmetry, responsible for the topological protection of the non-interacting system, but break U(1) symmetry. Such terms transform the Luttinger liquid model for the edge into the sine-Gordon model, see Ref. [Gia03] for details. As B is increased, this results in a phase transition to the state with broken U(1) symmetry at the edge and gapped edge excitations.

In either of the scenarios, the phase transitions from the TnT to the TT phases occur at the magnetic fields B *lower* than the single-particle phase transition point B^* and are thus *interaction-induced topological* quantum phase transitions.

These presented results are derived in the following sections, which are structured as follows. In Sec. 8.1, we obtain the phase diagram of the bulk alone. The edge is added in Sec. 8.2 and the ground state configurations of the whole system are analyzed. Special attention is paid to the edge charge excitations, which we study in Sec. 8.3. In the TnT phase, the σ-model can be mapped to a Luttinger liquid model for the edge excitations. This is derived in Sec. 8.4. We finish in Sec. 8.5 by establishing the U(1) symmetry as the requirement for topological protection.

8.1. Bulk Phase Diagram

We begin our analysis of the derived U(1)-symmetric σ-model, given in Eqs. (7.20)-(7.23) and (7.29). As a start, we neglect the edge and obtain the mean-field phase diagram of the bulk alone. This is done by minimizing the energy function $\mathcal{E}(n_z)$ (7.23). The minimum isospin configuration is denoted as $\boldsymbol{n}^\infty$ and referred to as the *bulk ground state*. The corresponding minimal energy is defined by

$$\mathcal{E}^\infty \equiv \mathcal{E}(n_z^\infty = \cos\theta^\infty) = \min_{n_z} \mathcal{E}(n_z). \tag{8.1}$$

In the case $u > 0$ of the easy-plane anisotropy we consider, minimizing $\mathcal{E}(n_z)$ within the interval $-1 \leq n_z \leq 1$ gives the phases

$$\boldsymbol{n}^\infty = \boldsymbol{n}_z = (0,0,1), \quad u < h_z, \tag{8.2}$$
$$\boldsymbol{n}^\infty = \boldsymbol{n}^*(\varphi_0) = (\sin\theta^*\cos\varphi_0, \sin\theta^*\sin\varphi_0, \cos\theta^*), \quad -u < h_z < u, \tag{8.3}$$
$$\boldsymbol{n}^\infty = -\boldsymbol{n}_z = (0,0,-1), \quad h_z < -u, \tag{8.4}$$

plotted in Fig. 8.1. The respective energy minima are

$$\begin{aligned} \mathcal{E}^\infty &= \mathcal{E}(n_z = +1) = \frac{u}{2} - h_z, \quad u < h_z, \\ \mathcal{E}^\infty &= \mathcal{E}(n_z^*) = -\frac{h_z^2}{2u}, \quad -u < h_z < u, \\ \mathcal{E}^\infty &= \mathcal{E}(n_z = -1) = \frac{u}{2} + h_z, \quad h_z < -u. \end{aligned} \tag{8.5}$$

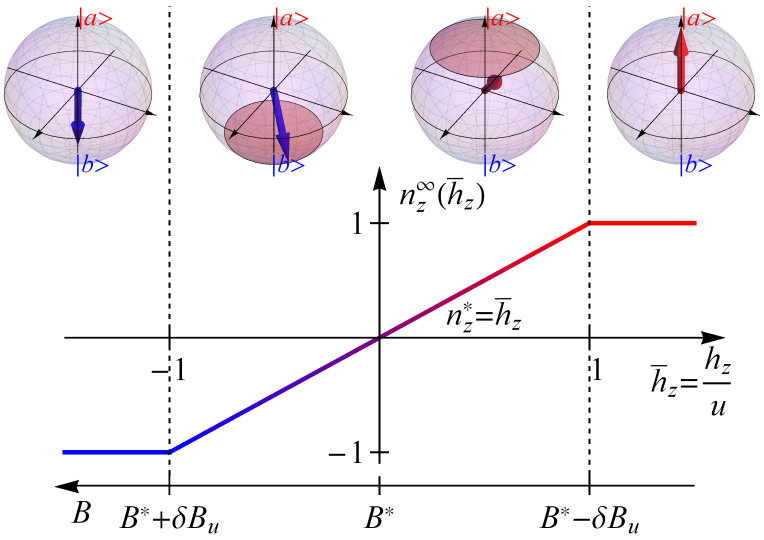


Figure 8.1. Bulk phase diagram for the U(1)-symmetric σ-model obtained by minimizing the energy function $\mathcal{E}(n_z)$ (7.23). The dependence $n_z^\infty(\bar{h}_z)$ of the optimal isospin projection on the normalized field $\bar{h}_z = h_z/u$ is shown. The color code represents the weight of the a (red) and b (blue) LL states. The Bloch spheres depict the corresponding isospin orders (8.2), (8.3), and (8.4). Reprinted figure with permission from Ref. [KJT16]. Copyright (2016) by the APS.

The phases $\boldsymbol{n}^\infty = \pm\boldsymbol{n}_z$ in the limits $u < h_z$ and $h_z < -u$ are fully polarized along the z direction of the field h_z. Following from the definition of the isospin in Eqs. (7.1) and (7.2), these two cases correspond to the occupation of either the $|+\boldsymbol{n}_z\rangle = |a\rangle$ or $|-\boldsymbol{n}_z\rangle = |b\rangle$ LL. The corresponding Slater-determinant ground state $\Psi(\boldsymbol{n})$ (7.19) is thus the same as in the non-interacting system.

In the intermediate phase $\boldsymbol{n}^\infty = \boldsymbol{n}^*(\varphi_0)$ for $-u < h_z < u$, the isospin has the optimal projection

$$n_z^* = \cos\theta^* = \bar{h}_z \tag{8.6}$$

on the z direction and arbitrary orientation in the xy plane, parameterized by the angle φ_0. Here we introduced the dimensionless field

$$\bar{h}_z = \frac{h_z}{u} \tag{8.7}$$

normalized by the anisotropy energy u. Thus, electrons are in a coherent superposition

$$|\boldsymbol{n}^*(\varphi_0)\rangle = \mathrm{e}^{-\mathrm{i}\frac{\varphi_0}{2}} \cos\frac{\theta^*}{2}|a\rangle + \mathrm{e}^{\mathrm{i}\frac{\varphi_0}{2}} \sin\frac{\theta^*}{2}|b\rangle \tag{8.8}$$

of the two LL states in this intermediate phase. Its appearance is the first important distinction from the non-interacting picture. The U(1) symmetry is therefore preserved in the $\boldsymbol{n}^\infty = \pm\boldsymbol{n}_z$ phases, but spontaneously broken in the intermediate $\boldsymbol{n}^\infty = \boldsymbol{n}^*(\varphi_0)$ phase. The corresponding phase transitions are of second order, the transition points $h_z = \pm u$ correspond to the values

$$B^* \mp \delta B_u, \quad \delta B_u \equiv \frac{u}{|\partial_B h_z(B^*)|} \tag{8.9}$$

of the magnetic field, respectively.

Comparing to the literature, an analogous phase diagram was obtained in Ref. [PSH16] for a double-layer system. Additionally, the region $0 \leq h_z$ reproduces the phase diagram of the $\nu = 0$ state in graphene discussed in Refs. [Her07, Kha12c, Kha12a], where the $\boldsymbol{n}^\infty = \boldsymbol{n}_z$ and $\boldsymbol{n}^\infty = \boldsymbol{n}^*(\varphi_0)$ phases correspond to the F and CAF phases, respectively. In this analogy, the isospin $\boldsymbol{n}$ would correspond to the spin polarization of one of the sublattices of the honeycomb crystal lattice of graphene.

8.2. System with an Edge: Ground States

Next, we take the effect of the edge into account and obtain the ground state configurations of the OP $\boldsymbol{n}(\boldsymbol{r})$ for the combined system. Such configurations, denoted by $\boldsymbol{n}_0(\boldsymbol{r})$, minimize the energy functional (7.22),

$$\mathbb{E}[\boldsymbol{n}_0] = \min_{\boldsymbol{n}} \mathbb{E}[\boldsymbol{n}], \tag{8.10}$$

under the boundary condition constraint (7.29). $\boldsymbol{n}_0(\boldsymbol{r})$ is a stationary point of the energy functional, fulfilling the stationary-point equations

$$\frac{\delta \mathbb{E}[\theta, \varphi]}{\delta \theta} = \rho[-\nabla^2 \theta + \tfrac{1}{2} \sin 2\theta (\nabla \varphi)^2] + \partial_\theta \mathcal{E}(\theta) = 0, \tag{8.11}$$

$$\frac{\delta \mathbb{E}[\theta, \varphi]}{\delta \varphi} = -\rho \nabla (\sin^2 \theta \nabla \varphi) = 0 \tag{8.12}$$

written in terms of the spherical angles θ and φ defined in Eq. (7.2). Throughout, we will denote the energy dependence $\mathcal{E}(\theta) = \mathcal{E}(n_z = \cos\theta)$ on θ and n_z by the same function. In the presence of the edge, the translational symmetry along y direction is still preserved. Hence, the ground state configuration is y-independent, $\boldsymbol{n}_0(x, y) \equiv \boldsymbol{n}_0(x)$. Changes of the isospin with y would only result in the rise of the gradient energy. Away from the edge, in the asymptotic limit $x \to -\infty$, the ground state configurations must approach the constant value

$$\boldsymbol{n}_0(x \to -\infty) = \boldsymbol{n}^\infty \tag{8.13}$$

of the bulk ground state order $\boldsymbol{n}^\infty$. Depending on h_z, this is one of the orders (8.2), (8.3), or (8.4) that minimize $\mathcal{E}(n_z)$, as obtained in the previous section. Therefore, whenever $\boldsymbol{n}^\infty$ differs from the boundary order (7.29)

$$\boldsymbol{n}_0(x = 0) = -\boldsymbol{n}_z, \tag{8.14}$$

$\boldsymbol{n}_0(x)$ is a spatially inhomogeneous domain-wall configuration along x that connects these two orders. Due to the U(1) symmetry of the energy function $\mathcal{E}(n_z)$ and boundary condition (7.29), the angle $\varphi(x) \equiv \varphi_0$ in the spherical parametrization (7.2) of $\boldsymbol{n}_0(x)$ is arbitrary and constant. Any spatial change in φ would result in an additional rise of the gradient energy. Therefore any constant φ automatically satisfies the stationary point equation (8.12).

In conclusion, the ground state configuration has the form

$$\boldsymbol{n}_0(x|\varphi_0) = (\sin\theta_0(x)\cos\varphi_0, \sin\theta_0(x)\sin\varphi_0, \cos\theta_0(x)). \tag{8.15}$$

For these angles $\theta(x)$, dependent only on x, and the constant φ_0, the energy functional $\mathbb{E}[\boldsymbol{n}]$ (7.22) per unit length in the y direction reduces to

$$E^{1\mathrm{D}}[\theta(x); u, h_z] = \int_{-\infty}^{0} \frac{\mathrm{d}x}{s} E_x[\theta], \quad E_x[\theta] = \frac{\rho}{2}(\nabla_x \theta)^2 + \mathcal{E}(\theta). \tag{8.16}$$

The ground state configuration $\theta_0(x)$ minimizes this functional and thus satisfies its stationary point equation

$$-\rho\nabla_x^2\theta + \partial_\theta\mathcal{E}(\theta) = 0, \quad \theta(x=-\infty) = \theta^\infty \text{ and } \theta(x=0) = \pi, \tag{8.17}$$

which we supplemented by the boundary conditions following from Eqs. (8.13) and (8.14), where θ^∞ is the angle of the bulk order $\boldsymbol{n}^\infty$.

We solve this boundary problem by noticing the analogy of Eq. (8.17) with the Newton equation for a point particle in one dimension, where θ and x play the roles of coordinate and time, respectively. The equation has an integral of motion

$$\frac{\rho}{2}(\nabla_x\theta)^2 - \mathcal{E}(\theta) = -\mathcal{E}^\infty, \tag{8.18}$$

equivalent to the total energy of the effective particle. It can be obtained by multiplying Eq. (8.17) by $\nabla_x\theta$ and integrating once over x. The gradient term $\frac{\rho}{2}(\nabla_x\theta)^2$ in Eq. (8.18) plays the role of the kinetic energy, while $-\mathcal{E}(\theta)$ corresponds to the potential energy. The value of this integral of motion is set by its value $-\mathcal{E}^\infty$ (8.5) in the bulk, where $\nabla_x\theta \to 0$. An additional integration of Eq. (8.18) produces an implicit dependence of $\theta_0(x)$ on x, given by

$$-x = \int_{\theta_0}^{\pi} \frac{\mathrm{d}\theta}{\sqrt{\frac{2}{\rho}\left[\mathcal{E}(\theta) - \mathcal{E}^\infty\right]}}. \tag{8.19}$$

The functional form (7.23) of $\mathcal{E}(n_z)$ allows for explicit integration of Eq. (8.19) in terms of elementary functions and subsequent inversion. We find the explicit solutions

$$\theta_0(x) = \arccos\left[1 - \frac{2(\bar{h}_z - 1)}{\bar{h}_z \cosh^2\left(\sqrt{\bar{h}_z - 1}\bar{x}\right) - 1}\right], \quad u < h_z, \tag{8.20}$$

$$\theta_0(x) = 2\arctan\left[\sqrt{\frac{1-\bar{h}_z}{1+\bar{h}_z}}\frac{1}{\tanh\left(-\frac{\sqrt{1-\bar{h}_z^2}}{2}\bar{x}\right)}\right], \quad -u < h_z < u, \tag{8.21}$$

$$\theta_0(x) = \pi, \quad h_z < -u, \tag{8.22}$$

which are plotted in Fig. 8.2. Here, $\bar{h}_z$ is the dimensionless field defined in Eq. (8.7) and

$$\bar{x} = \frac{x}{l_u}, \quad l_u = \sqrt{\frac{\rho}{u}} \tag{8.23}$$

is the dimensionless coordinate, normalized by the length scale l_u set by the anisotropy energy u. The solutions, therefore, have the scaling form $\theta_0(x) = \theta_0(x; u, h_z) = \theta_0\left(\bar{x}; \bar{h}_z\right)$. They minimize the functional $E^{1\mathrm{D}}[\theta]$ (8.16), while the respective isospin configurations $\boldsymbol{n}_0(x|\varphi_0)$ (8.15) minimize the functional $\mathbb{E}[\boldsymbol{n}]$ (7.22) in the presence of the edge, described by the boundary condition (8.14).

As shown in Fig. 8.2, in the $\boldsymbol{n}^\infty = \boldsymbol{n}_z$ and $\boldsymbol{n}^\infty = \boldsymbol{n}^*(\varphi_0)$ phases, $\theta_0(x)$ grows monotonically from the bulk value θ^∞ at $x = -\infty$ to π at the edge $x = 0$. In the $\boldsymbol{n}^\infty = -\boldsymbol{n}_z$ phase on the other hand, the bulk and edge orders are the same and $\theta_0(x) \equiv \pi$ is a constant. The solutions

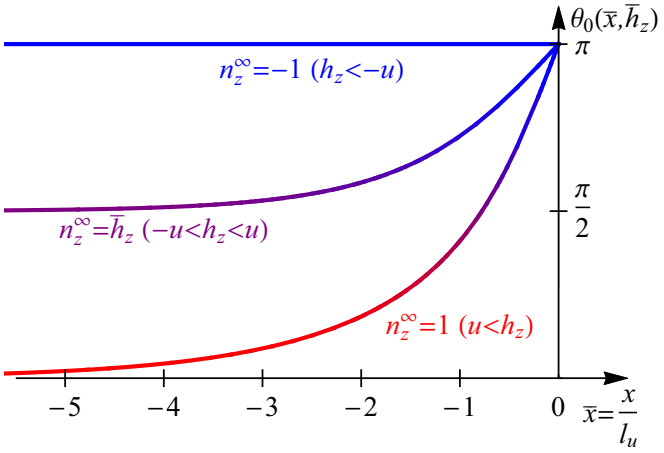


Figure 8.2. The angle functions $\theta_0(x) = \theta_0(\bar{x}, \bar{h}_z)$ (8.20), (8.21), and (8.22) for the ground state solution (8.15) of a system with an edge. Three cases according to the three phases (8.2), (8.3), and (8.4) are shown. At $x \to -\infty$, the solutions approach the values $\theta^\infty(\bar{h}_z)$ for the bulk ground state, while at $x = 0$, they satisfy the boundary condition (8.17) imposed by the edge. The functions are color-coded according to the $n_z = \cos\theta_0(x)$ projection, depicting the weight of the a (red) and b (blue) LL states. Reprinted figure with permission from Ref. [KJT16]. Copyright (2016) by the APS.

$\theta_0(x)$ approach the asymptotic bulk value θ^∞ exponentially over the length scales

$$\frac{l_u}{\sqrt{\bar{h}_z - 1}} = \sqrt{\frac{\rho}{h_z - u}}, \quad u < h_z, \tag{8.24}$$

$$\frac{l_u}{\sqrt{1 - \bar{h}_z^2}} = \sqrt{\frac{\rho u}{u^2 - h_z^2}}, \quad -u < h_z < u. \tag{8.25}$$

At both phase transitions $h_z = \pm u$, these length scales become infinite, changing the effective asymptotic scaling behavior. Exactly at the $h_z = u$ phase transition, e.g., the solution approaches the bulk value $\theta^\infty = 0$ as a power law $\theta_0(x) \approx -2/\bar{x}$, which follows from the form

$$\theta_0(x) = 2\arctan\left(-\frac{1}{\bar{x}}\right), \quad h_z = u, \tag{8.26}$$

obtained from both Eqs. (8.20) and (8.21) in the limits $h_z \to u \pm 0^+$. We emphasize that even exactly at the phase transition $h_z = u$, the domain wall has the spatial scale l_u, although the bulk asymptotic value is approached according to a power law, and not exponentially. Close to the other transition at $h_z = -u$, the bulk value θ^∞ is close to π, and the solution simplifies to

$$\theta_0(x) = \pi - \sqrt{2(1 + \bar{h}_z)}\tanh\left(-\sqrt{\frac{1 + \bar{h}_z}{2}}\bar{x}\right), \quad h_z = -u + 0^+. \tag{8.27}$$

Next, we consider the topological properties of the obtained solution. Here, the φ degeneracy of the isospin solutions $\boldsymbol{n}_0(x|\varphi_0)$ (8.15) will be most important. In accordance to the possible bulk phases (8.2), (8.3), and (8.4), we have the following three cases:

1) In the $\boldsymbol{n}^\infty = \boldsymbol{n}_z$ phase realized for $u < h_z$, the ground state solution $\boldsymbol{n}_0(x|\varphi_0)$ is degenerate according to the arbitrary angle φ_0. Yet in the bulk for $\boldsymbol{n}^\infty = \boldsymbol{n}_z$, φ_0 is undefined and the U(1)

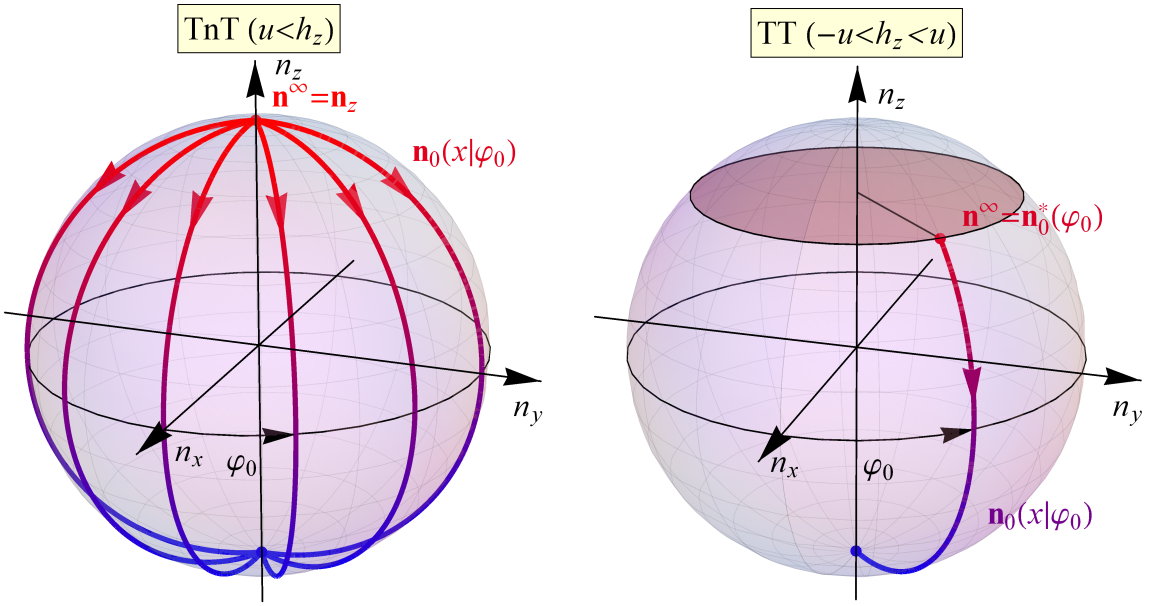


Figure 8.3. The ground state geodesic paths $\boldsymbol{n}_0(x|\varphi_0)$ (8.15) on the isospin Bloch sphere for the $\boldsymbol{n}^\infty = \boldsymbol{n}_z$ (left) and $\boldsymbol{n}^\infty = \boldsymbol{n}_0^*(\varphi_0)$ (right) phases, realized for $u < h_z$ and $-u < h_z < u$, respectively. Reprinted figures with permission from Ref. [KJT16]. Copyright (2016) by the APS.

symmetry is not spontaneously broken there. Thus, this degeneracy occurs (at the mean-field level) in the domain wall at the edge, not in the bulk.

2) In the $\boldsymbol{n}^\infty = \boldsymbol{n}^*(\varphi_0)$ phase realized for $-u < h_z < u$, the ground state configuration $\boldsymbol{n}_0(x|\varphi_0)$ is degenerate again. Here, the arbitrary angle φ_0 is the same as the one of the asymptotic bulk configuration $\boldsymbol{n}^*(\varphi_0)$. This degeneracy describes therefore the spontaneous breaking of U(1) symmetry in the bulk and there is no extra degeneracy at the edge.

3) In the $\boldsymbol{n}^\infty = -\boldsymbol{n}_z$ phase realized for $h_z < -u$, the bulk and edge orders are exactly the same. The ground state solution for the system with an edge is a constant $\boldsymbol{n}_0(x) \equiv -\boldsymbol{n}_z$ and thus nondegenerate, as φ_0 is undefined.

These properties can also be illustrated with the help of the Bloch sphere, see Fig. 8.3. The ground state domain wall configurations $\boldsymbol{n}_0(x|\varphi_0)$ can be visualized as geodesic paths connecting the bulk $\boldsymbol{n}^\infty$ and edge $-\boldsymbol{n}_z$ orders, parameterized by the coordinate x. In the $\boldsymbol{n}^\infty = \boldsymbol{n}_z$ phase, the bulk and edge orientations are exactly opposite, and there is an infinite number of geodesics, distinguished by their angle φ_0. This is different in the $\boldsymbol{n}^\infty = \boldsymbol{n}^*(\varphi_0)$ phase. Here, for a given angle φ_0 in the bulk, the geodesic connecting $\boldsymbol{n}^*$ and $-\boldsymbol{n}_z$ is unique: it is a path in the vertical plane of the constant φ_0. As we show in the next Sec. 8.3, these degeneracy properties of the ground state solutions are key to the properties of the charge edge excitations.

8.2.1. Domain-wall Energy

We end this section with an explicit calculation of the ground state energy of the system with an edge. It is convenient to split this energy in a contribution of the asymptotic bulk and one

from the domain wall. We thus define the energy quantities

$$\begin{aligned} dE^{\mathrm{1D}}[\theta(x);u,h_z] &\equiv \int_{-\infty}^{0} \frac{\mathrm{d}x}{s}\, dE_x[\theta(x);u,h_z] = E^{\mathrm{1D}}[\theta(x);u,h_z] - E^{\mathrm{1D}\infty}(u,h_z), \\ dE_x[\theta(x);u,h_z] &\equiv E_x[\theta(x);u,h_z] - \mathcal{E}^{\infty}(u,h_z), \\ E^{\mathrm{1D}\infty}(u,h_z) &\equiv \int_{-\infty}^{0} \frac{\mathrm{d}x}{s}\, \mathcal{E}^{\infty}(u,h_z) \end{aligned} \tag{8.28}$$

in terms of $E^{\mathrm{1D}}[\theta(x);u,h_z]$ (8.16). The corresponding ground-state energy per unit length in the y direction equals

$$dE_0^{\mathrm{1D}}(u,h_z) \equiv \min_{\theta(x)} dE^{\mathrm{1D}}[\theta(x);u,h_z] = dE^{\mathrm{1D}}[\theta_0(x;u,h_z);u,h_z] \tag{8.29}$$

and can be referred to as the *domain-wall energy*. It is not extensive in the x direction and describes the energy associated only with the domain-wall isospin texture at the edge. Exploiting the integral of motion (8.18), the domain-wall energy can be expressed as

$$dE_0^{\mathrm{1D}}(u;h_z) = 2\int_{-\infty}^{0} \frac{\mathrm{d}x}{s}[\mathcal{E}(\theta_0(x)) - \mathcal{E}^{\infty}] \tag{8.30}$$

and calculated explicitly using the expressions (8.20) and (8.21) for the ground state solutions. For the $\boldsymbol{n}^{\infty} = \boldsymbol{n}_z$ phase (the same can be done for the $\boldsymbol{n}^{\infty} = \boldsymbol{n}_0^*(\varphi)$ phase), we obtain

$$dE_0^{\mathrm{1D}}(u,h_z) = 2\frac{l_u}{s} u F(\bar{h}_z), \tag{8.31}$$

where

$$F(\bar{h}_z) = -F_2(\bar{h}_z) + \bar{h}_z F_1(\bar{h}_z) = \bar{h}_z \arcsin\frac{1}{\sqrt{\bar{h}_z}} + \sqrt{\bar{h}_z - 1}, \quad u < h_z, \tag{8.32}$$

$$F_2(\bar{h}_z) = \frac{1}{2}\int_{-\infty}^{0} \mathrm{d}\bar{x}\, \sin^2\theta_0(\bar{x}) = \bar{h}_z \arcsin\frac{1}{\sqrt{\bar{h}_z}} - \sqrt{\bar{h}_z - 1}, \tag{8.33}$$

$$F_1(\bar{h}_z) = \int_{-\infty}^{0} \mathrm{d}\bar{x}\,[1 - \cos\theta_0(\bar{x})] = 2\arcsin\frac{1}{\sqrt{\bar{h}_z}} \tag{8.34}$$

are dimensionless functions of the normalized field $\bar{h}_z$. The latter two arise from the anisotropy $\frac{u}{2}(n_z^2 - 1)$ and Zeeman $-h_z(n_z - 1)$ contributions to $\mathcal{E}(\theta_0(x)) - \mathcal{E}^{\infty}$, respectively. The function $F(\bar{h}_z)$ is plotted in Fig. 8.4. Importantly, as we will see below in Secs. 8.3.1 and 8.4, the dependence (8.31) of the domain-wall energy on parameters u and h_z defines not only the ground state, but also the properties of the low-energy edge excitations of the $\boldsymbol{n}^{\infty} = \boldsymbol{n}_z$ phase.

8.3. System with an Edge: Charge Excitations

Having established the properties of the ground states of the system with an edge, we now turn to the analysis of its charge excitations. As discussed in Sec. 3.2.3, QHFM systems support charge excitations which are described by the configurations of the OP with nonzero

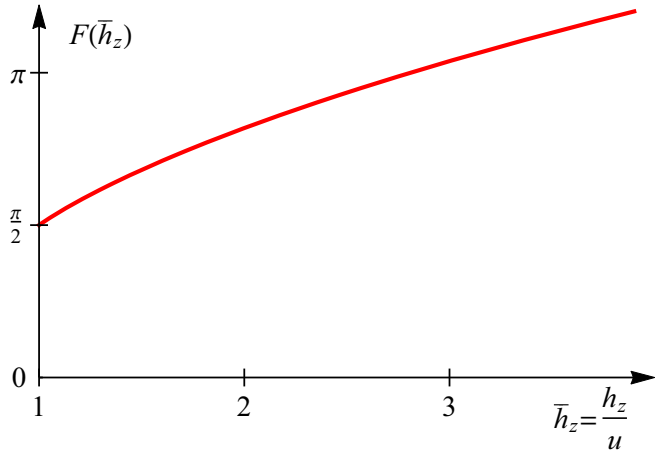


Figure 8.4. The function $F(\bar{h}_z)$ (8.32), which determines the dependence of the ground state domain-wall energy $dE_0^{1D}(u, h_z)$ (8.31) on the normalized field $\bar{h}_z = h_z/u$ in the $\boldsymbol{n}^\infty = \boldsymbol{n}_z$ phase for $u < h_z$. Reprinted figure with permission from Ref. [KJT16]. Copyright (2016) by the APS.

topological charge, equal to the electric charge. For our QHTI system with an isospin-$\frac{1}{2}$ OP, the charge density of a configuration $\boldsymbol{n}(\boldsymbol{r})$ is given by

$$\begin{aligned}\kappa[\boldsymbol{n}](\boldsymbol{r}) &= \frac{1}{4\pi}(\boldsymbol{n}\cdot[\nabla_x\boldsymbol{n}\times\nabla_y\boldsymbol{n}]) \\ &= \frac{1}{4\pi}\sin\theta(\nabla_x\theta\nabla_y\varphi - \nabla_y\theta\nabla_x\varphi).\end{aligned} \tag{8.35}$$

The total net charge of the configuration is then calculated by the integral

$$q[\boldsymbol{n}] = \int \mathrm{d}^2\boldsymbol{r}\,\kappa[\boldsymbol{n}](\boldsymbol{r}). \tag{8.36}$$

This topological charge is the invariant of the mapping realized by $\boldsymbol{n}(\boldsymbol{r})$ from the coordinate space of $\boldsymbol{r}$ to the 2D Bloch sphere. It can be visualized as the number of times the sphere is wound as the space of $\boldsymbol{r}$ is explored. In the following, we will be interested in the excitations with integer charge q, whose boundary values are the same as in the ground state. In principle, excitations with non-integer charge q are also possible, if the ground state has broken continuous symmetry. This is indeed the case for the intermediate phase $\boldsymbol{n}^\infty = \boldsymbol{n}^*(\varphi_0)$, but beyond the scope of our analysis here. We define the difference

$$\delta\mathbb{E}[\boldsymbol{n}] \equiv \mathbb{E}[\boldsymbol{n}] - \mathbb{E}[\boldsymbol{n}_0] \tag{8.37}$$

between the energy of a configuration $\boldsymbol{n}(\boldsymbol{r})$ and of the ground state $\boldsymbol{n}_0(x)$ as the *excitation energy* of $\boldsymbol{n}(\boldsymbol{r})$. The configuration $\boldsymbol{n}^q(\boldsymbol{r})$ of charge q will be designated with a superscript. Among all of these charge-q configurations, $\boldsymbol{n}_0^q(\boldsymbol{r})$ is the one that minimizes the energy (7.22),

$$\mathbb{E}[\boldsymbol{n}_0^q] = \min_{\boldsymbol{n}^q}\mathbb{E}[\boldsymbol{n}^q]. \tag{8.38}$$

Clearly, the ground state configuration $\boldsymbol{n}_0(x|\varphi_0)$ has zero charge, as the charge density $\kappa[\boldsymbol{n}_0] \equiv 0$, and hence $\boldsymbol{n}_0 = \boldsymbol{n}_0^{q=0}$. The excitation energy (8.37) of the minimal-energy configuration $\boldsymbol{n}_0^q(\boldsymbol{r})$,

$$\Delta^q \equiv \min_{\boldsymbol{n}^q}\delta\mathbb{E}[\boldsymbol{n}^q] = \delta\mathbb{E}[\boldsymbol{n}_0^q] = \mathbb{E}[\boldsymbol{n}_0^q] - \mathbb{E}[\boldsymbol{n}_0], \tag{8.39}$$

will be called the *gap* of charge-q excitations. The minimum

$$\Delta = \min_{q \neq 0} \Delta^q \tag{8.40}$$

among all $q \neq 0$ defines the gap of excitations of any charge. Typically, Δ^q is a growing function of q, and therefore the unit-charge $q = \pm 1$ excitations, for which the Bloch sphere is covered once, determine the gap, $\Delta = \Delta^{q=\pm 1}$. The minimal-energy charge-q configurations $\boldsymbol{n}_0^q(\boldsymbol{r})$ satisfy the same stationary-point Eqs. (8.11) and (8.12) as the ground state $\boldsymbol{n}_0(x|\varphi_0)$, since they describe a *local* minimum in the configuration space. In order not to contain extensive contributions to the excitation energy (8.39), proportional to the size of the system, the configurations $\boldsymbol{n}_0^q(\boldsymbol{r})$ must asymptotically approach the ground state configuration $\boldsymbol{n}_0(x|\varphi_0)$,

$$\boldsymbol{n}_0^q(x, y \to \pm\infty) \to \boldsymbol{n}_0(x|\varphi_0), \quad \boldsymbol{n}_0^q(x \to -\infty, y) \to \boldsymbol{n}^\infty. \tag{8.41}$$

The finite-size region where most of the winding of the Bloch sphere occurs, $\boldsymbol{n}_0^q(\boldsymbol{r})$ clearly differs from $\boldsymbol{n}_0(x|\varphi_0)$ and the charge density $\kappa[\boldsymbol{n}_0^q](\boldsymbol{r})$ is located, can be referred to as the *core* of the charge excitation.

For our QHTI system with an edge, one should distinguish between bulk and edge excitations. In the bulk, skyrmions are the low-energy charge excitations, as introduced in Sec. 3.2.3. Their core is located deep in the bulk, away from the domain wall at the edge, such that its effect can be neglected. For the edge excitations on the other hand, the core is located in the domain wall. This makes it clear, on a qualitative level at least, that the energy of the edge charge excitation will generally be smaller than that of the bulk skyrmion. Since in the ground state $\boldsymbol{n}_0(x|\varphi_0)$ some changes in the isospin orientation are already present, less additional changes are required for $\boldsymbol{n}_0^q(\boldsymbol{r})$ to wind the whole Bloch sphere. Hence, the smaller energy cost. As the general properties of bulk skyrmions are well-understood, see references in Sec. 3.2.3, we concentrate on the edge excitations in the following.

8.3.1. Gapless Edge Excitations of the $n^\infty = n_z$ Phase

Let us begin with the $\boldsymbol{n}^\infty = \boldsymbol{n}_z$ phase with preserved U(1) symmetry in the bulk. According to Sec. 8.2, the bulk isospin orientation in this phase is exactly opposite to the edge isospin. There is an infinite number of geodesics $\boldsymbol{n}_0(x|\varphi_0)$, parameterized by the angle φ_0, connecting these two orientations, as depicted in Fig. 8.3. This degeneracy allows us, following Refs. [FI99, FB06], to construct a charge excitation by winding the angle φ_0 in the y direction along the edge. The corresponding ansatz $(\theta(x), \varphi(y))$ for $\boldsymbol{n}(\boldsymbol{r})$ decouples the stationary point equations (8.11) and (8.12), where the latter reduces to

$$-\nabla_y^2 \varphi(y) = 0. \tag{8.42}$$

We introduce sample boundaries along the y direction at $y = \pm\frac{L_y}{2}$ and impose the periodic boundary conditions

$$\boldsymbol{n}(x, y = +\frac{L_y}{2}) = \boldsymbol{n}(x, y = -\frac{L_y}{2}). \tag{8.43}$$

The solution to Eq. (8.42) is then given by

$$\varphi_0^q(y) = \varphi_0 + 2\pi q \frac{y}{L_y}, \quad \nabla_y \varphi_0^q = \frac{2\pi q}{L_y}, \quad u < h_z, \tag{8.44}$$

where q is an integer. This q is indeed the topological charge of the configuration, as can be confirmed from either the geometric considerations or explicit calculation using Eqs. (8.35) and (8.36). The energy functional then becomes

$$\begin{aligned}\mathbb{E}[\theta(x),\varphi_0^q(y)] &= L_y\int_{-\infty}^{0}\frac{\mathrm{d}x}{s}\left\{\frac{\rho}{2}\left[(\nabla_x\theta)^2+\sin^2\theta(\nabla_y\varphi_0^q)^2\right]+\mathcal{E}\left(\theta(x)\right)\right\}\\ &= L_y\left\{dE^{1\mathrm{D}}\left[\theta(x);u-\rho(\nabla_y\varphi_0^q)^2,h_z\right]+E^{1\mathrm{D}\infty}\right\}.\end{aligned} \tag{8.45}$$

We notice that the gradient term $\frac{\rho}{2}\sin^2\theta(\nabla_y\varphi_0^q)^2$ has the shape of the anisotropy energy $\frac{u}{2}(n_z^2-1)$. Therefore the functional (8.45) has the same form as the one for the ground state, Eq. (8.16), with the redefined anisotropy energy $u-\rho(\nabla_y\varphi_0^q)^2$ and can be expressed in terms of the quantities in Eq. (8.28). The functional (8.45) is thus minimized by the modified ground state configuration

$$\theta_0^q(x)=\theta_0(x;u-\rho(\nabla_y\varphi_0^q)^2,h_z),\ \ u<h_z, \tag{8.46}$$

and the gap (8.39) of charge-q edge excitations of the $\boldsymbol{n}^\infty=\boldsymbol{n}_z$ phase is expressed as

$$\Delta^q = L_y\left[dE_0^{1\mathrm{D}}(u-\rho\left(2\pi q/L_y\right)^2,h_z)-dE_0^{1\mathrm{D}}(u,h_z)\right],\quad u<h_z, \tag{8.47}$$

in terms of the domain-wall energy (8.31). The corresponding isospin configuration has the form

$$\boldsymbol{n}_0^q(x,y)=(\sin\theta_0^q(x)\cos\varphi_0^q(y),\sin\theta_0^q(x)\sin\varphi_0^q(y),\cos\theta_0^q(x)),\quad u<h_z, \tag{8.48}$$

and is shown in Fig. 8.5 for $q=1$. The gap (8.47) is finite only due to the finite size L_y of the sample in the y direction. It thus vanishes in the limit $L_y\to\infty$, where the phase $\boldsymbol{n}^\infty=\boldsymbol{n}_z$ supports gapless edge charge excitations, similar to the findings of Refs. [FI99, FB06]. The leading term in the large-size limit $L_y\to\infty$ can be obtained as an expansion

$$\begin{aligned}\Delta^q &\approx -L_y\partial_u dE_0^{1\mathrm{D}}(u,h_z)\rho\left(\frac{2\pi q}{L_y}\right)^2\\ &= (2\pi q)^2\frac{l_u}{L_y}F_2(\bar{h}_z)\varepsilon_\odot\sim q^2\frac{l_u}{L_y}\varepsilon_\odot,\quad u<h_z,\end{aligned} \tag{8.49}$$

where the energy

$$\varepsilon_\odot\equiv\frac{\rho}{s}\sim\frac{e_*^2}{l_{B_z}} \tag{8.50}$$

is associated with the gradient term. It is due to the SU(2)-symmetric interactions and thus set by the Coulomb energy. The leading order in equation (8.49) can also be obtained in a simpler way by just ignoring the effect of the winding on $\theta_0^q(x)$, thus setting $\theta_0^q(x)\approx\theta_0(x;u,h_z)$. The isospin OP is then $\boldsymbol{n}_0^q(\boldsymbol{r})\approx\boldsymbol{n}_0(x|\varphi_0^q(y))$ and the gap contains only the gradient term

$$\Delta^q\approx\int\frac{\mathrm{d}^2\boldsymbol{r}}{s}\frac{\rho}{2}\sin^2\theta_0(x)(\nabla\varphi_0^q(y))^2, \tag{8.51}$$

which agrees with Eq. (8.49).

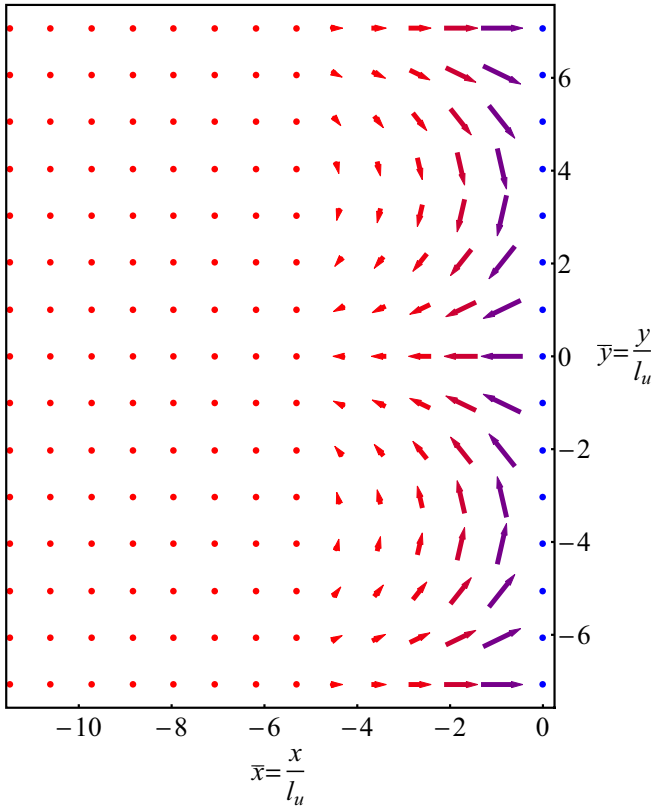


Figure 8.5. The unit-charge isospin configuration $\boldsymbol{n}_0^1(\boldsymbol{r})$ in the TnT $\boldsymbol{n}^\infty = \boldsymbol{n}_z$ phase with preserved U(1) symmetry, obtained numerically for $\bar{h}_z = 1.02$. The numerical solution agrees well with the exact analytical expressions (8.44), (8.46), and (8.48). The arrows depict the 2D components $(n_x(\boldsymbol{r}), n_y(\boldsymbol{r}))$, while the red-blue color code represents $n_z(\boldsymbol{r})$, in accord with Figs. 8.1 and 8.2. The paths on the Bloch sphere for varying $-\infty < x \leq 0$ and several constant values of y are shown in Fig. 8.3 (left). Reprinted figure with permission from Ref. [KJT16]. Copyright (2016) by the APS.

8.3.2. Gapped Edge Excitations of the $\boldsymbol{n}^\infty = -\boldsymbol{n}_z$ and $\boldsymbol{n}^\infty = \boldsymbol{n}^*(\varphi_0)$ Phases

The above construction of the gapless charge excitations in the $\boldsymbol{n}^\infty = \boldsymbol{n}_z$ phase is possible due to two conditions realized *in the ground state* $\boldsymbol{n}_0(x|\varphi_0)$. They are 1) preserved U(1) symmetry in the bulk and 2) continuous degeneracy of the ground state solution at the edge. In the other two phases we discuss now, one of these conditions is violated. In the intermediate phase $\boldsymbol{n}^\infty = \boldsymbol{n}^*(\varphi_0)$, the U(1) symmetry is spontaneously broken in the bulk and, for a given bulk order characterized by the angle φ_0, the ground state solution $\boldsymbol{n}_0(x|\varphi_0)$ is unique. In the $\boldsymbol{n}^\infty = -\boldsymbol{n}_z$ phase, on the other hand, the U(1) symmetry is preserved in the bulk, but the bulk and edge orientations are exactly the same and the ground state solution is just a constant $\boldsymbol{n}_0(x) \equiv -\boldsymbol{n}_z$. These crucial differences in the ground state edge configurations of the phases are visualized in Fig. 8.3. As a result, the ground state solution is unique in both phases for a given bulk order and an analogous construction of the gapless charge excitations is not possible. The edge charge excitations are therefore gapped. In the intermediate $\boldsymbol{n}^\infty = \boldsymbol{n}_0(\varphi_0)$ phase, the typical edge charge configuration $\boldsymbol{n}_0^1(\boldsymbol{r})$ has the form shown in Fig. 8.6, see Sec. 8.3.4 for details on the numerical calculations.

Let us consider the size of the gaps. For the phase $\boldsymbol{n}^\infty = -\boldsymbol{n}_z$, the lowest energy integer-

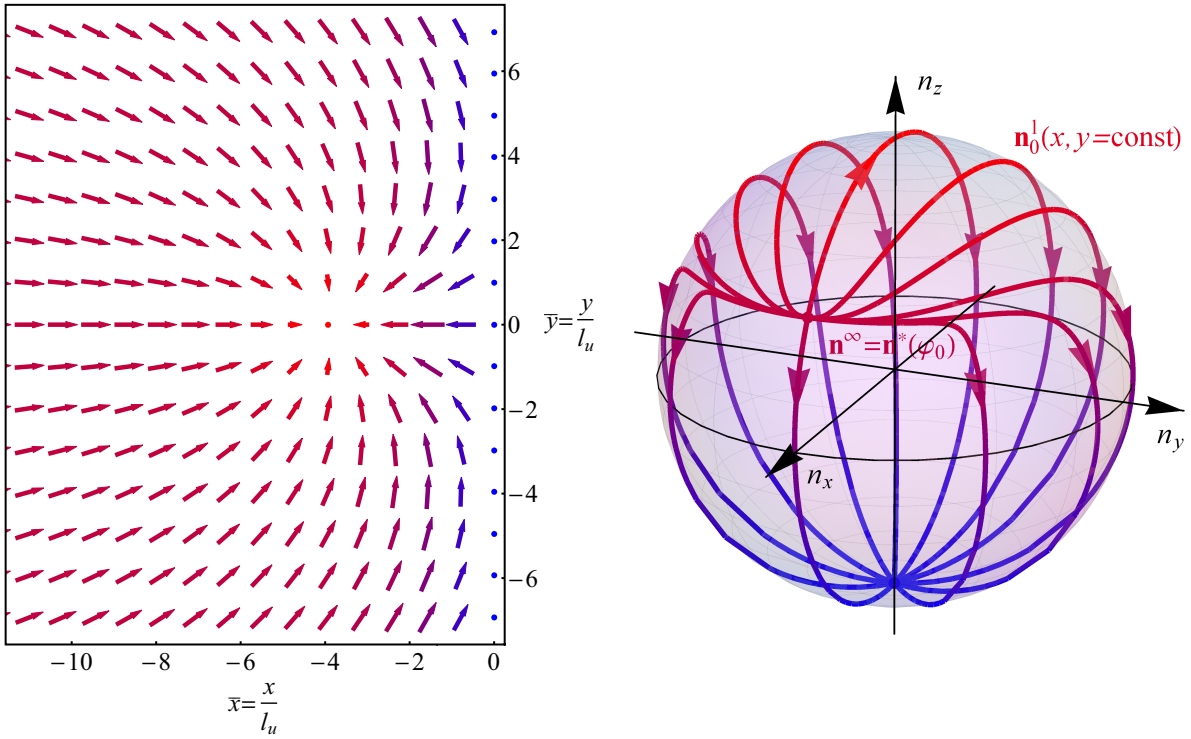


Figure 8.6. The isospin unit-charge configuration $\boldsymbol{n}_0^1(\boldsymbol{r})$ in the TT $\boldsymbol{n}^\infty = \boldsymbol{n}^*(\varphi_0)$ phase with spontaneously broken U(1) symmetry at $-u < h_z < u$, obtained numerically for $\bar{h}_z = 0.5$. (left) The arrows depict the 2D components $(n_x(\boldsymbol{r}), n_y(\boldsymbol{r}))$, while the red-blue color code represents $n_z(\boldsymbol{r})$. (right) The paths on the Bloch sphere $\boldsymbol{n}_0^1(x, y)$ for varying $-\infty < x \leq 0$ and several constant values of y. Reprinted figures with permission from Ref. [KJT16]. Copyright (2016) by the APS.

charge excitations are the bulk skyrmions, with the core infinitely far away from the edge. This is due to the fact that the *bulk* skyrmions are the minimal-energy configurations in an *infinite* sample among all charged configurations with the constraint $\boldsymbol{n}(r \to \infty) = -\boldsymbol{n}_z$. For our considered half-infinite sample where the boundary condition $\boldsymbol{n}(x = 0, y) = -\boldsymbol{n}_z$, Eq. (7.29), is imposed, the possible set of configurations is additionally constrained. This can only result in an increase of the excitation energy compared to that of the bulk skyrmions, for which such a constraint is absent. Thus placing the core closer to the edge can only lead to energetically less favorable configurations. As a result, in the $\boldsymbol{n}^\infty = -\boldsymbol{n}_z$ phase the charge excitations have the largest energy among all three phases, given by that of the bulk skyrmion,

$$\Delta^q = \Delta^q_{\text{sk}}, \quad h_z < -u. \tag{8.52}$$

In the σ-model we study, the term describing the Coulomb self-interaction of the charge density $\kappa[\boldsymbol{n}](\boldsymbol{r})$ (8.35) is neglected, as it is of higher order in the gradients. For this approximation, the skyrmion energy is given by

$$\Delta^q_{\text{sk}} = 4\pi\varepsilon_\odot|q| \tag{8.53}$$

and its size is formally zero due to the presence of the energy $\mathcal{E}(n_z)$ of the SU(2)-asymmetric terms, as found in Ref. [SKKR93].

Lastly, we consider the intermediate $\boldsymbol{n}^\infty = \boldsymbol{n}^*(\varphi_0)$ phase. Its gap $\Delta^q(h_z)$ of the edge charge excitations, Eq. (8.39), can be qualitatively understood from a continuity argument.

Since the second-order transitions at $h_z = \pm u$ are continuous and the intermediate phase continuously interpolates between the $\boldsymbol{n}^\infty = \boldsymbol{n}_z$ and $\boldsymbol{n}^\infty = -\boldsymbol{n}_z$ phases, $\Delta^q(h_z)$ monotonically and continuously grows upon decreasing h_z in the range $-u < h_z < u$, starting from zero at $h_z = u$ and reaching its maximal value of Δ^q_{sk} for $h_z = -u$. As the bulk phase is controlled solely by the normalized dimensionless field $\bar{h}_z$ (8.7), the gap has the scaling form

$$\Delta^q(h_z) = \varepsilon_\odot \bar{\Delta}^q(\bar{h}_z) \tag{8.54}$$

with $\bar{\Delta}^q(\bar{h}_z)$ a dimensionless function of $\bar{h}_z$.

An analogous continuous growth of the edge excitation gap was earlier predicted for the CAF phase of the $\nu = 0$ state in graphene, originally employing a similar continuity argument in Ref. [Kha12a] and within a simplified picture of single-particle edge excitations in Ref. [Kha12b]. The corresponding transport behavior, well consistent with the presented scenario, was shortly after observed in both bilayer [MDY+13] and monolayer [YSYH+13] graphene. More recently, an analytical estimate for the edge excitation gap of the CAF phase was obtained in Ref. [MSF14] within a low-energy theory approach analogous to the one employed by us. The estimate we make below is in accordance with that result.

8.3.3. Intermediate Phase close to the Phase Transition $h_z = u$: Analytical Estimates

In the intermediate phase $\boldsymbol{n}^\infty = \boldsymbol{n}^*(\varphi_0)$ close to the phase transition $h_z = u$, the gap $\bar{\Delta}^q(\bar{h}_z) \ll 1$ is small and can be estimated analytically. Analogous to the construction for the $\boldsymbol{n}^\infty = \boldsymbol{n}_z$ phase, we consider for the intermediate phase the configuration $\boldsymbol{n}_0(x|\varphi^q(y))$ obtained from the ground state configuration (8.15) and (8.21) by slowly winding, at scales much larger than the domain-wall width l_u, the angle $\varphi^q(y)$ q times as the y direction is spanned. The concrete shape of $\varphi^q(y)$ is to be optimized. The excitation energy $\delta\mathbb{E}[\boldsymbol{n}_0(x|\varphi^q(y))]$ of such a configuration would contain only the gradient contribution, analogous to the expansion Eq. (8.49). However, unlike the $\boldsymbol{n}^\infty = \boldsymbol{n}_z$ phase, the bulk asymptotic angle $\theta^\infty = \theta^* \neq 0$ (8.6) is nonzero in the intermediate phase. Thus the integral $\int_{-\infty}^0 \mathrm{d}x \ldots$ would not be constrained to the domain-wall region of size l_u, but also contain an extensive contribution proportional to the size of the sample in the x direction. Additionally, due to the winding of $\varphi^q(y)$, the asymptotic value $\boldsymbol{n}_0(x = -\infty|\varphi^q(y)) = \boldsymbol{n}^*(\varphi^q(y))$ would differ from $\boldsymbol{n}^q(x = -\infty, y) = \boldsymbol{n}^*(\varphi_0)$ of the bulk ground state. Qualitatively, the charge-q configuration $\boldsymbol{n}^q(\boldsymbol{r})$ must have the form shown in Fig. 8.6 for $q = 1$.

Nonetheless, for a given $\varphi^q(y)$, the energy in the domain-wall region is minimized well by the configuration $\boldsymbol{n}_0(x|\varphi^q(y))$. Thus to obtain a global solution, the configuration in the bulk needs to be modified. The proximity to the phase transition allows us to efficiently separate the domain-wall and bulk contributions as follows. First, we define the deviation

$$\delta\bar{h}_z = \frac{\delta h_z}{u} = \bar{h}_z - 1, \quad \delta h_z = h_z - u, \tag{8.55}$$

which is negative and small, $|\delta\bar{h}_z| \ll 1$, in the considered regime. There, the asymptotic bulk order $\boldsymbol{n}^\infty = \boldsymbol{n}^*(\varphi_0)$ deviates only a little from $\boldsymbol{n}_z$. Using Eq. (8.6), the optimal angle in the bulk ground state is given by

$$\theta^{*2} \approx 2|\delta\bar{h}_z| \ll 1. \tag{8.56}$$

In regions where the isospin $\boldsymbol{n}(\boldsymbol{r})$ is close to $\boldsymbol{n}_z$, such that its angle $\theta(\boldsymbol{r}) \lesssim \theta^*$, the energetics is governed by this smaller scale $|\delta h_z| \ll u$. The associated spatial scale

$$l_{\delta h_z} \equiv \sqrt{\frac{\rho}{|\delta h_z|}} = \frac{l_u}{\sqrt{|\delta \bar{h}_z|}} \gg l_u \tag{8.57}$$

follows from Eq. (8.25) and is much larger than the domain-wall width l_u. We choose a length scale x_0 in between,

$$l_{\delta h_z} \gg x_0 \gg l_u, \tag{8.58}$$

such that the ground state configuration $\boldsymbol{n}_0(x = -x_0|\varphi_0)$ at $x = -x_0$ is already close to its bulk asymptotic value, $\theta_0(x = -x_0) \approx \theta^*$. We emphasize that even exactly at the phase transition $h_z = u$, the domain wall width is l_u, only the bulk value is approached as a power law and not exponentially, see Eq. (8.26).

Now we come to the separation. For the sought charge-q configuration $\boldsymbol{n}^q(\boldsymbol{r})$, we consider the above ground state configuration with the adiabatically changing angle $\varphi^q(y)$ only in the region up to the distance x_0 from the edge:

$$\boldsymbol{n}^q(\boldsymbol{r}) \approx \boldsymbol{n}_0(x|\varphi^q(y)), \quad -x_0 < x \leq 0. \tag{8.59}$$

Due to the condition $l_{\delta h_z} \gg x_0$, the contribution to the excitation energy from the region $-x_0 < x < 0$ is not extensive and can be approximated as

$$\begin{aligned} \delta\mathbb{E}[\boldsymbol{n}^q]_{(-x_0,0)} &= \int_{-x_0<x<0} \frac{\mathrm{d}^2\boldsymbol{r}}{s} \left(E[\boldsymbol{n}^q] - E[\boldsymbol{n}_0]\right) \\ &\approx \#\varepsilon_\odot l_u \int_{-\infty}^{+\infty} \mathrm{d}y \, [\nabla_y \varphi(y)]^2. \end{aligned} \tag{8.60}$$

Here and below, # indicates undetermined numerical factors that are beyond the accuracy of the considered approximation.

In the remaining bulk region $x < -x_0$, the configuration $\boldsymbol{n}^q(\boldsymbol{r})$ has to connect the boundary values $\boldsymbol{n}^q(x = -\infty, y) = \boldsymbol{n}^*(\varphi_0)$ and $\boldsymbol{n}^q(x = -x_0, y) = \boldsymbol{n}_0(x|\varphi^q(y)) \approx \boldsymbol{n}^*(\varphi^q(y))$. In the whole region between, the isospin $\boldsymbol{n}^q(\boldsymbol{r})$ is close to $\boldsymbol{n}_z$ and therefore varies over spatial scales on the order of $l_{\delta h_z}$ or greater. The connection is made by trapping a vortex of charge q in this bulk region $x < -x_0$, as depicted by the red dot in Fig. 8.6 (left). On the one side of the vortex, closer to the edge, the phase $\varphi^q(x, y)$ winds q times along the $x = -x_0$ boundary, while the angle $\theta \approx \theta^*$ remains almost constant. On the other side, the winding is gone. The leading contribution to the energy of such a vortex configuration comes from the region outside of its core, where the latter is defined as the area where the isospin covers the solid angle $\theta \lesssim \theta^*$. This contribution is logarithmic; to obtain it, we consider the radial form $\theta(r, \phi) = \theta^*$, $\varphi(r, \phi) = q\phi$, where $\boldsymbol{r} = r(\cos\phi, \sin\phi)$, relative to the center of the vortex in the bulk region $x < -x_0$, given by the point at which $\boldsymbol{n}_0^q(\boldsymbol{r}) = \boldsymbol{n}_z$ in Fig. 8.6. We find for the energy

$$\begin{aligned} \delta\mathbb{E}[\boldsymbol{n}^q]_{(-\infty,-x_0)} &= \int_{x<-x_0} \frac{\mathrm{d}^2\boldsymbol{r}}{s} \left(E[\boldsymbol{n}^q] - E[\boldsymbol{n}_0]\right) \approx \int \frac{r\mathrm{d}r\,\mathrm{d}\phi}{s} \frac{\rho}{2} \sin^2\theta (\nabla\varphi)^2 \\ &\approx \int_{|q|l_{\delta h_z}}^{l} \frac{2\pi r \mathrm{d}r}{s} \frac{\rho}{2} \theta^{*2} \frac{q^2}{r^2} = \frac{2\pi\rho}{s} |\delta\bar{h}_z| q^2 \ln \frac{l}{|q| l_{\delta h_z}}. \end{aligned} \tag{8.61}$$

The lower limit is determined by the size $|q|l_{\delta h_z}$ of the vortex core. The upper limit $l \gg |q|l_{\delta h_z}$ is, in our case, set by the distance from the vortex core to the edge. The same scale l has to

match the extent of the winding region of $\varphi^q(y)$ in the domain wall along the y direction. We estimate $\nabla_y \varphi^q(y) \sim q/l$ and add the domain-wall (8.60) and bulk (8.61) contributions. The excitation energy of this constructed configuration is then given by

$$\begin{aligned} \delta\mathbb{E}[\boldsymbol{n}^q] &= \delta\mathbb{E}[\boldsymbol{n}^q]_{(-\infty,-x_0)} + \delta\mathbb{E}[\boldsymbol{n}^q]_{(-x_0,0)} \\ &\approx \varepsilon_\odot q^2 \left(2\pi |\delta\bar{h}_z| \ln \frac{l}{|q| l_{|\delta h_z|}} + \# \frac{l_u}{l} \right). \end{aligned} \tag{8.62}$$

The dimension l is the only remaining variational parameter, which we use to minimize this energy. The leading terms of the asymptotics expansion are then

$$\bar{\Delta}^q(\bar{h}_z \to 1 - 0^+) = \pi q^2 |\delta\bar{h}_z| \ln \frac{C}{q^2 |\delta\bar{h}_z|} + \mathcal{O}(|\delta\bar{h}_z|) \tag{8.63}$$

for the gap of charge edge excitations in the intermediate phase $\boldsymbol{n}^\infty = \boldsymbol{n}^*(\varphi_0)$ close to the phase transition $h_z = u$. The minimum (8.63) of Eq. (8.62) is reached at the optimal length

$$l^* \equiv \frac{l_u}{|\delta\bar{h}_z|} = \frac{l_{\delta h_z}}{\sqrt{|\delta\bar{h}_z|}}. \tag{8.64}$$

The numerical factor $C \sim 1$ cannot be determined within the accuracy of the considered logarithmic approximation. For unit charge $q = \pm 1$, the estimate (8.63) agrees with that of Ref. [MSF14].

8.3.4. Numerical Calculations

We complement the analytical studies from above by explicit numerical calculations of the unit-charge $q = 1$ edge excitations. The configurations $\boldsymbol{n}_0^1(\boldsymbol{r})$ with the energy minimum (8.38) within the $q = 1$ sector are found by solving the discretized version of the stationary-point Eqs. (8.11) and (8.12). For this we use a variant of the multi-grid relaxation methods for boundary value problems presented in Ref. [WHPF92]. Our relaxation scheme preserves the topological charge of the configuration, such that the initial trial configuration of a given charge converges to the minimal-energy configuration within that charge sector. The calculations are performed for a finite-size system $(x, y) \in (-L_x, 0) \times (-\frac{L_y}{2}, +\frac{L_y}{2})$ with dimensions L_x and L_y. For all sizes L_y indicated in Fig. 8.7, except the largest one $(L_x, L_y) = (120, 160)\sqrt{2} l_u$, we used square samples with $L_x = L_y$. The gap $\bar{\Delta}^1(\bar{h}_z)$ is then calculated using a discretized version of Eq. (8.39).

In Fig. 8.7, we plot the gap $\bar{\Delta}^1(\bar{h}_z)$ as a function of the normalized field $\bar{h}_z$ for several different system sizes $L_{x,y}/l_u$. In the region $u < h_z$ of the $\boldsymbol{n}^\infty = \boldsymbol{n}_z$ phase, the numerically calculated gap $\bar{\Delta}^1(\bar{h}_z)$ accurately agrees with the exact analytic dependence (8.47) on $\bar{h}_z$ and L_y. This confirms that the edge charge excitations are gapless in the infinite-size limit $L_y \to \infty$. The typical edge charge configuration $\boldsymbol{n}_0^1(\boldsymbol{r})$ in the $\boldsymbol{n}^\infty = \boldsymbol{n}_z$ phase has the form shown in Fig. 8.5, in full agreement with the analytical expressions (8.44), (8.46), and (8.48).

In the region $-u < h_z < u$ of the intermediate phase $\boldsymbol{n}^\infty = \boldsymbol{n}^*(\varphi_0)$, the gap becomes independent of the sample dimensions $L_{x,y}$, as the size of the configurations shrinks with decreasing $\bar{h}_z$ and is at some point smaller than $L_{x,y}$. Close to the transition point $h_z = u$ in the intermediate phase, for large enough $L_{x,y}/l_u$, the numerical data points fit to the analytical estimate (8.63). We obtain for the largest-size sample $(L_x, L_y) = (120, 160)\sqrt{2} l_u$,

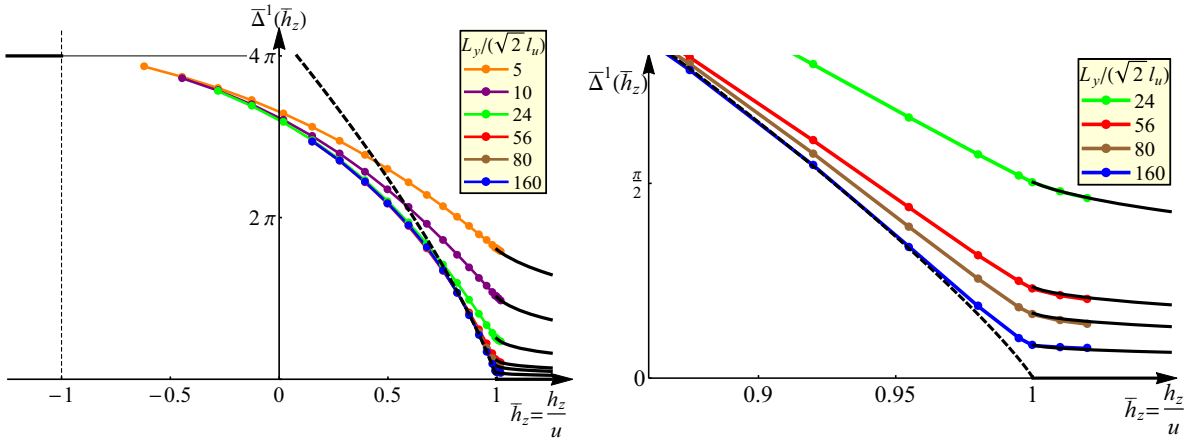


Figure 8.7. (left) The dependence of the gap $\bar{\Delta}^1(\bar{h}_z)$, Eq. (8.54), of the edge charge excitations $\boldsymbol{n}_0^1(\boldsymbol{r})$ on the normalized field $\bar{h}_z$, calculated numerically for various sample sizes. The black solid lines in the $1 < \bar{h}_z$ region are the exact gap dependencies (8.47) for a finite-size system. The black horizontal solid line for $\bar{h}_z < -1$ is the analytical value (8.53) of the gap, given by the energy of a free skyrmion. The black dashed line in the $-1 < \bar{h}_z < 1$ region shows the asymptotic gap dependence (8.63) with the fitted parameter $\ln C = 4.27$. (right) Zoomed-in region around the phase transition point $\bar{h}_z = 1$. Reprinted figures with permission from Ref. [KJT16]. Copyright (2016) by the APS.

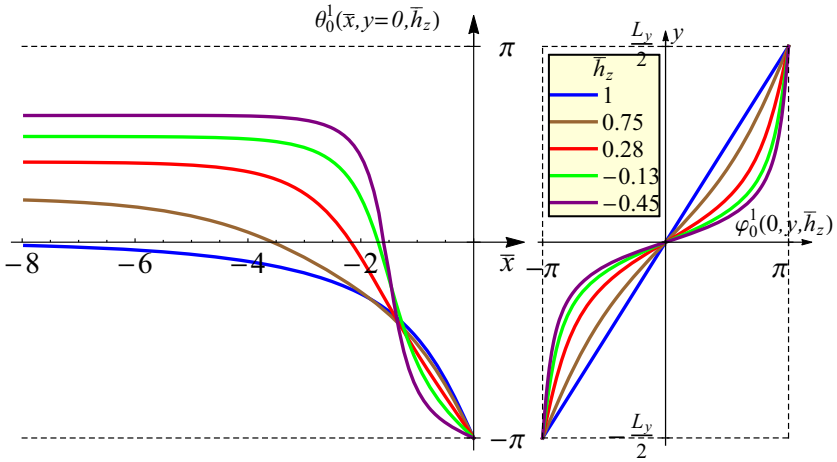


Figure 8.8. The angle functions $\theta_0^1(x, y = 0)$ and $\varphi_0^1(x = -0, y)$ of the unit-charge edge configurations $\boldsymbol{n}_0^1(\boldsymbol{r})$ in the intermediate phase $\boldsymbol{n}^\infty = \boldsymbol{n}^*(\varphi_0)$ for $-u < h_z < u$, obtained numerically for a $L_x = L_y = 10\sqrt{2}l_u$ sample for various values of h_z. The size of the configuration decreases with decreasing h_z. Reprinted figure with permission from Ref. [KJT16]. Copyright (2016) by the APS.

hence with the smallest size effects, $\ln C = 4.27$ by fitting to the data points $(\bar{h}_z, \bar{\Delta}^1(\bar{h}_z)) = (0.955, 1.058), (0.920, 1.715), (0.875, 2.484)$. Considering the opposite phase transition point $h_z = -u$, the available numerical data for $\bar{\Delta}^1(\bar{h}_z)$ visually extrapolates well to the value $\bar{\Delta}^1(\bar{h}_z = -1) = 4\pi$, Eqs. (8.52) and (8.53), of the bulk skyrmion.

The typical edge charge configuration $\boldsymbol{n}_0^1(\boldsymbol{r})$ in the $\boldsymbol{n}^\infty = \boldsymbol{n}^*(\varphi_0)$ phase is shown in Fig. 8.6. The corresponding angle functions $\theta_0^1(x, 0)$ and $\varphi_0^1(x = -0, y)$ of such a $\boldsymbol{n}_0^1(\boldsymbol{r})$ are plotted for various $\bar{h}_z$ in Fig. 8.8. The regions where these functions vary determine the size of the charge excitation. As h_z decreases in the range $-u < h_z < u$, this size monotonically decreases,

becoming smaller than l_u and approaching zero for $h_z \to -u + 0^+$. These behaviors of the gap and the size of the excitation with decreasing h_z in the $-u < h_z < u$ region are in accord with the general arguments of Sec. 8.3.2. For our considered model with neglected Coulomb self-interaction of the charge density (8.35), skyrmions in the $\boldsymbol{n}^\infty = -\boldsymbol{n}_z$ phase for $h_z < -u$ have size zero.

As a closing remark, we point out that while the properties of the edge excitations of the three phases are different, their bulk charge excitations are qualitatively the same. The bulk charge gap is finite in all three phases and never closes during the transformation from the TnT $\boldsymbol{n}^\infty = \boldsymbol{n}_z$ to the TT $\boldsymbol{n}^\infty = -\boldsymbol{n}_z$ phase with decreasing h_z. As rigorous mathematical definitions of topological phases in interacting systems are currently an actively researched subject, we adopt an intuitive nomenclature, referring to the phases with gapless and gapped edge excitations as TnT and TT phases, respectively.

8.4. Helical Luttinger Liquid

In the previous Sec. 8.3, it was demonstrated that the $\boldsymbol{n}^\infty = \boldsymbol{n}_z$ phase at $u < h_z$ is characterized by gapless charge edge excitations. Their dynamics can be described by an effective low-energy theory, the Helical Luttinger Liquid, which we now derive.

8.4.1. Derivation

The nondegenerate bulk ground state $\boldsymbol{n}^\infty = \boldsymbol{n}_z$ has a gapped excitation spectrum. The low-energy theory for the edge is then valid at energies below this gap $h_z - u$, i.e, as long as the bulk is not excited. This criterion will be established more rigorously below. The gapless edge excitations originate from the degeneracy of the (mean-field) ground state solution $\boldsymbol{n}_0(x;\varphi_0)$ at the edge, characterized by an arbitrary angle φ_0. In order to model the dynamics of these excitations, we now include slow variations of the angle φ_0 in space and time and perform a gradient expansion. As in the original Lagrangian $\mathbb{L}[\theta,\varphi]$ (7.20)-(7.23) the variables $\varphi(\boldsymbol{r};t)$ and $\theta(\boldsymbol{r};t)$ are coupled, also the dynamics of $\theta(\boldsymbol{r};t)$ needs to be considered. However, the deviations from the ground state configurations due to a slowly varying $\varphi(\boldsymbol{r};t)$ will be small, and so we expand the Lagrangian $\mathbb{L}[\theta,\varphi]$ in deviations $\delta\theta(\boldsymbol{r};t)$ about $\theta_0(x)$,

$$\theta(\boldsymbol{r};t) = \theta_0(x) + \delta\theta(\boldsymbol{r};t), \quad \delta\theta(x=0,y;t) = 0, \tag{8.65}$$

with the appropriate boundary condition at the edge. As a start, we assume a general, but slow, dependence of $\varphi(\boldsymbol{r},t)$ on $\boldsymbol{r}$ and t. As we are interested in the leading order in gradients of $\varphi(\boldsymbol{r};t)$, it is sufficient to expand all terms to the lowest necessary order in $\delta\theta(\boldsymbol{r};t)$. For the kinetic term (7.21), this gives

$$K[\delta\theta,\varphi] = \frac{\dot{\varphi}}{2}[\cos\theta_0 + \delta\theta\partial_\theta\cos\theta_0 + \mathcal{O}(\delta\theta^2)] \to \frac{\dot{\varphi}}{2}\delta\theta\partial_\theta\cos\theta_0. \tag{8.66}$$

The term $\frac{\dot{\varphi}}{2}\cos\theta_0$ is an inconsequential full time derivative and we drop it. For now, the derivative $\partial_\theta\cos\theta_0 = -\sin\theta_0$ is kept as it is. Considering the gradient term in Eq. (7.22),

$$\frac{\rho}{2}\sin^2\theta(\nabla\varphi)^2 = \frac{\rho}{2}[\sin^2\theta_0(x) + \mathcal{O}(\delta\theta)](\nabla\varphi)^2 \to \frac{\rho}{2}\sin^2\theta_0(x)(\nabla\varphi)^2, \tag{8.67}$$

keeping only the zero-order term is sufficient. The remaining terms

$$\int\frac{\mathrm{d}^2\boldsymbol{r}}{s}\left\{\frac{\rho}{2}(\nabla\theta)^2 + \mathcal{E}(\theta)\right\} = \int\frac{\mathrm{d}^2\boldsymbol{r}}{s}\left\{E[\boldsymbol{n}_0] + \frac{1}{2}\delta\theta\hat{U}[\delta\theta] + \mathcal{O}(\delta\theta^3)\right\} \to \int\frac{\mathrm{d}^2\boldsymbol{r}}{s}\frac{1}{2}\delta\theta\hat{U}[\delta\theta] \tag{8.68}$$

in Eq. (7.22) do not depend on $\varphi(\boldsymbol{r};t)$ and need to be expanded to quadratic order in $\delta\theta(\boldsymbol{r};t)$. Here, we dropped the zero-order ground state energy $\mathbb{E}[\boldsymbol{n}_0]$ and defined

$$\hat{U} = -\rho\nabla_y^2 + \hat{U}_x, \quad \hat{U}_x = -\rho\nabla_x^2 + \partial_\theta^2\mathcal{E}(\theta_0(x)), \quad \partial_\theta^2\mathcal{E}(\theta) = -u\cos 2\theta + h_z\cos\theta. \tag{8.69}$$

Collecting these leading terms (8.66), (8.67), and (8.68), we approximate the initial Lagrangian as

$$\begin{aligned} \mathbb{L}[\theta,\varphi] \to \mathbb{L}'[\delta\theta,\varphi] &= \int \frac{\mathrm{d}^2\boldsymbol{r}}{s}\, L'[\delta\theta,\varphi], \\ -L'[\delta\theta,\varphi] &\equiv -\frac{\dot{\varphi}}{2}\delta\theta\partial_\theta\cos\theta_0 + \frac{1}{2}\delta\theta\hat{U}[\delta\theta] + \frac{\rho}{2}\sin^2\theta_0(\nabla\varphi)^2. \end{aligned} \tag{8.70}$$

Its structure allows for further approximations. First, we observe that in both terms containing the $\varphi(\boldsymbol{r};t)$ variable, the function $\sin\theta_0(x)$ is present, constraining them to the domain-wall region of size l_u. This makes it reasonable to split the field into an x-independent average $\Phi(y;t) \equiv \langle\varphi(x,y;t)\rangle_x$ and fluctuations $\delta_x\varphi(x,y;t)$ along x like

$$\varphi(x,y;t) = \Phi(y;t) + \delta_x\varphi(x,y;t), \tag{8.71}$$

with $\langle\delta_x\varphi(x,y;t)\rangle_x = 0$. There is a certain freedom in the definition of the used average. Since the parametrization by the spherical angles becomes degenerate at $\theta = 0,\pi$, a meaningful average requires a weight function that takes this into account. We choose for the latter $\sin\theta_0(x)$ and define the average as

$$\langle f(x)\rangle_x \equiv \int_{-\infty}^{0} \mathrm{d}x\, \sin\theta_0(x) f(x). \tag{8.72}$$

Other choices could also be used, and the exact definition of the average will not be essential in the following. At low energies, $\varphi(x,y;t)$ varies over spatial scales exceeding the domain wall size l_u. Thus the fluctuations of the field $\varphi(x,y;t)$ in the x direction across the domain wall, described by $\delta_x\varphi(x,y;t)$, will produce a parametrically smaller contribution than those in the y direction along the domain wall, described by $\Phi(y;t)$. To the leading order we therefore neglect $\delta_x\varphi(x,y;t)$ and the field

$$\varphi(x,y;t) \to \Phi(y;t) \tag{8.73}$$

may be approximated by a quasi-1D field $\Phi(y;t)$. The corresponding gradient term becomes $(\nabla\varphi)^2 \to (\nabla_y\Phi)^2$.

Next, we analyze the properties of the operator $\hat{U}_x$ in Eq. (8.69). It has the form of the Hamiltonian for a Schrödinger particle in 1D with the potential energy $\partial_\theta^2\mathcal{E}(\theta_0(x))$, plotted in Fig. 8.9, and the hardwall boundary condition from Eq. (8.65). In the bulk,

$$\partial_\theta^2\mathcal{E}(\theta_0(x=-\infty)=0) = h_z - u > 0, \tag{8.74}$$

which represents the mass of the isospin wave in an infinite system. The minimum

$$\partial_\theta^2\mathcal{E}(\theta_0(0)=\pi) = -h_z - u < 0 \tag{8.75}$$

is reached at the edge and is negative. This potential energy $\partial_\theta^2\mathcal{E}(\theta_0(x))$ curve has a well in the region of the domain wall, and for $h_z < 4u$ it also has a barrier, as depicted in Fig. 8.9. From this analogy with the quantum-mechanical problem, we conclude that the operator $\hat{U}_x$

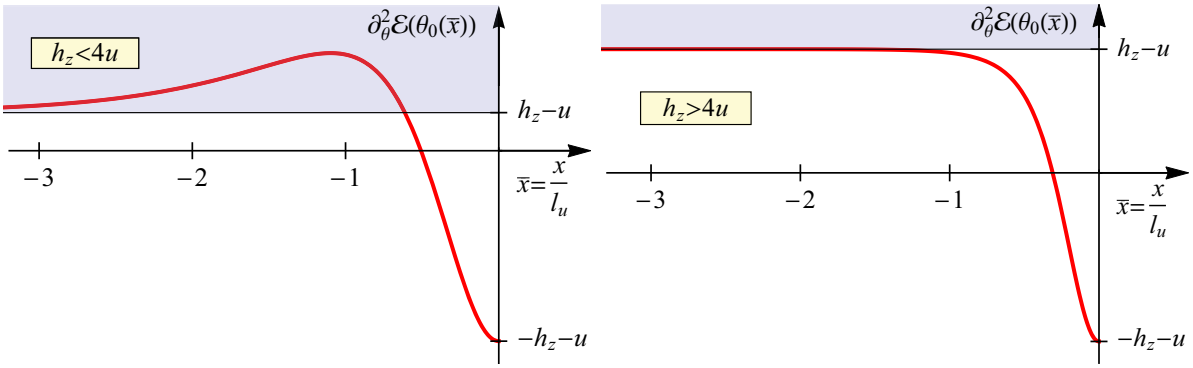


Figure 8.9. The function $\partial_\theta^2\mathcal{E}(\theta_0(x))$, which serves as an effective potential energy in the operator $\hat{U}_x$ defined in Eq. (8.69). The latter describes the quadratic fluctuations $\delta\theta$ about the ground state configuration $\theta_0(x)$ (8.20) in the $\boldsymbol{n}^\infty = \boldsymbol{n}_z$ phase for $u < h_z$. The operator $\hat{U}_x$ has only a continuous spectrum, marked by the shaded region, that starts at the bulk asymptotic value $\partial_\theta^2\mathcal{E}(\theta_0(x=-\infty)) = h_z - u$, equal to the gap of the bulk isospin wave excitations. There are no discrete bound states. Reprinted figures with permission from Ref. [KJT16]. Copyright (2016) by the APS.

has a continuous eigenvalue spectrum in the energy range $(h_z - u, +\infty)$. It *could* also have discrete levels in the range $(-h_z - u, h_z - u)$, corresponding to states bound within the well. However, we have checked numerically that there are no such discrete eigenvalues. In fact, only the positive range $(0, h_z - u)$ requires a check for bound states, as negative eigenvalues are prohibited. Since $\hat{U}_x$ is a quadratic form of the expansion about the minimum-energy configuration, it is a positive-definite operator.

Thus, $\hat{U}_x$ has only a continuous "massive" eigenvalue spectrum that starts from $h_z - u > 0$. The eigenfunctions are extended and can be viewed as bulk isospin waves, modified at the edge. Due to these properties, the operator $-\rho\nabla_y^2$ in $\hat{U}$ (8.69) may be neglected compared to $\hat{U}_x$,

$$\hat{U} = -\rho\nabla_y^2 + \hat{U}_x \to \hat{U}_x. \tag{8.76}$$

The approximations (8.73) and (8.76) allow us to further simplify the Lagrangian (8.70) as

$$\begin{aligned} \mathbb{L}'[\delta\theta, \varphi] \to \mathbb{L}''[\delta\theta, \Phi] &= \int \frac{\mathrm{d}^2\boldsymbol{r}}{s}\, L''[\delta\theta, \Phi], \\ -L''[\delta\theta, \Phi] &\equiv -\frac{\dot{\Phi}}{2}\delta\theta\partial_\theta\cos\theta_0 + \frac{1}{2}\delta\theta\hat{U}_x[\delta\theta] + \frac{\rho}{2}\sin^2\theta_0(\nabla_y\Phi)^2. \end{aligned} \tag{8.77}$$

This Lagrangian (8.77) is a second-order functional polynomial in $\delta\theta(\boldsymbol{r};t)$. We consider the configuration $\delta\theta_0[\Phi](\boldsymbol{r};t)$ that minimizes $-\mathbb{L}''[\delta\theta, \Phi]$ with respect to $\delta\theta(\boldsymbol{r};t)$ for a given $\Phi(y;t)$. It satisfies the stationary-point equation

$$-\frac{\delta}{\delta(\delta\theta)}\int \mathrm{d}t\, \mathbb{L}''[\delta\theta, \Phi] = -\frac{\dot{\Phi}}{2}\partial_\theta\cos\theta_0 + \hat{U}_x[\delta\theta] = 0. \tag{8.78}$$

This is a pure differential-in-x equation, and its solution can thus be formally written as

$$\delta\theta_0[\Phi](\boldsymbol{r};t) = \partial\theta_0(x)\frac{\dot{\Phi}(y;t)}{2}, \quad \partial\theta_0(x) = \hat{U}_x^{-1}[\partial_\theta\cos\theta_0]. \tag{8.79}$$

In terms of this minimum configuration and the deviation $\delta\tilde{\theta}(\boldsymbol{r};t) = \delta\theta(\boldsymbol{r};t) - \delta\theta_0[\Phi](\boldsymbol{r};t)$ from it, the Lagrangian (8.77) can be rewritten identically as

$$-L''[\delta\theta_0[\Phi]+\delta\tilde{\theta},\Phi] = -\frac{1}{2}\left(\frac{\dot{\Phi}}{2}\right)^2 \partial\theta_0(x)\partial_\theta\cos\theta_0(x) + \frac{\rho}{2}\sin^2\theta_0(x)(\nabla_y\Phi)^2 + \frac{1}{2}\delta\tilde{\theta}\hat{U}_x[\delta\tilde{\theta}]. \quad (8.80)$$

Essentially, this procedure amounts to completing the square of the quadratic polynomial in the functional sense, decoupling the field $\Phi(y;t)$ and the variable $\delta\tilde{\theta}(\boldsymbol{r};t)$ of the operator $\hat{U}_x$. Since the latter has a gap $h_z - u$, the contribution of $\delta\tilde{\theta}(\boldsymbol{r};t)$ can be neglected below this energy scale. Thus by setting $\delta\tilde{\theta}(\boldsymbol{r};t) \to 0$ in Eq. (8.80), we find the 1D effective Lagrangian

$$\mathbb{L}^{\mathrm{1D}}[\Phi] \equiv \mathbb{L}''[\delta\theta_0[\Phi],\Phi]. \quad (8.81)$$

The y-dependent 1D field $\Phi(y;t)$ is the only remaining variable. This allows us to separate the integrals over x and y and write the final Lagrangian for the edge excitations of the $\boldsymbol{n}^\infty = \boldsymbol{n}_z$ phase in the form

$$\mathbb{L}^{\mathrm{1D}}[\Phi] = \int \mathrm{d}y\, L^{\mathrm{1D}}[\Phi], \quad L^{\mathrm{1D}}[\Phi] = \frac{1}{8\pi\mathcal{K}}\left[\frac{1}{v}\dot{\Phi}^2 - v(\nabla_y\Phi)^2\right]. \quad (8.82)$$

We recognize in Eq. (8.82) the Lagrangian of a Luttinger liquid, compare to Eqs. (3.60) and (3.61) in Sec. 3.2.4, with the phase field $\Phi(y;t)$ at the edge being the collective bosonic variable. The parameters v and $\mathcal{K}$ are given by

$$\frac{1}{8\pi\mathcal{K}v} \equiv \frac{l_u}{su}F_t(\bar{h}_z), \quad F_t(\bar{h}_z) \equiv \frac{u}{8}\int_{-\infty}^{0}\mathrm{d}\bar{x}\,\partial_\theta\cos\theta_0(x)\partial\theta_0(x), \quad (8.83)$$

$$\frac{v}{8\pi\mathcal{K}} \equiv \frac{\rho l_u}{s}F_y(\bar{h}_z), \quad F_y(\bar{h}_z) = \frac{1}{2}\int_{-\infty}^{0}\mathrm{d}\bar{x}\sin^2\theta_0(\bar{x};\bar{h}_z). \quad (8.84)$$

Due to the scaling form $\theta_0(x) = \theta_0(\bar{x};\bar{h}_z)$ of the ground state solution (8.20), the parameters can be expressed in terms of the dimensionless functions $F_{t,y}(\bar{h}_z)$. We recognize that $F_y(\bar{h}_z) = F_2(\bar{h}_z)$ defined in Eq. (8.33), yet without additional insights, calculating $F_t(\bar{h}_z)$ would require first finding the solution $\partial\theta_0(x)$ to the differential equation (8.78) and then calculating the integral in Eq. (8.83). Instead, below we provide a more elegant and streamlined way of deriving the low-energy model (8.82). It not only allows us to obtain the explicit expression for $F_t(\bar{h}_z)$, but also uncovers the origin of the low-energy model in the degenerate ground state solution. The above derivation is nonetheless useful for justifying the employed approximations.

A more elegant Derivation

We make two crucial observations about the general structure of the Lagrangian (7.20)-(7.23). First, the kinetic term $\frac{\dot{\varphi}}{2}\cos\theta$ in Eq. (7.21) has the same structure as the isospin Zeeman term $-h_z\cos\theta$, such that $\frac{\dot{\varphi}}{2}$ plays the role of an additional, time- and coordinate-dependent, Zeeman field. Secondly, as already noticed in Sec. 8.3.1, the gradient term $\frac{\rho}{2}\sin^2\theta(\nabla\varphi)^2$ has the form of the anisotropy $\frac{u}{2}(n_z^2-1)$, so $-\rho(\nabla\varphi)^2$ can be seen as an additional anisotropy energy. These two observations allow us the rewrite the Lagrangian density (7.20) identically as

$$-L[\theta,\varphi;u,h_z] = E[\theta,\varphi;u,h_z+\frac{\dot{\varphi}}{2}] = dE_x[\theta;u-\rho(\nabla\varphi)^2,h_z+\frac{\dot{\varphi}}{2}]+\frac{\rho}{2}(\nabla_y\theta)^2+\mathcal{E}^\infty(u,h_z), \quad (8.85)$$

using the definitions from Eq. (8.28). We perform the two key approximations (8.73) and (8.76) from above of considering only-y-dependent configurations $\varphi(\boldsymbol{r};t) \to \Phi(y;t)$ and neglecting the gradient terms $\nabla_y\theta$. The Lagrangian density per unit length in the y direction

$$-\int \frac{\mathrm{d}x}{s}\, L[\theta,\varphi;u,h_z] \to dE^{1\mathrm{D}}[\theta;u-\rho(\nabla_y\Phi)^2,h_z+\frac{\dot\Phi}{2}]+E^{1\mathrm{D}\infty}(u,h_z) \tag{8.86}$$

then becomes equivalent to the functional (8.16) for the ground state with modified parameters. This functional (8.86) is minimized with respect to $\theta(\boldsymbol{r};t)$ by the modified ground state configuration $\theta_0(x;u-\rho(\nabla_y\Phi)^2,h_z+\frac{\dot\Phi}{2})$ given in Eq. (8.20). The minimum relative to the total ground state energy

$$-\tilde L^{1\mathrm{D}}[\Phi] = dE_0^{1\mathrm{D}}\left(u-\rho(\nabla_y\Phi)^2,h_z+\frac{\dot\Phi}{2}\right)-dE^{1\mathrm{D}}(u,h_z) \tag{8.87}$$

can be expressed in terms of the domain wall energy (8.31). Its expansion up to quadratic order in the derivatives yields the form

$$\begin{aligned}-\tilde L^{1\mathrm{D}}[\Phi] &\approx -\partial_u dE_0^{1\mathrm{D}}(u,h_z)\rho(\nabla_y\Phi)^2+\partial_{h_z}dE_0^{1\mathrm{D}}(u,h_z)\frac{\dot\Phi}{2}+\frac{1}{2}\partial_{h_z}^2 dE_0^{1\mathrm{D}}(u,h_z)\left(\frac{\dot\Phi}{2}\right)^2\\ &= -L^{1\mathrm{D}}[\Phi]+\partial_{h_z}dE_0^{1\mathrm{D}}(u,h_z)\frac{\dot\Phi}{2}\end{aligned} \tag{8.88}$$

of the Luttinger liquid Lagrangian (8.82). As in Eq. (8.66), the term linear in $\dot\Phi$ is an inconsequential full time derivative and will be dropped. This allows us to express the parameters

$$\frac{1}{8\pi\mathcal{K}v} = -\frac{1}{8}\partial_{h_z}^2 dE_0^{1\mathrm{D}}(u,h_z),\quad \frac{v}{8\pi\mathcal{K}} = -\rho\partial_u dE_0^{1\mathrm{D}}(u,h_z) \tag{8.89}$$

of the Luttinger liquid in terms of the derivatives of the domain-wall energy and thus calculate the coefficients (8.83) and (8.84) explicitly as

$$F_t(\bar h_z) = -\frac{1}{4}\partial_{\bar h_z}^2 F(\bar h_z) = \frac{1}{8}\frac{1}{\bar h_z\sqrt{\bar h_z-1}}, \tag{8.90}$$

$$F_y(\bar h_z) = F_2(\bar h_z) = \bar h_z \arcsin\frac{1}{\sqrt{\bar h_z}} - \sqrt{\bar h_z-1}. \tag{8.91}$$

The functions are plotted in Fig. 8.10. Their asymptotic expressions at the transition point and at large h_z are

$$F_t(\bar h_z) = \frac{1}{8}\begin{cases}\frac{1}{\sqrt{\bar h_z-1}}, & h_z\to u+0^+,\\ \frac{1}{\bar h_z^{\frac{3}{2}}}, & h_z\gg u,\end{cases}\qquad F_y(\bar h_z) = \begin{cases}\frac{\pi}{2}, & h_z\to u+0^+,\\ \frac{2}{3}\frac{1}{\sqrt{\bar h_z}}, & h_z\gg u.\end{cases} \tag{8.92}$$

As a side remark, we find that the function $\delta\theta_0[\Phi](\boldsymbol{r};t)$, Eq. (8.79), can be identified as the linear term of the expansion of $\theta_0(x;u,h_z+\frac{\dot\Phi}{2})$ in $\frac{\dot\Phi}{2}$. We thus have $\partial\theta_0(x) = \partial_{h_z}\theta_0(x;u,h_z)$, and inserting this form into Eq. (8.83) for $F_t(\bar h_z)$, we see that the integrand is a full derivative with respect to h_z. Thus the coefficient

$$F_t(\bar h_z) = \frac{1}{8}\int_{-\infty}^{0}\mathrm{d}\bar x\,\partial_{\bar h_z}\cos\theta_0(\bar x;\bar h_z) = -\frac{1}{4}\partial_{\bar h_z}F_1(\bar h_z), \tag{8.93}$$

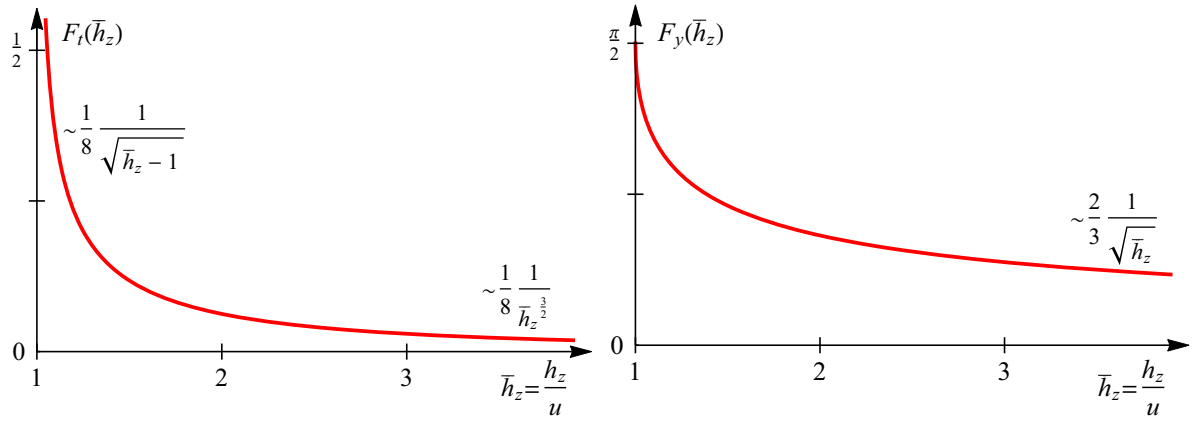


Figure 8.10. The functions $F_t(\bar{h}_z)$, Eqs. (8.83) and (8.90), and $F_y(\bar{h}_z)$, Eqs. (8.84) and (8.91). They determine the coefficients of the time- and coordinate-derivative terms in the Luttinger liquid model (8.82) for the edge excitations of the TnT $\boldsymbol{n}^\infty = \boldsymbol{n}_z$ phase. The asymptotic functions (8.92) at the transition point $\bar{h}_z = 1$ and for large $\bar{h}_z \gg 1$ are indicated. Reprinted figures with permission from Ref. [KJT16]. Copyright (2016) by the APS.

can alternatively be expressed in terms of $F_1(\bar{h}_z)$ (8.34).

This more elegant method of derivation of the Luttinger liquid model (8.82) as an expansion of the modified domain-wall energy also allows us to determine the restrictions on the allowed magnitude of fluctuations. The domain-wall energy $dE_0^{1D}(u, h_z)$ (8.31) contains a nonanalytic square-root dependence $\sqrt{\bar{h}_z - 1}$. As a result, the power expansion (8.88) is valid as long as the fluctuation energies

$$\dot{\Phi},\ \rho(\nabla_y\Phi)^2 \ll h_z - u \tag{8.94}$$

are much smaller than the deviation $h_z - u$ from the transition point. This deviation $h_z - u$ is exactly the gap of the neutral bulk excitations, given by isospin waves. In this regard, we caution about using the unexpanded functional (8.87) at fluctuation energies $\dot{\Phi}$ and $\rho(\nabla_y\Phi)^2$ comparable to $h_z - u$. The approximations made in the derivation of this expression amount to neglecting other isospin configurations, such as bulk excitations, which become relevant at energies $\sim h_z - u$.

In the next Sec. 8.4.2, we analyze the main properties of the obtained Luttinger liquid model.

8.4.2. Analysis

The Luttinger liquid model (8.82) describes the edge excitations of the $\boldsymbol{n}^\infty = \boldsymbol{n}_z$ phase, realized for $u < h_z$. Its collective variable is the angle $\Phi(y;t)$ varying with time t and coordinate y along the edge. In leading order, the isospin texture associated with $\Phi(y;t)$ can be approximated by the deformed ground state configuration $\boldsymbol{n}_0(x|\Phi(y;t))$ as defined in Eqs. (8.15) and (8.20). The field $\Phi(y;t)$ thus corresponds to the polarization in the xy isospin plane. Simultaneously, this edge isospin texture carries electric charge. According to Eq. (8.35), the charge density per unit length in the y direction reads

$$\begin{aligned}\kappa^{1D}[\Phi](y;t) &\equiv \int_{-\infty}^{0} \mathrm{d}x\, \kappa[\boldsymbol{n}_0(x|\Phi(y;t))] = \frac{1}{4\pi}\int_{-\infty}^{0} \mathrm{d}x\, \sin\theta_0(x)\nabla_x\theta_0(x)\nabla_y\Phi(y;t)\\ &= \frac{1}{2\pi}\nabla_y\Phi(y;t).\end{aligned} \tag{8.95}$$

The associated electric current in the y direction equals

$$j^{1\mathrm{D}}[\Phi](y;t) = -\frac{1}{2\pi}\dot{\Phi}(y;t), \tag{8.96}$$

as follows from the continuity equation

$$\dot{\kappa}^{1\mathrm{D}} + \nabla_y j^{1\mathrm{D}} = 0. \tag{8.97}$$

Therefore, the single field $\Phi(y;t)$ carries both isospin and charge degrees of freedom, locked to each other, and the Luttinger liquid (8.82) represents an isospin helical liquid. For further reading on interacting 1D edge channels, see Refs. [XM06, WBZ06, SJJ10, SRvOG12, BDRT12, CBD+12, LOB12, GCT14, KGCM14].

The low-energy edge theory is fully characterized by two parameters: the velocity v of the linear gapless excitation spectrum $\omega = vk$, and the dimensionless parameter $\mathcal{K}$ describing the effective strength of interactions, see Sec. 3.2.4 for details and references. In a generic Luttinger liquid, $\mathcal{K} = 1$ corresponds to a non-interacting system. $0 < \mathcal{K} < 1$ is the range of repulsive interactions, the stronger the interactions, the smaller $\mathcal{K}$, and $1 < \mathcal{K}$ is the range of attractive interactions (irrelevant for the considered system), the stronger the interactions, the larger $\mathcal{K}$. From Eqs. (8.83) and (8.84), these two parameters are expressed in terms of the functions $F_{t,y}(\bar{h}_z)$ (8.90) and (8.91) as

$$v(\bar{h}_z) = \sqrt{\rho u}\sqrt{\frac{F_y(\bar{h}_z)}{F_t(\bar{h}_z)}}, \quad \mathcal{K}(\bar{h}_z) = \frac{u}{8\pi\varepsilon_\odot}\frac{1}{\sqrt{F_t(\bar{h}_z)F_y(\bar{h}_z)}}. \tag{8.98}$$

Their asymptotic expressions at the phase transition point $h_z = u$ and at large h_z follow from Eq. (8.92) as

$$v(\bar{h}_z) = \sqrt{\rho u}\begin{cases} \sqrt{4\pi}(\bar{h}_z - 1)^{\frac{1}{4}}, & h_z \to u + 0^+, \\ \sqrt{\frac{16}{3}}\sqrt{\bar{h}_z}, & h_z \gg u, \end{cases} \tag{8.99}$$

$$\mathcal{K}(\bar{h}_z) = \frac{u}{8\pi\varepsilon_\odot}\begin{cases} \sqrt{\frac{16}{\pi}}(\bar{h}_z - 1)^{\frac{1}{4}}, & h_z \to u + 0^+, \\ \sqrt{12}\,\bar{h}_z, & h_z \gg u. \end{cases} \tag{8.100}$$

The dependence of the parameters $v(\bar{h}_z)$ and $\mathcal{K}(\bar{h}_z)$ on the normalized field $\bar{h}_z$ is plotted in Fig. 8.11 and results from the behavior of the functions $F_y(\bar{h}_z)$ and $F_t(\bar{h}_z)$ as shown in Fig. 8.10. Both the velocity $v(\bar{h}_z)$ and interaction parameter $\mathcal{K}(\bar{h}_z)$ are growing functions of the field $\bar{h}_z$, i.e., they increase as the magnetic field B decreases. Sufficiently far from the phase transition $h_z = u$, their scaling given in Eqs. (8.99) and (8.100) is simply

$$v \sim \sqrt{\rho h_z} \sim e_*^2\sqrt{\frac{h_z}{\varepsilon_\odot}} \ll e_*^2, \quad \mathcal{K} \sim \frac{h_z}{\varepsilon_\odot} \ll 1. \tag{8.101}$$

These estimates are correct in the whole range where the QHFM theory is valid. There, the energy scales $u, h_z \ll \varepsilon_\odot$ of the SU(2)-asymmetric terms are much smaller than $\varepsilon_\odot$ (8.50) of the SU(2)-symmetric interactions. Thus, the interaction parameter $\mathcal{K} \ll 1$ is small and the Luttinger liquid is *strongly interacting*. Moreover, both the velocity $v \sim \sqrt{\rho u}(\bar{h}_z - 1)^{\frac{1}{4}}$ and the interaction parameter $\mathcal{K} \sim \frac{u}{\varepsilon_\odot}(\bar{h}_z - 1)^{\frac{1}{4}}$ approach zero at the transition point, $h_z \to u + 0^+$.

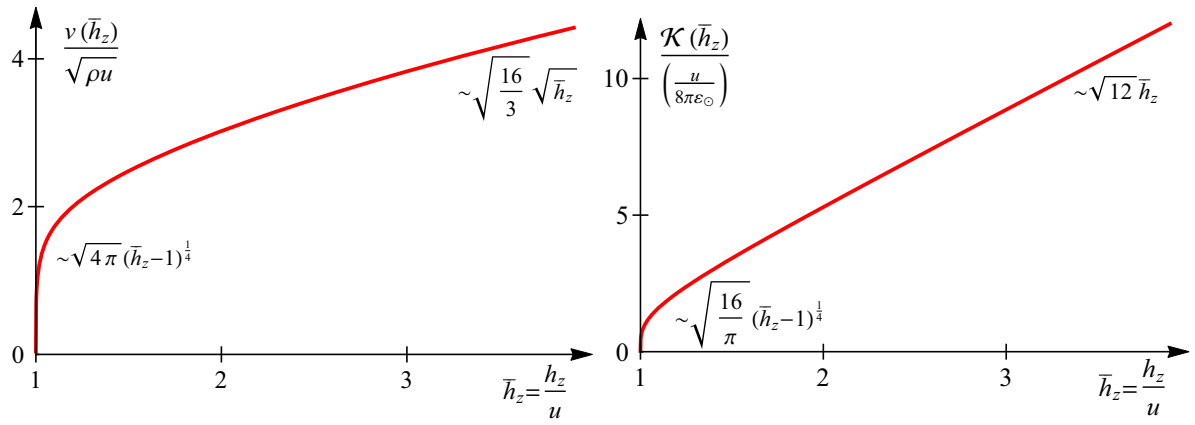


Figure 8.11. The velocity $v(\bar{h}_z)$ and interaction parameter $\mathcal{K}(\bar{h}_z)$, Eq. (8.98), of the Luttinger liquid model (8.82) for the edge excitations of the TnT $\boldsymbol{n}^\infty = \boldsymbol{n}_z$ phase. The asymptotic functions (8.99) and (8.100) at the transition point $\bar{h}_z = 1$ and at large $\bar{h}_z \gg 1$ are indicated. Reprinted figures with permission from Ref. [KJT16]. Copyright (2016) by the APS.

The Luttinger liquid therefore becomes *infinitely* strongly interacting at the transition point. We note though that the energy window of applicability of this low-energy theory narrows accordingly, see Eq. (8.94). For larger fluctuations, the Luttinger liquid model becomes invalid.

The Luttinger liquid models for the edge excitations in the form of Eq. (8.82) were obtained in Ref. [PSH16] for the double-layer system with inverted band structure and in Ref. [TSFM16] for the F phase of the $\nu = 0$ state in graphene. However, in Ref. [PSH16] the expression for the coefficient $F_t(\bar{h}_z)$ of the time-derivative term does not diverge at the phase transition $h_z \to u + 0^+$. This divergence is physical and to be expected, since the 1D model should fail at the phase transition, where the bulk modes become gapless. In Ref. [TSFM16], only the scaling of the parameters $F_{t,y}(\bar{h}_z)$, $v(\bar{h}_z)$, $\mathcal{K}(\bar{h}_z)$ at the phase transition $h_z \to u + 0^+$ was determined, which does agree with our asymptotic results. We additionally obtain explicit analytical expressions (8.83), (8.84), and (8.98) for them at all $u < h_z$.

The helical Luttinger liquid model for the edge excitations above was derived in a controlled way in the QHFM regime, realized in the vicinity of the topological phase transition at $B = B^*$. However, the single-particle TnT phase is present in the whole range $0 \le B < B^*$ of magnetic fields. It is safe to argue that if the Luttinger liquid persists in the region of strong effective interactions near the transition point, it also exists in the whole range $0 \le B < B^* - \delta B_u$ of magnetic fields. Let us consider the bare velocity $v_0(B)$ of the counter-propagating single-particle edge states at the crossing point of their energy curves, $\epsilon(p) = h_z$, in Fig. 7.2. It is largest at zero field $B = 0$ and monotonically decreases down to zero $v_0(B \to B^* - 0^+) = 0$ at the single-particle phase transition point, as shown schematically in Fig. 8.12. The effective strength of the interactions *at the edge* can be characterized by the dimensionless parameter $e_*^2/v_0(B)$, which is roughly equivalent to the ratio $\varepsilon_\odot/h_z$ characterizing the strength of interactions *in the bulk*. The effective strengths of interactions in the bulk and at the edge are thus in accord with each other.

Our basic assumption of weak bare Coulomb interactions, see Sec. 7.2.1, can be formulated as $v_0(B = 0) \gg e_*^2$. Thus, at zero or small fields $B \ll B^*$, the edge states are weakly interacting and the corresponding low-energy theory for collective excitations can be derived using the standard bosonization procedure, see Ref. [Gia03], and will have the form of a Luttinger liquid.

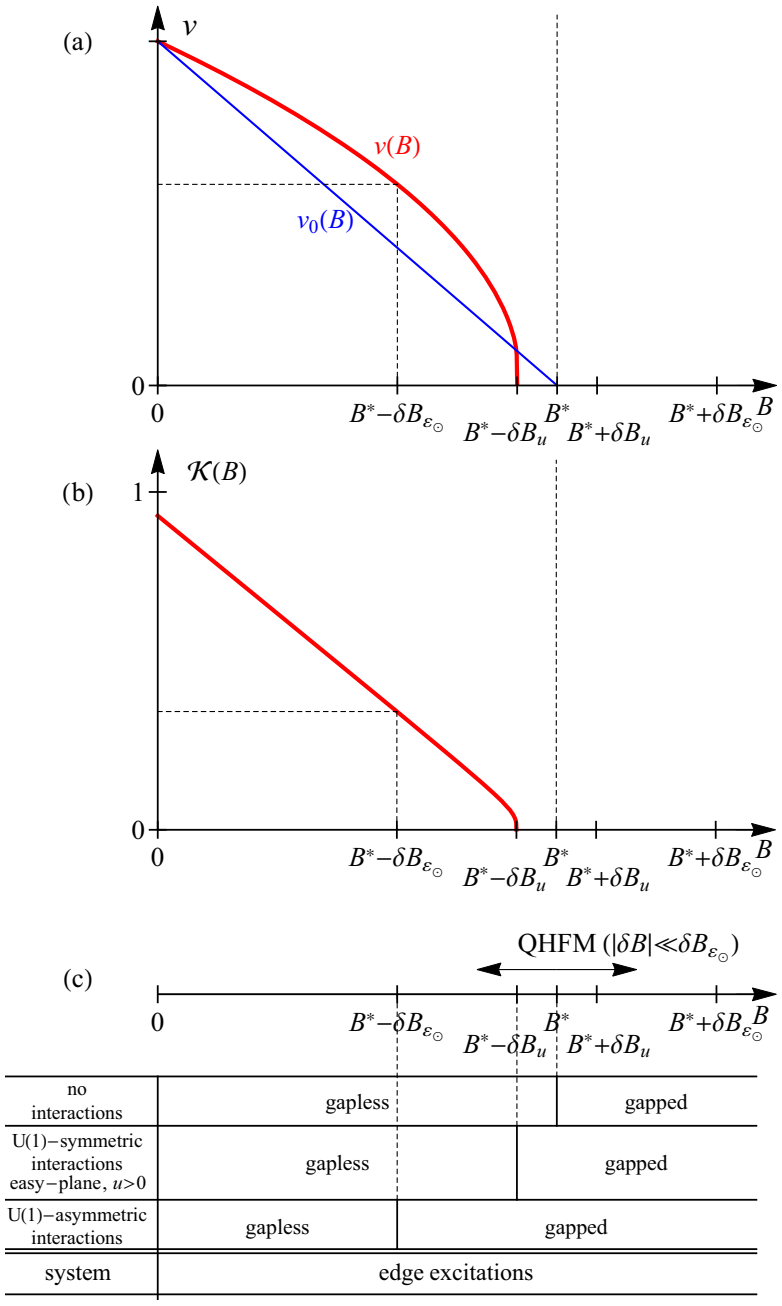


Figure 8.12. Schematic plots of the magnetic-field dependencies of (a) the velocity $v(B)$ and (b) the interaction parameter $\mathcal{K}(B)$ of the helical Luttinger liquid describing the edge excitations of the TnT phase of a QHTI for U(1)-symmetric interactions with easy-plane anisotropy $u > 0$. At small fields $B \ll B^*$, the effective interactions are weak, $\mathcal{K}(B) \approx 1$, and the velocity $v(B) \approx v_0(B)$ is close to its bare value. In the QHFM regime close to the single-particle topological phase transition, the effective interactions are strong, $\mathcal{K}(B) \ll 1$. The effective interactions are thus highly tunable. (c) Summarized properties of the edge excitations. The TT phase with gapped edge excitations ensues at lower magnetic fields upon switching on the interactions and lowering their symmetry. Reprinted figure with permission from Ref. [KJT16]. Copyright (2016) by the APS.

The velocity $v(B=0) \approx v_0(B=0)$ of the collective excitations will be close to the bare velocity and the interaction parameter $\mathcal{K} \approx 1$ will be close to unity, corresponding to weak effective interactions. In the QHFM regime on the other hand, the effective interactions in the Luttinger liquid are strong, with $v(B) \ll e_*^2$ and $\mathcal{K}(B) \ll 1$ from Eq. (8.101). The regime of intermediate effective interaction strength with $v(B) \sim e_*^2$ and $\mathcal{K}(B)$ is close to neither 0 nor 1, occurs at the verge of the QHFM regime, at fields $B = B^* - \delta B_{\varepsilon_\odot}$ with

$$\delta B_{\varepsilon_\odot} \equiv \frac{\varepsilon_\odot}{|\partial_B h_z(B^*)|}, \tag{8.102}$$

where $h_z(B) \sim \varepsilon_\odot(B)$. In this intermediate-strength regime, the weakly and strongly interacting Luttinger liquids must continuously connect.

We thus conclude that the helical Luttinger liquid describing the edge excitations of a QHTI persists in almost the whole range of the non-interacting TnT phase, $0 \leq B < B^* - \delta B_u$, and is highly tunable. The magnetic field allows one to adjust the effective interactions between weak at lower B and infinitely strong at higher B, close to the topological transition at $B = B^* - \delta B_u$. The corresponding behavior of the velocity $v(B)$ and interaction parameter $\mathcal{K}(B)$ in the whole range is plotted schematically in Fig. 8.12 (a) and (b). However, we remind that the derived Luttinger liquid theory for the edge was obtained under the specific assumption of U(1) isospin symmetry, defined in Eqs. (7.14) and (7.15), of the many-body Hamiltonian $\hat{H}$ (7.8). This U(1) symmetry is inherited by the Lagrangian (8.82) of the Luttinger liquid, which is invariant under rotations of the angle field

$$\Phi(y;t) \to \Phi(y;t) + \phi. \tag{8.103}$$

We now turn to the analysis of the effects that break U(1) symmetry, introduced in Secs. 7.2.3 and 7.3.3.

8.4.3. Broken U(1) Symmetry

As we discussed in Sec. 7.2.3, there are two categories of U(1)-asymmetric terms. One directly breaks the physical symmetry and exists already at the single-particle level, given by the term $\mathcal{E}_{1\varnothing}(\boldsymbol{n})$ in Eq. (7.31). The other one preserves the physical symmetry and can arise only from interactions, described by $\mathcal{E}_{2\varnothing}(\boldsymbol{n})$ in Eq. (7.34). If small, both can be incorporated into the Luttinger liquid model (8.82).

Broken physical Symmetry: Single-particle Effect

Proceeding along the same line as in Sec. 8.4.1, we expand the term

$$\begin{aligned} \mathcal{E}_{1\varnothing}(\boldsymbol{n}) &= -\,h_\perp[\sin\theta_0(x) + \mathcal{O}(\delta\theta)]\cos(\varphi - \varphi_{1\varnothing}) \\ &\to -\,h_\perp \sin\theta_0(x)\cos(\varphi - \varphi_{1\varnothing}) \end{aligned} \tag{8.104}$$

of the first category about the ground state, see Eq. (8.65), and keep only the zero-order term in $\delta\theta$. Due to the constraining function $\sin\theta_0(x)$, the approximation $\varphi(x,y;t) \to \Phi(y;t)$, Eq. (8.73), may then be used. This yields the contribution

$$-\,\mathbb{L}_{1\varnothing}^{\mathrm{1D}}[\Phi] = -h_\perp^{\mathrm{1D}} \int \mathrm{d}y\, \cos(\Phi(y;t) - \varphi_{1\varnothing}), \tag{8.105}$$

to the 1D Lagrangian of the edge excitations, where

$$h_\perp^{\mathrm{1D}} = \frac{h_\perp l_u}{s} F_\perp(\bar{h}_z), \quad F_\perp(\bar{h}_z) = \int_{-\infty}^{0} \mathrm{d}\bar{x}\, \sin\theta_0(\bar{x};\bar{h}_z) = \ln\frac{\sqrt{\bar{h}_z}+1}{\sqrt{\bar{h}_z}-1}. \tag{8.106}$$

The Lagrangian $\mathbb{L}^{\mathrm{1D}}[\Phi]+\mathbb{L}^{\mathrm{1D}}_{1\varnothing}[\Phi]$ describes the sine-Gordon model, the properties of which are well-studied, see Ref. [Gia03] for details. Its excitation spectrum is gapped at

$$\mathcal{K} < 2, \tag{8.107}$$

i.e., including all repulsive interactions $0 < \mathcal{K} < 1$, the non-interacting case $\mathcal{K} = 1$, and the range $1 < \mathcal{K} < 2$ of attractive interactions. Thus, if the physical symmetry is broken, there is no topological protection already at the single-particle level $\mathcal{K} = 1$. As a result, in the presence of repulsive interactions $\mathcal{K} < 1$, the system is in a TT phase with broken U(1) symmetry and gapped edge excitations for all magnetic fields $B \geq 0$.

Preserved physical, but broken U(1) Symmetry: Interaction Effect

Similarly, we expand the term of the second category about the ground state and keep only the zeroth-order term in $\delta\theta$ like

$$\begin{aligned}\mathcal{E}_{2\varnothing}(\boldsymbol{n}) =& \frac{1}{2} u_{2\varnothing}[\sin^2\theta_0(x) + \mathcal{O}(\delta\theta)]\cos 2(\varphi-\varphi_{2\varnothing}) \\ \to& \frac{1}{2} u_{2\varnothing}\sin^2\theta_0(x)\cos 2(\varphi-\varphi_{2\varnothing}).\end{aligned} \tag{8.108}$$

After the substitution $\varphi(x,y;t) \to \Phi(y;t)$, we obtain the contribution

$$-\mathbb{L}^{\mathrm{1D}}_{2\varnothing}[\Phi] = u^{\mathrm{1D}}_{2\varnothing}\int \mathrm{d}y\, \cos 2(\Phi(y;t)-\varphi_{2\varnothing}), \quad u^{\mathrm{1D}}_{2\varnothing} = \frac{u_{2\varnothing} l_u}{s} F_y(\bar{h}_z), \tag{8.109}$$

to the 1D Lagrangian of the edge excitations. Therefore, for preserved physical symmetry (considering inversion as an example), the edge is described by the Lagrangian $\mathbb{L}^{\mathrm{1D}}[\Phi]+\mathbb{L}^{1D}_{2\varnothing}[\Phi]$. This is also a sine-Gordon model, but Eq. (8.109) differs from Eq. (8.105) by the numerical factor 2 instead of 1 in the cosine argument. As a result, the edge ground state breaks U(1) symmetry (the field acquires a finite expectation value $\langle\Phi\rangle \neq 0$) and the edge excitations become gapped in a different range of interaction strength, namely already for

$$\mathcal{K} < \frac{1}{2}. \tag{8.110}$$

According to Sec. 8.4.2, the quantum phase transition at the field $B_{2\varnothing} \sim B^* - \delta B_{\varepsilon_\odot}$ such that $\mathcal{K}(B_{2\varnothing}) = \frac{1}{2}$ occurs at the verge of the QHFM regime, where $h_z(B_{2\varnothing}) \sim \varepsilon_\odot(B_{2\varnothing})$. The system has gapless edge excitations for lower fields $B < B_{2\varnothing}$ and gapped excitations for all higher fields $B > B_{2\varnothing}$.

8.5. Role of U(1) Symmetry for Topological Protection

One of the key questions raised in the studies of topological systems is how interactions affect the topological properties. A common understanding is that a non-interacting TI is generally

not guaranteed to remain such in the presence of interactions. Our presented results allow us to specify the symmetry requirements for the topological protection in the presence of interactions. The QHTIs we consider are protected by some physical symmetry, see Sec. 7.1. Due to this symmetry, two relevant LLs, plotted in Fig. 7.1, have different transformation properties and are thus not coupled at the single-particle level. As a result, the projected *single-particle* Hamiltonian within these two LLs possesses U(1) symmetry with respect to uniaxial isospin rotations, see Eqs. (7.2) and (7.15).

Summarizing the results of Secs. 8.3 and 8.4, we conclude that this *effective continuous* U(1) symmetry is responsible for the topological protection *in the presence of interactions* in this QHTI class of systems. As long as the U(1) symmetry is preserved, the TnT phase with gapless edge charge excitations persists for any effective strength of interactions, even in the strongly interacting QHFM regime, realized in the vicinity of the single-particle topological phase transition.

However, we find that the physical symmetry alone is generally not sufficient to protect the TnT phase in the presence of interactions. The interactions preserving the physical symmetry can still break U(1) symmetry and thereby destroy the TnT phase for strong effective interactions. "Preserved U(1) symmetry" means here that *both* (i) the bulk ground state and (ii) the interacting projected Hamiltonian are U(1)-symmetric. Accordingly, the two mechanisms by which U(1) symmetry can be broken correspond to violation of one of these conditions:

1) First, as demonstrated in Sec. 8.1, even the U(1)-symmetric interactions (condition (ii) satisfied) with the "right" properties, like easy-plane anisotropy, can result in a bulk ground state with spontaneously broken U(1) symmetry (condition (i) violated): the $\boldsymbol{n}^\infty = \boldsymbol{n}_0^*(\varphi_0)$ phase. As shown in Sec. 8.3, this phase has gapped edge charge excitations and is thus TT.

2) Secondly, as analyzed in Secs. 7.2.3 and 7.3.3, interactions can (depending on the physical symmetry) contain terms that preserve the physical symmetry, but break the effective U(1) symmetry (condition (ii) violated). These terms result in a phase transition to the TT phase with broken U(1) symmetry at the edge and gapped excitation spectrum, see Sec. 8.4.3 for details. At the same time, the U(1)-symmetric bulk ground state $\boldsymbol{n}^\infty = \boldsymbol{n}_z$ may still persist beyond this transition (condition (i) satisfied).

We mention that similar types of interactions have earlier been considered in the studies of helical Luttinger liquids for the edge of 2D TIs protected by $\boldsymbol{\mathcal{T}}$ symmetry, see Refs. [XM06, WBZ06, SJJ10, SRvOG12, BDRT12, CBD+12, LOB12, GCT14, KGCM14]. The transition from the TnT phase to either of these TT phases occurs upon increasing the magnetic field B, as the single-particle phase transition point B^* is approached and the effective interactions get stronger. Yet the transition point is at fields *lower* than B^*, namely at $B^* - \delta B_u$ [Eq. (8.9)] and $B_{2\varnothing} \sim B^* - \delta B_{\varepsilon_\odot}$ [Eq. (8.102)], respectively, as summarized in Fig. 8.12(c). Thus, these are *interaction-induced topological quantum phase transitions*, enabled by the tunability of the effective interactions by the magnetic field as presented in Sec. 8.4.2.

Conclusion

In this Chap. III, we studied the effect of electron interactions on the topological properties of QHTIs. Due to the crossing of LLs at the single-particle topological phase transition, its vicinity is automatically the regime of strong effective interactions. As an appealing theoretical aspect, such a system can be studied in a controlled way within the framework of QHFMism.

We found that the properties of the charge excitations of the phases $\boldsymbol{n}^\infty = \pm\boldsymbol{n}_z$ with

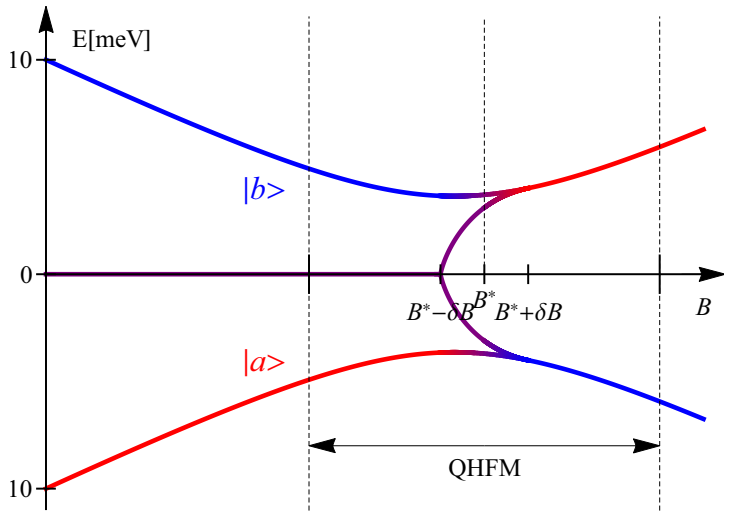


Figure 8.13. Sketch of the LL structure of a QHTI, based on Fig. 7.1 for HgTe QW parameters, modified by the U(1)-symmetric Coulomb interactions. The bulk charge excitation gap between the states $|a\rangle$ and $|b\rangle$ never closes. The edge states, depicted by the purple line, are gapless up to magnetic fields of $B^* - \delta B$, where the interaction driven topological phase transition occurs and the edge states gap out. For even larger magnetic fields, the bulk and edge charge excitation gaps coincide. Beware: the gap for neutral bulk excitations, see Eq. (8.94), is different from the one sketched here.

preserved U(1) symmetry remain qualitatively the same in the presence of strong interactions. The TnT phase $\boldsymbol{n}^\infty = \boldsymbol{n}_z$ has gapless edge excitations, while the TT phase $\boldsymbol{n}^\infty = -\boldsymbol{n}_z$ is gapped at the edge. However, collective charge excitations of the interacting system are microscopically quite different from the single-electron excitations. Spontaneous breaking of the U(1) symmetry leads in the intermediate phase $\boldsymbol{n}^\infty = \boldsymbol{n}^*(\varphi_0)$ to a controlled opening of the gap for the edge charge excitations. While the properties of the edge excitations of the three phases are different, their bulk charge excitations are qualitatively the same. The bulk charge gap is finite in all three phases and never closes during the transformation from the TnT $\boldsymbol{n}^\infty = \boldsymbol{n}_z$ to the TT $\boldsymbol{n}^\infty = -\boldsymbol{n}_z$ phase with increasing B_z, see Fig. 8.13.

Particular attention was paid to establishing the requirements for the topological protection in this interacting system. We find that this question is ultimately related to the effective symmetry of the system: the continuous U(1) symmetry is a necessary condition for the TnT phase to persist in the regime of strong effective interactions. If U(1) symmetry is preserved, the edge of the TnT phase is described by the helical Luttinger liquid. The effective interactions of this Luttinger liquid are highly tunable by the magnetic field B. They are weak (for weak bare Coulomb interactions) at small B and grow as B is increased, becoming strong in the QHFM regime in the vicinity of the single-particle topological phase transition.

The U(1) symmetry may be broken, however, either spontaneously or by the interactions that are explicitly U(1)-asymmetric. In either scenario, this eventually results in a phase transition to a TT phase with gapped edge excitations, which can be achieved by tuning the interaction strength via the magnetic field.

The tunability of interactions, the accessibility of the regime of strong effective interactions even in a system with weak bare interactions, and the possibility to realize interaction-induced topological phase transitions are among the properties that make QHTIs an attractive class of systems for investigating the interplay of interactions and topology, both theoretically and experimentally.

Part IV.

Hybrid Systems of Weyl Semimetals and 3D Topological Insulators

The study of topological properties in a semiconductor environment has become a strong and flourishing field in condensed matter physics. Topological insulators (TIs) are the standard materials in this context, well studied both theoretically and experimentally by now, see Sec. 2.4 for an introduction. Our own research in Chaps. II and III focused on weakly and strongly interacting two-dimensional (2D) TIs. Now, we widen our scope and also consider their semi-metallic counter-part, the Weyl semimetal (WSM), as introduced in Sec. 2.5.

WSMs were proposed to exist in condensed matter systems decades ago [NN81, NN83, Mur07]. However, only very recently with the prediction of concrete material realizations [WTVS11, WFF+15, HXB+15] the field has seen an enormous growth. The experimental proof of the existence of Weyl points and their corresponding surface states, called Fermi arcs, followed soon afterwards [XAB+15, LWF+15b, LXW+15, YLS+15, XDW+15, LWY+15]. Yet both for fundamental research and application purposes these "early" WSMs, such as the TaAs family of non-centrosymmetric monopnictides, are too complicated with many Weyl points (24 for TaAs) in the Brillouin zone (BZ). Simpler materials with eight [CXS+16, RJY+16, RAHM15, RJZ+16, SWY15, HMO+16, DWD+16, TWC+16, JJL+16, XAC+16] and four [KKE+16, BYS+16] Weyl points have been predicted and observed, where the latter is the minimal number of Weyl points for a system with time-reversal ($\boldsymbol{\mathcal{T}}$) symmetry. WSM materials with broken $\boldsymbol{\mathcal{T}}$ symmetry are more scarce, an example is $YbMnBi_2$ analyzed in Ref. [BEG+15]. They could realize the absolute minimum of two Weyl points, but here only theoretical proposals [BHB11, Cho11, XWW+11, BLQ14, WACB16] exist so far. Most of them rely on magnetically doped TIs or TI heterostructures.

TI and WSM are both topological phases that can be directly connected to each other through quantum phase transitions, as explained in Sec. 2.5 and the references therein. Here, we want to go a step further and study the question whether a system can be both in the TI and WSM phase at the same time, or at least support both corresponding surface states, 2D Dirac surface states and Fermi arcs, on the same surface.

Such a combined phase might exist in HgTe with applied compressive strain. The strain pushes the Γ_8 bands into one another, see Fig. 2.15, creating Dirac points. These are then split by breaking of inversion ($\boldsymbol{\mathcal{I}}$) symmetry through bulk inversion asymmetry (BIA) terms [RJY+16]. At the same time, the topological band inversion between the Γ_8 and the Γ_6 bands remains, leading to the conjecture that this system could have topological Dirac states and Fermi arcs on its surface. A different way to create such a hybrid surface state is placing a TI and WSM spatially adjacent to each other, possibly separated by a small, topologically trivial (TT) buffer layer. The separate surface states of TI and WSM will interact, e.g. by Coulomb interaction or tunneling due to a small overlap of wave functions, forming the hybrid surface dispersion relation. Previous related research on adjacent TI and WSM phases, presented in Ref. [GVB15], suggest that at such a shared surface both Dirac states and Fermi arcs can exist. However, they were found in different areas of k-space, mutually excluding one another such that they do not hybridize at all. Our approach differs from the one chosen in Ref. [GVB15] by considering only a small, perturbative coupling between the two phases. This ensures that both TI and WSM surface states survive and can interact with each other.

We focus in this chapter on an analytical study of the combined surface states generated from the hybridized TI and WSM. A simplified ansatz offers the possibility to calculate the surface Hamiltonian analytically, allowing for a detailed analysis of the surface physics. Depending on the symmetry of the assumed couplings, the surface dispersion relation shows quite different behavior. In the case of spin symmetry, two shifted Dirac nodes may emerge out of the combination of a single Dirac node and a Fermi arc. For spin-asymmetric coupling, the Fermi arc gaps out and spin-polarizes the former Dirac node.

9. Exotic Surface States

The research presented in this Sec. 9 and the corresponding Apps. B and C is published in Ref. [JT17]. We begin with a recap of the effective models for the separate phases of TIs and WSMs, see Secs. 2.4.2 and 2.5, and analyze their symmetry properties and surface states in Sec. 9.1. The coupling of the two Hamiltonians and the analytic form of the combined surface state is discussed in Sec. 9.2. Sec. 9.3 focuses on the different ways to influence and tune the hybrid surface dispersion relation. We end with the proposal of possible experimental realizations in Sec. 9.4.

9.1. Separate Models

The Hamiltonian (2.46) of the three-dimensional (3D) TI phase we will use was originally derived for the Bi_2Se_3 family of materials in Ref. [ZLQ+09, LQZ+10]. It contains four bands and serves as a minimal, but general, TI model. The Weyl Hamiltonian (2.57) considered in the following originates from Refs. [YLR11, MKT17]. It contains two bands and models an $\boldsymbol{\mathcal{I}}$ symmetric type I or II WSM with two Weyl points. We simplify the models as far as possible without losing too much versatility. It is important to retain terms quadratic in momentum for the introduction of the surface in the z direction. This is done via hardwall boundary conditions on a half space $z \leq 0$ or $z \geq 0$.

9.1.1. Topological Insulator

The effective Hamiltonian for a 3D TI is given by the 4x4 matrix

$$H_{TI} = \begin{pmatrix} \mathcal{M}(k)\tau_3 + \mathcal{B}k_z\tau_2 + \mathcal{C}\tau_0 & i\mathcal{A}k_-\tau_1 \\ -i\mathcal{A}^* k_+\tau_1 & \mathcal{M}(k)\tau_3 + \mathcal{B}k_z\tau_2 + \mathcal{C}\tau_0 \end{pmatrix} \tag{9.1}$$

with $\mathcal{M}(k) = M_0 + M_1(k_\parallel^2 + k_z^2)$, $k_\parallel^2 = k_x^2 + k_y^2$ and $k_\pm = k_x \pm ik_y = k_\parallel e^{\pm i\phi_k}$. In the original derivation for Bi_2Se_3, the Pauli matrices $\boldsymbol{\tau}$ describe an orbital degree of freedom. H_{TI} is written in a spin-up/down basis, represented by the Pauli matrices $\boldsymbol{\sigma}$ in the following. The coupling $\mathcal{A} = |\mathcal{A}|\, e^{i\phi_\mathcal{A}}$ can in principle be complex, with the angle $\phi_\mathcal{A}$. For fulfilled inversion condition $M_0 M_1 < 0$, the model is in the strong TI phase.

We define the $\boldsymbol{\mathcal{I}}$ operator $P_{TI} = \sigma_0 \otimes \tau_3$ and $\boldsymbol{\mathcal{T}}$ operator $T_{TI} = i\sigma_2 \otimes \tau_0 K$ with K the complex conjugation operator. H_{TI} is symmetric under both operations, fulfilling

$$P_{TI}^\dagger H_{TI}(-k)\, P_{TI} = H_{TI}(k)\,, \qquad T_{TI}^\dagger H_{TI}(-k)\, T_{TI} = H_{TI}(k)\,. \tag{9.2}$$

The bulk dispersion relation is double degenerate and given by

$$E_{TI} = \mathcal{C} \pm \sqrt{|\mathcal{A}|^2 k_\parallel^2 + \mathcal{B}^2 k_z^2 + \mathcal{M}(k)^2}. \tag{9.3}$$

Based on the method described in App. B, the surface states can be calculated analytically. We assume opposite surfaces to be well separated, which offers the possibility to treat them individually. Thus in the calculation we only consider one of them via hardwall boundary conditions at $z = 0$. The surface wave function is then given by

$$\Psi(z) = \frac{1}{\sqrt{2}}\left(e^{ik_{z,1}z} - e^{ik_{z,2}z}\right)\begin{pmatrix} \pm\frac{i\eta\mathcal{A}k_-}{|\mathcal{A}|k_\parallel}\psi_\eta \\ \psi_\eta \end{pmatrix} \tag{9.4}$$

with the vector $\psi_\eta = \frac{1}{\sqrt{2}}(1,\ \eta)^{\mathrm{T}}$ and the inverse localization length of $ik_{z,\frac{1}{2}} = \frac{1}{2M_1}\Big[-\eta\mathcal{B} \pm \sqrt{4M_1(M_0 + M_1k_\parallel^2) + \mathcal{B}^2}\Big]$. The sign $\eta = \pm$ depends on the surface, $\eta = -\mathrm{sgn}(\mathcal{B}/M_1)$ (upper surface) or $\eta = \mathrm{sgn}(\mathcal{B}/M_1)$ (lower surface). The existence condition for the surface state, see App. B, is

$$M_1(M_0 + M_1k_\parallel^2) < 0 \tag{9.5}$$

stressing the importance of being in the inverted regime.

The surface Hamiltonian (dispersion relation) is obtained from H_{TI} by projecting out the orbital (orbital & spin) degrees of freedom with the help of ψ_η ($\Psi(z)$). We find the usual Dirac form

$$H_{TI}^{sur} = \begin{pmatrix} \mathcal{C} & i\eta\mathcal{A}k_- \\ -i\eta\mathcal{A}^*k_+ & \mathcal{C} \end{pmatrix}, \quad E_{TI}^{sur} = \mathcal{C} \pm |\mathcal{A}|\, k_\parallel, \tag{9.6}$$

experiencing spin-momentum locking, with the angle $\phi_\mathcal{A} + \frac{\eta\pi}{2}$ between the spin projection and momentum vector in the x-y plane. The combined dispersion relations of the bulk and surface of the TI are shown in Fig. 9.1 (left).

9.1.2. Inversion symmetric Weyl Semimetal

A WSM exists in different flavors. On the one hand, one distinguishes type I and type II depending on preserved or broken Lorentz invariance at the Weyl points [XZZ15, SGW+15, SGT16]. Secondly, either $\boldsymbol{\mathcal{T}}$ or $\boldsymbol{\mathcal{I}}$ symmetry has to be broken to get from a Dirac to a Weyl semimetal. For all these phases minimal models have been proposed in the literature, see Sec. 2.5.

For simplicity, we focus on the model with broken $\boldsymbol{\mathcal{T}}$ and preserved $\boldsymbol{\mathcal{I}}$ symmetry, as it has the minimal number of one pair of Weyl points. The Hamiltonian is

$$H_W = t(k)\tau_3 + v_zk_z\tau_2 + v_yk_y\tau_1 + \gamma_W t\,(k_x^2 - k_W^2)\tau_0 \tag{9.7}$$

with $t(k) = t\,(k_\parallel^2 + k_z^2 - k_W^2)$ and the parameters t, $v_{y,z}$, γ_W and k_W. The degree of freedom described by the Pauli matrices $\boldsymbol{\tau}$ can be orbital, spin or a combination of the two, depending on the specific material realization. For the concrete form of the symmetry operations considered in the following we assume a spinless system, as it is done in Ref. [MKT17]. The two Weyl points are specified by $k_x = \pm k_W$, and the parameter γ_W leads to a tilting of the dispersion relation at the Weyl points. For $|\gamma_W| < 1$ one has a type I, otherwise a type II WSM. Expanding H_W around $k_x = \pm k_W$ yields a Hamiltonian with linearized Weyl form

$$H_W^{lin} = v_yk_y\tau_1 + v_zk_z\tau_2 \pm 2t\,k_Wk_x\,(\tau_3 + \gamma_W\tau_0)\,. \tag{9.8}$$

The Hamiltonian H_W (9.7) fulfills the symmetry conditions

$$P_W^\dagger H_W(-k)\,P_W = H_W(k)\,, \quad T_W^\dagger H_W(-k)\,T_W \neq H_W(k) \tag{9.9}$$

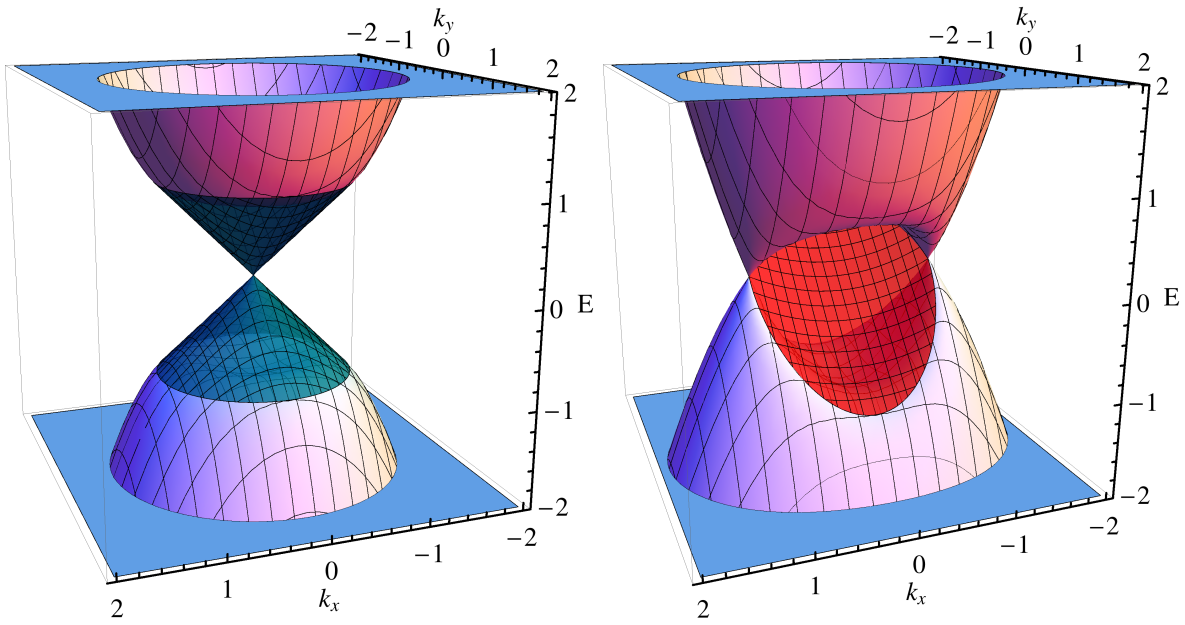


Figure 9.1. (left) Bulk and surface dispersion relations, Eqs. (9.3) and (9.6), of the TI model. The surface band is plotted in cyan. Parameters: $\mathcal{C} = 0$, $M_0 = -1$, $M_1 = 1$, $\mathcal{A} = 1$, $\mathcal{B} = 1$, $k_z = 0$. (right) Bulk and upper surface band structure, Eqs. (9.10) and (9.13), of the Weyl model. The surface band is colored in red. Parameters: $\gamma_W = \frac{1}{4}$, $k_W = 1$, $t = 1$, $v_y = 1$, $v_z = 1$, $k_z = 0$. Reprinted figures with permission from Ref. [JT17]. Copyright (2017) by the APS.

with the $\boldsymbol{\mathcal{I}}$ operator $P_W = \tau_3$ and $\boldsymbol{\mathcal{T}}$ operator $T_W = \tau_0 K$ with K the complex conjugation operator. Hence, parity is preserved and $\boldsymbol{\mathcal{T}}$ symmetry broken. The bulk dispersion relation is then given by

$$E_W = \gamma_W t\left(k_x^2 - k_W^2\right) \pm \sqrt{v_y^2 k_y^2 + v_z^2 k_z^2 + t(k)^2}. \tag{9.10}$$

The surface states can be calculated analytically based on the method discussed in the App. B. This results in the wave function

$$\Psi(z) = \left(e^{ik_{z,1}z} - e^{ik_{z,2}z}\right)\psi_\eta \tag{9.11}$$

with $\psi_\eta = \frac{1}{\sqrt{2}}(1,\ \eta)^{\mathrm{T}}$ and inverse localization length $ik_{z,\frac{1}{2}} = \frac{1}{2t}\left[-\eta v_z \pm \sqrt{4t^2(k_\parallel^2 - k_W^2) + v_z^2}\right]$. The sign $\eta = \pm$ depends on the surface; $\eta = -\mathrm{sgn}(v_z/t)$ (upper surface) or $\eta = \mathrm{sgn}(v_z/t)$ (lower surface). The existence condition for the surface state is

$$k_\parallel^2 < k_W^2 \tag{9.12}$$

such that Fermi arcs can only exist between the Weyl points. Hence, the surface dispersion relation yields the known Fermi arc spectrum

$$E_W^{sur} = \gamma_W t\left(k_x^2 - k_W^2\right) + \eta v_y k_y. \tag{9.13}$$

The combined dispersion relations of the bulk and surface of the WSM are shown in Fig. 9.1 (right).

9.2. Coupled System

The Hamiltonians and surface wave functions of the TI and WSM phases discussed in Sec. 9.1 are very similar. Thus, we conjecture that also the combined system may have surface states which can be calculated by the simplified method described in App. B. This will allow us to discuss the surface physics analytically. In this section, we define the combined Hamiltonian and discuss the couplings allowed by symmetry under the assumptions that certain symmetries are preserved. The surface state Hamiltonian and wave function are derived and the limitations due to the approximated calculation method are discussed.

The combined Hamiltonian of the TI and WSM phases is defined by

$$H_{WTI} = \begin{pmatrix} H_{TI} & H_C \\ H_C^\dagger & H_W \end{pmatrix} \tag{9.14}$$

with the coupling H_C. Such a coupling can be regarded as a tunneling Hamiltonian approach where H_C (weakly) couples the two entities H_{TI} (9.1) and H_W (9.7). A similar approach has been considered in Ref. [BPF+15] to combine topological systems of different kinds with each other and study their emerging physics. The combined symmetry operator for $\boldsymbol{\mathcal{I}}$ symmetry is now given by

$$P_{WTI} = \begin{pmatrix} P_{TI} & 0 \\ 0 & P_W \end{pmatrix}. \tag{9.15}$$

As $\boldsymbol{\mathcal{T}}$ symmetry is already broken in the subsystem of the WSM, it will also be absent in the combined system. The study of the effect on the TI of such a breaking of $\boldsymbol{\mathcal{T}}$ symmetry via coupling, applicable e.g. in the setup of spatially separate Weyl and TI phases as depicted in Fig. 9.7 (b), is one of the goals of this paper. The stability of gapless edge states to $\boldsymbol{\mathcal{T}}$ symmetry breaking perturbations such as magnetic fields [MCW+15] and considerable Coulomb interaction, see Chap. III about quantum Hall topological insulators (QHTIs), is an active research topic and has been experimentally observed in 2D. It is proposed that crystalline symmetries such as $\boldsymbol{\mathcal{I}}$ or rotational symmetries protect the gapless edge states in the absence of $\boldsymbol{\mathcal{T}}$ symmetry, as discussed in Sec. 2.4.3 for topological crystalline insulators (TCIs). Since $\boldsymbol{\mathcal{I}}$ symmetry is preserved in our system, we conjecture that the use of the gapless TI model can be justified even in a $\boldsymbol{\mathcal{T}}$ breaking environment.

Applying the $\boldsymbol{\mathcal{I}}$ operator to the Hamiltonian, following Eqs. (9.2) and (9.9), yields restrictions for the allowed couplings, in the assumption that this symmetry is not broken. As the symmetry operator is block-diagonal, these restrictions do not depend on H_W or H_{TI}. For an $\boldsymbol{\mathcal{I}}$ symmetric system, the couplings proportional to τ_3 and τ_0 have to be even in momentum, while the ones proportional to τ_2 and τ_1 have to be odd in momentum. We choose the following representation

$$H_{C,IS} = \begin{pmatrix} H_{c,IS} \\ \tilde{H}_{c,IS} \end{pmatrix}, \quad H_{c,IS} = d(k_\parallel)\tau_3 + c_1 k_+ \tau_2 + b_1 k_+ \tau_1 + a(k_\parallel)\tau_0, \tag{9.16}$$

where $d(k_\parallel) = d_0 + d_2 k_\parallel^2$ and $a(k_\parallel) = a_0 + a_2 k_\parallel^2$. $\tilde{H}_{c,IS}$ has the same structure. This choice ensures the preservation of parity for the combined system. The size of the terms depends on the concrete experimental realization, where the best candidate materials for our proposal have yet to be identified. In the case of two spatially separate Weyl and TI systems, as depicted in Fig. 9.7 (b), the coupling parameters can be calculated from the overlap of the wave functions

of the different materials. As an example, this is done in Ref. [MPER12] for a bilayer HgTe quantum well (QW) system by fitting a $\boldsymbol{k}\cdot\boldsymbol{p}$ model to experimentally obtained band structures. In general, all symmetry allowed couplings can be relevant for the following discussion.

In our analysis below however, coupling terms proportional to k_z are not considered, for simplicity. This is a physically reasonable assumption at least for the surface states, if one assumes them to be 2D, perfectly localized in the z-direction. Close to the Weyl points or the TI bulk band edge, where the surface states delocalize, a k_z dependent coupling should be taken into account. The ansatz we will consider is thus

$$\Psi(z) = \left(e^{ik_{z,1}z} - e^{ik_{z,2}z}\right) \begin{pmatrix} L_1(k_\pm)\,\psi_{\eta_{TI}} \\ L_2(k_\pm)\,\psi_{\eta_{TI}} \\ L_3(k_\pm)\,\psi_{\eta_W} \end{pmatrix} \tag{9.17}$$

with $\psi_\eta = \frac{1}{\sqrt{2}}(1,\ \eta)^{\mathrm{T}}$. This is a special case of the general form of the surface wave function $\Psi_g(z) = \sum_j a_j e^{ik_{z,j}z}\psi(k_\pm, k_{z,j})$, $j \in \{1,...,6\}$. Its choice is motivated by the ability to obtain analytical solutions for the surface states. Physically it means that we only consider solutions were the TI and WSM surface states have the same exponential localization with the same localization length. This implies that phase transitions of the subsystems, such as normal insulator (NI) to TI or NI to WSM, can not be discussed separatly in this treatment. However, for a system deep in the TI and WSM phase, the simplification should not alter the essential physics. We have checked numerically that small differences in the localization lengths of the subsystems do not alter the surface dispersion relations in a qualitative way, see App. C.

Projecting the Hamiltonian (9.14) on the surface, the eigenvalue equation separates into simpler problems

$$H^{sur}_{WTI}\begin{pmatrix} L1 \\ L2 \\ L3 \end{pmatrix} = E^{sur}_{WTI}\begin{pmatrix} L1 \\ L2 \\ L3 \end{pmatrix}, \quad H^{k_z}_{WTI}\begin{pmatrix} L1 \\ L2 \\ L3 \end{pmatrix} = 0 \tag{9.18}$$

with the Hamiltonians

$$H^{sur}_{WTI} = \begin{pmatrix} \mathcal{C} & i\eta\mathcal{A}k_- & a(k_\parallel) + \eta b_1 k_+ \\ -i\eta\mathcal{A}^* k_+ & \mathcal{C} & \tilde{a}(k_\parallel) + \eta\tilde{b}_1 k_+ \\ a(k_\parallel)^* + \eta b_1^* k_- & \tilde{a}(k_\parallel)^* + \eta\tilde{b}_1^* k_- & \gamma_W t\,(k_x^2 - k_W^2) + \eta v_y k_y \end{pmatrix}, \tag{9.19}$$

$$\underset{part.\ diag.}{\Rightarrow} \begin{pmatrix} \mathcal{C} + |\mathcal{A}|\,k_\parallel & 0 & \tilde{H}_c + e^{i\phi_k^{\mathcal{A}}} H_c \\ 0 & \mathcal{C} - |\mathcal{A}|\,k_\parallel & \tilde{H}_c - e^{i\phi_k^{\mathcal{A}}} H_c \\ \tilde{H}_c^* + e^{-i\phi_k^{\mathcal{A}}} H_c^* & \tilde{H}_c^* - e^{-i\phi_k^{\mathcal{A}}} H_c^* & \gamma_W t\,(k_x^2 - k_W^2) + \eta v_y k_y \end{pmatrix}, \tag{9.20}$$

$$H^{k_z}_{WTI} = \begin{pmatrix} \mathcal{M}(k) - i\eta\mathcal{B}k_z & 0 & d(k_\parallel) - i\eta c_1 k_+ \\ 0 & \mathcal{M}(k) - i\eta\mathcal{B}k_z & \tilde{d}(k_\parallel) - i\eta\tilde{c}_1 k_+ \\ d(k_\parallel)^* - i\eta c_1^* k_- & \tilde{d}(k_\parallel)^* - i\eta\tilde{c}_1^* k_- & t(k) - i\eta v_z k_z \end{pmatrix} \tag{9.21}$$

for $\eta = \eta_{TI} = \eta_W$. In Eq. (9.20), we partially diagonalize the Hamiltonian and define $\phi_k^{\mathcal{A}} = \phi_k - \phi_{\mathcal{A}} - \eta\frac{\pi}{2}$, $H_c = a(k_\parallel) + \eta b_1 k_+$ and $\tilde{H}_c = \tilde{a}(k_\parallel) + \eta\tilde{b}_1 k_+$. This will help in the interpretation of the surface dispersion relation in terms of coupled Dirac cone and Fermi arc. In the case of $\eta = \eta_{TI} = -\eta_W$, one has to replace in Eqs. (9.19) - (9.21) $a(k_\parallel) \leftrightarrow d(k_\parallel)$, $b_1 \leftrightarrow ic_1$, $v_y \to -v_y$ and $v_z \to -v_z$. We will focus in the following on the former, $\eta_{TI} = \eta_W$, case.

Taking $(\ L1 \quad L2 \quad L3\ )^T$ as the same eigenvector in Eq. (9.18), the latter can only be fulfilled by further restrictions on the parameters. We choose a locking between some of the TI and

the WSM parameters, i.e. $t(k) = \zeta\mathcal{M}(k)$ and $v_z = \zeta\mathcal{B}$ with ζ a constant (set to 1 in the following). This ensures the same localization length for the two subsystems. Additionally, the couplings c_1 and $d(k_\parallel)$ are set to be 0 for simplicity. Therefore the total coupling does not change the original orbital character of the TI and WSM surface states, being eigenstates of the τ_1 matrix with fixed eigenvalue + or −. With regard to these restrictions, we have checked that the neglected couplings can be considered numerically with only quantitative changes to the surface dispersion relations, see App. C.

In total, this leads to the same quadratic equation for k_z as in the pure TI case, $ik_{z,\frac{1}{2}} = \frac{1}{2M_1}\left[-\eta\mathcal{B} \pm \sqrt{4M_1(M_0 + M_1k_\parallel^2) + \mathcal{B}^2}\right]$. The existence condition is again

$$M_1\left(M_0 + M_1k_\parallel^2\right) < 0 \tag{9.22}$$

and the (unnormalized) eigenvectors are given by

$$\begin{pmatrix} L1 \\ L2 \\ L3 \end{pmatrix} = \begin{pmatrix} \left(E_{WTI}^{sur} - \mathcal{C}\right) H_c + i\eta\mathcal{A}k_-\tilde{H}_c \\ \left(E_{WTI}^{sur} - \mathcal{C}\right) \tilde{H}_c - i\eta\mathcal{A}^*k_+H_c \\ \left(E_{WTI}^{sur} - \mathcal{C}\right)^2 - |\mathcal{A}|^2 k_\parallel^2 \end{pmatrix}. \tag{9.23}$$

The eigenenergies E_{WTI}^{sur} are too lengthy to state them here, but can also be derived analytically.

The obtained solution leads to the possibility to tune bulk and surface dispersion relations rather independently. Parameters M_i and $\mathcal{B}$ influence the surface dispersion relation only indirectly via the existence condition and finite γ_W parameter, while they strongly influence the bulk band structure as will be shown in the next section. Tuning the coupling constants v_y, $\mathcal{A}$, $a(k_\parallel)$, b_1 and their relative phases will still provide a rich parameter space to be explored below.

9.3. Surface Dispersion Relation

In this section, we discuss the influence of the different coupling parameters on the combined surface states of TIs and WSMs. Depending on the choice of symmetries of the coupling, observed phenomenas are the generation of additional Dirac points in the dispersion relation or the spin polarization of certain surface bands.

9.3.1. Uncoupled Scenario

Beginning with the uncoupled case, $H_{C,IS} = 0$, the dispersion relations of the surface and bulk states are shown in Fig. 9.2. The black lines denote the bulk dispersion relation, cyan (from blue (green) for spin up (down)) and red stand for the TI and WSM surface states, respectively. The two black dots give the position of the bulk Weyl points. We note that the surface states originate at the bulk states, but cross them unaffectedly. Together with the fact that one can tune the bulk gap M_0 without changing the surface dispersion relation (aside from the existence condition), we find the possibility to discuss the bulk and surface dispersion relations rather separately from each other. It will always be possible to increase the bulk gap and the distance between the two Weyl points such that the interesting surface physics happens in regions of the BZ where no bulk state is located. Therefore, we will focus in the following on tuning of the surface dispersion relation only. In numerical calculations, see App. C, purely exponentially decaying surface states do not coexist with bulk states at the same energy and momenta. This is due to finite hybridization between the bulk and surface states.

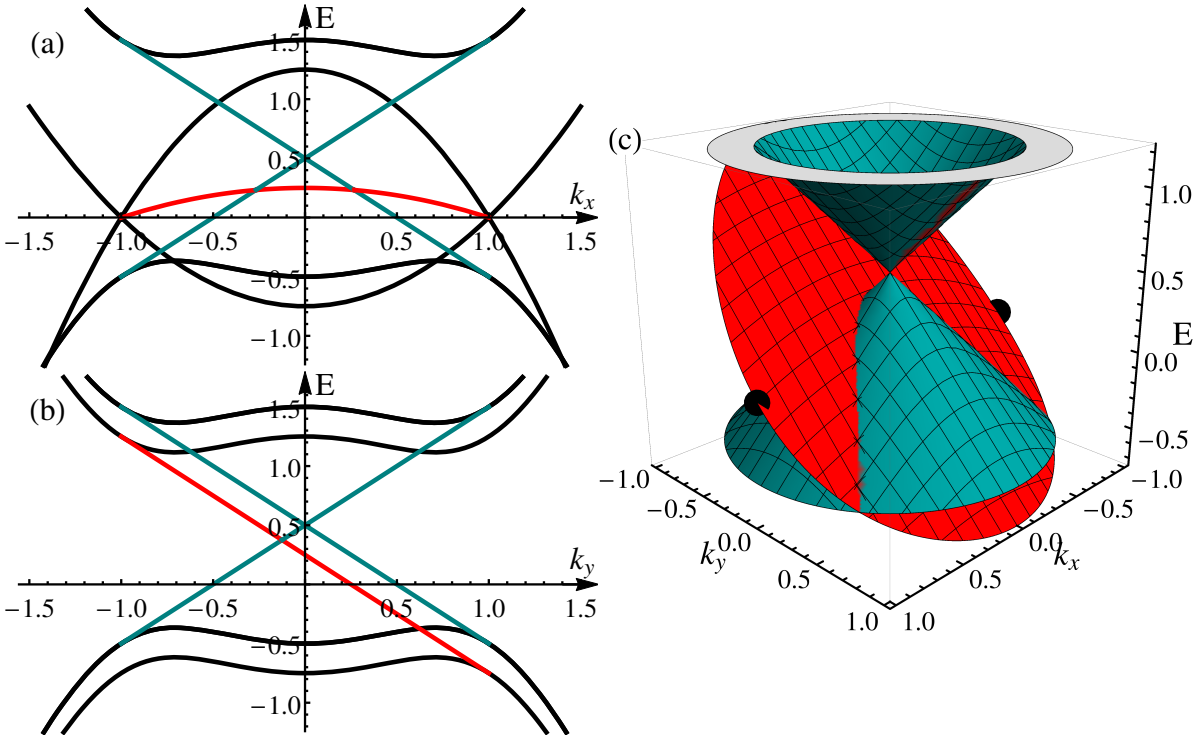


Figure 9.2. (a) and (b) Bulk and upper surface dispersion relations of the uncoupled TI and WSM models. Color code: Black lines for the bulk states, red for WSM character and cyan for the TI character of the surface states. (c) 3D plot of the surface dispersion relation. The two black dots denote the position of the bulk Weyl points. Parameters: $\mathcal{C} = \frac{1}{2}$, $M_0 = -1$, $M_1 = 1$, $\mathcal{B} = 1$, $k_z = 0$, $\gamma_W = -\frac{1}{4}$, $\mathcal{A} = 1$, $v_y = 1$, $a(k_\parallel) = b_1 = 0$, $H_c = \tilde{H}_c$. Reprinted figures with permission from Ref. [JT17]. Copyright (2017) by the APS.

9.3.2. Real, spin-symmetric Coupling: Creation of additional Dirac Points

A straight-forward way to couple TI and WSM is a real and spin-symmetric coupling via $a(k_\parallel) > 0$ or $b_1 > 0$ with $H_c = \tilde{H}_c$. This kind of coupling leads generally to two Dirac points in the combined surface dispersion relation, as plotted in Fig. 9.3. One Dirac point is just shifted by the coupling to the Weyl surface state. The other one is created out of the Weyl and Dirac states along a momentum direction where there is no coupling between these two bands. Under the assumption that both spin species couple equally strong to the WSM, $|\tilde{H}_c| = |H_c|$, there is always such a momentum direction ϕ_k where one part (hole or electron) of the Dirac cone is not coupled to the WSM surface state, while the other part is maximally coupled, see Eq. (9.20) above. For the lower, hole-like cone, using the parameters in Fig. 9.3, this direction is $\phi_k = -\frac{\pi}{2}$, thus the negative k_y axis with $k_x = 0$. The dispersion relation is then $E = \mathcal{C} + \mathcal{A}k_y$ corresponding to the cyan line in Fig. 9.3 (a) which crosses the other two straight lines.

Considering finite couplings $b_1 \neq 0$ instead of $a_0 \neq 0$ gives only quantitative differences in the dispersion relations (not shown). The Dirac point generation is unaffected, except for the special case where the Dirac point and Fermi arc cross only at $k_x = k_y = 0$. As in this case the coupling for $b_1 \neq 0$ is absent in this point, no second Dirac point is generated.

A perturbative calculation can provide some insight into both kinds of Dirac points. We take the surface Hamiltonian, Eq. (9.19), and treat one band as a perturbation to the other

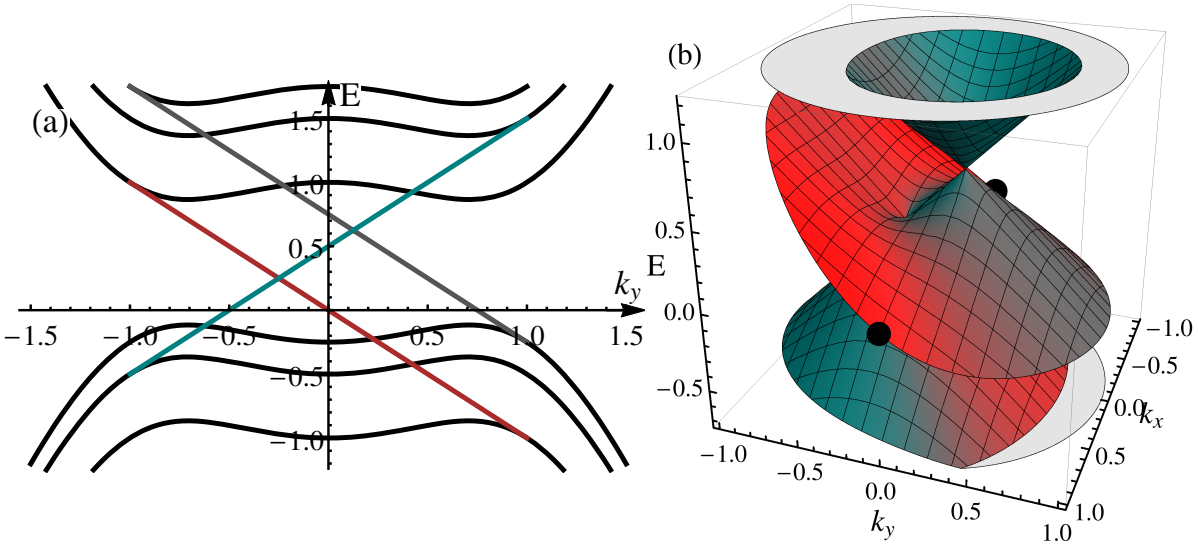


Figure 9.3. (a) Bulk and upper surface dispersions relation of the TI and WSM model with real, spin-symmetric coupling. Color code: Black lines for the bulk states, red for WSM character and cyan for the TI character of the surface states. (b) 3D plot of the surface dispersion relation. Two Dirac points are visible. Parameters: $\mathcal{C} = \frac{1}{2}$, $M_0 = -1$, $M_1 = 1$, $\mathcal{B} = 1$, $k_z = 0$, $\gamma_W = -\frac{1}{4}$, $\mathcal{A} = 1$, $v_y = 1$, $a(k_\parallel) = \frac{1}{4}$, $b_1 = 0$, $H_c = \tilde{H}_c$. Reprinted figures with permission from Ref. [JT17]. Copyright (2017) by the APS.

two. For the shifted Dirac point one directly finds in 2nd order perturbation theory in the coupling

$$H_D^1 = \begin{pmatrix} \mathcal{C} & i\eta\mathcal{A}k_- \\ -i\eta\mathcal{A}^* k_+ & \mathcal{C} \end{pmatrix} + \frac{1}{\mathcal{C} - \gamma_W\left(M_0 + M_1 k_x^2\right) - \eta v_y k_y} \begin{pmatrix} |H_c|^2 & H_c\tilde{H}_c^* \\ \tilde{H}_c H_c^* & |\tilde{H}_c|^2 \end{pmatrix}. \quad (9.24)$$

Evidently, a difference in the absolute values of the coupling between the Weyl system and the different spin species of the TI system will open a gap. In the limit of spin degeneracy, where $H_c = \tilde{H}_c$, we insert the coupling from Eq. (9.16), expand Eq. (9.24) for small momenta and find

$$H_D^1 = \begin{pmatrix} \mathcal{C} & i\eta\mathcal{A}k_- \\ -i\eta\mathcal{A}^* k_+ & \mathcal{C} \end{pmatrix} + \frac{|a_0|^2}{\mathcal{C} - \gamma_W M_0} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \mathcal{O}\left(k_\pm\right), \quad (9.25)$$

corresponding to a Dirac cone shifted in energy and momentum by the coupling. For real $\mathcal{A}$, the shift occurs in the k_y direction as shown in Fig. 9.3.

The creation of the second Dirac point can be understood from a similar calculation. The perturbative Hamiltonian for this Dirac point is given by

$$H_D^2 = \begin{pmatrix} \mathcal{C} - |\mathcal{A}|\, k_\parallel & \frac{1}{\sqrt{2}} H_c (1 - e^{i\phi_k^{\mathcal{A}}}) \\ \frac{1}{\sqrt{2}} H_c^* (1 - e^{-i\phi_k^{\mathcal{A}}}) & \gamma_W\left(M_0 + M_1 k_x^2\right) + \eta v_y k_y - |H_c|^2 \frac{1+\cos(\phi_k^{\mathcal{A}})}{\mathcal{C}+|\mathcal{A}|k_\parallel - \gamma_W(M_0+M_1 k_x^2) - \eta v_y k_y} \end{pmatrix}. \quad (9.26)$$

The off-diagonal elements vanish along the momentum direction $\phi_k = \phi_{\mathcal{A}} + \eta\frac{\pi}{2}$. Thus, Weyl and Dirac surface states are uncoupled in one point. This point becomes the new Dirac point, and setting the diagonal elements of Eq. (9.26) equal, this gives its precise value k_D. For

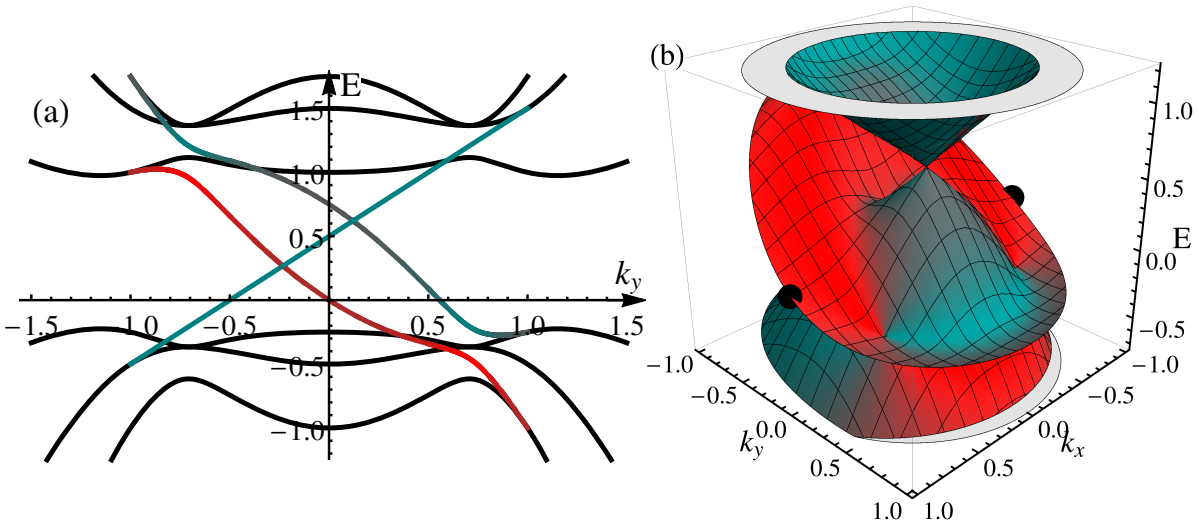


Figure 9.4. (a) Bulk and upper surface dispersion relations of the TI and WSM model with a coupling that changes sign. Color code: Black lines for the bulk states, red for WSM character and cyan for the TI character of the surface states. (b) 3D plot of the surface dispersion relation. Four Dirac points are visible. Parameters: $\mathcal{C} = \frac{1}{2}$, $M_0 = -1$, $M_1 = 1$, $\mathcal{B} = 1$, $k_z = 0$, $\gamma_W = -\frac{1}{4}$, $\mathcal{A} = 1$, $v_y = 1$, $a_0 = \frac{1}{4}$, $a_2 = -\frac{1}{2}$, $b_1 = 0$, $H_c = \tilde{H}_c$. Reprinted figures with permission from Ref. [JT17]. Copyright (2017) by the APS.

the parameters used in Fig. 9.3, the Dirac point k_D is on the negative k_y axis, with the corresponding Hamiltonian

$$H_D^2 = \begin{pmatrix} \mathcal{C} + \mathcal{A}\,(k_D + k_y) & \frac{1}{\sqrt{2}} k_x \left(b_1 + i\frac{a_0}{k_D}\right) \\ \frac{1}{\sqrt{2}} k_x \left(b_1 - i\frac{a_0}{k_D}\right) & \mathcal{C} + \mathcal{A}\,(k_D - k_y) + 2 f_{di}(k_x, k_y) \end{pmatrix} \tag{9.27}$$

including the distortion $f_{di}(k_x, k_y) = (\mathcal{A} - v_y)\, k_y + \frac{2b_1(a_0 k_x - x_1 k_D k_y) + \mathcal{A} k_D k_y (\mathcal{A} - v_y)}{\mathcal{C} - (\mathcal{A} - v_y) k_D - \gamma_W M_0}$ and $k_D = \frac{v_y(\mathcal{C} - \gamma_W M_0) - \sqrt{2a_0^2(\mathcal{A}^2 - v_y^2 - 2b_1^2) + (\mathcal{A}^2 - 2b_1^2)(\mathcal{C} - \gamma_W M_0)^2}}{\mathcal{A}^2 - v_y^2 - 2b_1^2}$. The Dirac point is stable for any real combination of spin-symmetric couplings. A finite distortion $f_{di} \neq 0$ tilts the Dirac cone but does not open a gap.

The number of Dirac points in the surface dispersion relation can be extended further by a coupling that changes sign as a function of k_x and k_y, e.g. by setting $a_0 > 0$ and $a_2 < 0$ or a combination of $a_0 \neq 0$ and $b_1 \neq 0$. The positions in k-space where the coupling is zero and TI and WSM surface state intersect will then harbor additional Dirac points, see Fig. 9.4.

9.3.3. Spin-asymmetric Coupling: Creation of Gaps & Spin Polarization

The spin-up and spin-down TI bands do not need to have the same coupling to the WSM. If the absolute values are different, $|H_c| \neq |\tilde{H}_c|$, the Dirac points in the surface dispersion relation are gapped out, see Eq. (9.24) and Fig. 9.5. This leaves the bulk Weyl points, however, unaffected. The resulting surface bands are partly spin polarized, as shown in Fig. 9.5. The weaker coupled spin-up electrons form a band with the Weyl surface state at intermediate energies, while the stronger coupled spin-down electrons are pushed into the upper and lower bands. The resulting surface dispersion relation is still, overall, gapless.

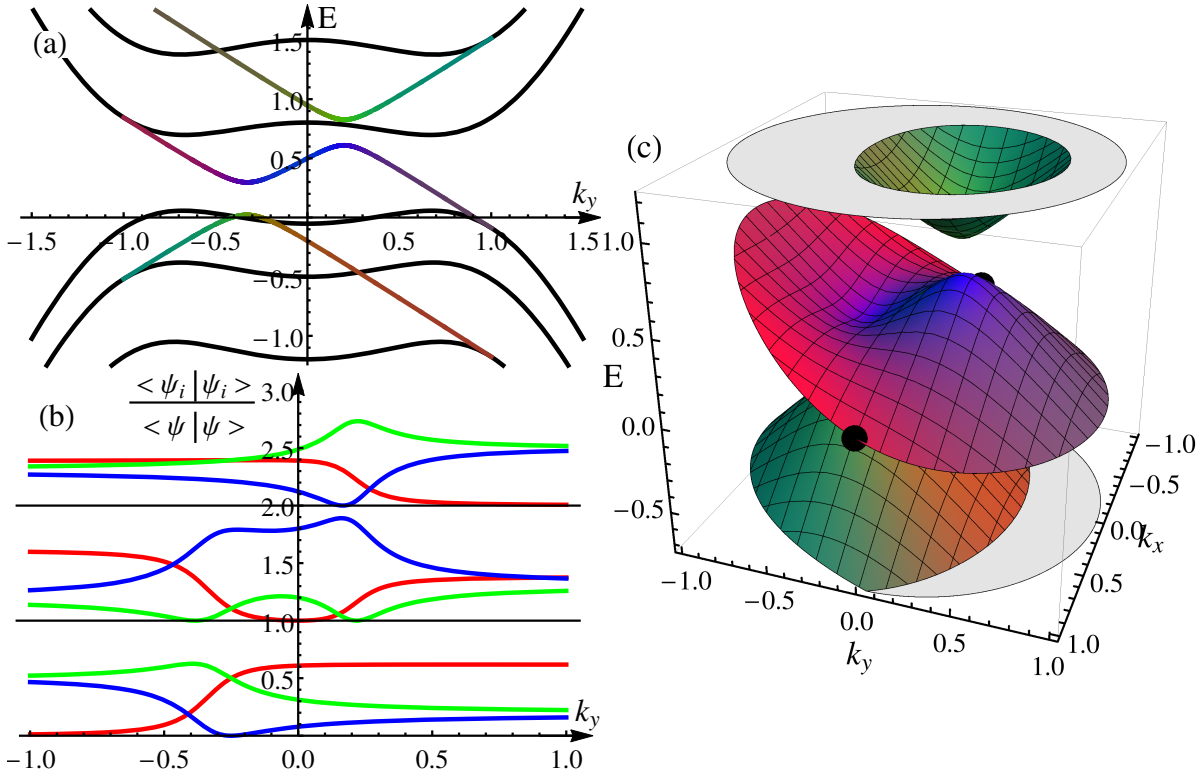


Figure 9.5. (a) Bulk and upper surface dispersion relations of the TI and WSM model with a spin-asymmetric coupling. Color code: Black lines for the bulk states, red for WSM character and blue (green) for the TI spin up (down) character of the surface states. (b) Character of the three surface bands, shifted for clarity. (c) 3D plot of the surface dispersion relation. All Dirac points are gapped. Parameters: $\mathcal{C} = \frac{1}{2}$, $M_0 = -1$, $M_1 = 1$, $\mathcal{B} = 1$, $k_z = 0$, $\gamma_W = -\frac{1}{4}$, $\mathcal{A} = 1$, $v_y = 1$, $a(k_\parallel) = \frac{1}{4}$, $\tilde{a}(k_\parallel) = \frac{2}{4}$, $b_1 = \tilde{b}_1 = 0$. Reprinted figures with permission from Ref. [JT17]. Copyright (2017) by the APS.

Considering a finite $b_1 \neq 0$ instead of an a_0 coupling, only the lower Dirac point will split. As the upper one is located at $k_x = k_y = 0$ for a pure momentum dependent coupling, the effective coupling between the WSM and TI surface states is zero here.

9.3.4. Phase-shifted Coupling: Moving Dirac Points, tilting Dispersion Relation

Including complex coupling constants, this offers additional ways to alter the bulk and surface spectrum. In general, the dispersion relation will look much less symmetric compared to the previous, real couplings. Assuming $\tilde{H}_c = H_c$, one can directly conclude from the Hamiltonian in Eq. (9.20) that a complex coupling $\mathcal{A} = i$ will lead to two Dirac points lying on the k_x, rather than on the k_y axis as discussed in Sec. 9.3.2. This is confirmed in Fig. 9.6. One also sees that the bulk Weyl points lie not on the k_x axis, but are rotated by the complex coupling. Yet the rotation is much smaller than the $\pi/2$ rotation of the surface Dirac points.

The same effect is obtained by a complex phase difference between the couplings H_c and $\tilde{H}_c$. It can even undo the rotation induced by $\mathcal{A} = i$. Note also that in the spin-symmetric case, already for real and finite $a(k_\parallel)$ and b_1 the Weyl points are rotated away from the k_x axis. Here, the effective coupling is complex, with a phase changing with $k_\pm$. Supplementing this with a complex $a(k_\parallel)$, this can again lead to points where the effective coupling is zero,

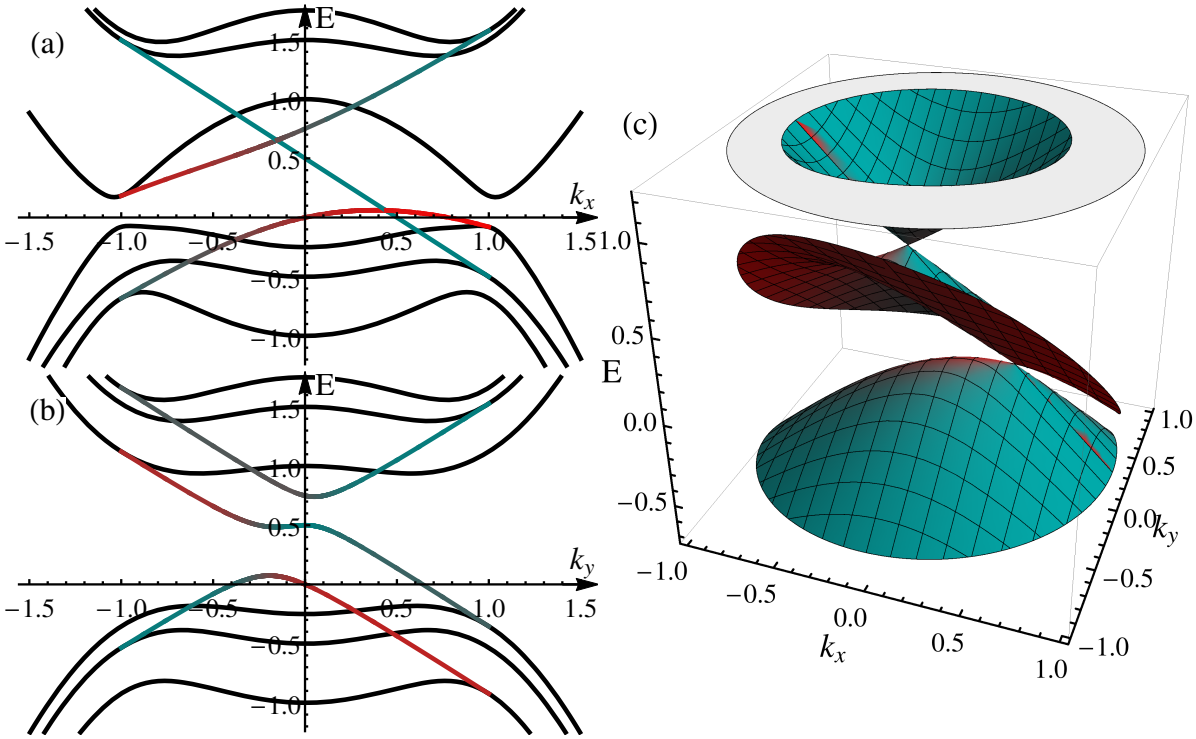


Figure 9.6. (a) and (b) Bulk and upper surface dispersion relations of the TI and WSM model with a complex coupling. Color code: Black lines for the bulk states, red for WSM character and cyan for the TI character of the surface states. (c) 3D plot of the surface dispersion relation. Two Dirac points on the k_x axis are visible. Parameters: $\mathcal{C} = \frac{1}{2}$, $M_0 = -1$, $M_1 = 1$, $\mathcal{B} = 1$, $k_z = 0$, $\gamma_W = -\frac{1}{4}$, $\mathcal{A} = i$, $v_y = 1$, $a(k_\parallel) = \frac{1}{4}$, $b_1 = 0$, $H_c = \tilde{H}_c$. Reprinted figures with permission from Ref. [JT17]. Copyright (2017) by the APS.

resulting in additional Dirac points like in Sec. 9.3.2.

In conclusion, the four parameters v_y, $\mathcal{A}$, $a(k_\parallel)$ and b_1 offer almost endless possibilities to tune and control the surface band structure, of which we presented only the most basic ones above. Especially in the combination of the parameters lies still uncharted potential.

9.4. Experimental Realization

We propose two ways to realize the physics of hybrid TI and WSM phases in an experimental setup. First, a material that naturally is in this combined phase will have corresponding surface states, as depicted in Fig. 9.7 (a). Compressively strained HgTe is here a candidate material: the compressive strain pushes the Γ_8 bands against each other creating Weyl points [RJY+16]. In addition, the topological band inversion between the Γ_8 and Γ_6 bands prevails, see Fig. 2.15 for a plot of the band structure. A difference to our calculation is the preserved $\mathcal{T}$ symmetry, leading to eight Weyl points in HgTe instead of two. However, if HgTe is doped with Mn, the number of Weyl points could be reduced by a (partial) magnetic ordering.

The second realization consists of a WSM in contact with a 3D TI, possibly separated by a thin buffer layer as depicted in Fig. 9.7 (b). This should lead to a hybrid surface state at the joint boundary. The finite coupling H_C could be provided by tunneling or Coulomb

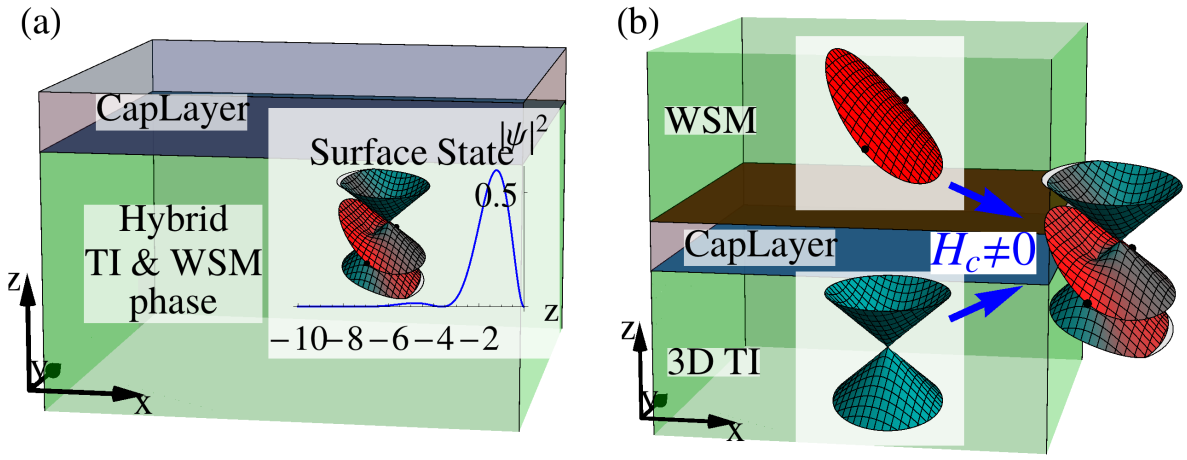


Figure 9.7. Possible experimental realizations. (a) Bulk materials being in the combined 3D TI and WSM phase will naturally have hybrid surface states. (b) A heterostructure where TI and WSM phases are adjacent to each other will exhibit hybrid surface states for finite coupling $H_C \neq 0$, provided e.g. by Coulomb interaction or tunneling.

interaction. While this surface state is not exactly of the form of the ansatz in Eq. (9.17), the surface Hamiltonian, Eq. (9.19), should still be valid with the modification $\eta = \eta_{TI} = -\eta_W$. As several proposals of $\mathcal{T}$-symmetry-broken WSM with two Weyl points are based on magnetically doped 3D TI materials [BHB11, Cho11, XWW+11, BLQ14], the fabrication of the described hybrid system should be technically feasible.

Conclusion

We have analyzed a hybrid system composed of a 3D TI coupled to an $\mathcal{I}$ symmetric, $\mathcal{T}$-symmetry-broken WSM. In the spirit of a tunnel coupling approach between the two topological phases, the use of a simplified ansatz made it possible to find an analytical solution for the surface states. The resulting surface Hamiltonian, Eq. (9.19), is a major result of this paper.

The dispersion relation of the hybrid system shows different phenomena depending on the assumed coupling between WSM and TI. Preserved spin symmetry, e.g., leads to the creation of additional Dirac points in the surface dispersion relation. Breaking of spin symmetry on the other hand, this opens gaps and induces spin polarization in the former Dirac surface cone.

As an experimental realization we have presented both strained HgTe, which might naturally be in the discussed hybrid phase, and a heterostructure of TI and WSM. In the latter case, the joint boundary would harbor the interesting hybrid surface state.

There are several directions how to proceed with this research. Looking for measurable consequences, e.g. in transport or spectroscopy, of the new hybrid surface states should be the most immediate one. We expect, for instance, that different Dirac points will give rise to different minima in the conductivity, similar to the graphene case [Kat06, TTT+06]. An extension to $\mathcal{T}$ symmetric WSM is another one. For this, one should use a 4x4 Hamiltonian for the WSM, which offers the possibility of more involved Fermi arcs on the surface, e.g. including spin polarization along the arcs [LMQ+15, XWW+16]. $TaIrTe_4$ [KKE+16, BYS+16] with its four Weyl points could be a candidate material for a hybrid system of this kind.

10. Conclusion & Outlook

We began our analysis of interacting topological systems in Chap. II, where we have examined the influence of Coulomb interaction on a 2D TI. The system there is described by the Bernevig-Hughes-Zhang (BHZ) model, applicable e.g. to HgTe QWs. The bulk physics is governed by an interplay between Dirac and Schrödinger fermions. This gives rise to a new interband plasmon in the Random Phase Approximation (RPA) excitation spectrum in the intrinsic limit. Its visibility in the doped regime is facilitated by the broken particle-hole (p-h) symmetry. Importantly, the plasmon occurs in the right parameter regime for experiments and should be observable by Raman spectroscopy or electron loss spectroscopy. We emphasize that this new plasmon is expected in all BHZ like models with an interplay of linear and quadratic band structures. The interacting plasmon spectrum shows clear differences depending on the topology of the system. Even more suitable for experimental detection could be the optical conductivity, which can diverge in the topologically non-trivial (TnT) phase, offering a way to distinguish the topological phase of 2D TIs from a bulk measurement.

Our studies can be continued in several directions. One is the inclusion of the one-dimensional (1D) edge states appearing in the TnT phase. As the RPA is exact in 1D systems as stressed in Ref. [LDJ92], it offers the possibility to study the interplay between bulk and edge collective charge excitations on the same footing. This should amplify the difference between the TnT and TT phases that we already observed. In another research direction, one could work out the possible benefits of our new plasmons to the field of plasmonics. A starting point could be the analysis of plasmon-mediated Coulomb drag, as done in Ref. [SJ14] for graphene.

The limit of strongly interacting systems has been studied in Chap. III. Here we have introduced QHTIs: topological systems with weak bare interactions, that can become strongly interacting in a strong magnetic field, e.g. around the single-particle topological phase transition. The topological properties are preserved due to some extra symmetry, like $\boldsymbol{\mathcal{I}}$ or a spin rotation symmetry. We have used the well established framework of quantum Hall ferromagnetism (QHFMism), where the dynamics of the system can be described in a nonlinear σ-model for the collective order parameter (OP). We have identified a U(1) symmetry as the necessary condition for a preserved TnT phase in the presence of interactions. If this symmetry is present, the edge can be described by a helical Luttinger liquid. Here, the interaction strength $\mathcal{K}$ and velocity v are tunable by the magnetic field B, all the way from weakly interacting at low fields to divergent interaction strength at the phase transition.

As an outlook, future research should focus on measureable consequences in QHTI materials. The interaction parameter $\mathcal{K}(B)$ and velocity $v(B)$ of the Luttinger edge are quite difficult to measure in dc transport, as discussed in Ref. [SS95]. They usually appear in power laws like $I \propto V^{\alpha(\mathcal{K})}$ for an edge with impurities or setups where the current directly tunnels into the edge, see Ref. [FB06]. A better way is time-resolved edge transport as discussed in Refs. [ZHKE93, MTT+17]. Here, $\mathcal{K}(B)$ and $v(B)$ as a function of the magnetic field are directly accessible. For these measurements, one should look for QHTI systems that are easily accessed in experiments. A more theory focused direction of future research would be the extension of the effective edge theory beyond the point of the topological phase transition. The low-energy physics in this

regime is expected to be governed by an entanglement of neutral bulk and edge excitations, promising rich new physics to explore.

In the last Chap. IV, we have examined a hybrid system of a 3D TI coupled to an $\mathcal{I}$-symmetric, $\mathcal{T}$-symmetry-broken WSM. A main result is the analytical surface Hamiltonian of the combined system, describing the combination of a Fermi arc and a Dirac cone. For preserved spin symmetry, additional Dirac points are created in the surface dispersion relation. Breaking the spin symmetry leads to the opening of gaps and induces a spin polarization in the surface state. As experimental realizations we have considered strained HgTe, which might even be naturally in this hybrid phase. Another one are heterostructures of TIs and WSMs, with a hybrid surface state at the joint boundary.

In our analysis, we have focused on the surface dispersion relation. An obvious proceeding of our work would be the search for measurable signatures of the discussed hybrid physics in transport or spectroscopy. The identification of experimentally realistic systems is a related research direction. Here, an extension of the analysis to $\mathcal{T}$-symmetric WSM systems is a possibility, as they are more abundant in experiments so far.

Part V.

Appendix

A. Excitation spectra: The Dirac and 2DEG regimes

In this appendix, we examine the excitation spectra of the Dirac and the two dimensional electron gas (2DEG) model. They serve as reference for the discussion of the BHZ model in Sec. 6.3, which smoothly interpolates between them.

A.1. 2DEG

In the 2DEG limit, only intraband excitations are possible. The polarization function has a well-known analytical form, cf. Ref. [GV05], and therefore we can easily plot the non-interacting excitation spectrum in Fig. A.1 (a). Π^{Im} is peaked for $q, \omega \to 0$ closely to the

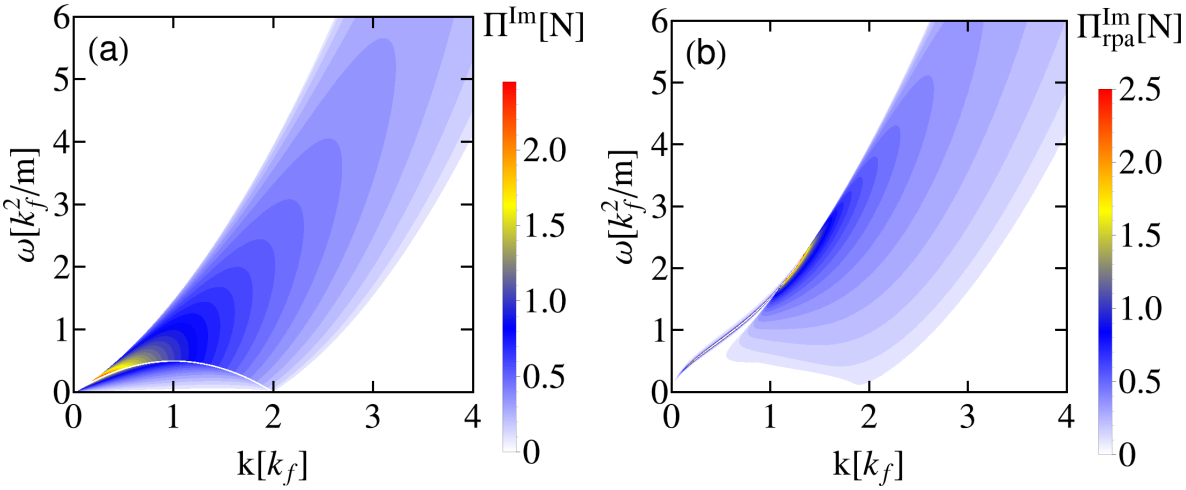


Figure A.1. Excitation spectra of a 2DEG. (a) Imaginary part of Π^R with $N = \frac{g_s m}{2\pi\hbar^2}$ and g_s the degeneracy factor. (b) $\Pi^{\mathrm{Im}}_{\mathrm{rpa}}$ for $r_s = 2$, with $v_q = \frac{r_s k_f}{Nq}$ the Coulomb interaction. We add an artificial damping in the region of $\Pi^{\mathrm{Im}} = 0$ to make the plasmons visible. Reprinted figures with permission from Ref. [JMT14b]. Copyright (2014) by the APS.

upper boundary of the spectrum. It decays to zero instead for large momenta and frequencies like $\Pi^{\mathrm{Im}} \propto q^{-1}$, if one considers a fixed ratio $\omega \propto q^2$ within the single-particle excitation (SPE) region.

The interacting spectrum is shown in Fig. A.1 (b). An intraband plasmon appears with the usual $\sqrt{q}$ dispersion for $q \to 0$. It absorbs all of the spectral weight in this limit, thus $\Pi^{\mathrm{Im}}_{\mathrm{rpa}}$ is suppressed in the SPE region. For intermediate momenta, the plasmon dispersion enters and runs through the SPE region. As a result, the plasmon decays and the peak in the spectrum broadens. For even larger momenta and frequencies, the interacting and non-interacting spectra agree qualitatively.

A.2. Dirac

The Dirac excitation spectrum comprises at finite doping both inter- and intraband excitations. The polarization function again has a well-known analytical expression, see Refs. [WSSG06, HD07], of which we plot the non-interacting spectrum Π^{Im} in Fig. A.2 (a). The interband

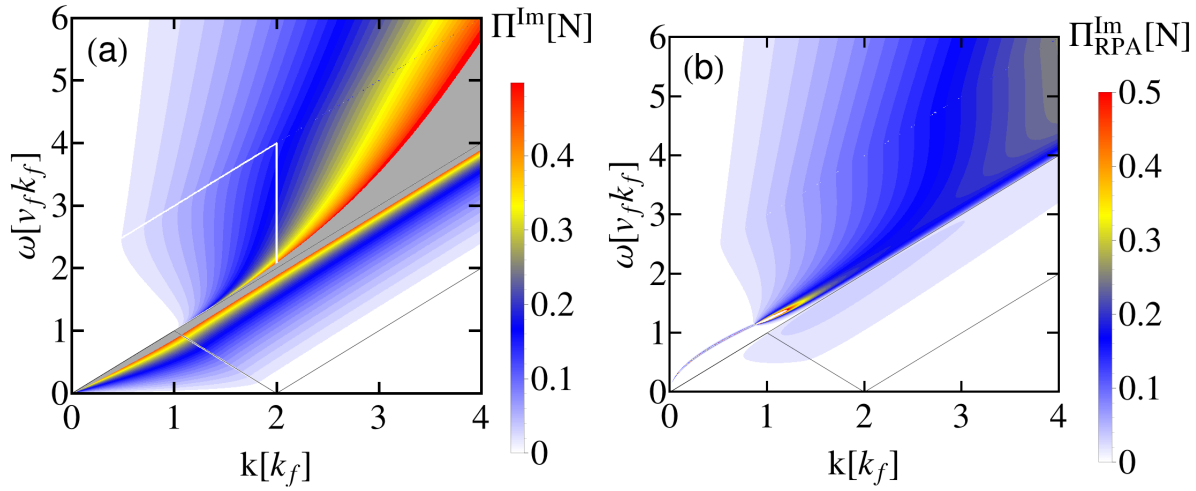


Figure A.2. Excitation spectra of a Dirac system. (a) Π^{Im} with $N = \frac{g_s k_f}{\hbar v_f}$ and g_s the degeneracy factor. In the gray area, the color scale is exceeded due to the divergency of Π^{Im}. (b) $\Pi^{\mathrm{Im}}_{\mathrm{rpa}}$ for $r_s = 2\pi g_s \alpha = 4\pi \cdot 0.6$, with $v_q = \frac{r_s k_f}{Nq}$ the Coulomb interaction. We add an artificial damping in the regions of $\Pi^{\mathrm{Im}} = 0$ to make the plasmons visible. Reprinted figures with permission from Ref. [JMT14b]. Copyright (2014) by the APS.

excitations occur for higher energies $\omega > v_f q$, while for intraband excitations less energy is needed, $\omega < v_f q$. Both excitation spectra touch at $v_f q = \omega$, where they diverge. This divergency is due to the perfect nesting condition of the unbounded, purely linear Dirac model. Only the Fermi-blockade suppresses the interband transitions in Π^{Im} for $q < 2k_f$ and cures the divergency, see Fig. A.2 (a) for $\omega > v_f q$. In the limit of high energies, one finds a ω^{-1} decay.

The interacting spectrum $\Pi^{\mathrm{Im}}_{\mathrm{rpa}}$ is plotted in Fig. A.2 (b) for an interaction strength $\alpha = 0.6$. Similar to the 2DEG, all of the intraband spectral weight is absorbed by a plasmon in the limit $q \to 0$ and the divergence at $v_f q = \omega$ is cured. Interestingly, for sufficient large interaction strength α, the plasmon decays in the interband spectrum. In this case, for larger momenta and frequencies, the intraband polarization does not recover the non-interacting value, as it does for the 2DEG, but remains significantly reduced. Therefore single-particle intraband excitations are blocked altogether for all momenta and frequencies in this limit. The missing spectral weight goes into a charge resonance at higher frequencies in the interband spectrum, as analyzed in Ref. [SNC08]. Yet, this resonance is not a solution of the plasmon equation and therefore not a plasmon, as stated in Ref. [SNPS15].

B. Hardwall Boundary Condition 2x2

In this appendix, we recap a simple method for calculating exponentially localized boundary states of a 2x2 Hamiltonian, following Ref. [LQZ+10] and references therein. We introduce hardwall boundary conditions on a half space $z \leq 0$ or $z \geq 0$. Thus, the surface state is localized at $z = 0$ and decays either in direction $z \to -\infty$ (upper surface) or $z \to +\infty$ (lower surface). The state should fulfill the eigenvalue equation

$$H\Psi(z) = E\Psi(z) \tag{B.1}$$

with $H = \left[h_4\left(k_\parallel^2 + k_z^2\right) + h_3\right]\tau_3 + h_2 k_z \tau_2 + h_1(k_\pm)\tau_1 + h_0\tau_0$ and h_j being real constants or functions of $k_\pm$. The Hamiltonian can represent a TI or WSM depending on the chosen h_j.

The general ansatz for the eigenstate is

$$\Psi_g(z) = \sum_{j\in\{1,2\}} a_j \mathrm{e}^{ik_{z,j}z}\psi\left(k_\pm, k_{z,j}\right), \tag{B.2}$$

which could be used to solve for the surface states of Eq. (B.1) in the usual manner. Yet, due to the specific structure of our Hamiltonian, we can choose a simplified version of the ansatz, given by

$$\Psi(z) = \left(\mathrm{e}^{ik_{z,1}z} - \mathrm{e}^{ik_{z,2}z}\right)\psi\left(k_\pm\right). \tag{B.3}$$

Here, the relative sign ensures that the wave function vanishes at $z = 0$. This ansatz offers the possibility to separate Eq. (B.1) into two parts

$$\left[h_1(k_\pm)\tau_1 + h_0\tau_0\right]\Psi(z) = E\tau_0\Psi(z), \tag{B.4}$$

$$\left[\left[h_4\left(k_\parallel^2 + k_z^2\right) + h_3\right]\tau_3 + h_2 k_z \tau_2\right]\Psi(z) = 0. \tag{B.5}$$

Eq. (B.4) is independent of k_z and can be solved for the surface dispersion relation E, while the solution of Eq. (B.5) defines the two quantized values of k_z needed for the surface eigenstate. Following this procedure, $\psi\left(k_\pm\right) = f\left(k_\pm\right)\psi_\pm$ is taken to be proportional to the eigenstate of the τ_1 Pauli matrix, $\tau_1\psi_\pm = \pm\psi_\pm$, with

$$\psi_\pm = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ \pm 1 \end{pmatrix} \tag{B.6}$$

and $f\left(k_\pm\right) = 1$. Using $\tau_2\psi_\pm = \mp i\psi_\mp$ and $\tau_3\psi_\pm = \psi_\mp$, Eq. (B.5) reduces to the quadratic equation

$$h_4\left(k_\parallel^2 + k_z^2\right) + h_3 - \eta i h_2 k_z = 0 \tag{B.7}$$

with $\eta = \pm$ the sign inherited from $\psi_\pm$. Solving for k_z, we find the two solutions

$$ik_{z,\frac{1}{2}} = \frac{1}{2h_4}\left[-\eta h_2 \pm \sqrt{4h_4\left(h_3 + h_4 k_\parallel^2\right) + h_2^2}\right]. \tag{B.8}$$

In order to obtain a wave function, exponentially decaying of the form of Eq. (B.3), both $ik_{z,\frac{1}{2}}$ need a real part of the same sign. For real h_j, this gives us the existence condition

$$h_4 \left(h_3 + h_4 k_\parallel^2 \right) < 0. \tag{B.9}$$

Depending on the sign of h_2/h_4 and the direction in which the wave function should decay, $z \to +\infty$ or $z \to -\infty$, one chooses the corresponding eigenstate $\psi_\pm$, fixing

$$\eta = -\mathrm{sgn}\left(\frac{h_2}{h_4}\right), \text{ top;} \quad \eta = \mathrm{sgn}\left(\frac{h_2}{h_4}\right), \text{ bottom.} \tag{B.10}$$

The surface dispersion relations and wave functions are then given by

$$E^{sur} = h_0 + \eta h_1(k_\pm), \quad \Psi(z) = \left(\mathrm{e}^{ik_{z,1}z} - \mathrm{e}^{ik_{z,2}z} \right) \psi_\eta. \tag{B.11}$$

The localization length is $l_c = \max\left\{ \left| 1/\Re\left(ik_{z,\frac{1}{2}} \right) \right| \right\}$.

The surface solution described in this section fulfills the eigenvalue Eq. (B.1) and is thus a valid, non-perturbative eigenstate of the Hamiltonian. Calculating the surface state with the general ansatz (B.2), this gives the same dispersion relation as for the simplified ansatz (B.3) for the TI model in Sec. 9.1.1.

C. Numerical Validation of the Approximate Solution Method

In Sec. 9 we use an analytical method to calculate the localized boundary states, described in App. B. It requires certain restrictions on the parameters of the coupled TI-WSM Hamiltonian, such as the same localization length for both TI and WSM phase and half of the symmetry allowed couplings to be zero, see Sec. 9.2.

These constraints might seem quite restrictive. In order to proof the general applicability of our results, we have checked numerically that the neglected couplings have no qualitative effect on the surface band structure if kept reasonably small. The same is true for variations that alter the localization lengths of the subsystems. The numerical method is similar to the analytical approach: We solve for exponentially localized surface wave functions on the half space $z \leq 0$ or $z \geq 0$ with hardwall boundary conditions at $z = 0$. The difference to the approximate solution is the use of the full ansatz for the wave function, i.e.

$$\Psi_g(z) = \sum_{j\in\{1,\dots,6\}} a_j \mathrm{e}^{ik_{z,j}z}\psi(k_\pm, k_{z,j}). \tag{C.1}$$

As an example, we take the case of the generation of the second Dirac point, discussed in Sec. 9.3.2 and depicted in Fig. 9.3. Besides the finite $a_0 = \frac{1}{4}$, we add an additional coupling $d_0 = \frac{1}{8}$ or change the localization length of the Weyl Hamiltonian by setting $v_z = \frac{3}{4} \neq \mathcal{B} = 1$ and $t = \frac{5}{4} \neq M_1 = 1$. The latter choice leads then to differing localization lengths of the separate systems of

$$ik_{z,TI} = \frac{1}{2}\left[\frac{\mathcal{B}}{M_1} \pm \sqrt{4\left(k_\parallel^2 - 1\right) + \left(\frac{\mathcal{B}}{M_1}\right)^2}\right], \tag{C.2}$$

$$ik_{z,WSM} = \frac{1}{2}\left[\frac{v_z}{t} \pm \sqrt{4\left(k_\parallel^2 - 1\right) + \left(\frac{v_z}{t}\right)^2}\right]. \tag{C.3}$$

The resultant dispersion relations are shown in Fig. C.1, depicted by blue dots. The analytical solution for the $a_0 = \frac{1}{4}$ coupling alone is displayed as a continuous surface.

First we notice that the numerical and analytical solution agree very well and show no qualitative difference. The added coupling $d_0 = \frac{1}{8}$ in Fig. C.1 (a) has almost no effect, the same was found for a finite $c_1 = \frac{1}{8}$. The changed localization length in Fig. C.1 (b) shifts a bit the lower Dirac point, but does not open a gap. The major difference between the analytical and numerical solutions is the restriction of the surface solution to energies and momenta where no bulk state exists. This becomes especially clear for the upper and lower parts of the Dirac cone in Fig. C.1, and is due to hybridization between the bulk and surface states. It prevents the existence of purely exponentially localized surface wave functions in this parameter range.

We conclude that the physical results and conclusions of Sec. 9 are valid beyond the restrictions on allowed couplings and localization lengths which are necessary to keep the analytical form of the equations simple.

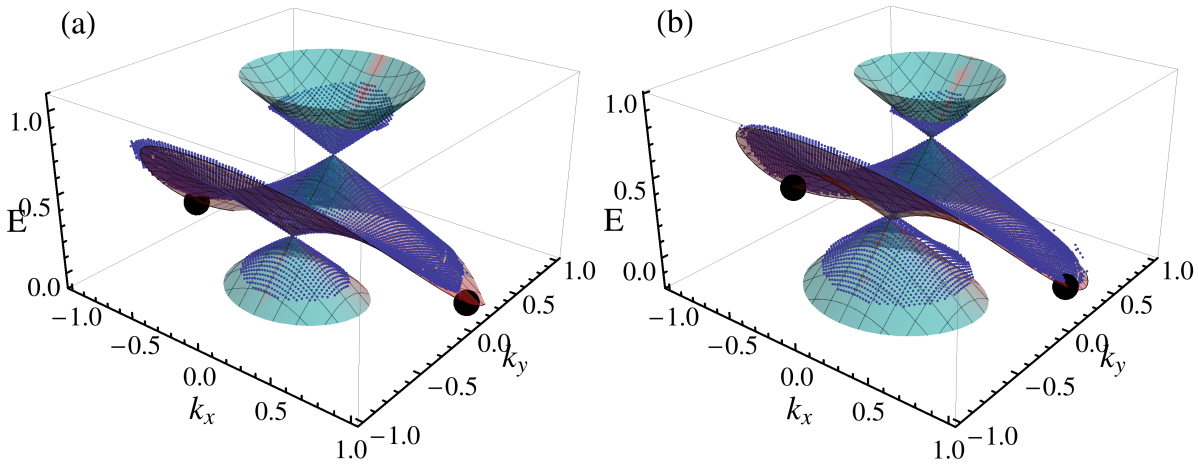


Figure C.1. Dispersion of the upper surface of the combined TI and WSM models. The continuous surface is the analytical solution from Fig. 9.3, the blue dots represent the full numerical solution. In (a) the additional coupling $d_0 = \frac{1}{8}$ was considered in the numerical solution, in (b) the altered parameters $v_z = \frac{3}{4} \neq \mathcal{B} = 1$ and $t = \frac{5}{4} \neq M_1 = 1$. The large black dots denote the position of the Weyl notes of the analytical solution. Reprinted figures with permission from Ref. [JT17]. Copyright (2017) by the APS.

Acknowledgments

The work presented in this thesis benefited considerably from the input of countless people with whom I have worked and discussed over the last couple of years. Foremost, I would like to thank Prof. Dr. Björn Trauzettel for his relaxed way of supervision and mentoring. His policy of open doors gave me the chance to grow, personally and professionally, while having an open ear to turn to whenever needed. I am especially grateful for the possibility of attending and giving talks at international schools and conferences.

For dealing with me on a day to day basis and teaching me all the tricks of their trade, I thank my two postdoc supervisors, Dr. Paolo Michetti and Dr. Maxim Kharitonov. I learned a lot from you, the easy or sometimes the hard way. Thanks also for the cactus, it has survived all my accidental attempts to murder it so far ...

Then there is all the rest of TP4, who made my stay in Würzburg such a joyful time. I especially thank Jan Böttcher for all the short physics discussions, Florian Geissler for being my co-admin and fellow PhD student, Moritz Fuchs for sharing his passion about martial arts, Dietrich Rothe for sharing all his admin and Debian knowledge with us, and of course Rolf Reinthaler for being Rolf, a guy with an incredible sense of humor, taking care of most of the things related to our group. I also thank my two Bachelor students, Ferdinand Schulz and Vanessa Werner, for their dedication, proving that very different characters can be successful in theoretical physics. And of course I thank Nelly Meyer for taking care of most of the administrative stuff in a delightful manner, always reminding us to get back the money as soon as possible.

Outside of our little TP4 group, I thank Tobias Stauber for inviting me to Madrid, and of course acknowledge financial support, coming from the DFG (SPP1666 and the DFG-JST research unit *Topotronics*) as well as from the Helmholtz Foundation (VITI) and the ENB Graduate School on "Topological Insulators".

All the time spent on research would not have been as productive and enjoyable without a good compensation outside of the university. I thus thank Franz Scheiner, Robert Sparks, Harald Lutz and Bernd Altenhöfer for teaching me their individual martial art and thereby changing my life in a profound manner.

Last but not least, I thank my family and Carina for the moral support and patience, asking only sporadicly about the remaining time till handing in my dissertation.

Bibliography

[ABC+15] A. Amaricci, J. C. Budich, M. Capone, B. Trauzettel, and G. Sangiovanni. First-Order Character and Observable Signatures of Topological Quantum Phase Transitions. *Physical Review Letters*, 114(18):185701, may 2015. doi:10.1103/PhysRevLett.114.185701, 1411.7390. [p. 112]

[AKL99] D. P. Arovas, A. Karlhede, and D. Lilliehöök. SU(N) quantum Hall skyrmions. *Physical Review B*, 59(20):13147, may 1999. doi:10.1103/PhysRevB.59.13147, cond-mat/9811097. [p. 49]

[ALL06] D. A. Abanin, P. A. Lee, and L. S. Levitov. Spin-Filtered Edge States and Quantum Hall Effect in Graphene. *Physical Review Letters*, 96(17):176803, may 2006. doi:10.1103/PhysRevLett.96.176803, cond-mat/0602645. [p. 113]

[And72] P. W. Anderson. More Is Different. *Science*, 177(4047):393, aug 1972. doi:10.1126/science.177.4047.393. [p. 1]

[And06] T. Ando. Screening Effect and Impurity Scattering in Monolayer Graphene. *Journal of the Physical Society of Japan*, 75(7):074716, jul 2006. doi:10.1143/JPSJ.75.074716. [pp. 62, 91]

[And13] Y. Ando. Topological Insulator Materials. *Journal of the Physical Society of Japan*, 82(10):102001, oct 2013. doi:10.7566/JPSJ.82.102001, 1304.5693. [pp. 23, 32]

[BDRT12] J. C. Budich, F. Dolcini, P. Recher, and B. Trauzettel. Phonon-Induced Backscattering in Helical Edge States. *Physical Review Letters*, 108(8):086602, feb 2012. doi:10.1103/PhysRevLett.108.086602, 1109.5188. [pp. 112, 146, 151]

[BEG+15] S. Borisenko, D. Evtushinsky, Q. Gibson, A. Yaresko, T. Kim, M. N. Ali, B. Buechner, M. Hoesch, and R. J. Cava. Time-Reversal Symmetry Breaking Type-II Weyl State in YbMnBi2. *arXiv*, 1507.04847, jul 2015. 1507.04847. [p. 154]

[BF07] L. Brey and H. Fertig. Elementary electronic excitations in graphene nanoribbons. *Physical Review B*, 75(12):125434, mar 2007. doi:10.1103/PhysRevB.75.125434, cond-mat/0701787. [p. 62]

[BH13] B. A. Bernevig and T. L. Hughes. *Topological Insulators and Topological Superconductors*. Princeton University Press, 41 William Street, Princeton, New Jersey 08540, 2013. [pp. 10, 13, 15, 18, 21, 23, 33]

[BHB11] A. A. Burkov, M. D. Hook, and L. Balents. Topological nodal semimetals. *Physical Review B*, 84(23):235126, 2011. doi:10.1103/PhysRevB.84.235126, 1110.1089. [pp. 154, 166]

[BHTD14] E. Barnes, E. H. Hwang, R. E. Throckmorton, and S. Das Sarma. Effective field theory, three-loop perturbative expansion, and their experimental implications in graphene many-body effects. *Physical Review B*, 89(23):235431, jun 2014. doi:10.1103/PhysRevB.89.235431, 1401.7011. [p. 63]

[BHZ06] B. A. Bernevig, T. L. Hughes, and S.-C. Zhang. Quantum spin Hall effect and topological phase transition in HgTe quantum wells. *Science*, 314(5806):1757, dec 2006. doi:10.1126/science.1133734, cond-mat/0611399. [pp. 24, 25, 27]

[BLQ14] D. Bulmash, C.-X. Liu, and X.-L. Qi. Prediction of a Weyl semimetal in Hg1-x-y Cd x Mn y. *Physical Review B*, 89(8):081106, feb 2014. doi:10.1103/PhysRevB.89.081106, 1309.6327. [pp. 154, 166]

[BLT+11] B. Büttner, C. X. Liu, G. Tkachov, E. G. E. G. Novik, C. Brüne, H. Buhmann, E. M. Hankiewicz, P. Recher, B. Trauzettel, S. C. Zhang, and L. W. Molenkamp. Single valley Dirac fermions in zero-gap HgTe quantum wells. *Nature Physics*, 7(5):418, may 2011. doi:10.1038/nphys1914, 1009.2248. [pp. 26, 68, 71]

[BMAVF10] S. M. Badalyan, A. Matos-Abiague, G. Vignale, and J. Fabian. Beating of Friedel oscillations induced by spin-orbit interaction. *Physical Review B*, 81(20):205314, may 2010. doi:10.1103/PhysRevB.81.205314, 0911.5632. [p. 94]

[BPBP+07] Y. Barlas, T. Pereg-Barnea, M. Polini, R. Asgari, and A. H. MacDonald. Chirality and Correlations in Graphene. *Physical Review Letters*, 98(23):236601, jun 2007. doi:10.1103/PhysRevLett.98.236601, cond-mat/0701257. [p. 62]

[BPF+15] Y. Baum, T. Posske, I. C. Fulga, B. B. Trauzettel, and A. Stern. Coexisting Edge States and Gapless Bulk in Topological States of Matter. *Physical Review Letters*, 114(13):136801, mar 2015. doi:10.1103/PhysRevLett.114.136801, 1412.0021. [p. 158]

[BS17] Y. Baum and A. Stern. Nonlocal Coulomb drag in Weyl semimetals. *Physical Review B*, 95(7):075141, feb 2017. doi:10.1103/PhysRevB.95.075141, 1612.00018. [p. 59]

[BT13] J. C. Budich and B. Trauzettel. From the adiabatic theorem of quantum mechanics to topological states of matter. *physica status solidi (RRL) - Rapid Research Letters*, 7(1-2):109, feb 2013. doi:10.1002/pssr.201206416, 1210.6672. [p. 23]

[BTM14] J. C. Budich, B. Trauzettel, and P. Michetti. Time Reversal Symmetric Topological Exciton Condensate in Bilayer HgTe Quantum Wells. *Physical Review Letters*, 112(14):146405, apr 2014. doi:10.1103/PhysRevLett.112.146405, 1311.2043. [pp. 59, 60]

[Bud11] J. Budich. You cannot comb the hedgehog smooth: Differential topology in condensed matter physics. 2011. [pp. 8, 10, 21, 23]

[BYS+16] I. Belopolski, P. Yu, D. S. Sanchez, Y. Ishida, T.-R. Chang, S. S. Zhang, S.-Y. Xu, D. Mou, H. Zheng, G. Chang, G. Bian, H.-T. Jeng, T. Kondo, A. Kaminski, H. Lin, Z. Liu, S. Shin, and M. Z. Hasan. A minimal, "hydrogen atom" version of an inversion-breaking Weyl semimetal. *arXiv*, 1610.02013, oct 2016. 1610.02013. [pp. 154, 166]

[BZY+17] Z. Bi, R. Zhang, Y.-Z. You, A. Young, L. Balents, C.-x. Liu, and C. Xu. Bilayer Graphene as a Platform for Bosonic Symmetry-Protected Topological States. *Physical Review Letters*, 118(12):126801, mar 2017. doi:10.1103/PhysRevLett.118.126801, 1602.03190. [p. 123]

[CBD+12] F. Crépin, J. C. Budich, F. Dolcini, P. Recher, and B. Trauzettel. Renormalization group approach for the scattering off a single Rashba impurity in a helical liquid. *Physical Review B*, 86(12):121106, sep 2012. doi:10.1103/PhysRevB.86.121106, 1205.0374. [pp. 112, 146, 151]

[CGP+09] A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim. The electronic properties of graphene. *Reviews of Modern Physics*, 81(1):109, jan 2009. doi:10.1103/RevModPhys.81.109, 0709.1163. [p. 62]

[Cho11] G. Y. Cho. Possible topological phases of bulk magnetically doped Bi2Se3: turning a topological band insulator into the Weyl semimetal. *arXiv*, 1110.1939, oct 2011. 1110.1939. [pp. 154, 166]

[CTSR16] C.-k. Chiu, J. C. Y. Teo, A. P. Schnyder, and S. Ryu. Classification of topological quantum matter with symmetries. *Reviews of Modern Physics*, 88(3):035005, aug 2016. doi:10.1103/RevModPhys.88.035005, 1505.03535. [pp. 20, 32]

[CTV12] V. Cvetkovic, R. E. Throckmorton, and O. Vafek. Electronic multicriticality in bilayer graphene. *Physical Review B*, 86(7):075467, aug 2012. doi:10.1103/PhysRevB.86.075467, 1206.0288. [p. 112]

[CWS12] J.-c. Chen, J. Wang, and Q.-f. Sun. Effect of magnetic field on electron transport in HgTe/CdTe quantum wells: Numerical analysis. *Physical Review B*, 85(12):125401, mar 2012. doi:10.1103/PhysRevB.85.125401. [p. 62]

[CXS+16] G. Chang, S.-Y. Xu, D. S. Sanchez, S.-M. Huang, C.-C. Lee, T.-R. Chang, G. Bian, H. Zheng, I. Belopolski, N. Alidoust, H.-T. Jeng, A. Bansil, H. Lin, and M. Z. Hasan. A strongly robust type II Weyl fermion semimetal state in Ta3S2. *Science Advances*, 2(6):e1600295, jun 2016. doi:10.1126/sciadv.1600295, 1512.08781. [p. 154]

[DL13] S. Das Sarma and Q. Li. Intrinsic plasmons in two-dimensional Dirac materials. *Physical Review B*, 87(23):235418, jun 2013. doi:10.1103/PhysRevB.87.235418, 1305.0825. [pp. 62, 84, 85, 87]

[DOL+13] P. Di Pietro, M. Ortolani, O. Limaj, A. Di Gaspare, V. Giliberti, F. Giorgianni, M. Brahlek, N. Bansal, N. Koirala, S. Oh, P. Calvani, and S. Lupi. Observation of Dirac plasmons in a topological insulator. *Nature Nanotechnology*, 8(8):556, jul 2013. doi:10.1038/nnano.2013.134, 1307.5974. [pp. 58, 62, 71, 81]

[DR16] V. Dwivedi and S. T. Ramamurthy. Connecting the dots: Time-reversal symmetric Weyl semimetals with tunable Fermi arcs. *Physical Review B*, 94(24):245143, dec 2016. doi:10.1103/PhysRevB.94.245143, 1608.01313. [p. 37]

[DSGC10] M. Dzero, K. Sun, V. Galitski, and P. Coleman. Topological Kondo Insulators. *Physical Review Letters*, 104(10):106408, mar 2010. doi:10.1103/PhysRevLett.104.106408, 0912.3750. [p. 112]

[DWD+16] K. Deng, G. Wan, P. Deng, K. Zhang, S. Ding, E. Wang, M. Yan, H. Huang, H. Zhang, Z. Xu, J. Denlinger, A. Fedorov, H. Yang, W. Duan, H. Yao, Y. Wu, S. Fan, H. Zhang, X. Chen, and S. Zhou. Experimental observation of topological Fermi arcs in type-II Weyl semimetal MoTe2. *Nature Physics*, 12(12):1105, sep 2016. doi:10.1038/nphys3871, 1603.08508. [p. 154]

[DXGC16] M. Dzero, J. Xia, V. Galitski, and P. Coleman. Topological Kondo Insulators. *Annual Review of Condensed Matter Physics*, 7(1):249, mar 2016. doi:10.1146/annurev-conmatphys-031214-014749, 1506.05635. [p. 112]

[Eis13] J. P. Eisenstein. Exciton Condensation in Bilayer Quantum Hall Systems. *arXiv*, 1306.0584, jun 2013. 1306.0584. [p. 59]

[EM04] J. P. Eisenstein and A. H. MacDonald. Bose-Einstein condensation of excitons in bilayer electron systems. *Nature*, 432(7018):691, dec 2004. doi:10.1038/nature03081. [p. 59]

[Eul35] L. Euler. Solutio problematis ad geometriam situs pertinentis. *Commentarii academiae scientiarum Petropolitanae 8, 1741, pp. 128-140, see also: http://eulerarchive.maa.org//pages/E053.html, Newman's "World of Mathematics" and Biggs, Lloyd & Wilson's "Graph Theory 1736-1936"*, 1735. [p. 7]

[FB97] M. Franco and L. Brey. Phase diagram of a quantum Hall ferromagnet edge, spin-textured edges, and collective excitations. *Physical Review B*, 56(16):10383, oct 1997. doi:10.1103/PhysRevB.56.10383. [p. 123]

[FB06] H. Fertig and L. Brey. Luttinger Liquid at the Edge of Undoped Graphene in a Strong Magnetic Field. *Physical Review Letters*, 97(11):116805, sep 2006. doi:10.1103/PhysRevLett.97.116805, cond-mat/0604260. [pp. 123, 132, 133, 167]

[FGB12] C. Fang, M. J. Gilbert, and B. A. Bernevig. Bulk topological invariants in noninteracting point group symmetric insulators. *Physical Review B*, 86(11):115112, sep 2012. doi:10.1103/PhysRevB.86.115112, 1207.5767. [p. 32]

[FHJK95] K. Flensberg, B. Y.-K. Hu, A.-P. Jauho, and J. M. Kinaret. Linear-response theory of Coulomb drag in coupled electron systems. *Physical Review B*, 52(20):14761, nov 1995. doi:10.1103/PhysRevB.52.14761. [p. 59]

[FI99] V. Fal'ko and S. Iordanskii. Topological Defects and Goldstone Excitations in Domain Walls between Ferromagnetic Quantum Hall Liquids. *Physical Review Letters*, 82(2):402, 1999. doi:10.1103/PhysRevLett.82.402, cond-mat/9901053. [pp. 123, 132, 133]

[FK07] L. Fu and C. L. Kane. Topological insulators with inversion symmetry. *Physical Review B*, 76(4):045302, 2007. doi:10.1103/PhysRevB.76.045302, cond-mat/0611341. [p. 17]

[FM13] M. Franz and L. Molenkamp, editors. *Topological Insulators*, volume 6 of *Contemporary Concepts of Condensed Matter Science*. Elsevier, 2013. [pp. 6, 10, 12, 13, 15, 18, 21, 23, 28, 29, 32, 33, 39]

[FRA+12] Z. Fei, a. S. Rodin, G. O. Andreev, W. Bao, a. S. McLeod, M. Wagner, L. M. Zhang, Z. Zhao, M. Thiemens, G. Dominguez, M. M. Fogler, A. H. C. Neto, C. N. Lau, F. Keilmann, and D. N. Basov. Gate-tuning of graphene plasmons revealed by infrared nano-imaging. *Nature*, 487(7405):82, jun 2012. doi:10.1038/nature11253, 1202.4993. [p. 62]

[FW71] A. L. Fetter and J. D. Walecka. *Quantum Theory of Many-particle Systems*. McGraw-Hill Book Company, 1971. [pp. 44, 70]

[GC11] P. Goswami and S. Chakravarty. Quantum Criticality between Topological and Band Insulators in 3+1 Dimensions. *Physical Review Letters*, 107(19):196803, 2011. doi:10.1103/PhysRevLett.107.196803, 1101.2210. [p. 42]

[GCT14] F. Geissler, F. Crépin, and B. Trauzettel. Random Rashba spin-orbit coupling at the quantum spin Hall edge. *Physical Review B*, 89(23):235136, jun 2014. doi:10.1103/PhysRevB.89.235136, 1403.1082. [pp. 112, 146, 151]

[GFM08] S. Gangadharaiah, A. Farid, and E. Mishchenko. Charge Response Function and a Novel Plasmon Mode in Graphene. *Physical Review Letters*, 100(16):166802, apr 2008. doi:10.1103/PhysRevLett.100.166802, 0710.0622. [pp. 62, 82, 85]

[GGK+12] R. V. Gorbachev, a. K. Geim, M. I. Katsnelson, K. S. Novoselov, T. Tudorovskiy, I. V. Grigorieva, a. H. MacDonald, S. V. Morozov, K. Watanabe, T. Taniguchi, and L. a. Ponomarenko. Strong Coulomb drag and broken symmetry in double-layer graphene. *Nature Physics*, 8(12):896, oct 2012. doi:10.1038/nphys2441, 1206.6626. [p. 59]

[GGMS02] E. Gorbar, V. Gusynin, V. Miransky, and I. Shovkovy. Magnetic field driven metal-insulator phase transition in planar systems. *Physical Review B*, 66(4):045108, jul 2002. doi:10.1103/PhysRevB.66.045108, cond-mat/0202422. [p. 91]

[Gia03] T. Giamarchi. *Quantum Physics in One Dimension*. Oxford University Press, 2003. [pp. 43, 52, 54, 124, 147, 150]

[Gir99] S. M. Girvin. The Quantum Hall Effect: Novel Excitations And Broken Symmetries. In *Aspects topologiques de la physique en basse dimension. Topological aspects of low dimensional systems*, number July 1998, pages 53–175. Springer Berlin Heidelberg, Berlin, Heidelberg, jul 1999. cond-mat/9907002. [p. 49]

[GN07] A. K. Geim and K. S. Novoselov. The rise of graphene. *Nature Materials*, 6(3):183, mar 2007. doi:10.1038/nmat1849, cond-mat/0702595. [p. 62]

[GPN12] A. N. Grigorenko, M. Polini, and K. S. Novoselov. Graphene plasmonics. *Nature Photonics*, 6(11):749, nov 2012. doi:10.1038/nphoton.2012.262, 1301.4241. [pp. 57, 58, 68]

[GV05] G. Giuliani and G. Vignale. *Quantum Theory of the Electron Liquid*. Cambridge University Press, 2005. [pp. 44, 91, 95, 171]

[GVB15] A. G. Grushin, J. W. F. Venderbos, and J. H. Bardarson. Coexistence of Fermi arcs with two-dimensional gapless Dirac states. *Physical Review B*, 91(12):121109, mar 2015. doi:10.1103/PhysRevB.91.121109, 1501.00905. [p. 154]

[HA13] M. Hohenadler and F. F. Assaad. Correlation effects in two-dimensional topological insulators. *Journal of Physics: Condensed Matter*, 25(14):143201, apr 2013. doi:10.1088/0953-8984/25/14/143201, 1211.1774. [p. 112]

[HBD14] J. Hofmann, E. Barnes, and S. Das Sarma. Why Does Graphene Behave as a Weakly Interacting System? *Physical Review Letters*, 113(10):105502, sep 2014. doi:10.1103/PhysRevLett.113.105502, 1405.7036. [p. 63]

[HD07] E. H. Hwang and S. Das Sarma. Dielectric function, screening, and plasmons in two-dimensional graphene. *Physical Review B*, 75(20):205418, may 2007. doi:10.1103/PhysRevB.75.205418, cond-mat/0610561. [pp. 62, 84, 172]

[Her07] I. F. Herbut. SO(3) symmetry between Néel and ferromagnetic order parameters for graphene in a magnetic field. *Physical Review B*, 76(8):085432, aug 2007. doi:10.1103/PhysRevB.76.085432, 0705.4039. [pp. 123, 126]

[HK10] M. Z. Hasan and C. L. Kane. Colloquium: Topological insulators. *Reviews of Modern Physics*, 82(4):3045, nov 2010. doi:10.1103/RevModPhys.82.3045, 1002.3895. [p. 23]

[HKLN99] T. H. Hansson, A. Karlhede, J. M. Leinaas, and U. Nilsson. Field theory for partially polarized quantum Hall states. *Physical Review B*, 60(7):4866, aug 1999. doi:10.1103/PhysRevB.60.4866, cond-mat/981028. [p. 123]

[HLL+12] T. H. Hsieh, H. Lin, J. Liu, W. Duan, A. Bansi, and L. Fu. Topological crystalline insulators in the SnTe material class. *Nature Communications*, 3:982, jul 2012. doi:10.1038/ncomms1969, 1202.1003. [p. 32]

[HMO+16] L. Huang, T. M. McCormick, M. Ochi, Z. Zhao, M.-t. Suzuki, R. Arita, Y. Wu, D. Mou, H. Cao, J. Yan, N. Trivedi, and A. Kaminski. Spectroscopic evidence for a type II Weyl semimetallic state in MoTe2. *Nature Materials*, 15(11):1155, jul 2016. doi:10.1038/nmat4685, 1603.06482. [p. 154]

[HPV12] P. Hosur, S. a. Parameswaran, and A. Vishwanath. Charge Transport in Weyl Semimetals. *Physical Review Letters*, 108(4):046602, jan 2012. doi:10.1103/PhysRevLett.108.046602, 1109.6330. [p. 42]

[HQ13] P. Hosur and X. Qi. Recent developments in transport phenomena in Weyl semimetals. *Comptes Rendus Physique*, 14(9-10):857, nov 2013. doi:10.1016/j.crhy.2013.10.010, 1309.4464. [p. 33]

[HQW+08] D. Hsieh, D. Qian, L. Wray, Y. Xia, Y. S. Hor, R. J. Cava, and M. Z. Hasan. A topological Dirac insulator in a quantum spin Hall phase. *Nature*, 452(7190):970, apr 2008. doi:10.1038/nature06843, 0910.2420. [p. 29]

[HXB+15] S.-M. Huang, S.-Y. Xu, I. Belopolski, C.-C. Lee, G. Chang, B. Wang, N. Alidoust, G. Bian, M. Neupane, C. Zhang, S. Jia, A. Bansil, H. Lin, and M. Z.

Hasan. A Weyl Fermion semimetal with surface Fermi arcs in the transition metal monopnictide TaAs class. *Nature Communications*, 6:7373, jun 2015. doi:10.1038/ncomms8373, 1501.00755. [p. 154]

[IN12] H. Isobe and N. Nagaosa. Theory of a quantum critical phenomenon in a topological insulator: (3+1)-dimensional quantum electrodynamics in solids. *Physical Review B*, 86(16):165127, oct 2012. doi:10.1103/PhysRevB.86.165127, 1205.2427. [p. 42]

[IN13] H. Isobe and N. Nagaosa. Renormalization group study of electromagnetic interaction in multi-Dirac-node systems. *Physical Review B*, 87(20):205138, may 2013. doi:10.1103/PhysRevB.87.205138, 1303.2822. [p. 42]

[JGH+11] L. Ju, B. Geng, J. Horng, C. Girit, M. Martin, Z. Hao, H. a. Bechtel, X. Liang, A. Zettl, Y. R. Shen, and F. Wang. Graphene plasmonics for tunable terahertz metamaterials. *Nature Nanotechnology*, 6(10):630, sep 2011. doi:10.1038/nnano.2011.146. [p. 62]

[JJL+16] Q.-D. Jiang, H. Jiang, H. Liu, Q.-F. Sun, and X. C. Xie. Chiral wave-packet scattering in Weyl semimetals. *Physical Review B*, 93(19):195165, may 2016. doi:10.1103/PhysRevB.93.195165, 1601.07297. [p. 154]

[JMT14a] S. Juergens, P. Michetti, and B. Trauzettel. Plasmons due to the Interplay of Dirac and Schrödinger Fermions. *Physical Review Letters*, 112(7):076804, feb 2014. doi:10.1103/PhysRevLett.112.076804, 1309.6083. [pp. 2, 63, 66, 83, 84, 88]

[JMT14b] S. Juergens, P. Michetti, and B. Trauzettel. Screening properties and plasmons of Hg(Cd)Te quantum wells. *Physical Review B*, 90(11):115425, sep 2014. doi:10.1103/PhysRevB.90.115425, 1406.7754. [pp. 2, 48, 63, 74, 78, 79, 80, 86, 89, 92, 93, 96, 97, 98, 99, 100, 101, 102, 103, 105, 106, 107, 108, 171, 172]

[JS93] A.-P. Jauho and H. Smith. Coulomb drag between parallel two-dimensional electron systems. *Physical Review B*, 47(8):4420, feb 1993. doi:10.1103/PhysRevB.47.4420, cond-mat/9205001. [p. 59]

[JT17] S. Juergens and B. Trauzettel. Exotic surface states in hybrid structures of topological insulators and Weyl semimetals. *Physical Review B*, 95(8):085313, feb 2017. doi:10.1103/PhysRevB.95.085313, 1612.02176. [pp. 2, 34, 155, 157, 161, 162, 163, 164, 165, 176]

[JXH16] S. Jia, S.-Y. Xu, and M. Z. Hasan. Weyl semimetals, Fermi arcs and chiral anomalies. *Nature Materials*, 15(11):1140–1144, oct 2016. doi:10.1038/nmat4787, 1612.00416. [p. 33]

[Kat06] M. I. Katsnelson. Zitterbewegung, chirality, and minimal conductivity in graphene. *The European Physical Journal B*, 51(2):157, may 2006. doi:10.1140/epjb/e2006-00203-1, cond-mat/0512337. [p. 166]

[KBW+08] M. König, H. Buhmann, L. W. Molenkamp, T. Hughes, C.-X. Liu, X.-L. Qi, and S.-C. Zhang. The Quantum Spin Hall Effect: Theory and Experiment. *Journal of the Physical Society of Japan*, 77(3):031007, mar 2008. doi:10.1143/JPSJ.77.031007, 0801.0901. [pp. 25, 26]

[KCG11] F. H. L. Koppens, D. E. Chang, and F. J. GarciÌĄa de Abajo. Graphene Plasmonics: A Platform for Strong Light-Matter Interactions. *Nano Letters*, 11(8):3370, aug 2011. doi:10.1021/nl201771h, 1104.2068. [pp. 57, 58]

[KDP80] K. V. Klitzing, G. Dorda, and M. Pepper. New Method for High-Accuracy Determination of the Fine-Structure Constant Based on Quantized Hall Resistance. *Physical Review Letters*, 45(6):494, aug 1980. doi:10.1103/PhysRevLett.45.494, 1011.1669. [p. 21]

[KDS11] I. Knez, R.-R. Du, and G. Sullivan. Evidence for Helical Edge Modes in Inverted InAs/GaSb Quantum Wells. *Physical Review Letters*, 107(13):136603, sep 2011. doi:10.1103/PhysRevLett.107.136603, 1105.0137. [p. 26]

[KE08] M. Kharitonov and K. Efetov. Electron screening and excitonic condensation in double-layer graphene systems. *Physical Review B*, 78(24):241401, dec 2008. doi:10.1103/PhysRevB.78.241401, 0808.2164. [p. 60]

[KGCM14] N. Kainaris, I. V. Gornyi, S. T. Carr, and A. D. Mirlin. Conductivity of a generic helical liquid. *Physical Review B*, 90(7):075118, aug 2014. doi:10.1103/PhysRevB.90.075118, 1404.3129. [pp. 112, 146, 151]

[Kha12a] M. Kharitonov. Canted Antiferromagnetic Phase of the ν=0 Quantum Hall State in Bilayer Graphene. *Physical Review Letters*, 109(4):046803, jul 2012. doi:10.1103/PhysRevLett.109.046803, 1105.5386. [pp. 123, 126, 136]

[Kha12b] M. Kharitonov. Edge excitations of the canted antiferromagnetic phase of the ν=0 quantum Hall state in graphene: A simplified analysis. *Physical Review B*, 86(7):075450, aug 2012. doi:10.1103/PhysRevB.86.075450, 1206.0724. [pp. 123, 136]

[Kha12c] M. Kharitonov. Phase diagram for the ν=0 quantum Hall state in monolayer graphene. *Physical Review B*, 85(15):155439, apr 2012. doi:10.1103/PhysRevB.85.155439, 1103.6285. [pp. 123, 126]

[KJT16] M. Kharitonov, S. Juergens, and B. Trauzettel. Interplay of topology and interactions in quantum Hall topological insulators: U(1) symmetry, tunable Luttinger liquid, and interaction-induced phase transitions. *Physical Review B*, 94(3):035146, jul 2016. doi:10.1103/PhysRevB.94.035146, 1603.07048. [pp. 2, 112, 114, 115, 116, 125, 128, 129, 131, 134, 135, 139, 142, 145, 147, 148]

[KKE+16] K. Koepernik, D. Kasinathan, D. V. Efremov, S. Khim, S. Borisenko, B. Büchner, and J. van den Brink. TaIrTe: A ternary type-II Weyl semimetal. *Physical Review B*, 93(20):201101, may 2016. doi:10.1103/PhysRevB.93.201101, 1603.04323. [pp. 154, 166]

[KKW+13] H.-J. Kim, K.-S. Kim, J.-F. Wang, M. Sasaki, N. Satoh, A. Ohnishi, M. Kitaura, M. Yang, and L. Li. Dirac versus Weyl Fermions in Topological Insulators: Adler-Bell-Jackiw Anomaly in Transport Phenomena. *Physical Review Letters*, 111(24):246603, dec 2013. doi:10.1103/PhysRevLett.111.246603, 1307.6990. [p. 36]

[KLS99] A. Karlhede, K. Lejnell, and S. L. Sondhi. Dynamics of the compact, ferromagnetic ν=1 edge. *Physical Review B*, 60(23):15948, dec 1999. doi:10.1103/PhysRevB.60.15948, cond-mat/9907175. [p. 123]

[KM05a] C. L. Kane and E. J. Mele. Quantum Spin Hall Effect in Graphene. *Physical Review Letters*, 95(22):226801, nov 2005. doi:10.1103/PhysRevLett.95.226801, cond-mat/0411737. [p. 23]

[KM05b] C. L. Kane and E. J. Mele. Z_{2} Topological Order and the Quantum Spin Hall Effect. *Physical Review Letters*, 95(14):146802, sep 2005. doi:10.1103/PhysRevLett.95.146802, cond-mat/0506581. [p. 23]

[Koh85] M. Kohmoto. Topological invariant and the quantization of the Hall conductance. *Annals of Physics*, 160(2):343, apr 1985. doi:10.1016/0003-4916(85)90148-4. [p. 21]

[KPU08] V. Kotov, V. Pereira, and B. Uchoa. Polarization charge distribution in gapped graphene: Perturbation theory and exact diagonalization analysis. *Physical Review B*, 78(7):075433, aug 2008. doi:10.1103/PhysRevB.78.075433, 0806.1228. [pp. 77, 79]

[KUP+12] V. N. Kotov, B. Uchoa, V. M. Pereira, F. Guinea, and a. H. Castro Neto. Electron-Electron Interactions in Graphene: Current Status and Perspectives. *Reviews of Modern Physics*, 84(3):1067, jul 2012. doi:10.1103/RevModPhys.84.1067, 1012.3484. [pp. 62, 63, 68]

[KWB+07] M. Konig, S. Wiedmann, C. Brune, A. Roth, H. Buhmann, L. W. Molenkamp, X.-L. Qi, and S.-C. Zhang. Quantum Spin Hall Insulator State in HgTe Quantum Wells. *Science*, 318(5851):766, nov 2007. doi:10.1126/science.1148047, 0710.0582. [pp. 2, 24, 26]

[LA14] T. Low and P. Avouris. Graphene Plasmonics for Terahertz to Mid-Infrared Applications. *ACS Nano*, 8(2):1086, feb 2014. doi:10.1021/nn406627u, 1403.2799. [p. 63]

[LAF12] Y. Lemonik, I. Aleiner, and V. I. Fal'ko. Competing nematic, antiferromagnetic, and spin-flux orders in the ground state of bilayer graphene. *Physical Review B*, 85(24):245451, jun 2012. doi:10.1103/PhysRevB.85.245451, 1203.4608. [p. 112]

[LDJ92] Q. P. Li, S. Das Sarma, and R. Joynt. Elementary excitations in one-dimensional quantum wires: Exact equivalence between the random-phase approximation and the Tomonaga-Luttinger model. *Physical Review B*, 45(23):13713, jun 1992. doi:10.1103/PhysRevB.45.13713. [p. 167]

[Lig58] M. J. Lighthill. *Introduction to Fourier Analysis and Generalized Functions*. Cambridge University Press, 1958. [p. 93]

[LKS99] K. Lejnell, A. Karlhede, and S. L. Sondhi. Effective-action studies of quantum Hall spin textures. *Physical Review B*, 59(15):10183, apr 1999. doi:10.1103/PhysRevB.59.10183, cond-mat/9809267. [p. 123]

[LL14] Y.-M. Lu and D.-H. Lee. Inversion symmetry protected topological insulators and superconductors. *arXiv*, 1403.5558, mar 2014. 1403.5558. [p. 32]

[LLC16] H. Liu, W. E. Liu, and D. Culcer. Coulomb drag in topological insulator films. *Physica E: Low-dimensional Systems and Nanostructures*, 79:72, may 2016. doi:10.1016/j.physe.2015.11.027, 1601.02291. [p. 59]

[LMQ+15] B. Q. Lv, S. Muff, T. Qian, Z. D. Song, S. M. Nie, N. Xu, P. Richard, C. E. Matt, N. C. Plumb, L. X. Zhao, G. F. Chen, Z. Fang, X. Dai, J. H. Dil, J. Mesot, M. Shi, H. M. Weng, and H. Ding. Observation of Fermi-Arc Spin Texture in TaAs. *Physical Review Letters*, 115(21):217601, nov 2015. doi:10.1103/PhysRevLett.115.217601, 1510.07256. [p. 166]

[LOB12] N. Lezmy, Y. Oreg, and M. Berkooz. Single and multiparticle scattering in helical liquid with an impurity. *Physical Review B*, 85(23):235304, jun 2012. doi:10.1103/PhysRevB.85.235304, 1201.6197. [pp. 112, 146, 151]

[LQZ+10] C.-X. Liu, X.-L. Qi, H. Zhang, X. Dai, Z. Fang, and S.-C. Zhang. Model Hamiltonian for topological insulators. *Physical Review B*, 82(4):045122, 2010. doi:10.1103/PhysRevB.82.045122, 1005.1682. [pp. 29, 30, 31, 155, 173]

[LWF+15a] T. Li, P. Wang, H. Fu, L. Du, K. A. Schreiber, X. Mu, X. Liu, G. Sullivan, G. A. Csáthy, X. Lin, and R.-r. Du. Observation of a Helical Luttinger Liquid in InAs/GaSb Quantum Spin Hall Edges. *Physical Review Letters*, 115(13):136804, sep 2015. doi:10.1103/PhysRevLett.115.136804, 1507.08362. [p. 26]

[LWF+15b] B. Q. Lv, H. M. Weng, B. B. Fu, X. P. Wang, H. Miao, J. Ma, P. Richard, X. C. Huang, L. X. Zhao, G. F. Chen, Z. Fang, X. Dai, T. Qian, and H. Ding. Experimental Discovery of Weyl Semimetal TaAs. *Physical Review X*, 5(3):031013, jul 2015. doi:10.1103/PhysRevX.5.031013, 1502.04684. [p. 154]

[LWY+15] L. Lu, Z. Wang, D. Ye, L. Ran, L. Fu, J. D. Joannopoulos, and M. Solja i. Experimental observation of Weyl points. *Science*, 349(6248):622, aug 2015. doi:10.1126/science.aaa9273, 1502.03438. [p. 154]

[LXW+15] B. Q. Lv, N. Xu, H. M. Weng, J. Z. Ma, P. Richard, X. C. Huang, L. X. Zhao, G. F. Chen, C. E. Matt, F. Bisti, V. N. Strocov, J. Mesot, Z. Fang, X. Dai, T. Qian, M. Shi, and H. Ding. Observation of Weyl nodes in TaAs. *Nature Physics*, 11(9):724, aug 2015. doi:10.1038/nphys3426, 1503.09188. [p. 154]

[Mah00] G. D. Mahan. *Many-particle Physics*. Kluwer Academic / Plenum Publishers, New York, 3rd edition, 2000. [p. 44]

[MBSM08] H. Min, R. Bistritzer, J.-J. Su, and A. H. MacDonald. Room-temperature superfluidity in graphene bilayers. *Physical Review B*, 78(12):121401, sep 2008. doi:10.1103/PhysRevB.78.121401, 0802.3462. [p. 60]

[MCW+15] E. Y. Ma, M. R. Calvo, J. Wang, B. Lian, M. Mühlbauer, C. Brüne, Y.-T. Cui, K. Lai, W. Kundhikanjana, Y. Yang, M. Baenninger, M. König, C. Ames, H. Buhmann, P. Leubner, L. W. Molenkamp, S.-C. Zhang, D. Goldhaber-Gordon, M. A. Kelly, and Z.-X. Shen. Unexpected edge conduction in mercury telluride

quantum wells under broken time-reversal symmetry. *Nature Communications*, 6(May):7252, may 2015. doi:10.1038/ncomms8252, 1212.6441. [p. 158]

[MDY+13] P. Maher, C. R. Dean, a. F. Young, T. Taniguchi, K. Watanabe, K. L. Shepard, J. Hone, and P. Kim. Evidence for a spin phase transition at charge neutrality in bilayer graphene. *Nature Physics*, 9(3):154, 2013. doi:10.1038/nphys2528, 1212.3846. [p. 136]

[MK08] S. Murakami and S.-i. Kuga. Universal phase diagrams for the quantum spin Hall systems. *Physical Review B*, 78(16):165313, oct 2008. doi:10.1103/PhysRevB.78.165313, 0806.3309. [p. 34]

[MKT17] T. M. McCormick, I. Kimchi, and N. Trivedi. Minimal models for topological Weyl semimetals. *Physical Review B*, 95(7):075133, feb 2017. doi:10.1103/PhysRevB.95.075133, 1604.03096. [pp. 37, 155, 156]

[MMY+95] K. Moon, H. Mori, K. Yang, S. M. Girvin, A. H. MacDonald, L. Zheng, D. Yoshioka, and S.-C. Zhang. Spontaneous interlayer coherence in double-layer quantum Hall systems: Charged vortices and Kosterlitz-Thouless phase transitions. *Physical Review B*, 51(8):5138, feb 1995. doi:10.1103/PhysRevB.51.5138, cond-mat/9407031. [pp. 49, 50]

[MNH+16] P. J. W. Moll, N. L. Nair, T. Helm, A. C. Potter, I. Kimchi, A. Vishwanath, and J. G. Analytis. Transport evidence for Fermi-arc-mediated chirality transfer in the Dirac semimetal Cd3As2. *Nature*, 535(7611):266, jul 2016. doi:10.1038/nature18276, 1505.02817. [p. 37]

[MPER12] P. Michetti, P. H. Penteado, J. C. Egues, and P. Recher. Helical edge states in multiple topological mass domains. *Semiconductor Science and Technology*, 27(12):124007, dec 2012. doi:10.1088/0268-1242/27/12/124007, 1209.2313. [pp. 66, 159]

[MSF14] G. Murthy, E. Shimshoni, and H. a. Fertig. Collective edge modes near the onset of a graphene quantum spin Hall state. *Physical Review B*, 90(24):241410, dec 2014. doi:10.1103/PhysRevB.90.241410, 1408.7064. [pp. 123, 136, 138]

[MSF16] G. Murthy, E. Shimshoni, and H. A. Fertig. Collective bulk and edge modes through the quantum phase transition in graphene at ν=0. *Physical Review B*, 93(4):045105, jan 2016. doi:10.1103/PhysRevB.93.045105, 1510.04255. [p. 123]

[MT17] P. Michetti and C. Timm. Electron-hole pairing of Fermi-arc surface states in a Weyl semimetal bilayer. *Physical Review B*, 95(12):125435, mar 2017. doi:10.1103/PhysRevB.95.125435, 1612.08612. [p. 60]

[MTT+17] T. Müller, R. Thomale, B. Trauzettel, E. Bocquillon, and O. Kashuba. Dynamical transport measurement of the Luttinger parameter in helical edges states of 2D topological insulators. *Arxiv*, 1701.03050, jan 2017. 1701.03050. [p. 167]

[Mur07] S. Murakami. Phase transition between the quantum spin Hall and insulator phases in 3D: Emergence of a topological gapless phase. *New Journal of Physics*, 9:356, 2007. doi:10.1088/1367-2630/9/9/356, 0710.0930. [pp. 34, 35, 154]

[NL16] B. N. Narozhny and A. Levchenko. Coulomb drag. *Reviews of Modern Physics*, 88(2):025003, may 2016. doi:10.1103/RevModPhys.88.025003, 1505.07468. [p. 59]

[NN81] H. Nielsen and M. Ninomiya. A no-go theorem for regularizing chiral fermions. *Physics Letters B*, 105(2-3):219, oct 1981. doi:10.1016/0370-2693(81)91026-1. [pp. 33, 154]

[NN83] H. Nielsen and M. Ninomiya. The Adler-Bell-Jackiw anomaly and Weyl fermions in a crystal. *Physics Letters B*, 130(6):389, 1983. doi:10.1016/0370-2693(83)91529-0. [pp. 33, 154]

[NP66] P. Nozieres and D. Pines. *The Theory Of Quantum Liquids*. Perseus Books, 1966. [p. 72]

[NT13] N. Nagaosa and Y. Tokura. Topological properties and dynamics of magnetic skyrmions. *Nature Nanotechnology*, 8(12):899, dec 2013. doi:10.1038/nnano.2013.243. [pp. 50, 51]

[OM14] R. Okugawa and S. Murakami. Dispersion of Fermi arcs in Weyl semimetals and their evolutions to Dirac cones. *Physical Review B*, 89(23):235315, jun 2014. doi:10.1103/PhysRevB.89.235315, 1402.7145. [pp. 34, 35, 37]

[Ozb06] E. Ozbay. Plasmonics: Merging Photonics and Electronics at Nanoscale Dimensions. *Science*, 311(5758):189, jan 2006. doi:10.1126/science.1114849. [p. 57]

[PB10] D. Pesin and L. Balents. Mott physics and band topology in materials with strong spin-orbit interaction. *Nature Physics*, 6(5):376, may 2010. doi:10.1038/nphys1606, 0907.2962. [p. 112]

[PKV14] A. C. Potter, I. Kimchi, and A. Vishwanath. Quantum oscillations from surface Fermi arcs in Weyl and Dirac semimetals. *Nature Communications*, 5(May):5161, oct 2014. doi:10.1038/ncomms6161, 1402.6342. [p. 37]

[PPA+10] R. E. V. Profumo, M. Polini, R. Asgari, R. Fazio, and a. H. MacDonald. Electron-electron interactions in decoupled graphene layers. *Physical Review B*, 82(8):085443, aug 2010. doi:10.1103/PhysRevB.82.085443, 1004.4335. [p. 59]

[PRK+15] C. Pauly, B. Rasche, K. Koepernik, M. Liebmann, M. Pratzer, M. Richter, J. Kellner, M. Eschbach, B. Kaufmann, L. Plucinski, C. M. Schneider, M. Ruck, J. van den Brink, and M. Morgenstern. Subnanometre-wide electron channels protected by topology. *Nature Physics*, 11(4):338, mar 2015. doi:10.1038/nphys3264, 1501.05919. [p. 28]

[PSH16] D. I. Pikulin, P. G. Silvestrov, and T. Hyart. Confinement-deconfinement transition due to spontaneous symmetry breaking in quantum Hall bilayers. *Nature Communications*, 7(814):10462, jan 2016. doi:10.1038/ncomms10462, 1504.05154. [pp. 123, 126, 147]

[Pya09] P. K. Pyatkovskiy. Dynamical polarization, screening, and plasmons in gapped graphene. *Journal of Physics: Condensed Matter*, 21(2):025506, jan 2009. doi:10.1088/0953-8984/21/2/025506, 0808.0931. [p. 77]

[QZ11] X.-L. Qi and S.-C. Zhang. Topological insulators and superconductors. *Reviews of Modern Physics*, 83(4):1057, oct 2011. doi:10.1103/RevModPhys.83.1057, 1008.2026. [pp. 19, 23, 52]

[RAHM15] T. Rauch, S. Achilles, J. Henk, and I. Mertig. Spin Chirality Tuning and Topological Semimetals in Strained HgTe x S 1-x. *Physical Review Letters*, 114(23):236805, jun 2015. doi:10.1103/PhysRevLett.114.236805. [p. 154]

[RIR+13] B. Rasche, A. Isaeva, M. Ruck, S. Borisenko, V. Zabolotnyy, B. Büchner, K. Koepernik, C. Ortix, M. Richter, and J. van den Brink. Stacked topological insulator built from bismuth-based graphene sheet analogues. *Nature Materials*, 12(5):422, mar 2013. doi:10.1038/nmat3570, 1303.2193. [p. 28]

[RJY+16] J. Ruan, S.-k. Jian, H. Yao, H. Zhang, S.-c. Zhang, and D. Xing. Symmetry-protected ideal Weyl semimetal in HgTe-class materials. *Nature Communications*, 7:11136, apr 2016. doi:10.1038/ncomms11136, 1511.08284. [pp. 154, 165]

[RJZ+16] J. Ruan, S.-K. Jian, D. Zhang, H. Yao, H. Zhang, S.-C. Zhang, and D. Xing. Ideal Weyl Semimetals in the Chalcopyrites CuTlSe2, AgTlTe2, AuTlTe2, and ZnPbAs2. *Physical Review Letters*, 116(22):226801, jun 2016. doi:10.1103/PhysRevLett.116.226801, 1603.01279. [p. 154]

[SGT16] G. Sharma, P. Goswami, and S. Tewari. Chiral anomaly and longitudinal magnetotransport in type-II Weyl semimetals. *arXiv*, 1608.06625, aug 2016. 1608.06625. [pp. 37, 156]

[SGW+15] A. A. Soluyanov, D. Gresch, Z. Wang, Q. Wu, M. Troyer, X. Dai, and B. A. Bernevig. Type-II Weyl semimetals. *Nature*, 527(7579):495, nov 2015. doi:10.1038/nature15768, 1507.01603. [pp. 37, 156]

[She13] S.-Q. Shen. *Topological Insulators*. Springer, 2013. doi:10.1007/978-3-642-32858-9. [pp. 10, 12, 13, 15, 18, 20, 21, 23, 28, 29]

[Shu86] K. W. K. Shung. Dielectric function and plasmon structure of stage-1 intercalated graphite. *Physical Review B*, 34(2):979, jul 1986. doi:10.1103/PhysRevB.34.979. [p. 62]

[SJ14] A. A. Shylau and A.-P. Jauho. Plasmon-mediated Coulomb drag between graphene waveguides. *Physical Review B*, 89(16):165421, apr 2014. doi:10.1103/PhysRevB.89.165421, 1404.6136. [pp. 59, 167]

[SJJ10] A. Ström, H. Johannesson, and G. I. Japaridze. Edge Dynamics in a Quantum Spin Hall State: Effects from Rashba Spin-Orbit Interaction. *Physical Review Letters*, 104(25):256804, jun 2010. doi:10.1103/PhysRevLett.104.256804, 1004.2777. [pp. 112, 146, 151]

[SKKR93] S. L. Sondhi, A. Karlhede, S. A. Kivelson, and E. H. Rezayi. Skyrmions and the crossover from the integer to fractional quantum Hall effect at small Zeeman energies. *Physical Review B*, 47(24):16419, 1993. doi:10.1103/PhysRevB.47.16419. [pp. 49, 135]

[SMAF12] B. Scharf, A. Matos-Abiague, and J. Fabian. Magnetic properties of HgTe quantum wells. *Physical Review B*, 86(7):075418, aug 2012. doi:10.1103/PhysRevB.86.075418, 1207.4578. [pp. 27, 28]

[SMAŽF15] B. Scharf, A. Matos-Abiague, I. Žutić, and J. Fabian. Probing topological transitions in HgTe/CdTe quantum wells by magneto-optical measurements. *Physical Review B*, 91(23):235433, jun 2015. doi:10.1103/PhysRevB.91.235433, 1502.05605. [p. 23]

[SMF09] B. Seradjeh, J. Moore, and M. Franz. Exciton Condensation and Charge Fractionalization in a Topological Insulator Film. *Physical Review Letters*, 103(6):066402, aug 2009. doi:10.1103/PhysRevLett.103.066402, 0902.1147. [p. 60]

[SNC08] J. Sabio, J. Nilsson, and A. H. Castro Neto. f-sum rule and unconventional spectral weight transfer in graphene. *Physical Review B*, 78(7):075410, aug 2008. doi:10.1103/PhysRevB.78.075410, 0806.1684. [pp. 72, 74, 88, 172]

[SNKT09] M. Schmidt, E. Novik, M. Kindermann, and B. Trauzettel. Optical manipulation of edge-state transport in HgTe quantum wells in the quantum Hall regime. *Physical Review B*, 79(24):241306, jun 2009. doi:10.1103/PhysRevB.79.241306, 0901.0621. [pp. 68, 71]

[SNPS15] T. Stauber, D. Noriega-Pérez, and J. Schliemann. Universal absorption of two-dimensional systems. *Physical Review B*, 91(11):115407, mar 2015. doi:10.1103/PhysRevB.91.115407, 1412.5835. [p. 172]

[SPM12] I. Sodemann, D. a. Pesin, and a. H. MacDonald. Interaction-enhanced coherence between two-dimensional Dirac layers. *Physical Review B*, 85(19):195136, may 2012. doi:10.1103/PhysRevB.85.195136, 1203.3594. [pp. 62, 85]

[SRF+09] A. P. Schnyder, S. Ryu, A. Furusaki, A. W. W. Ludwig, V. Lebedev, and M. Feigel'man. Classification of Topological Insulators and Superconductors. In *AIP Conference Proceedings*, volume 1134, page 10. AIP, 2009. [p. 20]

[SRFL08] A. P. Schnyder, S. Ryu, A. Furusaki, and A. W. W. Ludwig. Classification of topological insulators and superconductors in three spatial dimensions. *Physical Review B*, 78(19):195125, nov 2008. doi:10.1103/PhysRevB.78.195125, 0803.2786. [p. 20]

[SRvOG12] T. L. Schmidt, S. Rachel, F. von Oppen, and L. I. Glazman. Inelastic Electron Backscattering in a Generic Helical Edge Channel. *Physical Review Letters*, 108(15):156402, apr 2012. doi:10.1103/PhysRevLett.108.156402, 1201.0278. [pp. 112, 146, 151]

[SS95] I. Safi and H. J. Schulz. Transport in an inhomogeneous interacting one-dimensional system. *Physical Review B*, 52(24):R17040, dec 1995. doi:10.1103/PhysRevB.52.R17040, cond-mat/9505079. [p. 167]

[SS14] K. Shiozaki and M. Sato. Topology of crystalline insulators and superconductors. *Physical Review B*, 90(16):165114, oct 2014. doi:10.1103/PhysRevB.90.165114, 1403.3331. [p. 32]

[SSG16] K. Shiozaki, M. Sato, and K. Gomi. Topology of nonsymmorphic crystalline insulators and superconductors. *Physical Review B*, 93(19):195413, may 2016. doi:10.1103/PhysRevB.93.195413, 1403.3331. [p. 32]

[SSS13] A. Scholz, T. Stauber, and J. Schliemann. Plasmons and screening in a monolayer of MoS2. *Physical Review B*, 88(3):035135, jul 2013. doi:10.1103/PhysRevB.88.035135, 1306.1666. [p. 63]

[Sta14] T. Stauber. Plasmonics in Dirac systems: from graphene to topological insulators. *Journal of Physics: Condensed Matter*, 26(12):123201, mar 2014. doi:10.1088/0953-8984/26/12/123201, 1310.4296. [pp. 58, 63, 69]

[Ste67] F. Stern. Polarizability of a Two-Dimensional Electron Gas. *Physical Review Letters*, 18(14):546, apr 1967. doi:10.1103/PhysRevLett.18.546. [p. 87]

[SWY15] Y. Sun, S.-C. Wu, and B. Yan. Topological surface states and Fermi arcs of the noncentrosymmetric Weyl semimetals TaAs, TaP, NbAs, and NbP. *Physical Review B*, 92(11):115428, sep 2015. doi:10.1103/PhysRevB.92.115428, 1508.06649. [p. 154]

[TKNdN82] D. J. Thouless, M. Kohmoto, M. P. Nightingale, and M. den Nijs. Quantized Hall Conductance in a Two-Dimensional Periodic Potential. *Physical Review Letters*, 49(6):405, aug 1982. doi:10.1103/PhysRevLett.49.405. [p. 21]

[TRS+12] Y. Tanaka, Z. Ren, T. Sato, K. Nakayama, S. Souma, T. Takahashi, K. Segawa, and Y. Ando. Experimental realization of a topological crystalline insulator in SnTe. *Nature Physics*, 8(11):800, sep 2012. doi:10.1038/nphys2442, 1304.0430. [p. 32]

[TSFM16] P. Tikhonov, E. Shimshoni, H. A. Fertig, and G. Murthy. Emergence of helical edge conduction in graphene at the ν=0 quantum Hall state. *Physical Review B*, 93(11):115137, mar 2016. doi:10.1103/PhysRevB.93.115137, 1512.07825. [pp. 123, 147]

[TTT+06] J. Tworzydło, B. Trauzettel, M. Titov, A. Rycerz, and C. W. J. Beenakker. Sub-Poissonian Shot Noise in Graphene. *Physical Review Letters*, 96(24):246802, jun 2006. doi:10.1103/PhysRevLett.96.246802, cond-mat/0603315. [p. 166]

[TWC+16] A. Tamai, Q. S. Wu, I. Cucchi, F. Y. Bruno, S. Riccò, T. K. Kim, M. Hoesch, C. Barreteau, E. Giannini, C. Besnard, A. A. Soluyanov, and F. Baumberger. Fermi Arcs and Their Topological Character in the Candidate Type-II Weyl Semimetal MoTe2. *Physical Review X*, 6(3):031021, aug 2016. doi:10.1103/PhysRevX.6.031021, 1604.08228. [p. 154]

[Vaf06] O. Vafek. Thermoplasma Polariton within Scaling Theory of Single-Layer Graphene. *Physical Review Letters*, 97(26):266406, dec 2006. doi:10.1103/PhysRevLett.97.266406, cond-mat/0605642. [pp. 62, 85, 87]

[vMOS16] G. van Miert, C. Ortix, and C. M. Smith. Topological origin of edge states in two-dimensional inversion-symmetric insulators and semimetals. *2D Materials*, 4(1):015023, nov 2016. doi:10.1088/2053-1583/4/1/015023, 1606.03232. [pp. 20, 32]

[Vol03] G. E. Volovik. *The Universe in a Helium Droplet*. Oxford University Press, 2003. [p. 1]

[WACB16] Z. Wang, A. Alexandradinata, R. J. Cava, and B. A. Bernevig. Hourglass fermions. *Nature*, 532(7598):189, apr 2016. doi:10.1038/nature17410, 1602.05585. [p. 154]

[WBZ06] C. Wu, B. A. Bernevig, and S.-C. Zhang. Helical Liquid and the Edge of Quantum Spin Hall Systems. *Physical Review Letters*, 96(10):106401, mar 2006. doi:10.1103/PhysRevLett.96.106401, cond-mat/0508273. [pp. 19, 24, 112, 146, 151]

[WFF+15] H. Weng, C. Fang, Z. Fang, B. A. Bernevig, and X. Dai. Weyl Semimetal Phase in Noncentrosymmetric Transition-Metal Monophosphides. *Physical Review X*, 5(1):011029, mar 2015. doi:10.1103/PhysRevX.5.011029, 1501.00060. [p. 154]

[WHFZ12] Z. Wang, N. Hao, Z.-G. Fu, and P. Zhang. Excitonic condensation for the surface states of topological insulator bilayers. *New Journal of Physics*, 14(6):063010, jun 2012. doi:10.1088/1367-2630/14/6/063010, 1106.5838. [p. 60]

[WHPF92] W. T. V. William H. Press, Saul A. Teukolsky and B. P. Flannery. *Numerical Recipes in C*. Cambridge University Press, 1992. [p. 138]

[wik17a] https://en.wikipedia.org/wiki/Topology. Accessed 21.03.2017. [pp. 6, 8]

[wik17b] https://en.wikipedia.org/wiki/Seven_Bridges_of_Königsberg. Accessed 21.03.2017. [p. 7]

[WSSG06] B. Wunsch, T. Stauber, F. Sols, and F. Guinea. Dynamical polarization of graphene at finite doping. *New Journal of Physics*, 8(12):318, dec 2006. doi:10.1088/1367-2630/8/12/318, cond-mat/0610630. [pp. 62, 95, 172]

[WTVS11] X. Wan, A. M. Turner, A. Vishwanath, and S. Y. Savrasov. Topological semimetal and Fermi-arc surface states in the electronic structure of pyrochlore iridates. *Physical Review B*, 83(20):205101, 2011. doi:10.1103/PhysRevB.83.205101, 1007.0016. [p. 154]

[XAB+15] S.-Y. Xu, N. Alidoust, I. Belopolski, Z. Yuan, G. Bian, T.-R. Chang, H. Zheng, V. N. Strocov, D. S. Sanchez, G. Chang, C. Zhang, D. Mou, Y. Wu, L. Huang, C.-C. Lee, S.-M. Huang, B. Wang, A. Bansil, H.-T. Jeng, T. Neupert, A. Kaminski, H. Lin, S. Jia, and M. Zahid Hasan. Discovery of a Weyl fermion state with Fermi arcs in niobium arsenide. *Nature Physics*, 11(9):748, aug 2015. doi:10.1038/nphys3437, 1504.01350. [p. 154]

[XAC+16] S.-y. Xu, N. Alidoust, G. Chang, H. Lu, B. Singh, I. Belopolski, D. Sanchez, X. Zhang, G. Bian, H. Zheng, M.-a. Husanu, Y. Bian, S.-m. Huang, C.-h. Hsu, T.-r. Chang, H.-t. Jeng, A. Bansil, V. N. Strocov, H. Lin, S. Jia, and M. Z. Hasan. Discovery of Lorentz-violating Weyl fermion semimetal state in LaAlGe materials. *arXiv*, 1603.07318, mar 2016. 1603.07318. [p. 154]

[XDW+15] D.-F. Xu, Y.-P. Du, Z. Wang, Y.-P. Li, X.-H. Niu, Q. Yao, D. Pavel, Z.-A. Xu, X.-G. Wan, and D.-L. Feng. Observation of Fermi Arcs in Non-Centrosymmetric Weyl Semi-Metal Candidate NbP. *Chinese Physics Letters*, 32(10):107101, oct 2015. doi:10.1088/0256-307X/32/10/107101, 1509.03847. [p. 154]

[XM06] C. Xu and J. E. Moore. Stability of the quantum spin Hall effect: Effects of interactions, disorder, and Z2 topology. *Physical Review B*, 73(4):045322, jan 2006. doi:10.1103/PhysRevB.73.045322, cond-mat/0508291. [pp. 24, 112, 146, 151]

[XQH+09] Y. Xia, D. Qian, D. Hsieh, L. Wray, A. Pal, H. Lin, A. Bansil, D. Grauer, Y. S. Hor, R. J. Cava, and M. Z. Hasan. Observation of a large-gap topological-insulator class with a single Dirac cone on the surface. *Nature Physics*, 5(6):398, jun 2009. doi:10.1038/nphys1274, 0908.3513. [p. 29]

[XWW+11] G. Xu, H. Weng, Z. Wang, X. Dai, and Z. Fang. Chern Semimetal and the Quantized Anomalous Hall Effect in HgCr2Se4. *Physical Review Letters*, 107(18):186806, oct 2011. doi:10.1103/PhysRevLett.107.186806, 1106.3125. [pp. 154, 166]

[XWW+16] N. Xu, Z. J. Wang, A. P. Weber, A. Magrez, P. Bugnon, H. Berger, C. E. Matt, J. Z. Ma, B. B. Fu, B. Q. Lv, N. C. Plumb, M. Radovic, E. Pomjakushina, K. Conder, T. Qian, J. H. Dil, J. Mesot, H. Ding, and M. Shi. Discovery of Weyl semimetal state violating Lorentz invariance in MoTe2. *arXiv*, 1604.02116, apr 2016. 1604.02116. [p. 166]

[XZZ15] Y. Xu, F. Zhang, and C. Zhang. Structured Weyl Points in Spin-Orbit Coupled Fermionic Superfluids. *Physical Review Letters*, 115(26):265304, dec 2015. doi:10.1103/PhysRevLett.115.265304, 1411.7316. [pp. 37, 156]

[YDM06] K. Yang, S. Das Sarma, and a. MacDonald. Collective modes and skyrmion excitations in graphene SU(4) quantum Hall ferromagnets. *Physical Review B*, 74(7):075423, aug 2006. doi:10.1103/PhysRevB.74.075423. [p. 49]

[YF17] B. Yan and C. Felser. Topological Materials: Weyl Semimetals. *Annual Review of Condensed Matter Physics*, 8(1):annurev–conmatphys–031016–025458, mar 2017. doi:10.1146/annurev-conmatphys-031016-025458, 1611.04182. [p. 33]

[YLR11] K.-Y. Yang, Y.-M. Lu, and Y. Ran. Quantum Hall effects in a Weyl semimetal: Possible application in pyrochlore iridates. *Physical Review B*, 84(7):075129, 2011. doi:10.1103/PhysRevB.84.075129, 1105.2353. [pp. 37, 155]

[YLS+15] L. X. Yang, Z. K. Liu, Y. Sun, H. Peng, H. F. Yang, T. Zhang, B. Zhou, Y. Zhang, Y. F. Guo, M. Rahn, D. Prabhakaran, Z. Hussain, S.-K. Mo, C. Felser, B. Yan, and Y. L. Chen. Weyl semimetal phase in the non-centrosymmetric compound TaAs. *Nature Physics*, 11(9):728, aug 2015. doi:10.1038/nphys3425, 1507.00521. [p. 154]

[YMB+96] K. Yang, K. Moon, L. Belkhir, H. Mori, S. M. Girvin, A. H. MacDonald, L. Zheng, and D. Yoshioka. Spontaneous interlayer coherence in double-layer quantum Hall

systems: Symmetry-breaking interactions, in-plane fields, and phase solitons. *Physical Review B*, 54(16):11644, oct 1996. doi:10.1103/PhysRevB.54.11644, cond-mat/9605153. [p. 49]

[YMZ+94] K. Yang, K. Moon, L. Zheng, A. H. MacDonald, S. M. Girvin, D. Yoshioka, and S.-C. Zhang. Quantum ferromagnetism and phase transitions in double-layer quantum Hall systems. *Physical Review Letters*, 72(5):732, jan 1994. doi:10.1103/PhysRevLett.72.732, cond-mat/9310071. [p. 49]

[YN14] B.-J. Yang and N. Nagaosa. Classification of stable three-dimensional Dirac semimetals with nontrivial topology. *Nature Communications*, 5:4898, 2014. doi:10.1038/ncomms5898, 1404.0754. [p. 42]

[YSYH+13] A. F. Young, J. D. Sanchez-Yamagishi, B. Hunt, S. H. Choi, K. Watanabe, T. Taniguchi, R. C. Ashoori, and P. Jarillo-Herrero. Tunable symmetry breaking and helical edge transport in a graphene quantum spin Hall state. *Nature*, 505(7484):528, dec 2013. doi:10.1038/nature12800, 1307.5104. [pp. 113, 136]

[ZHKE93] N. B. Zhitenev, R. J. Haug, K. V. Klitzing, and K. Eberl. Time-resolved measurements of transport in edge channels. *Physical Review Letters*, 71(14):2292, oct 1993. doi:10.1103/PhysRevLett.71.2292. [p. 167]

[ZLQ+09] H. Zhang, C.-X. Liu, X.-L. Qi, X. Dai, Z. Fang, and S.-C. Zhang. Topological insulators in Bi2Se3, Bi2Te3 and Sb2Te3 with a single Dirac cone on the surface. *Nature Physics*, 5(6):438, 2009. doi:10.1038/nphys1270. [pp. 29, 155]