Master's thesis

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

Stein Kåre Lundeberg Fosstveit

Spin and Thermal Hall Effects in the Strongly Correlated Kane-Mele-Hubbard Model

Master's thesis in Applied Physics and Mathematics
Supervisor: Alireza Qaiumzadeh
July 2022

NTNU
Norwegian University of
Science and Technology

Stein Kåre Lundeberg Fosstveit

Spin and Thermal Hall Effects in the Strongly Correlated Kane-Mele-Hubbard Model

Master's thesis in Applied Physics and Mathematics
Supervisor: Alireza Qaiumzadeh
July 2022

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

NTNU
Norwegian University of
Science and Technology

Abstract

We study the properties of an extended Kane-Mele-Hubbard model on the honeycomb lattice when next nearest neighbor out-of-plane Dzyaloshinskii-Moriya interactions are taken into account. Using Schwinger boson mean field theory, we construct qualitative phase diagrams and find two magnetically ordered phases and a quantum spin liquid phase in the ground state of the model. The quantum spin liquid phase is found to have significant Abelian Berry curvatures in the presence of a magnetic field, and displays both the spin Nernst effect and the thermal Hall effect even though its bands have zero Chern numbers. Dzyaloshinskii-Moriya interactions are found to open partial band gaps in the magnetically ordered phases in the mean field Ansatz employed, leading to ill-defined Berry curvatures due to the extensive band overlaps in the first Brillouin zone. The dynamical properties of the quantum spin liquid phase are unchanged by Dzyaloshinskii-Moriya interactions, which means they do not influence the transport properties. Strong Dzyaloshinskii-Moriya interactions are found to enlarge the quantum spin liquid region of the phase diagram, however, making it stable for stronger next nearest neighbor interactions.

Sammendrag

Vi studerer egenskapene til en utvidet Kane-Mele-Hubbard-modell på honeycombgitteret når Dzyaloshinskii-Moriya-vekselvirkninger for nest nærmeste naboer er rettet ut av planet. Ved hjelp av middelfeltteorien for Schwingerbosoner konstruerer vi kvalitative fasediagrammer og finner to magnetisk ordnede faser og en kvantespinnvæskefase i grunntilstanden til modellen. Kvantespinnvæskefasen viser seg å ha betydelige abelske Berry-kurvaturer i nærvær av et magnetisk felt, og utviser både spinn-Nernst-effekten og den termiske Hall-effekten selv om båndene har Chern-tall lik null. Dzyaloshinskii-Moriya-vekselvirkningene viser seg å åpne delvise båndgap i de magnetisk ordnede fasene i middelfeltansatzen som benyttes, noe som fører til udefinerte Berry-kurvaturer på grunn av de omfattende båndoverlappingene i den første Brillouin-sonen. De dynamiske egenskapene til kvantespinnvæskefasen endres ikke av Dzyaloshinskii-Moriya-vekselvirkningene, som betyr at de ikke påvirker transportegenskapene. Sterke Dzyaloshinskii-Moriya-vekselvirkninger viser seg likevel å forstørre kvantespinnvæskeområdet i fasediagrammet, som gjør det stabilt for sterkere vekselvirkninger mellom nest nærmeste naboer.

Acknowledgments

I wish to express my deepest gratitude to my supervisor, Alireza Qaiumzadeh, for his invaluable guidance and support through this journey. Completing the master's thesis would not have been possible without his help, and I will greatly miss working with him.

I am also grateful to my family and friends for their unconditional love and emotional support.

Contents

1 Introduction

Spintronics is the study of the spin, as well as the charge, of the electron in solid-state devices. The dawn of spintronics came in the late 1980s with Grünberg [1] and Fert's [2] discovery of the giant magnetoresistance (GMR), which has found uses in e.g. hard-disk read-heads and electrical current measurement [3]. Since then, spintronics has become a large field of research, with the goal of delivering novel solutions for fast and low-power information and communication technologies. Spintronic devices have largely been realized in ferromagnetic systems so far, with antiferromagnets playing a less prominent role in technical applications. Louis Néel, who first predicted the existence of antiferromagnets in the first half of the 20th century, called them "interesting, but useless", failing to see how they could be used in practice.

Today, the situation is different, but there are still aspects of antiferromagnetic systems to which Néel's famous words apply. One such interesting, but "useless" aspect is the phase of matter known as a quantum spin liquid. Originally proposed by Anderson [4] as the ground state of the Heisenberg model on the triangular lattice, quantum spin liquids arise from magnetic frustration, where the system geometry prohibits localized quantum spins from satisfying their mutual exchange interactions. As a result, the ground state is massively degenerate, and the spins fluctuate even when the temperature tends to zero. Unlike in a traditional ferro- or antiferromagnet, where there is symmetry breaking and an emergent order parameter, the quantum spin liquid has no long-range magnetic order at low temperature. Instead, quantum spin liquids are characterized by their topological order, quantum entanglement, fractionalized excitations and gauge field fluctuations [5].

The reason why quantum spin liquids can be deemed "useless" is that it is very hard to detect them experimentally. A first step towards overcoming this challenge is to gain a comprehensive theoretical understanding of models that harbor quantum spin liquid phases. One such model is the Kane-Mele-Hubbard model on the honeycomb lattice studied by Ushakov [6] and Vaezi et al. [7]. Several realizable honeycomb materials have been suggested as candidates for quantum spin liquids [8], encouraging investigations into how the quantum spin liquid phase in the Kane-Mele-Hubbard model may be probed experimentally. The main objective of this thesis is therefore to explore the theoretical possibility of detecting this quantum spin liquid phase by its characteristic transport signatures.

The thesis is structured as follows: Section 2 introduces the concept of a spin Hamiltonian and how it can be expressed in second quantized form through fermionization and bosonization procedures. In Section 3, we then demonstrate how these procedures can be used to obtain the low-lying excitation spectrum of the XY spin chain model. Section 4 introduces the geometric Berry phase and the associated Berry curvature, which is linked to the topology of quantum mechanical systems through the Chern number. We study the phase diagram of the Chern insulator to demonstrate how these concepts present themselves in condensed matter physics. In Section 5, we develop the linear response theory for quantum mechanical systems. We derive the Kubo formula, and show how it may be used to compute response functions using real time and Matsubara Green's functions. In Section 6, we use the linear response theory to compute the Hall conductivity in electron systems, which is a topological invariant. We also present topology-dependent conductivity formulas for the spin Nernst effect and thermal Hall effect in bosonic spin systems. In Section 7, we obtain

the ground state of the Kane-Mele-Hubbard model in the presence of a magnetic field using Schwinger boson mean field theory. Finally, in Section 8, we present and discuss the dynamical properties and phase diagrams of the model and calculate spin Nernst and thermal Hall conductivities in the predicted phases.

It should be noted that parts of the thesis are included in order to gain a better understanding of the theory presented, and are not directly related to the final results. The thesis is a continuation of the specialization project report [9], which contained Sections 2, 3, 5 and 7. The contents of these Sections may have been altered in the final thesis.

2 Spin Hamiltonians in the Second Quantization

When we study magnetic systems using a classical approach, the usual starting point is a spin Hamiltonian function. The spins are modeled as vectors on a lattice, and the spin Hamiltonian describes how the spins interact with each other and with external fields. A well known example of a spin Hamiltonian is the classical Heisenberg model:

$$H = J \sum_{\langle i,j \rangle} \vec{S}_i \cdot \vec{S}_j, \tag{2.1}$$

where J represents the spin exchange interaction and the sum is taken over nearest neighbors. The sign of the spin coupling determines the nature of the magnetic system. If $J < 0$, it is energetically favorable for nearest neighbor spins to be parallel, making the system ferromagnetic, while $J > 0$ favors anti-parallel nearest neighbor spins, making the system antiferromagnetic.

In quantum mechanics, observables are represented by operators. To study the quantum nature of magnetic systems, we therefore replace the spin vector components with spin operators. This complicates the analytical treatment of the spin Hamiltonian, which becomes an operator, and makes it far more difficult to obtain the ground state of the spin system. In the following, we present some techniques that can help us with such an analysis. We introduce the concepts of fermionization and bosonization, which allow us to write spin Hamiltonians in second quantized form, and consider how such Hamiltonians can be diagonalized to identify the system's low-lying energy excitation spectrum.

2.1 Fermionization and Bosonization

In the second quantization, the energy spectrum of a quantum mechanical system is described by creation and annihilation operators, denoted by $\hat{a}^\dagger$ and $\hat{a}$, respectively. Suppose that $\{|n\rangle\}$ constitutes a set of orthonormalized single-particle eigenstates of the Hamiltonian $\hat{H}$ with corresponding eigenvalues E_n. Then, $\hat{a}$ and $\hat{a}^\dagger$ act on $|n\rangle$ in the following way:

$$\hat{a}\,|n\rangle = \sqrt{n}\,|n-1\rangle\,, \tag{2.2}$$

$$\hat{a}^\dagger\,|n\rangle = \sqrt{n+1}\,|n+1\rangle\,. \tag{2.3}$$

From (2.2) and (2.3) it follows that $\hat{a}^\dagger\hat{a}$ is the number operator $\hat{n}$:

$$\hat{n}\,|n\rangle = \hat{a}^\dagger\hat{a}\,|n\rangle = n\,|n\rangle\,. \tag{2.4}$$

We see from (2.4) that the eigenvalues of $\hat{n}$ describe the degree of excitation of the system. Consider now the spin operators $\hat{S}^x$, $\hat{S}^y$ and $\hat{S}^z$. We take the z-axis to be the quantization axis, such that the spin state $|s, m_s\rangle$, where s is the spin quantum number and m_s is the spin projection quantum number, is an eigenstate of $\hat{S}^z$:

$$\hat{S}^z\,|s, m_s\rangle = \hbar m_s\,|s, m_s\rangle\,. \tag{2.5}$$

We then define the lowering and raising operators

$$\hat{S}^\pm = \hat{S}^x \pm i\hat{S}^y, \tag{2.6}$$

which can be shown to have the following effects when acting on the spin state $|s, m_s\rangle$:

$$\hat{S}^- |s, m_s\rangle = \hbar\sqrt{s(s+1) - m_s(m_s-1)}\, |s, m_s - 1\rangle\,, \tag{2.7}$$

$$\hat{S}^+ |s, m_s\rangle = \hbar\sqrt{s(s+1) - m_s(m_s+1)}\, |s, m_s + 1\rangle\,. \tag{2.8}$$

We see from (2.7) and (2.8) that $\hat{S}^-$ lowers m_s and $\hat{S}^+$ raises m_s. Thus, $\hat{S}^-$ and $\hat{S}^+$ are analogous to $\hat{a}$ and $\hat{a}^\dagger$. However, creation and annihilation operators obey much simpler algebras than the spin operators. Spin operators acting on lattice sites i and j obey

$$[\hat{S}_i^\alpha, \hat{S}_j^\beta] = i\hbar\delta_{ij}\sum_\gamma \epsilon_{\alpha\beta\gamma}\hat{S}_i^\gamma \quad \text{and} \quad \{\hat{S}_i^\alpha, \hat{S}_j^\beta\} = \frac{\hbar^2}{2}\delta_{ij}\delta_{\alpha\beta}I; \quad \alpha, \beta, \gamma \in \{x, y, z\}, \tag{2.9}$$

where δ_{ij} is the Kronecker delta, $\epsilon_{\alpha\beta\gamma}$ is the Levi-Civita symbol and I is the identity operator. The algebras of the creation and annihilation operators depend on whether they are fermionic or bosonic. The anticommutator of two fermionic operators equals zero or one, and the same is true for the commutator of two bosonic operators. We define the statistical sign

$$\zeta = \begin{cases} -1 & \text{for fermions,} \\ +1 & \text{for bosons,} \end{cases} \tag{2.10}$$

which allows us to write anticommutators and commutators for operators $\hat{A}$ and $\hat{B}$ as

$$[\hat{A}, \hat{B}]_\zeta = \hat{A}\hat{B} - \zeta\hat{B}\hat{A} = \begin{cases} \{\hat{A}, \hat{B}\} & \text{for fermions,} \\ [\hat{A}, \hat{B}] & \text{for bosons.} \end{cases} \tag{2.11}$$

The fermionic anticommutation relations and bosonic commutation relations may then be written compactly as

$$[\hat{a}_i, \hat{a}_j^\dagger]_\zeta = \delta_{ij}, \tag{2.12}$$

$$[\hat{a}_i, \hat{a}_j]_\zeta = 0, \tag{2.13}$$

$$[\hat{a}_i^\dagger, \hat{a}_j^\dagger]_\zeta = 0. \tag{2.14}$$

The analogy to spin raising and lowering operators, along with these highly convenient ζ-commutation relations, motivates us to express spin operators as functions of creation and annihilation operators:

$$\hat{S}_i^\alpha = f_i^\alpha(\{\hat{a}_i\}, \{\hat{a}_i^\dagger\}); \quad \alpha \in \{x, y, z\}. \tag{2.15}$$

This procedure is called fermionization or bosonization, depending on whether the spin operators are mapped to fermionic or bosonic operators. This allows us to transform spin operator Hamiltonians into second quantized Hamiltonians, which are much easier to work with.

2.2 Diagonalization of Non-Interacting Hamiltonians

Once the spin Hamiltonian is expressed in terms of second quantized operators, we can diagonalize it to obtain the ground state dynamics of the spin system. We will now present a general diagonalization procedure for quadratic Hamiltonians, which describe non-interacting quasiparticles. We will then consider the special case where the quadratic Hamiltonian is bosonic, as this will be particularly valuable later in the thesis.

2.2.1 The General Procedure

A general quadratic Hamiltonian can be written as

$$\hat{H} = \sum_{i,j=1}^{N} \left[C_{ij}\hat{a}_i^\dagger \hat{a}_j + D_{ij}\hat{a}_i\hat{a}_j^\dagger + E_{ij}\hat{a}_i\hat{a}_j + F_{ij}\hat{a}_i^\dagger\hat{a}_j^\dagger \right] + G, \tag{2.16}$$

where $\{C_{ij},\ D_{ij},\ E_{ij},\ F_{ij}\}$ are a set of coefficients, G is a constant and $\{\hat{a}_i\}$ are fermionic or bosonic operators. This Hamiltonian can readily be written in matrix form. However, we can first utilize its hermiticity and natural symmetries to write it in a simpler form. We follow the steps of Ushakov [6]. The first step is to apply (2.12) to the second term in (2.16):

$$\begin{aligned} \hat{H} &= \sum_{i,j=1}^{N} \left[C_{ij}\hat{a}_i^\dagger \hat{a}_j + \zeta D_{ij}\hat{a}_j^\dagger\hat{a}_i + E_{ij}\hat{a}_i\hat{a}_j + F_{ij}\hat{a}_i^\dagger\hat{a}_j^\dagger + D_{ij}\delta_{ij} \right] + G \\ &= \sum_{i,j=1}^{N} \left[(C_{ij} + \zeta D_{ji})\hat{a}_i^\dagger \hat{a}_j + E_{ij}\hat{a}_i\hat{a}_j + F_{ij}\hat{a}_i^\dagger\hat{a}_j^\dagger \right] + \mathrm{Tr}\{\mathbf{D}\} + G, \end{aligned} \tag{2.17}$$

where $\mathbf{D}$ is an $N \times N$ matrix whose elements are D_{ij}. Since $\mathrm{Tr}\{\mathbf{D}\}$ and G are constants, their only effect is to redefine the reference energy, so we drop them in the following. Next, we demand that our Hamiltonian be Hermitian, $\hat{H} = \hat{H}^\dagger$. We take the Hermitian conjugate of (2.17) to obtain

$$\begin{aligned} \hat{H}^\dagger &= \sum_{i,j=1}^{N} \left[(C_{ij}^* + \zeta D_{ji}^*)\hat{a}_j^\dagger \hat{a}_i + E_{ij}^*\hat{a}_j^\dagger\hat{a}_i^\dagger + F_{ij}^*\hat{a}_j\hat{a}_i \right] \\ &= \sum_{i,j=1}^{N} \left[(C_{ji}^* + \zeta D_{ij}^*)\hat{a}_i^\dagger \hat{a}_j + E_{ji}^*\hat{a}_i^\dagger\hat{a}_j^\dagger + F_{ji}^*\hat{a}_i\hat{a}_j \right]. \end{aligned} \tag{2.18}$$

Defining $2A_{ij} = C_{ij} + \zeta D_{ji}$, the hermiticity condition requires that

$$A_{ij} = A_{ji}^* \quad \text{and} \quad F_{ij} = E_{ji}^*. \tag{2.19}$$

We define $B_{ij} = F_{ij}$ for alphabetical order, which means that $E_{ij} = B_{ji}^*$ by (2.19). Finally, we apply (2.13) and (2.14) to the last two summands in (2.17):

$$\begin{aligned} \sum_{i,j=1}^{N} \left[B_{ji}^*\hat{a}_i\hat{a}_j + B_{ij}\hat{a}_i^\dagger\hat{a}_j^\dagger \right] &= \sum_{i,j=1}^{N} \left[\zeta B_{ji}^*\hat{a}_j\hat{a}_i + \zeta B_{ij}\hat{a}_j^\dagger{\hat{a}_i}^\dagger \right] \\ &= \sum_{i,j=1}^{N} \left[\zeta B_{ij}^*\hat{a}_i\hat{a}_j + \zeta B_{ji}\hat{a}_i^\dagger\hat{a}_j^\dagger \right], \end{aligned} \tag{2.20}$$

which gives us the following relations:

$$B_{ij} = \zeta B_{ji} \quad \text{and} \quad B_{ij}^* = \zeta B_{ji}^*. \tag{2.21}$$

With these results we can write the general quadratic Hamiltonian in (2.16), up to an additive constant, as

$$\hat{H} = \sum_{i,j=1}^{N} \left[2A_{ij}\hat{a}_i^\dagger \hat{a}_j + B_{ij}\hat{a}_i^\dagger \hat{a}_j^\dagger + B_{ji}^* \hat{a}_i \hat{a}_j \right]. \tag{2.22}$$

We now show how this simplified Hamiltonian can be written in matrix form. We write (2.22) as

$$\begin{aligned}
\hat{H} &= \sum_{i,j=1}^{N} \left[A_{ij}\hat{a}_i^\dagger \hat{a}_j + A_{ji}\hat{a}_j^\dagger \hat{a}_i + \zeta A_{ji}\delta_{ij} - \zeta A_{ij}\delta_{ij} + B_{ij}\hat{a}_i^\dagger \hat{a}_j^\dagger + B_{ji}^* \hat{a}_i \hat{a}_j \right] \\
&= \sum_{i,j=1}^{N} \left[A_{ij}\hat{a}_i^\dagger \hat{a}_j + \zeta A_{ij}^*(\delta_{ij} + \zeta \hat{a}_j^\dagger \hat{a}_i) + B_{ij}\hat{a}_i^\dagger \hat{a}_j^\dagger + \zeta B_{ij}^* \hat{a}_i \hat{a}_j \right] - \zeta \operatorname{Tr}\{\mathbf{A}\} \\
&= \sum_{i,j=1}^{N} \left[A_{ij}\hat{a}_i^\dagger \hat{a}_j + \zeta A_{ij}^* \hat{a}_i \hat{a}_j^\dagger + B_{ij}\hat{a}_i^\dagger \hat{a}_j^\dagger + \zeta B_{ij}^* \hat{a}_i \hat{a}_j \right] - \zeta \operatorname{Tr}\{\mathbf{A}\},
\end{aligned} \tag{2.23}$$

where we used $\zeta^2 = 1$, (2.19) and (2.21) in the second equality and (2.12) in the third equality. $\mathbf{A}$ is an $N \times N$ matrix with elements A_{ij}. We define $\mathbf{B}$ to be the $N \times N$ matrix with elements B_{ij}, as well as the operator vector

$$\begin{bmatrix} \hat{\mathbf{a}} & \hat{\mathbf{a}}^\dagger \end{bmatrix}^\mathsf{T} = \begin{bmatrix} \hat{a}_1 & \dots & \hat{a}_N & \hat{a}_1^\dagger & \dots & \hat{a}_N^\dagger \end{bmatrix}^\mathsf{T}. \tag{2.24}$$

With these definitions, we can write (2.23) as:

$$\hat{H} = \begin{bmatrix} \hat{\mathbf{a}}^\dagger & \hat{\mathbf{a}} \end{bmatrix} \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \zeta\mathbf{B}^* & \zeta\mathbf{A}^* \end{bmatrix} \begin{bmatrix} \hat{\mathbf{a}} \\ \hat{\mathbf{a}}^\dagger \end{bmatrix} - \zeta \operatorname{Tr}\{\mathbf{A}\} = \begin{bmatrix} \hat{\mathbf{a}}^\dagger & \hat{\mathbf{a}} \end{bmatrix} \mathbb{H}_\zeta \begin{bmatrix} \hat{\mathbf{a}} \\ \hat{\mathbf{a}}^\dagger \end{bmatrix} - \zeta \operatorname{Tr}\{\mathbf{A}\}. \tag{2.25}$$

We now wish to diagonalize this Hamiltonian. If we can find some invertible matrix $\mathbb{T}$ such that

$$\hat{H} = \begin{bmatrix} \hat{\mathbf{a}}^\dagger & \hat{\mathbf{a}} \end{bmatrix} \mathbb{H}_\zeta \begin{bmatrix} \hat{\mathbf{a}} \\ \hat{\mathbf{a}}^\dagger \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{a}}^\dagger & \hat{\mathbf{a}} \end{bmatrix} (\mathbb{T}^\dagger)^{-1} (\mathbb{T}^\dagger \mathbb{H}_\zeta \mathbb{T}) \mathbb{T}^{-1} \begin{bmatrix} \hat{\mathbf{a}} \\ \hat{\mathbf{a}}^\dagger \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{c}}^\dagger & \hat{\mathbf{c}} \end{bmatrix} \mathbb{D} \begin{bmatrix} \hat{\mathbf{c}} \\ \hat{\mathbf{c}}^\dagger \end{bmatrix}, \tag{2.26}$$

where $\mathbb{D} = \mathbb{T}^\dagger \mathbb{H}_\zeta \mathbb{T}$ is a diagonal matrix and

$$\begin{bmatrix} \hat{\mathbf{c}} \\ \hat{\mathbf{c}}^\dagger \end{bmatrix} = \begin{bmatrix} \hat{c}_1 & \dots & \hat{c}_N & \hat{c}_1^\dagger & \dots & \hat{c}_N^\dagger \end{bmatrix}^\mathsf{T} = \mathbb{T}^{-1} \begin{bmatrix} \hat{\mathbf{a}} \\ \hat{\mathbf{a}}^\dagger \end{bmatrix}, \tag{2.27}$$

where $\{\hat{c}_i\}$ are operators of the same type as $\{\hat{a}_i\}$, we are done. Normally, this is straightforward: we let the diagonal matrix $\mathbb{D}$ contain the eigenvalues of the matrix and construct $\mathbb{T}$ from the corresponding eigenvectors. Here, the situation is complicated by the fact that $\mathbb{T}$ must be canonical. This means that the transformation (2.27), which is known as a Bogoliubov transformation, must preserve the fermionic or bosonic nature of the operators $\{\hat{a}_i\}$,

such that the new operators $\{\hat{c}_i\}$ also obey (2.12), (2.13) and (2.14). This changes the normal diagonalization procedure. In particular, the anti-linear operator $\mathbb{J}$ in $\mathbb{C}^{2N}$, defined by

$$\mathbb{J}\begin{bmatrix}\mathbf{u}\\ \mathbf{v}\end{bmatrix} = \begin{bmatrix}\mathbf{v}^*\\ \mathbf{u}^*\end{bmatrix}; \quad \mathbf{u},\mathbf{v}\in\mathbb{C}^N, \tag{2.28}$$

is essential in the derivation of the canonical procedure. We will not show the derivation, but instead present the results in Theorem 2.1. The derivation is due to Van Hemmen [10].

Theorem 2.1 Suppose that the Hamiltonian of a given system can be written as

$$\hat{H} = \begin{bmatrix}\hat{\mathbf{a}}^\dagger & \hat{\mathbf{a}}\end{bmatrix}\mathbb{H}_\zeta\begin{bmatrix}\hat{\mathbf{a}}\\ \hat{\mathbf{a}}^\dagger\end{bmatrix} - \zeta\,\mathrm{Tr}\{\mathbf{A}\}, \tag{2.29}$$

where

$$\mathbb{H}_\zeta = \begin{bmatrix}\mathbf{A} & \mathbf{B}\\ \zeta\mathbf{B}^* & \zeta\mathbf{A}^*\end{bmatrix} \tag{2.30}$$

is a diagonalizable $2N\times 2N$ matrix. Assume that the matrix

$$\mathbb{M} = \begin{bmatrix}\mathbf{A} & \mathbf{B}\\ -\mathbf{B}^* & -\mathbf{A}^*\end{bmatrix} \tag{2.31}$$

has N orthonormal eigenvectors $\{\mathbf{x}_1,\ldots,\mathbf{x}_N\}$ with eigenvalues $\lambda^{(i)} \geq 0$. Then $\{\mathbb{J}\mathbf{x}_1,\ldots,\mathbb{J}\mathbf{x}_N\}$ is also a set of orthonormal eigenvectors of $\mathbb{M}$ with eigenvalues $-\lambda^{(i)}$, and the set $\{\mathbf{x}_1,\ldots,\mathbf{x}_N,\mathbb{J}\mathbf{x}_1,\ldots,\mathbb{J}\mathbf{x}_N\}$ forms an orthonormal basis for $\mathbb{C}^{2N}$. The canonical matrix

$$\mathbb{T} = \begin{bmatrix}\mathbf{x}_1 & \ldots & \mathbf{x}_N & \mathbb{J}\mathbf{x}_1 & \ldots & \mathbb{J}\mathbf{x}_N\end{bmatrix} \tag{2.32}$$

diagonalizes $\mathbb{H}_\zeta$, and the Hamiltonian may be written as

$$\hat{H} = \sum_{i=1}^{N}\left[2\lambda^{(i)}\hat{c}_i^\dagger\hat{c}_i + \zeta\lambda^{(i)}\right] - \zeta\,\mathrm{Tr}\{\mathbf{A}\}, \tag{2.33}$$

where the operators $\{\hat{c}_i\}$ are given by

$$\begin{bmatrix}\hat{\mathbf{c}}\\ \hat{\mathbf{c}}^\dagger\end{bmatrix} = \mathbb{T}^{-1}\begin{bmatrix}\hat{\mathbf{a}}\\ \hat{\mathbf{a}}^\dagger\end{bmatrix}. \tag{2.34}$$

For Theorem 2.1 to be of any use, the assumption that $\mathbb{H}_\zeta$ is diagonalizable must hold true. As shown by Van Hemmen [10], fermionic matrices are diagonalizable if they are Hermitian, while bosonic matrices are diagonalizable if they are positive definite. A Hermitian matrix is positive definite if and only if all its eigenvalues are positive. As long as this is the case, Theorem 2.1 provides a powerful tool for diagonalizing Hamiltonians without having to perform the Bogoliubov transformations directly.

2.2.2 The Bosonic Procedure

While Theorem 2.1 is completely general and a very useful result, it requires that the Hamiltonian be on the form (2.25), where the operator vector contains the annihilation operators $\{\hat{a}_i\}$ and all of their creation operator counterparts $\{\hat{a}_i^\dagger\}$. In practice, however, we may arrive at a Hamiltonian where the operator vector has unique operator entries. By unique operator entries we mean that the operator vector does not contain both a specific operator and its conjugate operator. This means that the Hamiltonian matrix is half the size of $\mathbb{H}_\zeta$ in (2.25). Instead of extending the operator vector and the matrix to accommodate Theorem 2.1, we would like to develop a procedure that lets us diagonalize these smaller matrices directly. Here, we will consider only such bosonic matrices, as these are of most relevance in this thesis. As for the general procedure, we follow Ushakov [6]. Let the Hamiltonian of the system be given by

$$\hat{H} = \hat{\mathbf{b}}^\dagger \mathbb{H} \hat{\mathbf{b}}, \tag{2.35}$$

where

$$\hat{\mathbf{b}} = \begin{bmatrix} \hat{b}_1 & \dots & \hat{b}_N & \hat{b}_{N+1}^\dagger & \dots & \hat{b}_{2N}^\dagger \end{bmatrix}^\mathsf{T} \tag{2.36}$$

is the bosonic operator vector and

$$\mathbb{H} = \begin{bmatrix} \mathbf{M}_1 & \mathbf{M}_2 \\ \mathbf{M}_3 & \mathbf{M}_4 \end{bmatrix} \tag{2.37}$$

is a $2N \times 2N$ matrix comprised of four $N \times N$ block matrices $\mathbf{M}_i$. Diagonalizing the matrix amounts to finding a transformation matrix $\mathbb{T}$ such that

$$\hat{H} = \hat{\mathbf{b}}^\dagger \mathbb{H} \hat{\mathbf{b}} = \hat{\mathbf{b}}^\dagger (\mathbb{T}^\dagger)^{-1} (\mathbb{T}^\dagger \mathbb{H} \mathbb{T}) \mathbb{T}^{-1} \hat{\mathbf{b}} = \hat{\mathbf{c}}^\dagger \mathbb{D} \hat{\mathbf{c}}, \tag{2.38}$$

where $\mathbb{D} = \mathbb{T}^\dagger \mathbb{H} \mathbb{T}$ and

$$\hat{\mathbf{c}} = \begin{bmatrix} \hat{c}_1 & \dots & \hat{c}_N & \hat{c}_{N+1}^\dagger & \dots & \hat{c}_{2N}^\dagger \end{bmatrix}^\mathsf{T} = \mathbb{T}^{-1} \hat{\mathbf{b}}. \tag{2.39}$$

We now introduce the $2N \times 2N$ para-unit matrix

$$\sigma_3 = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}, \tag{2.40}$$

where I is the $N \times N$ identity matrix. We denote the para-unit matrix σ_3 because it is the $2N \times 2N$ extension of the third Pauli matrix. The Pauli matrices and some of their properties are given in Appendix A. The commutation relations for the operator vector entries $\{\hat{\mathbf{b}}_i\}$ may then be written

$$[\hat{\mathbf{b}}_i, \hat{\mathbf{b}}_j^\dagger] = (\sigma_3)_{ij}. \tag{2.41}$$

This must also hold for the transformed operator vector entries $\{\hat{\mathbf{c}}_i\}$. Inserting (2.39) into (2.41), it is straightforward to show that the transformation matrix $\mathbb{T}$ must satisfy

$$\mathbb{T} \sigma_3 \mathbb{T}^\dagger = \mathbb{T}^\dagger \sigma_3 \mathbb{T} = \sigma_3. \tag{2.42}$$

Such a matrix is called pseudo-unitary.[1] With this relation established we turn to the diagonal matrix $\mathbb{D}$. We write

$$\mathbb{D} = \mathbb{T}^\dagger \mathbb{H} \mathbb{T} = \text{diag}(\lambda^{(1)}, \ldots, \lambda^{(N)}, -\lambda^{(N+1)}, \ldots, -\lambda^{(2N)}) = \sigma_3 \mathbb{L}, \tag{2.43}$$

where

$$\mathbb{L} = \text{diag}(\lambda^{(1)}, \ldots, \lambda^{(N)}, \lambda^{(N+1)}, \ldots, \lambda^{(2N)}). \tag{2.44}$$

Using the pseudo-unitarity (2.42), (2.43) may be rewritten as

$$\mathbb{H} \mathbb{T} = \sigma_3 \mathbb{T} \mathbb{L}. \tag{2.45}$$

Writing $\mathbb{T}$ in terms of its column vectors $\mathbb{T}^{(i)}$, (2.45) may be expressed as:

$$\mathbb{H} \begin{bmatrix} \mathbb{T}^{(1)} & \ldots & \mathbb{T}^{(2N)} \end{bmatrix} = \sigma_3 \begin{bmatrix} \lambda^{(1)} \mathbb{T}^{(1)} & \ldots & \lambda^{(2N)} \mathbb{T}^{(2N)} \end{bmatrix}. \tag{2.46}$$

Multiplying (2.46) by σ_3 from the left gives us the eigenvalue equations

$$\sigma_3 \mathbb{H} \mathbb{T}^{(i)} = \mathbb{K} \mathbb{T}^{(i)} = \lambda^{(i)} \mathbb{T}^{(i)}, \tag{2.47}$$

where we introduced the dynamical matrix $\mathbb{K} = \sigma_3 \mathbb{H}$. We have assumed here that the matrix $\mathbb{H}$ is diagonalizable. It can be shown that a $2N \times 2N$ Hermitian bosonic matrix $\mathbb{H}$ can be diagonalized by a pseudo-unitary matrix $\mathbb{T}$ into a matrix $\mathbb{D}$ with all diagonal elements positive if and only if $\mathbb{H}$ is positive definite [11]. We summarize our findings in the following Theorem:

Theorem 2.2 Suppose that the Hamiltonian of a given system can be written in terms of a $2N \times 2N$ Hermitian matrix $\mathbb{H}$ as

$$\hat{H} = \hat{\mathbf{b}}^\dagger \mathbb{H} \hat{\mathbf{b}}, \tag{2.48}$$

where

$$\hat{\mathbf{b}} = \begin{bmatrix} \hat{b}_1 & \ldots & \hat{b}_N & \hat{b}_{N+1}^\dagger & \ldots & \hat{b}_{2N}^\dagger \end{bmatrix}^\mathsf{T} \tag{2.49}$$

is a bosonic operator vector. If $\mathbb{H}$ is positive definite, it can be diagonalized by a pseudo-unitary matrix $\mathbb{T}$ into a matrix

$$\mathbb{D} = \mathbb{T}^\dagger \mathbb{H} \mathbb{T} = \text{diag}(\lambda^{(1)}, \ldots, \lambda^{(N)}, -\lambda^{(N+1)}, \ldots, -\lambda^{(2N)}), \tag{2.50}$$

where $\{\lambda^{(i)}\}$ are the eigenvalues of the dynamical matrix $\mathbb{K} = \sigma_3 \mathbb{H}$. The diagonalized Hamiltonian is then given by

$$\hat{H} = \hat{\mathbf{c}}^\dagger \mathbb{D} \hat{\mathbf{c}} = \sum_{i=1}^{N} \lambda^{(i)} \hat{c}_i^\dagger \hat{c}_i - \sum_{i=N+1}^{2N} \lambda^{(i)} \hat{c}_i^\dagger \hat{c}_i - \sum_{i=N+1}^{2N} \lambda^{(i)}, \tag{2.51}$$

where $\{\hat{c}_i\}$ are the canonically transformed operator entries of

$$\hat{\mathbf{c}} = \mathbb{T}^{-1} \hat{\mathbf{b}}. \tag{2.52}$$

[1] A unitary matrix $\mathbb{U}$ satisfies $\mathbb{U}\mathbb{U}^\dagger = \mathbb{U}^\dagger \mathbb{U} = I$.

All terms in (2.51) are positive since the positive definiteness of $\mathbb{H}$ guarantees that $\mathbb{D}$ has only positive diagonal elements.

2.3 The Mean Field Approximation

Transformations of the type (2.15) do not always yield Hamiltonians that are quadratic in fermionic or bosonic operators. In particular, if the transformed spin operators are quadratic, the second quantized Hamiltonian will be quartic. The Hamiltonian then describes interactions between particles, and we cannot use Theorem 2.1, or Theorem 2.2 if it is bosonic, to diagonalize it. However, if we can approximate these interactions as a mean field felt by each particle, we effectively have a single-particle problem that we can solve. Suppose that the Hamiltonian of a given system is

$$\hat{H} = \hat{A}\hat{B}. \tag{2.53}$$

If the operators $\hat{A}$ and $\hat{B}$ deviate only slightly from their average values $\langle\hat{A}\rangle$ and $\langle\hat{B}\rangle$, they can be used in an approximation scheme. To this end, define the deviation operators

$$\Delta\hat{A} = \hat{A} - \langle\hat{A}\rangle, \tag{2.54}$$

$$\Delta\hat{B} = \hat{B} - \langle\hat{B}\rangle, \tag{2.55}$$

which allow us to rewrite (2.53) as

$$\hat{H} = \hat{A}\langle\hat{B}\rangle + \hat{B}\langle\hat{A}\rangle - \langle\hat{A}\rangle\langle\hat{B}\rangle + \Delta\hat{A}\Delta\hat{B}. \tag{2.56}$$

Since the deviations are assumed to be small, the last term in (2.56) may be neglected, which gives us the mean field Hamiltonian

$$\hat{H}_{\mathrm{MF}} = \hat{A}\langle\hat{B}\rangle + \hat{B}\langle\hat{A}\rangle - \langle\hat{A}\rangle\langle\hat{B}\rangle. \tag{2.57}$$

The next step is then to determine what the average values, the mean field parameters, actually are. We do this by demanding that the mean field free energy of our system be minimized with respect to each of the mean field parameters [12, p. 65-71]. From statistical mechanics, we know that the free energy is given by

$$F = -\frac{1}{\beta}\ln Z, \tag{2.58}$$

where $\beta = 1/k_B T$, with k_B being Boltzmann's constant and T the temperature, and Z is the partition function:

$$Z = \mathrm{Tr}\Big\{e^{-\beta\hat{H}}\Big\}. \tag{2.59}$$

At finite temperatures, the average value of an operator $\hat{A}$ is given by

$$\langle\hat{A}\rangle = \mathrm{Tr}\Big\{\hat{\rho}\hat{A}\Big\} = \frac{1}{Z}\,\mathrm{Tr}\Big\{e^{-\beta\hat{H}}\hat{A}\Big\}, \tag{2.60}$$

with the density matrix $\hat{\rho}$ defined as

$$\hat{\rho} = \frac{e^{-\beta \hat{H}}}{Z}. \tag{2.61}$$

With these definitions, minimizing the free energy of the Hamiltonian (2.57) with respect to $\langle \hat{A} \rangle$ yields

$$\begin{aligned} 0 &= \frac{d}{d\langle \hat{A} \rangle} F_{\mathrm{MF}} = \frac{d}{d\langle \hat{A} \rangle} \left(-\frac{1}{\beta} \ln Z_{\mathrm{MF}} \right) = \frac{1}{Z_{\mathrm{MF}}} \mathrm{Tr} \left\{ e^{-\beta \hat{H}_{\mathrm{MF}}} \frac{d}{d\langle \hat{A} \rangle} \hat{H}_{\mathrm{MF}} \right\} \\ &= \frac{1}{Z_{\mathrm{MF}}} \mathrm{Tr} \Big\{ e^{-\beta \hat{H}_{\mathrm{MF}}} (\hat{B} - \langle \hat{B} \rangle) \Big\} = \langle \hat{B} \rangle - \langle \hat{B} \rangle, \end{aligned} \tag{2.62}$$

where we used the definition of the average value (2.60) in the last equality. Minimizing the mean field free energy with respect to $\langle \hat{B} \rangle$ yields a similar equation. These equations determine the mean field parameters self-consistently, which is why they are known as self-consistency equations.

3 Fermionization and Bosonization of the XY Model

We will now apply the techniques presented in Section 2 to a concrete spin model to familiarize ourselves with them. The model consists of N spins on a 1D lattice, and each spin is separated from its neighbors by a distance a. We study the system using periodic boundary conditions. The spins interact with their nearest neighbors through exchange interactions J_x and J_y and with a transverse magnetic field B in the z-direction. The Hamiltonian of the system is given by [13]

$$\hat{H} = -\sum_{i=1}^{N} \left[J_x \hat{S}_i^x \hat{S}_{i+1}^x + J_y \hat{S}_i^y \hat{S}_{i+1}^y + BS\hat{S}_i^z \right]. \tag{3.1}$$

We will first fermionize this model using the Jordan-Wigner transformation, which gives us the exact energy spectrum. Then we will consider three bosonization techniques: the Holstein-Primakoff transformation, the Dyson-Maleev transformation and the Schwinger boson transformation. For reasons that will soon be clear, the latter is of particular interest.

3.1 The Jordan-Wigner Transformation

The Jordan-Wigner transformation is given by

$$\boxed{\begin{aligned} \hat{S}_i^+ &= \hbar \hat{f}_i^\dagger e^{i\pi \sum_{j=1}^{i-1} \hat{n}_j}, \\ \hat{S}_i^- &= \hbar \hat{f}_i e^{-i\pi \sum_{j=1}^{i-1} \hat{n}_j}, \\ \hat{S}_i^z &= \hbar \left(\hat{n}_i - \frac{1}{2} \right), \end{aligned}} \tag{3.2}$$

where $\{\hat{f}_i\}$ are fermionic operators and $\hat{n}_i$ is the number operator. These operators do not represent physical quasiparticles, but they allow us to solve for the exact eigenvalues of the Hamiltonian. Before we transform the spin Hamiltonian, it is useful to establish the operator relations given in Reference [14, p. 73]. Since fermionic operators follow anticommutation relations, we have

$$[e^{i\pi\hat{n}_j}, \hat{f}_i] = [e^{i\pi\hat{n}_j}, \hat{f}_i^\dagger] = 0 \tag{3.3}$$

for $j \neq i$. We also have

$$\{e^{i\pi\hat{n}_i}, \hat{f}_i\} = e^{i\pi\hat{n}_i}\hat{f}_i + \hat{f}_i e^{i\pi\hat{n}_i} = 0. \tag{3.4}$$

To see this, recall that the Pauli exclusion principle forbids two identical fermions from occupying the same quantum state. This means that the only allowed quantum states associated with lattice site i are $|0\rangle$ and $|1\rangle$, the unoccupied and occupied states. It is easy to see that $\{e^{i\pi\hat{n}_i}, \hat{f}_i\}$ gives zero when acting on $|0\rangle$. When $\{e^{i\pi\hat{n}_i}, \hat{f}_i\}$ acts on $|1\rangle$, we get

$$(e^{i\pi\hat{n}_i}\hat{f}_i + \hat{f}_i e^{i\pi\hat{n}_i})\,|1\rangle = e^{i\pi\hat{n}_i}\hat{f}_i\,|1\rangle + \hat{f}_i e^{i\pi\hat{n}_i}\,|1\rangle = e^{i\pi\hat{n}_i}\,|0\rangle - \hat{f}_i\,|1\rangle = |0\rangle - |0\rangle = 0. \tag{3.5}$$

Since the eigenvalues of $\hat{n}_i$ are 0 and 1, $e^{i\pi\hat{n}_i} = e^{-i\pi\hat{n}_i}$ is Hermitian, and we get

$$\{e^{i\pi\hat{n}_i}, \hat{f}_i^\dagger\} = 0 \tag{3.6}$$

by taking the conjugate of (3.4). From (3.4) and (3.6) it follows that

$$e^{i\pi\hat{n}_i}\hat{f}_i = \hat{f}_i; \quad \hat{f}_i e^{i\pi\hat{n}_i} = -\hat{f}_i; \quad e^{i\pi\hat{n}_i}\hat{f}_i^\dagger = -\hat{f}_i^\dagger; \quad \hat{f}_i^\dagger e^{i\pi\hat{n}_i} = \hat{f}_i^\dagger. \tag{3.7}$$

Using (2.6) and (3.2), we get

$$\begin{aligned}
\hat{S}_i^x \hat{S}_{i+1}^x &= \frac{\hbar^2}{4}\left[\hat{f}_i^\dagger e^{i\pi\sum_{j=1}^{i-1}\hat{n}_j} + \hat{f}_i e^{-i\pi\sum_{j=1}^{i-1}\hat{n}_j}\right]\left[\hat{f}_{i+1}^\dagger e^{i\pi\sum_{j=1}^{i}\hat{n}_j} + \hat{f}_{i+1} e^{-i\pi\sum_{j=1}^{i}\hat{n}_j}\right] \\
&= \frac{\hbar^2}{4}\left[\hat{f}_i^\dagger e^{i\pi\hat{n}_i}\hat{f}_{i+1}^\dagger + \hat{f}_i^\dagger e^{i\pi\hat{n}_i}\hat{f}_{i+1} + \hat{f}_i e^{i\pi\hat{n}_i}\hat{f}_{i+1}^\dagger + \hat{f}_i e^{i\pi\hat{n}_i}\hat{f}_{i+1}\right] \\
&= \frac{\hbar^2}{4}\left[\hat{f}_i^\dagger \hat{f}_{i+1}^\dagger + \hat{f}_i^\dagger \hat{f}_{i+1} - \hat{f}_i \hat{f}_{i+1}^\dagger - \hat{f}_i \hat{f}_{i+1}\right] \\
&= \frac{\hbar^2}{4}\left[\hat{f}_i^\dagger \hat{f}_{i+1}^\dagger + \hat{f}_i^\dagger \hat{f}_{i+1} + \hat{f}_{i+1}^\dagger \hat{f}_i + \hat{f}_{i+1}\hat{f}_i\right]
\end{aligned} \tag{3.8}$$

and

$$\begin{aligned}
\hat{S}_i^y \hat{S}_{i+1}^y &= -\frac{\hbar^2}{4}\left[\hat{f}_i^\dagger e^{i\pi\sum_{j=1}^{i-1}\hat{n}_j} - \hat{f}_i e^{-i\pi\sum_{j=1}^{i-1}\hat{n}_j}\right]\left[\hat{f}_{i+1}^\dagger e^{i\pi\sum_{j=1}^{i}\hat{n}_j} - \hat{f}_{i+1} e^{-i\pi\sum_{j=1}^{i}\hat{n}_j}\right] \\
&= -\frac{\hbar^2}{4}\left[\hat{f}_i^\dagger e^{i\pi\hat{n}_i}\hat{f}_{i+1}^\dagger - \hat{f}_i^\dagger e^{i\pi\hat{n}_i}\hat{f}_{i+1} - \hat{f}_i e^{i\pi\hat{n}_i}\hat{f}_{i+1}^\dagger + \hat{f}_i e^{i\pi\hat{n}_i}\hat{f}_{i+1}\right] \\
&= -\frac{\hbar^2}{4}\left[\hat{f}_i^\dagger \hat{f}_{i+1}^\dagger - \hat{f}_i^\dagger \hat{f}_{i+1} + \hat{f}_i \hat{f}_{i+1}^\dagger - \hat{f}_i \hat{f}_{i+1}\right] \\
&= \frac{\hbar^2}{4}\Big[- \hat{f}_i^\dagger \hat{f}_{i+1}^\dagger + \hat{f}_i^\dagger \hat{f}_{i+1} + \hat{f}_{i+1}^\dagger \hat{f}_i - \hat{f}_{i+1}\hat{f}_i\Big].
\end{aligned} \tag{3.9}$$

The Hamiltonian (3.1) may then be written as

$$\begin{aligned}
\hat{H} &= -\sum_{i=1}^{N}\Bigg[\frac{\hbar^2}{4}(J_x + J_y)(\hat{f}_{i+1}^\dagger \hat{f}_i + \hat{f}_i^\dagger \hat{f}_{i+1}) + \frac{\hbar^2}{4}(J_x - J_y)(\hat{f}_i^\dagger \hat{f}_{i+1}^\dagger + \hat{f}_{i+1}\hat{f}_i) \\
&\qquad + BS\hbar\left(\hat{f}_i^\dagger \hat{f}_i - \frac{1}{2}\right)\Bigg] \\
&= \frac{NBS\hbar}{2} - \sum_{i=1}^{N}\Bigg[\frac{\hbar^2}{4}(J_x + J_y)(\hat{f}_{i+1}^\dagger \hat{f}_i + \hat{f}_i^\dagger \hat{f}_{i+1}) + \frac{\hbar^2}{4}(J_x - J_y)(\hat{f}_i^\dagger \hat{f}_{i+1}^\dagger + \hat{f}_{i+1}\hat{f}_i) \\
&\qquad + BS\hbar \hat{f}_i^\dagger \hat{f}_i\Bigg].
\end{aligned} \tag{3.10}$$

We drop the constant term in the following. We see that the Hamiltonian contains products of operators acting on different lattice sites. Due to the periodicity of the lattice, we can decouple these products using the Fourier transform

$$\hat{f}_j = \frac{1}{\sqrt{N}}\sum_k e^{ikja}\hat{f}_k, \tag{3.11}$$

where the sum over k is taken over the first Brillouin zone of the 1D reciprocal lattice. Inserting (3.11) into (3.10) and using the orthogonality of complex exponentials

$$\sum_{j=1}^{N} e^{i(k-k')ja} = N\delta_{kk'}, \tag{3.12}$$

we obtain:

$$\begin{aligned}\hat{H} = -\sum_k \Bigg[&\frac{\hbar^2}{4}(J_x + J_y)(e^{-ika}\hat{f}_k^\dagger \hat{f}_k + e^{ika}\hat{f}_k^\dagger \hat{f}_k) + \frac{\hbar^2}{4}(J_x - J_y)(e^{ika}\hat{f}_k^\dagger \hat{f}_{-k}^\dagger + e^{-ika}\hat{f}_k \hat{f}_{-k}) \\ &+ BS\hbar \hat{f}_k^\dagger \hat{f}_k \Bigg].\end{aligned} \tag{3.13}$$

We can isolate the operator terms by making use of the following relations:

$$\sum_k (e^{-ika}\hat{f}_k^\dagger \hat{f}_k + e^{ika}\hat{f}_k^\dagger \hat{f}_k) = \sum_k (e^{-ika} + e^{ika})\hat{f}_k^\dagger \hat{f}_k = 2\sum_{k>0} \cos(ka)(\hat{f}_k^\dagger \hat{f}_k + \hat{f}_{-k}^\dagger \hat{f}_{-k}), \tag{3.14}$$

$$\sum_k e^{ika}\hat{f}_k^\dagger \hat{f}_{-k}^\dagger = \sum_{k>0} (e^{ika} - e^{-ika})\hat{f}_k^\dagger \hat{f}_{-k}^\dagger = 2i\sum_{k>0} \sin(ka)\hat{f}_k^\dagger \hat{f}_{-k}^\dagger, \tag{3.15}$$

where the sum is now taken only over the positive k values in the first Brillouin zone. The Hamiltonian may then be written

$$\begin{aligned}\hat{H} &= \sum_{k>0} \left[A_k(\hat{f}_k^\dagger \hat{f}_k + \hat{f}_{-k}^\dagger \hat{f}_{-k}) + B_k(\hat{f}_k^\dagger \hat{f}_{-k}^\dagger - \hat{f}_k \hat{f}_{-k})\right] \\ &= \sum_{k>0} \left[A_k(\hat{f}_k^\dagger \hat{f}_k + \hat{f}_{-k}^\dagger \hat{f}_{-k}) + \frac{B_k}{2}(\hat{f}_k^\dagger \hat{f}_{-k}^\dagger - \hat{f}_{-k}^\dagger \hat{f}_k^\dagger - \hat{f}_k \hat{f}_{-k} + \hat{f}_{-k}\hat{f}_k)\right],\end{aligned} \tag{3.16}$$

where we defined

$$A_k = -\frac{\hbar^2}{2}(J_x + J_y)\cos(ka) - BS\hbar, \tag{3.17}$$

$$B_k = -\frac{i\hbar^2}{2}(J_x - J_y)\sin(ka). \tag{3.18}$$

Letting labels k and $-k$ correspond to labels 1 and 2, we see that each term in the sum of (3.16) is of the form (2.22), with non-zero matrix elements

$$A_{11} = A_{22} = \frac{A_k}{2}; \quad B_{12} = \frac{B_k}{2}; \quad B_{21} = -\frac{B_k}{2}. \tag{3.19}$$

We can then use Theorem 2.1 to diagonalize the Hamiltonian. The matrix $\mathbb{M}$ is given by

$$\mathbb{M} = \begin{bmatrix} A_k/2 & 0 & 0 & B_k/2 \\ 0 & A_k/2 & -B_k/2 & 0 \\ 0 & -B_k^*/2 & -A_k/2 & 0 \\ B_k^*/2 & 0 & 0 & -A_k/2 \end{bmatrix}, \tag{3.20}$$

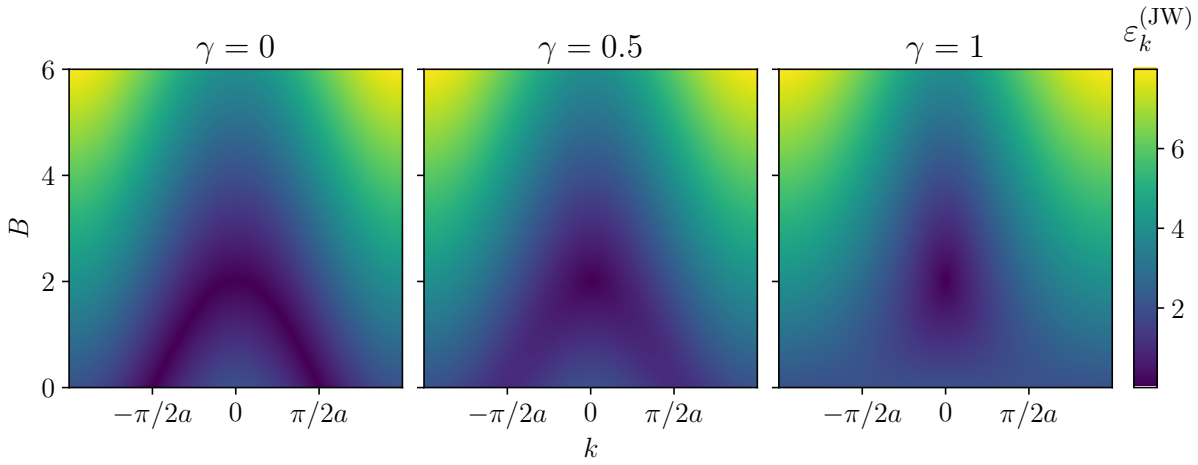


Figure 1: Energy dispersion for $\gamma = 0$, $\gamma = 0.5$ and $\gamma = 1$ obtained using the Jordan-Wigner transformation. The other parameters are $\hbar = 1$, $S = 1/2$ and $J = 4$.

and its eigenvalues are

$$\lambda_k^{(1)}, \lambda_k^{(2)} = \frac{1}{2}\sqrt{A_k^2 + |B_k|^2}, \tag{3.21}$$

$$\lambda_k^{(3)}, \lambda_k^{(4)} = -\frac{1}{2}\sqrt{A_k^2 + |B_k|^2}. \tag{3.22}$$

The diagonalized Hamiltonian becomes

$$\hat{H} = \sum_{k>0} \varepsilon_k^{(\mathrm{JW})} \left[\hat{\phi}_k^\dagger \hat{\phi}_k + \hat{\phi}_{-k}^\dagger \hat{\phi}_{-k} - 1 \right], \tag{3.23}$$

with the dispersion given by

$$\varepsilon_k^{(\mathrm{JW})} = \frac{\hbar^2 J}{2} \left[\left(\cos(ka) + \frac{2BS}{\hbar J} \right)^2 + \gamma^2 \sin^2(ka) \right]^{1/2}. \tag{3.24}$$

Here, J and γ have been introduced via

$$J_x = J\frac{1+\gamma}{2}; \quad J_y = J\frac{1-\gamma}{2}. \tag{3.25}$$

The energy dispersion (3.24) for three different values of γ is shown in Figure 1.

3.2 The Holstein-Primakoff Transformation

The Holstein-Primakoff transformation is given by

$$\boxed{\begin{aligned} \hat{S}_i^+ &= \hbar\sqrt{2S - \hat{b}_i^\dagger \hat{b}_i}\,\hat{b}_i, \\ \hat{S}_i^- &= \hbar\hat{b}_i^\dagger \sqrt{2S - \hat{b}_i^\dagger \hat{b}_i}, \\ \hat{S}_i^z &= \hbar(S - \hat{b}_i^\dagger \hat{b}_i), \end{aligned}} \tag{3.26}$$

where $\{\hat{b}_i\}$ are bosonic operators. These operators represent physical spin excitations known as magnons. Unlike the Jordan-Wigner transformation, the Holstein-Primakoff transformation cannot be used to obtain the exact energy spectrum of the XY model. However, if we consider the ground state to be highly ordered, such that the number of magnons is small compared to the magnitude of the spin S, we may expand the square roots in (3.26) to obtain an approximate result for the ground state energy. In the case of a strong magnetic field B, we expect the spins of the XY chain to align parallel to the field. However, if the field is weak, the spin couplings come into play and the ground state spin configuration may be tilted with respect to the z-axis. We will assume here that $J_x > |J_y|$. Then, due to the symmetry of the system, each spin will be tilted by the same angle θ in the xz-plane. In order to study both cases simultaneously, we will employ the spin rotation technique described by Haraldsen and Fishman [15]. First, we rotate the local reference frame of each spin to express the spin operators $\hat{S}_i^\alpha$ in terms of the local spin operators $\hat{\tilde{S}}_i^\alpha$:

$$\hat{S}_i^x = \hat{\tilde{S}}_i^x \cos(\theta) + \hat{\tilde{S}}_i^z \sin(\theta), \tag{3.27}$$

$$\hat{S}_i^y = \hat{\tilde{S}}_i^y, \tag{3.28}$$

$$\hat{S}_i^z = \hat{\tilde{S}}_i^z \cos(\theta) - \hat{\tilde{S}}_i^x \sin(\theta). \tag{3.29}$$

We then apply the Holstein-Primakoff transformation to the local spin operators. Using (2.6) and expanding the square roots in (3.26) to lowest order, the local spin operators in the x- and y-directions may be approximated as

$$\hat{\tilde{S}}_i^x \approx \hbar\sqrt{\frac{S}{2}}\big(\hat{b}_i + \hat{b}_i^\dagger\big), \tag{3.30}$$

$$\hat{\tilde{S}}_i^y \approx -i\hbar\sqrt{\frac{S}{2}}\big(\hat{b}_i - \hat{b}_i^\dagger\big). \tag{3.31}$$

Inserting (3.27), (3.28) and (3.29) into (3.1) then produces an approximate Hamiltonian that may be separated into four different terms:

$$\hat{H} \approx E_0 + \hat{H}_1 + \hat{H}_2 + \hat{H}_3, \tag{3.32}$$

where E_0 is the classical ground state energy and $\hat{H}_n$ is of order n in the bosonic operators:

$$E_0 = -N\hbar S^2[\hbar J_x \sin^2(\theta) + B\cos(\theta)], \tag{3.33}$$

$$\hat{H}_1 = -\sum_{i=1}^{N} \hbar\sqrt{2S^3}\sin(\theta)\left(\hbar J_x\cos(\theta) - \frac{B}{2}\right)(\hat{b}_i + \hat{b}_i^\dagger), \tag{3.34}$$

$$\begin{aligned}\hat{H}_2 = -\sum_{i=1}^{N}\Bigg[&\frac{\hbar^2 S}{2}(J_x\cos^2(\theta) - J_y)(\hat{b}_i\hat{b}_{i+1} + \hat{b}_i^\dagger\hat{b}_{i+1}^\dagger) + \frac{\hbar^2 S}{2}(J_x\cos^2(\theta) + J_y)(\hat{b}_i\hat{b}_{i+1}^\dagger + \hat{b}_i^\dagger\hat{b}_{i+1}) \\ &- [2\hbar^2 J_x S\sin^2(\theta) + BS\hbar\cos(\theta)]\hat{b}_i^\dagger\hat{b}_i\Bigg],\end{aligned} \tag{3.35}$$

$$\hat{H}_3 = \sum_{i=1}^{N} \hbar^2 J_x \sqrt{\frac{S}{2}} \sin(\theta)\cos(\theta)(\hat{b}_i\hat{b}_{i+1}^\dagger\hat{b}_{i+1} + \hat{b}_i^\dagger\hat{b}_{i+1}^\dagger\hat{b}_{i+1} + \hat{b}_i^\dagger\hat{b}_i\hat{b}_{i+1} + \hat{b}_i^\dagger\hat{b}_i\hat{b}_{i+1}^\dagger). \tag{3.36}$$

The lowest energy spin configuration is the one that minimizes the classical ground state energy E_0. Minimizing (3.33) with respect to θ yields the following equation:

$$\sin(\theta)\left(\hbar J_x \cos(\theta) - \frac{B}{2}\right) = 0, \tag{3.37}$$

from which we deduce that the ground state spin configuration is described by

$$\theta = \begin{cases} 0 & \text{if } B \geq 2\hbar J_x, \\ \arccos(B/2\hbar J_x) & \text{if } B < 2\hbar J_x. \end{cases} \tag{3.38}$$

We see from (3.37) that when the system is in the ground state, the non-physical first order Hamiltonian $\hat{H}_1$, which describes creation and annihilation of magnons from the vacuum, vanishes. We also note that the third order Hamiltonian $\hat{H}_3$ is present only when the spin system is tilted. We are interested in the quadratic Hamiltonian $\hat{H}_2$, however, which describes the non-interacting spin dynamics. As for the Jordan-Wigner transformation, we use the Fourier transform to decouple the operators acting on different lattice sites and obtain

$$\begin{aligned}\hat{H}_2 = \sum_{k>0} \Big[&[2\hbar^2 J_x S \sin^2(\theta) + BS\hbar\cos(\theta) - \hbar^2 S(J_x\cos^2(\theta) + J_y)\cos(ka)](\hat{b}_k^\dagger\hat{b}_k + \hat{b}_{-k}^\dagger\hat{b}_{-k}) \\ &+ \hbar^2 S(J_y - J_x\cos^2(\theta))\cos(ka)(\hat{b}_k\hat{b}_{-k} + \hat{b}_k^\dagger\hat{b}_{-k}^\dagger)\Big].\end{aligned} \tag{3.39}$$

Defining

$$A_k = 2\hbar^2 J_x S\sin^2(\theta) + BS\hbar\cos(\theta) - \hbar^2 S(J_x\cos^2(\theta) + J_y)\cos(ka), \tag{3.40}$$

$$B_k = \hbar^2 S(J_y - J_x\cos^2(\theta))\cos(ka), \tag{3.41}$$

we may write (3.39) on the form (2.22) as

$$\hat{H}_2 = \sum_{k>0}\left[A_k(\hat{b}_k^\dagger\hat{b}_k + \hat{b}_{-k}^\dagger\hat{b}_{-k}) + \frac{B_k}{2}(\hat{b}_k\hat{b}_{-k} + \hat{b}_{-k}\hat{b}_k + \hat{b}_k^\dagger\hat{b}_{-k}^\dagger + \hat{b}_{-k}^\dagger\hat{b}_k^\dagger)\right]. \tag{3.42}$$

With k and $-k$ corresponding to labels 1 and 2, we identify the non-zero matrix elements of each term in the sum as

$$A_{11} = A_{22} = \frac{A_k}{2}; \quad B_{12} = B_{21} = \frac{B_k}{2}. \tag{3.43}$$

The $\mathbb{M}$-matrix is then given by

$$\mathbb{M} = \begin{bmatrix} A_k/2 & 0 & 0 & B_k/2 \\ 0 & A_k/2 & B_k/2 & 0 \\ 0 & -B_k/2 & -A_k/2 & 0 \\ -B_k/2 & 0 & 0 & -A_k/2 \end{bmatrix}, \tag{3.44}$$

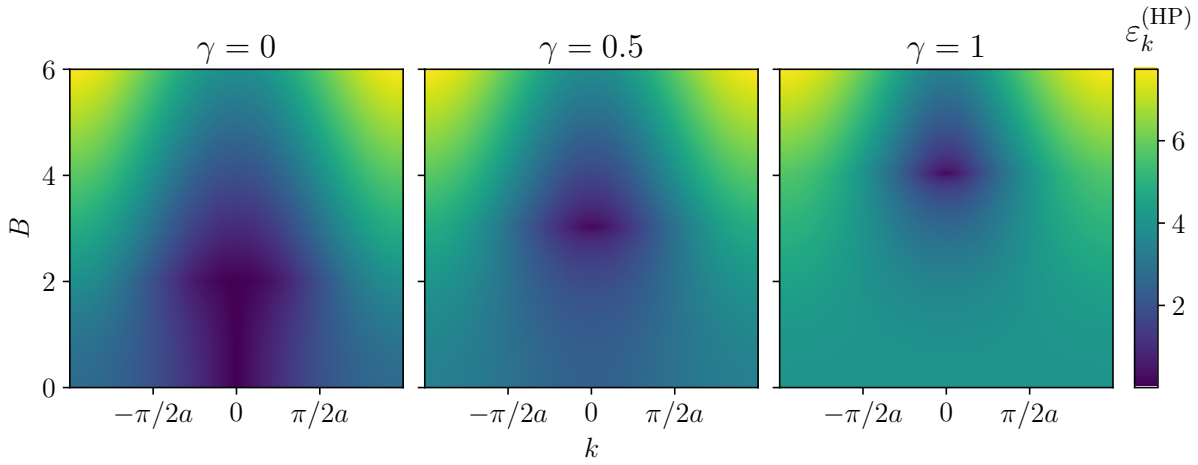


Figure 2: Energy dispersion for $\gamma = 0$, $\gamma = 0.5$ and $\gamma = 1$ obtained using the Holstein-Primakoff transformation. The other parameters are $\hbar = 1$, $S = 1/2$ and $J = 4$.

with eigenvalues

$$\lambda_k^{(1)}, \lambda_k^{(2)} = \frac{1}{2}[(A_k + B_k)(A_k - B_k)]^{1/2}, \tag{3.45}$$

$$\lambda_k^{(3)}, \lambda_k^{(4)} = -\frac{1}{2}[(A_k + B_k)(A_k - B_k)]^{1/2}. \tag{3.46}$$

Inserting the two ground state configuration angles (3.38) into A_k and B_k, we end up with the following diagonalized Hamiltonian:

$$\hat{H}_2 = \sum_{k>0} \varepsilon_k^{(\mathrm{HP})} \left[\hat{\beta}_k^\dagger \hat{\beta}_k + \hat{\beta}_{-k}^\dagger \hat{\beta}_{-k} - 1 \right], \tag{3.47}$$

where

$$\varepsilon_k^{(\mathrm{HP})} = \begin{cases} BS\hbar \left[\left(1 - \dfrac{2\hbar J_x}{B} \cos(ka) \right) \left(1 - \dfrac{2\hbar J_y}{B} \cos(ka) \right) \right]^{1/2} & \text{if } B \geq 2\hbar J_x, \\ 2S\hbar^2 J_x \left[\left(1 - \left(\dfrac{B}{2\hbar J_x} \right)^2 \cos(ka) \right) \left(1 - \dfrac{J_y}{J_x} \cos(ka) \right) \right]^{1/2} & \text{if } B < 2\hbar J_x. \end{cases} \tag{3.48}$$

The energy dispersion (3.48) for three different values of γ is shown in Figure 2. As we can see, there is qualitative agreement between Figures 1 and 2, indicating that the Holstein-Primakoff transformation produces a decent approximation of the exact ground state.

3.3 The Dyson-Maleev Transformation

The Dyson-Maleev transformation is given by

$$\boxed{\begin{aligned}\hat{S}_i^+ &= \hbar\sqrt{2S}\left(1-\frac{\hat{b}_i^\dagger\hat{b}_i}{2S}\right)\hat{b}_i,\\ \hat{S}_i^- &= \hbar\sqrt{2S}\hat{b}_i^\dagger,\\ \hat{S}_i^z &= \hbar(S-\hat{b}_i^\dagger\hat{b}_i),\end{aligned}} \tag{3.49}$$

where $\{\hat{b}_i\}$ are bosonic operators representing magnons. The Dyson-Maleev transformation is reminiscent of the Holstein-Primakoff transformation, and in fact, it can be shown that the Hamiltonians obtained using the two transformations are related by a non-linear operator transformation [16]. There are no operators in square roots in (3.49), so we avoid having to do the perturbative expansion that we did for the Holstein-Primakoff transformation. This means that the Hamiltonian obtained using the Dyson-Maleev transformation is exact. The downside to this transformation is that it is not Hermitian, which is evident from the definitions of $\hat{S}^+$ and $\hat{S}^-$. We consider the same physical situation that we did for the Holstein-Primakoff transformation. We therefore employ the spin rotation technique described previously for this case as well. By (2.6) and (3.49), the local spin operators in the x- and y-directions are given by

$$\hat{\tilde{S}}_i^x = \hbar\sqrt{\frac{S}{2}}\left(\hat{b}_i+\hat{b}_i^\dagger-\frac{\hat{b}_i^\dagger\hat{b}_i\hat{b}_i}{2S}\right), \tag{3.50}$$

$$\hat{\tilde{S}}_i^y = -i\hbar\sqrt{\frac{S}{2}}\left(\hat{b}_i-\hat{b}_i^\dagger-\frac{\hat{b}_i^\dagger\hat{b}_i\hat{b}_i}{2S}\right). \tag{3.51}$$

Since the spin operators $\hat{S}_i^\alpha$ are given by (3.27), (3.28) and (3.29), we see from (3.50) and (3.51) that the exact Hamiltonian will contain terms up to sixth order in the bosonic operators. We write

$$\hat{H} = E_0+\hat{H}_1+\hat{H}_2+\hat{H}_3+\hat{H}_4+\hat{H}_5+\hat{H}_6, \tag{3.52}$$

where, like before, E_0 is the classical ground state energy and $\hat{H}_n$ is of order n in the bosonic operators:

$$E_0 = -N\hbar S^2[\hbar J_x\sin^2(\theta)+B\cos(\theta)], \tag{3.53}$$

$$\hat{H}_1 = -\sum_{i=1}^{N}\hbar\sqrt{2S^3}\sin(\theta)\left(\hbar J_x\cos(\theta)-\frac{B}{2}\right)(\hat{b}_i+\hat{b}_i^\dagger), \tag{3.54}$$

$$\begin{aligned}\hat{H}_2 = -\sum_{i=1}^{N}\Bigg[&\frac{\hbar^2 S}{2}(J_x\cos^2(\theta)-J_y)(\hat{b}_i\hat{b}_{i+1}+\hat{b}_i^\dagger\hat{b}_{i+1}^\dagger)+\frac{\hbar^2 S}{2}(J_x\cos^2(\theta)+J_y)(\hat{b}_i\hat{b}_{i+1}^\dagger+\hat{b}_i^\dagger\hat{b}_{i+1})\\ &-[2\hbar^2 J_x S\sin^2(\theta)+BS\hbar\cos(\theta)]\hat{b}_i^\dagger\hat{b}_i\Bigg],\end{aligned} \tag{3.55}$$

$$\hat{H}_3 = \sum_{i=1}^{N} \Bigg[\hbar^2 J_x \sqrt{\frac{S}{2}} \sin(\theta)\cos(\theta)(\hat{b}_i\hat{b}_{i+1}^\dagger\hat{b}_{i+1} + \hat{b}_i^\dagger\hat{b}_{i+1}^\dagger\hat{b}_{i+1} + \hat{b}_i^\dagger\hat{b}_i\hat{b}_{i+1} + \hat{b}_i^\dagger\hat{b}_i\hat{b}_{i+1}^\dagger + \hat{b}_i^\dagger\hat{b}_i\hat{b}_i) - \frac{\hbar B}{2}\sqrt{\frac{S}{2}}\sin(\theta)\hat{b}_i^\dagger\hat{b}_i\hat{b}_i \Bigg], \tag{3.56}$$

$$\hat{H}_4 = \sum_{i=1}^{N} \Bigg[\frac{\hbar^2}{4}(J_x\cos^2(\theta) - J_y)(\hat{b}_i\hat{b}_{i+1}^\dagger\hat{b}_{i+1}\hat{b}_{i+1} + \hat{b}_i^\dagger\hat{b}_i\hat{b}_i\hat{b}_{i+1}) + \frac{\hbar^2}{4}(J_x\cos^2(\theta) + J_y)(\hat{b}_i^\dagger\hat{b}_{i+1}^\dagger\hat{b}_{i+1}\hat{b}_{i+1} + \hat{b}_i^\dagger\hat{b}_i\hat{b}_i\hat{b}_{i+1}^\dagger) - \hbar^2 J_x\sin^2(\theta)\hat{b}_i^\dagger\hat{b}_i\hat{b}_{i+1}^\dagger\hat{b}_{i+1} \Bigg], \tag{3.57}$$

$$\hat{H}_5 = -\sum_{i=1}^{N} \frac{\hbar^2 J_x}{\sqrt{8S}}\sin(\theta)\cos(\theta)(\hat{b}_i^\dagger\hat{b}_i\hat{b}_i\hat{b}_{i+1}^\dagger\hat{b}_{i+1} + \hat{b}_i^\dagger\hat{b}_i\hat{b}_{i+1}^\dagger\hat{b}_{i+1}\hat{b}_{i+1}), \tag{3.58}$$

$$\hat{H}_6 = -\sum_{i=1}^{N} \frac{\hbar^2}{8S}(J_x\cos^2(\theta) - J_y)\hat{b}_i^\dagger\hat{b}_i\hat{b}_i\hat{b}_{i+1}^\dagger\hat{b}_{i+1}\hat{b}_{i+1}. \tag{3.59}$$

We see that (3.53), (3.54) and (3.55) are identical to (3.33), (3.34) and (3.35), respectively. Thus we conclude that, to second order in the bosonic operators, the Dyson-Maleev transformation yields exactly the same results for the ground state configuration and spin dynamics as the Holstein-Primakoff transformation. We note, however, that (3.56) has terms proportional to $\hat{b}_i^\dagger\hat{b}_i\hat{b}_i$ which are not present in (3.36). This is because the expansion of the square roots in (3.26) was done to lowest order, and we neglected a third order bosonic operator term in the resulting spin operators since our main concern is the quadratic part of the Hamiltonian.[2] Note that the Hamiltonian terms that are odd in the bosonic operators are only present when the spin system is tilted.

3.4 The Schwinger Boson Transformation

The three transformations considered so far have all been of an anisotropic nature, as the z-direction is given special attention. For the XY model, this is advantageous, but there exist systems, such as quantum spin liquids, where no direction is special. This motivates us to find a more isotropic transformation. An example of this is the Schwinger boson transformation. The Schwinger boson transformation, defined for spin-1/2 systems,[3] is given by

$$\boxed{\hat{\vec{S}} = \frac{\hbar}{2}\begin{bmatrix}\hat{b}_1^\dagger & \hat{b}_2^\dagger\end{bmatrix}\vec{\sigma}\begin{bmatrix}\hat{b}_1 \\ \hat{b}_2\end{bmatrix},} \tag{3.60}$$

[2] Even if we had kept the third order bosonic operator term, the third order Holstein-Primakoff Hamiltonian would not be identical to the third order Dyson-Maleev Hamiltonian.

[3] It is possible to generalize this approach to other spin values, but we will focus on the $S = 1/2$ case here.

where $\vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ is the Pauli matrix vector. The Fock space of the Schwinger bosons is constrained by

$$\hat{b}_1^\dagger \hat{b}_1 + \hat{b}_2^\dagger \hat{b}_2 = \hat{n}_1 + \hat{n}_2 = 2S. \tag{3.61}$$

This means that there exists exactly one such boson for a spin-1/2 particle in a given state, which means that we may interpret these as spin up and spin down bosons. For this reason, we will write $\hat{b}_1 = \hat{b}_\uparrow$ and $\hat{b}_2 = \hat{b}_\downarrow$ in the remainder of the thesis. We will now apply the Schwinger boson transformation to the XY model. Using the definition of the Pauli matrices given in Appendix A, the components of the spin operator $\hat{\vec{S}}_i$ are given by

$$\hat{S}_i^x = \frac{\hbar}{2}(\hat{b}_{i\uparrow}^\dagger \hat{b}_{i\downarrow} + \hat{b}_{i\downarrow}^\dagger \hat{b}_{i\uparrow}), \tag{3.62}$$

$$\hat{S}_i^y = \frac{i\hbar}{2}(\hat{b}_{i\downarrow}^\dagger \hat{b}_{i\uparrow} - \hat{b}_{i\uparrow}^\dagger \hat{b}_{i\downarrow}), \tag{3.63}$$

$$\hat{S}_i^z = \frac{\hbar}{2}(\hat{b}_{i\uparrow}^\dagger \hat{b}_{i\uparrow} - \hat{b}_{i\downarrow}^\dagger \hat{b}_{i\downarrow}). \tag{3.64}$$

The spin component products in the Hamiltonian (3.1) are

$$\hat{S}_i^x \hat{S}_{i+1}^x = \frac{\hbar^2}{4}\left[\hat{b}_{i\uparrow}^\dagger \hat{b}_{i+1\uparrow}^\dagger \hat{b}_{i\downarrow}\hat{b}_{i+1\downarrow} + \hat{b}_{i\uparrow}^\dagger \hat{b}_{i+1\downarrow}^\dagger \hat{b}_{i\downarrow}\hat{b}_{i+1\uparrow} + \hat{b}_{i\downarrow}^\dagger \hat{b}_{i+1\uparrow}^\dagger \hat{b}_{i\uparrow}\hat{b}_{i+1\downarrow} + \hat{b}_{i\downarrow}^\dagger \hat{b}_{i+1\downarrow}^\dagger \hat{b}_{i\uparrow}\hat{b}_{i+1\uparrow}\right] \tag{3.65}$$

and

$$\hat{S}_i^y \hat{S}_{i+1}^y = -\frac{\hbar^2}{4}\left[\hat{b}_{i\uparrow}^\dagger \hat{b}_{i+1\uparrow}^\dagger \hat{b}_{i\downarrow}\hat{b}_{i+1\downarrow} - \hat{b}_{i\uparrow}^\dagger \hat{b}_{i+1\downarrow}^\dagger \hat{b}_{i\downarrow}\hat{b}_{i+1\uparrow} - \hat{b}_{i\downarrow}^\dagger \hat{b}_{i+1\uparrow}^\dagger \hat{b}_{i\uparrow}\hat{b}_{i+1\downarrow} + \hat{b}_{i\downarrow}^\dagger \hat{b}_{i+1\downarrow}^\dagger \hat{b}_{i\uparrow}\hat{b}_{i+1\uparrow}\right], \tag{3.66}$$

which means that a mean field approximation must be applied to these terms in order to make the Hamiltonian quadratic. Before we do so, we introduce some decoupling operators that make these expressions a bit simpler. Define

$$\begin{aligned} \hat{A}_{i,i+1} &= \hat{b}_{i\uparrow}\hat{b}_{i+1\uparrow} + \hat{b}_{i\downarrow}\hat{b}_{i+1\downarrow}, \\ \hat{B}_{i,i+1} &= \hat{b}_{i\uparrow}^\dagger \hat{b}_{i+1\uparrow} - \hat{b}_{i\downarrow}^\dagger \hat{b}_{i+1\downarrow}, \\ \hat{C}_{i,i+1} &= \hat{b}_{i\uparrow}^\dagger \hat{b}_{i+1\uparrow} + \hat{b}_{i\downarrow}^\dagger \hat{b}_{i+1\downarrow}. \end{aligned} \tag{3.67}$$

With these operators, (3.65) and (3.66) may be rewritten as

$$\hat{S}_i^x \hat{S}_{i+1}^x = \frac{\hbar^2}{4}\left[\hat{A}_{i,i+1}^\dagger \hat{A}_{i,i+1} - \hat{B}_{i,i+1}^\dagger \hat{B}_{i,i+1} + 1\right] \tag{3.68}$$

and

$$\hat{S}_i^y \hat{S}_{i+1}^y = -\frac{\hbar^2}{4}\left[\hat{A}_{i,i+1}^\dagger \hat{A}_{i,i+1} - \hat{C}_{i,i+1}^\dagger \hat{C}_{i,i+1} + 1\right]. \tag{3.69}$$

We drop the constant terms in the following. Setting $S = 1/2$ in (3.1), the Hamiltonian may be written

$$\begin{aligned} \hat{H}_0 = \sum_{i=1}^{N} \Bigg[&\frac{\hbar^2}{4}(J_x \hat{B}_{i,i+1}^\dagger \hat{B}_{i,i+1} - J_y \hat{C}_{i,i+1}^\dagger \hat{C}_{i,i+1} - (J_x - J_y)\hat{A}_{i,i+1}^\dagger \hat{A}_{i,i+1}) \\ &+ \frac{\hbar B}{4}(\hat{b}_{i\downarrow}^\dagger \hat{b}_{i\downarrow} - \hat{b}_{i\uparrow}^\dagger \hat{b}_{i\uparrow})\Bigg]. \end{aligned} \tag{3.70}$$

We apply the mean field approximation (2.57) to the Hamiltonian and obtain

$$\begin{aligned}\hat{H}_{0(\mathrm{MF})} = \sum_{i=1}^{N}\Bigg[\frac{\hbar^2}{4}\Big(&J_x[\langle\hat{B}_{i,i+1}\rangle^*\hat{B}_{i,i+1} + \langle\hat{B}_{i,i+1}\rangle\hat{B}^\dagger_{i,i+1}] \\ &- J_y[\langle\hat{C}_{i,i+1}\rangle^*\hat{C}_{i,i+1} + \langle\hat{C}_{i,i+1}\rangle\hat{C}^\dagger_{i,i+1}] \\ &- (J_x - J_y)[\langle\hat{A}_{i,i+1}\rangle^*\hat{A}_{i,i+1} + \langle\hat{A}_{i,i+1}\rangle\hat{A}^\dagger_{i,i+1}]\Big) + \frac{\hbar B}{4}(\hat{b}^\dagger_{i\downarrow}\hat{b}_{i\downarrow} - \hat{b}^\dagger_{i\uparrow}\hat{b}_{i\uparrow}) \\ &- \frac{\hbar^2}{4}\Big(J_x|\langle\hat{B}_{i,i+1}\rangle|^2 - J_y|\langle\hat{C}_{i,i+1}\rangle|^2 - (J_x - J_y)|\langle\hat{A}_{i,i+1}\rangle|^2\Big)\Bigg].\end{aligned} \tag{3.71}$$

Before moving on with this Hamiltonian, we have to make sure that the constraint (3.61) is satisfied. We do so by introducing Lagrange multipliers μ_i, which can be regarded as chemical potentials for the Schwinger bosons at each lattice site i. To include these multipliers in the treatment of the system, we add the term

$$\hat{H}_L = \hbar^2\sum_{i=1}^{N}\mu_i(\hat{n}_i - \kappa) \tag{3.72}$$

to the Hamiltonian (3.71), which enforces the constraint to be satisfied on average. Here, $\kappa = \langle\hat{n}_i\rangle$ is the mean field boson density at site i. Assuming that the chemical potential is site-independent, we may define the mean field parameters of the system as

$$\langle\hat{B}_{i,i+1}\rangle = \Delta_1; \quad \langle\hat{C}_{i,i+1}\rangle = \Delta_2; \quad \langle\hat{A}_{i,i+1}\rangle = \Delta_3; \quad \mu_i = \mu. \tag{3.73}$$

We now Fourier transform the Hamiltonian, starting off with the Lagrange term $\hat{H}_L$. Using (3.12) and the symmetry of the first Brillouin zone, we have

$$\begin{aligned}\sum_{i=1}^{N}\hat{n}_i &= \sum_{i=1}^{N}\sum_{\alpha}\hat{b}^\dagger_{i\alpha}\hat{b}_{i\alpha} = \sum_{i=1}^{N}(\hat{b}^\dagger_{i\uparrow}\hat{b}_{i\uparrow} + \hat{b}^\dagger_{i\downarrow}\hat{b}_{i\downarrow}) = \sum_{k}(\hat{b}^\dagger_{k\uparrow}\hat{b}_{k\uparrow} + \hat{b}^\dagger_{k\downarrow}\hat{b}_{k\downarrow}) \\ &= \frac{1}{2}\sum_{k}(\hat{b}^\dagger_{k\uparrow}\hat{b}_{k\uparrow} + \hat{b}^\dagger_{-k\uparrow}\hat{b}_{-k\uparrow} + \hat{b}^\dagger_{k\downarrow}\hat{b}_{k\downarrow} + \hat{b}^\dagger_{-k\downarrow}\hat{b}_{-k\downarrow}) \\ &= \frac{1}{2}\sum_{k}(\hat{b}^\dagger_{k\uparrow}\hat{b}_{k\uparrow} + \hat{b}_{-k\uparrow}\hat{b}^\dagger_{-k\uparrow} + \hat{b}^\dagger_{k\downarrow}\hat{b}_{k\downarrow} + \hat{b}_{-k\downarrow}\hat{b}^\dagger_{-k\downarrow} - 2).\end{aligned} \tag{3.74}$$

The Lagrange term may then be written

$$\begin{aligned}\hat{H}_L &= \hbar^2\sum_{k}\left[\frac{\mu}{2}(\hat{b}^\dagger_{k\uparrow}\hat{b}_{k\uparrow} + \hat{b}_{-k\uparrow}\hat{b}^\dagger_{-k\uparrow}) + \frac{\mu}{2}(\hat{b}^\dagger_{k\downarrow}\hat{b}_{k\downarrow} + \hat{b}_{-k\downarrow}\hat{b}^\dagger_{-k\downarrow}) - \mu(1+\kappa)\right] \\ &= \hbar^2\sum_{k}\left[\begin{bmatrix}\hat{b}^\dagger_{k\uparrow} & \hat{b}^\dagger_{k\downarrow} & \hat{b}_{-k\uparrow} & \hat{b}_{-k\downarrow}\end{bmatrix}\begin{bmatrix}\mu/2 & 0 & 0 & 0\\ 0 & \mu/2 & 0 & 0\\ 0 & 0 & \mu/2 & 0\\ 0 & 0 & 0 & \mu/2\end{bmatrix}\begin{bmatrix}\hat{b}_{k\uparrow}\\ \hat{b}_{k\downarrow}\\ \hat{b}^\dagger_{-k\uparrow}\\ \hat{b}^\dagger_{-k\downarrow}\end{bmatrix} - \mu(1+\kappa)\right].\end{aligned} \tag{3.75}$$

We now go back to $\hat{H}_{0(\mathrm{MF})}$. Using the definitions of the decoupling operators (3.67), we have

$$\sum_{i=1}^{N}\Big[\Delta_1^*\hat{B}_{i,i+1}+\Delta_1\hat{B}_{i,i+1}^\dagger\Big]=\sum_k\Big[\Delta_1^*e^{ika}+\Delta_1e^{-ika}\Big]\Big[\hat{b}_{k\uparrow}^\dagger\hat{b}_{k\uparrow}-\hat{b}_{k\downarrow}^\dagger\hat{b}_{k\downarrow}\Big], \tag{3.76}$$

$$\sum_{i=1}^{N}\Big[\Delta_2^*\hat{C}_{i,i+1}+\Delta_2\hat{C}_{i,i+1}^\dagger\Big]=\sum_k\Big[\Delta_2^*e^{ika}+\Delta_2e^{-ika}\Big]\Big[\hat{b}_{k\uparrow}^\dagger\hat{b}_{k\uparrow}+\hat{b}_{k\downarrow}^\dagger\hat{b}_{k\downarrow}\Big], \tag{3.77}$$

$$\sum_{i=1}^{N}\Big[\Delta_3^*\hat{A}_{i,i+1}+\Delta_3\hat{A}_{i,i+1}^\dagger\Big]=\sum_k\Big[\Delta_3^*e^{-ika}(\hat{b}_{k\uparrow}\hat{b}_{-k\uparrow}+\hat{b}_{k\downarrow}\hat{b}_{-k\downarrow}) \tag{3.78}$$

$$+\Delta_3e^{ika}(\hat{b}_{k\uparrow}^\dagger\hat{b}_{-k\uparrow}^\dagger+\hat{b}_{k\downarrow}^\dagger\hat{b}_{-k\downarrow}^\dagger)\Big], \tag{3.79}$$

which lets us write (3.71) as

$$\begin{aligned}\hat{H}_{0(\mathrm{MF})}=\sum_k\frac{\hbar^2}{4}\Big[&\Big[J_x(\Delta_1^*e^{ika}+\Delta_1e^{-ika})-J_y(\Delta_2^*e^{ika}+\Delta_2e^{-ika})-\frac{B}{\hbar}\Big]\hat{b}_{k\uparrow}^\dagger\hat{b}_{k\uparrow}\\&+\Big[-J_x(\Delta_1^*e^{ika}+\Delta_1e^{-ika})-J_y(\Delta_2^*e^{ika}+\Delta_2e^{-ika})+\frac{B}{\hbar}\Big]\hat{b}_{k\downarrow}^\dagger\hat{b}_{k\downarrow}\\&+(J_y-J_x)\Delta_3^*e^{-ika}(\hat{b}_{k\uparrow}\hat{b}_{-k\uparrow}+\hat{b}_{k\downarrow}\hat{b}_{-k\downarrow})\\&+(J_y-J_x)\Delta_3e^{ika}(\hat{b}_{k\uparrow}^\dagger\hat{b}_{-k\uparrow}^\dagger+\hat{b}_{k\downarrow}^\dagger\hat{b}_{-k\downarrow}^\dagger)\\&-J_x|\Delta_1|^2+J_y|\Delta_2|^2+(J_x-J_y)|\Delta_3|^2\Big].\end{aligned} \tag{3.80}$$

Defining

$$\eta_k=\frac{1}{4}\Big[J_x(\Delta_1^*e^{ika}+\Delta_1e^{-ika})-J_y(\Delta_2^*e^{ika}+\Delta_2e^{-ika})-\frac{B}{\hbar}\Big], \tag{3.81}$$

$$\xi_k=\frac{1}{4}\Big[-J_x(\Delta_1^*e^{ika}+\Delta_1e^{-ika})-J_y(\Delta_2^*e^{ika}+\Delta_2e^{-ika})+\frac{B}{\hbar}\Big], \tag{3.82}$$

$$\psi_k=\frac{1}{4}(J_y-J_x)\Delta_3^*e^{-ika}, \tag{3.83}$$

for brevity, the Hamiltonian $\hat{H}_{0(\mathrm{MF})}$ may be written more compactly as

$$\begin{aligned}\hat{H}_{0(\mathrm{MF})}=\hbar^2\sum_k\Big[&\eta_k\hat{b}_{k\uparrow}^\dagger\hat{b}_{k\uparrow}+\xi_k\hat{b}_{k\downarrow}^\dagger\hat{b}_{k\downarrow}+\psi_k(\hat{b}_{k\uparrow}\hat{b}_{-k\uparrow}+\hat{b}_{k\downarrow}\hat{b}_{-k\downarrow})+\psi_k^*(\hat{b}_{k\uparrow}^\dagger\hat{b}_{-k\uparrow}^\dagger+\hat{b}_{k\downarrow}^\dagger\hat{b}_{-k\downarrow}^\dagger)\\&-\frac{1}{4}(J_x|\Delta_1|^2-J_y|\Delta_2|^2-(J_x-J_y)|\Delta_3|^2)\Big].\end{aligned} \tag{3.84}$$

Assuming that all mean field parameters are real, (3.81) and (3.82) may be written as

$$\eta_k=\frac{1}{2}\Big[(J_x\Delta_1-J_y\Delta_2)\cos(ka)-\frac{B}{2\hbar}\Big], \tag{3.85}$$

$$\xi_k=-\frac{1}{2}\Big[(J_x\Delta_1+J_y\Delta_2)\cos(ka)-\frac{B}{2\hbar}\Big], \tag{3.86}$$

from which it is clear that $\eta_{-k} = \eta_k$ and $\xi_{-k} = \xi_k$. Using this property to rewrite the first two terms in (3.84) in the same way as in (3.74), we get

$$\begin{aligned}\hat{H}_{0(\mathrm{MF})} = \hbar^2 \sum_k \Bigg[& \frac{\eta_k}{2}(\hat{b}^\dagger_{k\uparrow}\hat{b}_{k\uparrow} + \hat{b}_{-k\uparrow}\hat{b}^\dagger_{-k\uparrow}) + \frac{\xi_k}{2}(\hat{b}^\dagger_{k\downarrow}\hat{b}_{k\downarrow} + \hat{b}_{-k\downarrow}\hat{b}^\dagger_{-k\downarrow}) \\ & + \psi_k(\hat{b}_{k\uparrow}\hat{b}_{-k\uparrow} + \hat{b}_{k\downarrow}\hat{b}_{-k\downarrow}) + \psi_k^*(\hat{b}^\dagger_{k\uparrow}\hat{b}^\dagger_{-k\uparrow} + \hat{b}^\dagger_{k\downarrow}\hat{b}^\dagger_{-k\downarrow}) \\ & - \frac{1}{2}(\eta_k + \xi_k) - \frac{1}{4}(J_x|\Delta_1|^2 - J_y|\Delta_2|^2 - (J_x - J_y)|\Delta_3|^2)\Bigg],\end{aligned} \tag{3.87}$$

which may be written in matrix form as

$$\begin{aligned}\hat{H}_{0(\mathrm{MF})} = \hbar^2 \sum_k \Bigg[& \begin{bmatrix}\hat{b}^\dagger_{k\uparrow} & \hat{b}^\dagger_{k\downarrow} & \hat{b}_{-k\uparrow} & \hat{b}_{-k\downarrow}\end{bmatrix} \begin{bmatrix}\eta_k/2 & 0 & \psi_k^* & 0 \\ 0 & \xi_k/2 & 0 & \psi_k^* \\ \psi_k & 0 & \eta_k/2 & 0 \\ 0 & \psi_k & 0 & \xi_k/2\end{bmatrix} \begin{bmatrix}\hat{b}_{k\uparrow} \\ \hat{b}_{k\downarrow} \\ \hat{b}^\dagger_{-k\uparrow} \\ \hat{b}^\dagger_{-k\downarrow}\end{bmatrix} \\ & + \frac{1}{2}J_y\Delta_2\cos(ka) - \frac{1}{4}(J_x|\Delta_1|^2 - J_y|\Delta_2|^2 - (J_x - J_y)|\Delta_3|^2)\Bigg].\end{aligned} \tag{3.88}$$

Here, we used the definitions of η_k and ξ_k to write the first constant in (3.87) in terms of J_y and Δ_2. The contribution from this term is exactly zero when the sum is taken over the first Brillouin zone. The total mean field Hamiltonian $\hat{H}_{\mathrm{MF}} = \hat{H}_L + \hat{H}_{0(\mathrm{MF})}$ is then given by

$$\begin{aligned}\hat{H}_{\mathrm{MF}} = \hbar^2 \sum_k \Bigg[& \begin{bmatrix}\hat{b}^\dagger_{k\uparrow} & \hat{b}^\dagger_{k\downarrow} & \hat{b}_{-k\uparrow} & \hat{b}_{-k\downarrow}\end{bmatrix} \begin{bmatrix}(\mu+\eta_k)/2 & 0 & \psi_k^* & 0 \\ 0 & (\mu+\xi_k)/2 & 0 & \psi_k^* \\ \psi_k & 0 & (\mu+\eta_k)/2 & 0 \\ 0 & \psi_k & 0 & (\mu+\xi_k)/2\end{bmatrix} \begin{bmatrix}\hat{b}_{k\uparrow} \\ \hat{b}_{k\downarrow} \\ \hat{b}^\dagger_{-k\uparrow} \\ \hat{b}^\dagger_{-k\downarrow}\end{bmatrix} \\ & - \frac{1}{4}(J_x|\Delta_1|^2 - J_y|\Delta_2|^2 - (J_x - J_y)|\Delta_3|^2) - \mu(1+\kappa)\Bigg].\end{aligned} \tag{3.89}$$

In order to diagonalize the Hamiltonian using Theorem 2.2, we have to ensure the positive definiteness of the matrix

$$\mathbb{H}(\vec{k}) = \begin{bmatrix}(\mu+\eta_k)/2 & 0 & \psi_k^* & 0 \\ 0 & (\mu+\xi_k)/2 & 0 & \psi_k^* \\ \psi_k & 0 & (\mu+\eta_k)/2 & 0 \\ 0 & \psi_k & 0 & (\mu+\xi_k)/2\end{bmatrix}. \tag{3.90}$$

Using symbolic computation, we find that the eigenvalues of $\mathbb{H}(\vec{k})$ are

$$\begin{aligned}\tilde{\lambda}^{(1)}_k, \tilde{\lambda}^{(2)}_k &= \frac{1}{2}(\mu+\eta_k) \pm |\psi_k|, \\ \tilde{\lambda}^{(3)}_k, \tilde{\lambda}^{(4)}_k &= \frac{1}{2}(\mu+\xi_k) \pm |\psi_k|,\end{aligned} \tag{3.91}$$

which we demand be positive. This places constraints on the chemical potential:

$$\begin{aligned} \mu &> 2|\psi_k| - \eta_k, \\ \mu &> 2|\psi_k| - \xi_k. \end{aligned} \tag{3.92}$$

With these conditions satisfied we obtain the energy dispersion by computing the eigenvalues of the dynamical matrix $\mathbb{K} = \sigma_3 \mathbb{H}_k$. Symbolic computation yields

$$\lambda_k^{(1)}, \lambda_k^{(2)} = \pm \frac{1}{2} \sqrt{(\mu + \eta_k)^2 - 4|\psi_k|^2}, \tag{3.93}$$

$$\lambda_k^{(3)}, \lambda_k^{(4)} = \pm \frac{1}{2} \sqrt{(\mu + \xi_k)^2 - 4|\psi_k|^2}. \tag{3.94}$$

Defining E_k^η and E_k^ξ as

$$E_k^\eta = \sqrt{(\mu + \eta_k)^2 - 4|\psi_k|^2}, \tag{3.95}$$

$$E_k^\xi = \sqrt{(\mu + \xi_k)^2 - 4|\psi_k|^2}, \tag{3.96}$$

Theorem 2.2 gives us the diagonalized mean field Hamiltonian:

$$\begin{aligned} \hat{H}_{\mathrm{MF}} = \hbar^2 \sum_k \Bigg[&\frac{E_k^\eta}{2} (\hat{\beta}_{k\uparrow}^\dagger \hat{\beta}_{k\uparrow} + \hat{\beta}_{-k\uparrow}^\dagger \hat{\beta}_{-k\uparrow}) + \frac{E_k^\xi}{2} (\hat{\beta}_{k\downarrow}^\dagger \hat{\beta}_{k\downarrow} + \hat{\beta}_{-k\downarrow}^\dagger \hat{\beta}_{-k\downarrow}) + \frac{1}{2}(E_k^\eta + E_k^\xi) \\ &- \frac{1}{4}(J_x |\Delta_1|^2 - J_y |\Delta_2|^2 - (J_x - J_y)|\Delta_3|^2) - \mu(1 + \kappa) \Bigg]. \end{aligned} \tag{3.97}$$

Using $E_{-k}^\eta = E_k^\eta$ and $E_{-k}^\xi = E_k^\xi$, (3.97) may be written more compactly as

$$\begin{aligned} \hat{H}_{\mathrm{MF}} = \hbar^2 \sum_k \Bigg[&E_k^\eta \hat{\beta}_{k\uparrow}^\dagger \hat{\beta}_{k\uparrow} + E_k^\xi \hat{\beta}_{k\downarrow}^\dagger \hat{\beta}_{k\downarrow} + \frac{1}{2}(E_k^\eta + E_k^\xi) \\ &- \frac{1}{4}(J_x |\Delta_1|^2 - J_y |\Delta_2|^2 - (J_x - J_y)|\Delta_3|^2) - \mu(1 + \kappa) \Bigg]. \end{aligned} \tag{3.98}$$

Clearly, the ground state energy E_0 of the system is given by the absence of β-bosons:

$$E_0 = \hbar^2 \sum_k \left[\frac{1}{2}(E_k^\eta + E_k^\xi) - \frac{1}{4}(J_x |\Delta_1|^2 - J_y |\Delta_2|^2 - (J_x - J_y)|\Delta_3|^2) - \mu(1 + \kappa) \right]. \tag{3.99}$$

Normally, the next step is to use the ground state energy to obtain the self-consistency equations, which we can solve for the mean field parameters. However, it turns out that the mean field approach used here is too naïve to produce correct physical results for the XY chain. The ground state of the XY chain is magnetically ordered, which in the Schwinger boson picture corresponds to the condensation of $\hat{b}_\uparrow$-bosons. However, the ground state obtained using mean field theory is described by the absence of bosons, which describes a magnetically disordered state. Cieplak and Turski [17] described the proper way of treating the XY model using the Schwinger boson transformation. They eliminate the $\hat{b}_\uparrow$ operators in

the transformed spin Hamiltonian, such that deviations in the magnetically ordered ground state are described by $\hat{b}_{\downarrow}$-excitations. Since we are not interested in accurately describing the dynamics of the XY chain here, we will not pursue this approach. We note instead that Schwinger boson mean field theory is ill-suited for the study of one-dimensional magnetically ordered systems.[4] Nevertheless, we have seen how we can use it in practice, which will be useful for the study of the two-dimensional, magnetically frustrated Kane-Mele-Hubbard model in the latter parts of this thesis.

[4]Mean field theories generally fall short in one-dimensional systems, where direct particle-particle interactions govern the dynamics.

4 The Berry Phase and Topological Band Theory

For a long time, the general consensus among physicists was that the phase of a quantum mechanical wave function is arbitrary and therefore has no impact on the observable physics of the system. This notion prevailed until 1984, when Berry [18] showed that a quantum state acquires a geometrical phase under cyclic evolutions in the parameter space of the Hamiltonian. This phase, named the Berry phase after its discoverer, cannot be removed by a gauge transformation and is therefore linked to physical observables. We now show how the Abelian Berry phase, valid for non-degenerate energy eigenstates, comes about and how it can be expressed in terms of the Berry curvature. The Berry curvature is shown to be linked to the topology of the quantum mechanical system via the Chern theorem. We then show how the Berry phase manifests itself in condensed matter physics by studying a simple topological insulator, the Chern insulator. This Section is based on results derived by Bernevig [19, p.6-32, p.91-99].

4.1 The Berry Phase

Consider a quantum mechanical system described by a Hamiltonian $\hat{H}(\vec{R})$, where $\vec{R} = \vec{R}(t)$ constitutes a set of time-dependent parameters. We are interested in how the system evolves when these parameters are varied adiabatically along a path $\mathcal{C}$ in parameter space. At each point $\vec{R}$ along the path $\mathcal{C}$, we introduce an instantaneous orthonormal basis of eigenstates $\{|n(\vec{R})\rangle\}$ of the Hamiltonian $\hat{H}(\vec{R})$:

$$\hat{H}(\vec{R})\,|n(\vec{R})\rangle = E_n(\vec{R})\,|n(\vec{R})\rangle\,. \tag{4.1}$$

Assume that the system is prepared in the initial state $|n(\vec{R}(0))\rangle$. Then, according to the adiabatic theorem, the system remains in an instantaneous eigenstate of the Hamiltonian as we vary $\vec{R}$ along $\mathcal{C}$, and the only degree of freedom we have is the phase of the wave function. If we write the instantaneous eigenstate of the Hamiltonian $\hat{H}(\vec{R}(t))$ at a later time t as $|\psi(t)\rangle = e^{-i\theta(t)}\,|n(\vec{R}(t))\rangle$, we can identify the phase $\theta(t)$ from the time-dependent Schrödinger equation

$$\hat{H}(\vec{R}(t))\,|\psi(t)\rangle = i\hbar\frac{d}{dt}\,|\psi(t)\rangle\,. \tag{4.2}$$

Using (4.1), we arrive at a differential equation for the phase:

$$E_n(\vec{R}(t))\,|n(\vec{R}(t))\rangle = \hbar\left(\frac{d}{dt}\theta(t)\right)|n(\vec{R}(t))\rangle + i\hbar\frac{d}{dt}\,|n(\vec{R}(t))\rangle\,. \tag{4.3}$$

Multiplying (4.3) by $\langle n(\vec{R}(t))|$ and using the orthonormality of the instantaneous eigenstates, we get

$$\frac{d}{dt}\theta(t) = \frac{1}{\hbar}E_n(\vec{R}(t)) - i\,\langle n(\vec{R}(t))|\,\frac{d}{dt}\,|n(\vec{R}(t))\rangle\,, \tag{4.4}$$

from which the phase is found to be

$$\theta(t) = \frac{1}{\hbar}\int_0^t dt'\;E_n(\vec{R}(t')) - i\int_0^t dt'\;\langle n(\vec{R}(t'))|\,\frac{d}{dt'}\,|n(\vec{R}(t'))\rangle\,. \tag{4.5}$$

The first term in (4.5) is the well known dynamical phase. The second term is new, and the negative of it is the announced Berry phase:

$$\gamma_n = i \int_0^t dt' \ \langle n(\vec{R}(t'))| \frac{d}{dt'} |n(\vec{R}(t'))\rangle \,. \tag{4.6}$$

We can remove the explicit time dependence from (4.6) by use of the chain rule, which allows us to rewrite the Berry phase as a line integral along the path $\mathcal{C}$ in parameter space:

$$\begin{aligned} \gamma_n &= i \int_0^t dt' \ \frac{d\vec{R}(t')}{dt'} \cdot \langle n(\vec{R}(t'))| \nabla_{\vec{R}} |n(\vec{R}(t'))\rangle = i \int_{\mathcal{C}} d\vec{R} \ \cdot \langle n(\vec{R})| \nabla_{\vec{R}} |n(\vec{R})\rangle \\ &= \int_{\mathcal{C}} d\vec{R} \ \cdot \vec{A}_n(\vec{R}), \end{aligned} \tag{4.7}$$

where we defined the Berry connection

$$\vec{A}_n(\vec{R}) = i \langle n(\vec{R})| \nabla_{\vec{R}} |n(\vec{R})\rangle \tag{4.8}$$

in the last equality. It is worth noting that the Berry phase is real, since $\langle n(\vec{R})| \nabla_{\vec{R}} |n(\vec{R})\rangle$ is purely imaginary, which can be seen by differentiating the orthonormality condition $\langle n(\vec{R})|n(\vec{R})\rangle = 1$. The question is then whether the Berry phase has any physical meaning; that is, whether or not it can be canceled through an astute choice of gauge. Under the gauge transformation

$$|n(\vec{R})\rangle \to e^{i\phi(\vec{R})} |n(\vec{R})\rangle \,, \tag{4.9}$$

the Berry connection transforms as

$$\vec{A}_n(\vec{R}) \to \vec{A}_n(\vec{R}) - \nabla_{\vec{R}} \phi(\vec{R}). \tag{4.10}$$

The change in the Berry phase due to the gauge transformation is

$$\Delta\gamma_n = -\int_{\mathcal{C}} d\vec{R} \ \cdot \nabla_{\vec{R}} \phi(\vec{R}) = \phi(\vec{R}(0)) - \phi(\vec{R}(t)). \tag{4.11}$$

Now, suppose that $\mathcal{C}$ is a closed curve in parameter space and let $t = T$ be the time taken to traverse it. At time T, we then return to the original parameter configuration, and the instantaneous eigenstate $|n(\vec{R}(T))\rangle$ must be equal to the initial state $|n(\vec{R}(0))\rangle$. This must also hold under gauge transformations, so

$$e^{i\phi(\vec{R}(0))} |n(\vec{R}(0))\rangle = e^{i\phi(\vec{R}(T))} |n(\vec{R}(T))\rangle = e^{i\phi(\vec{R}(T))} |n(\vec{R}(0))\rangle \,, \tag{4.12}$$

which means that

$$\phi(\vec{R}(0)) - \phi(\vec{R}(T)) = 2\pi m; \quad m \in \mathbb{Z}. \tag{4.13}$$

Thus we see that for a closed curve $\mathcal{C}$ in parameter space, the only way the Berry phase is canceled by a gauge transformation is if it is an integer multiple of 2π. Since the Berry phase is gauge invariant modulo 2π, the phase factor $e^{i\gamma_n}$ acquired by the wave function under the cyclic evolution in parameter space is gauge invariant and can therefore be related to physical observables.

4.2 The Berry Curvature and the Chern Number

In the following treatment, we will consider only closed curves $\mathcal{C}$. This allows us to apply Stokes' theorem to the Abelian Berry phase (4.7):

$$\gamma_n = \oint_{\mathcal{C}} d\vec{R} \cdot \vec{A}_n(\vec{R}) = \iint_{\mathcal{S}} \Omega_{n,ij} ds_i \wedge ds_j, \tag{4.14}$$

where $\mathcal{S}$ is the surface bounded by the curve $\mathcal{C}$ and $\Omega_{n,ij}$ is the antisymmetric Berry curvature, given in terms of the components $A_{n,i}$ of the Berry connection:

$$\Omega_{n,ij} = \partial_i A_{n,j} - \partial_j A_{n,i}, \tag{4.15}$$

where $\partial_i = \partial/\partial R^i$ is introduced for notational simplicity. The Berry curvature has several interesting properties. Unlike the Berry connection, which transforms as (4.10) under a gauge transformation, the Berry curvature is gauge invariant since the double derivatives cancel in (4.15). Furthermore, the sum of the Berry curvatures of all eigenstates is zero. To see this, we write out the components of the Berry curvature explicitly:

$$\begin{aligned} \Omega_{n,ij} &= i\big[\partial_i \langle n(\vec{R})| \partial_j |n(\vec{R})\rangle - \partial_j \langle n(\vec{R})| \partial_i |n(\vec{R})\rangle\big] \\ &= i\big[\langle \partial_i n(\vec{R})|\partial_j n(\vec{R})\rangle - \langle \partial_j n(\vec{R})|\partial_i n(\vec{R})\rangle\big]. \end{aligned} \tag{4.16}$$

The terms involving double derivatives cancel in the second equality. Inserting a complete set of states $\sum_m |m(\vec{R})\rangle \langle m(\vec{R})| = I$ into (4.16), we obtain

$$\Omega_{n,ij} = i \sum_{m\neq n} \big[\langle \partial_i n(\vec{R})|m(\vec{R})\rangle \langle m(\vec{R})|\partial_j n(\vec{R})\rangle - \langle \partial_j n(\vec{R})|m(\vec{R})\rangle \langle m(\vec{R})|\partial_i n(\vec{R})\rangle\big]. \tag{4.17}$$

The contribution from $m = n$ vanishes; since $\langle \partial_i n(\vec{R})|n(\vec{R})\rangle$ and $\langle n(\vec{R})|\partial_j n(\vec{R})\rangle$ are purely imaginary, their product multiplied by the imaginary unit is also purely imaginary. Therefore, this term has to be zero since the Berry phase is a real quantity. This also shows that we cannot have a finite Berry curvature for a single-band Hamiltonian. To calculate $\langle m(\vec{R})|\partial_i n(\vec{R})\rangle$, we first apply the differential operator to (4.1) to obtain an equation involving $|\partial_i n(\vec{R})\rangle$:

$$(\partial_i \hat{H}(\vec{R}))\,|n(\vec{R})\rangle + \hat{H}(\vec{R})\,|\partial_i n(\vec{R})\rangle = (\partial_i E_n(\vec{R}))\,|n(\vec{R})\rangle + E_n(\vec{R})\,|\partial_i n(\vec{R})\rangle\,. \tag{4.18}$$

Multiplying (4.18) by $\langle m(\vec{R})| \neq \langle n(\vec{R})|$, we get

$$\langle m(\vec{R})|\partial_i n(\vec{R})\rangle = \frac{\langle m(\vec{R})|\,(\partial_i \hat{H}(\vec{R}))\,|n(\vec{R})\rangle}{E_n(\vec{R}) - E_m(\vec{R})}. \tag{4.19}$$

A similar expression can be obtained for $\langle \partial_i n(\vec{R})|m(\vec{R})\rangle$ by differentiating the conjugate of (4.1) and multiplying by $|m(\vec{R})\rangle \neq |n(\vec{R})\rangle$. The Berry curvature (4.17) may then be written

$$\begin{aligned} \Omega_{n,ij} = i \sum_{m\neq n} \frac{1}{(E_n(\vec{R}) - E_m(\vec{R}))^2} \Big[&\langle n(\vec{R})|\,\partial_i \hat{H}(\vec{R}))\,|m(\vec{R})\rangle \langle m(\vec{R})|\,(\partial_j \hat{H}(\vec{R}))\,|n(\vec{R})\rangle \\ &- \langle n(\vec{R})|\,\partial_j \hat{H}(\vec{R}))\,|m(\vec{R})\rangle \langle m(\vec{R})|\,(\partial_i \hat{H}(\vec{R}))\,|n(\vec{R})\rangle \Big]. \end{aligned} \tag{4.20}$$

From (4.20) it is evident that

$$\sum_n \Omega_{n,ij} = 0, \tag{4.21}$$

and by (4.14) the same is true for the Berry phase: $\sum_n \gamma_n = 0$. We see from (4.20) that the Berry curvature is a property of the system's full energy spectrum, since the overlap $\langle n(\vec{R})| \, \partial_i \hat{H}(\vec{R})) \, |m(\vec{R})\rangle$ describes the interaction between the eigenstates $|n(\vec{R})\rangle$ and $|m(\vec{R})\rangle$ under the adiabatic evolution.

The Berry curvature is also related to the topological properties of the system. Let $\mathcal{S}$ be a closed, smooth manifold in parameter space, and let $\mathcal{C}$ be the closed curve separating $\mathcal{S}$ into two regions $\mathcal{S}_1$ and $\mathcal{S}_2$. The integral of the Berry curvature over the closed manifold $\mathcal{S}$ is then the sum of the integrals of the Berry curvature over these two regions:

$$\oiint_{\mathcal{S}} \Omega_{n,ij} ds_i \wedge ds_j = \iint_{\mathcal{S}_1} \Omega_{n,ij} ds_i \wedge ds_j + \iint_{\mathcal{S}_2} \Omega_{n,ij} ds_i \wedge ds_j. \tag{4.22}$$

We can now apply Stokes' theorem to the integrals on the right hand side of (4.22). Let the Berry connections in $\mathcal{S}_1$ and $\mathcal{S}_2$ be $\vec{A}_n^{(1)}(\vec{R})$ and $\vec{A}_n^{(2)}(\vec{R})$, respectively; since the Berry connection is gauge-dependent, it may not be the same in the two regions. Then, (4.22) may be written

$$\oiint_{\mathcal{S}} \Omega_{n,ij} ds_i \wedge ds_j = \oint_{\mathcal{C}} d\vec{R} \cdot \big[\vec{A}_n^{(1)}(\vec{R}) - \vec{A}_n^{(2)}(\vec{R})\big]. \tag{4.23}$$

By (4.10), the integrand on the right hand side of (4.23) is simply $\nabla_{\vec{R}} \phi(\vec{R})$, with $\phi(\vec{R})$ describing the possibly different gauges in the two regions. Our previous treatment of the Berry phase then tells us that the integral of the Berry curvature over a closed manifold $\mathcal{S}$ is an integer multiple of 2π. The integer in question is a topological invariant known as the Chern number:

$$C_n = \frac{1}{2\pi} \oiint_{\mathcal{S}} \Omega_{n,ij} ds_i \wedge ds_j. \tag{4.24}$$

To see why the Chern number is a topological invariant, note that the Abelian Berry curvature (4.20) is well-defined only if there are energy gaps separating the nth energy band from the other energy bands. If the bands are continuously deformed under an adiabatic evolution in parameter space in such a way that these gaps remain open, the Berry curvature changes continuously. But since the Chern number is integer-valued, its value must remain the same throughout the adiabatic evolution. However, if the deformation of the bands results in the closing and reopening of the gap between the nth band and a neighboring band, the adiabatic theorem no longer holds and the Chern number may change. The Chern number may therefore be used to distinguish different topological phases; a topologically trivial phase has zero Chern numbers, whereas phases with non-zero Chern numbers are topologically non-trivial.

4.3 The Chern Insulator

Our discussion so far has been very general, treating $\vec{R}$ as a set of generic parameters in an arbitrary-dimensional parameter space. In condensed matter physics, we are most often interested in the case where $\vec{R} = \vec{k}$ describes the momentum of the particles comprising

the system. For example, $\vec{k}$ can be the momentum of an electron in a crystal, which evolves adiabatically when the system is subjected to an applied electric field. We will now illustrate how the general, abstract ideas introduced in the previous Sections present themselves in condensed matter physics through a simple model of a topological insulator, the Chern insulator. The Chern insulator was first introduced by Haldane [20] on the honeycomb lattice. Here, we consider the simpler model of the Chern insulator on the square lattice. The Hamiltonian of the system is given by

$$\hat{H} = \sum_{\vec{k}} \hat{\mathbf{c}}^{\dagger}_{\vec{k}} \mathbb{H}(\vec{k}) \hat{\mathbf{c}}_{\vec{k}}, \tag{4.25}$$

where

$$\hat{\mathbf{c}}_{\vec{k}} = \begin{bmatrix} \hat{c}_{\vec{k}A} & \hat{c}_{\vec{k}B} \end{bmatrix}^{\mathsf{T}} \tag{4.26}$$

contains annihilation operators for spinless (or fully spin-polarized) electrons of momentum $\vec{k}$ in orbitals A and B. The Hamiltonian matrix is given by

$$\begin{aligned} \mathbb{H}(\vec{k}) &= \vec{d}(\vec{k}) \cdot \vec{\sigma} \\ &= t[\sin(k_x a)\sigma_1 + \sin(k_y a)\sigma_2 + (2 + M/t - \cos(k_x a) - \cos(k_y a))\sigma_3] \\ &= \begin{bmatrix} d_3 & d_1 - id_2 \\ d_1 + id_2 & -d_3 \end{bmatrix}, \end{aligned} \tag{4.27}$$

where t is the nearest neighbor hopping amplitude, a is the lattice constant and $\{d_i\} = \{d_i(\vec{k})\}$ are the components of $\vec{d}(\vec{k})$. Like before, $\vec{\sigma}$ is the Pauli matrix vector. The mass term M gives rise to an energy gap $2M$ between orbitals A and B.

We set $t = 1$ in the following. In the limit $\vec{k} \to 0$, (4.27) describes the physics close to the Dirac points that occur in certain electronic band structures, such as that of graphene. The Hamiltonian (4.27) is the generalization of this so-called Dirac Hamiltonian to a lattice. The lattice Dirac Hamiltonian is gapped everywhere, except at three distinct values of M, where the bands touch at certain points in the first Brillouin zone. These values are $M = 0$, $M = -2$, $M = -4$, for which the gap closes at $(k_x, k_y) = (0, 0)$, $(k_x, k_y) = (0, \pi/a), (\pi/a, 0)$ and $(k_x, k_y) = (\pi/a, \pi/a)$, respectively. These gap closings suggest that the system may exhibit different topological phases as the parameter M is varied. We now show that this is indeed the case by computing the Chern numbers of the bands.

The eigenvalues of the Hamiltonian are $\pm d$, where $d = \sqrt{d_1^2 + d_2^2 + d_3^2}$. The corresponding eigenvectors, which we denote $|\psi_+\rangle = |\psi_+(\vec{k})\rangle$ and $|\psi_-\rangle = |\psi_-(\vec{k})\rangle$ for bands A and B, respectively, may be written

$$|\psi_\pm\rangle = \frac{1}{\sqrt{2d(d \mp d_3)}} \begin{bmatrix} d_1 - id_2 \\ \pm d - d_3 \end{bmatrix}. \tag{4.28}$$

Notice how the normalizing prefactor is undefined at $d = d_3$ for $|\psi_+\rangle$. However, this is not the only way to write the eigenvectors; we could pick a different ratio of vector components when solving for the eigenvectors to obtain

$$|\psi_\pm\rangle = \frac{1}{\sqrt{2d(d \pm d_3)}} \begin{bmatrix} \pm d + d_3 \\ d_1 + id_2 \end{bmatrix}, \tag{4.29}$$

where the normalizing prefactor becomes undefined at $d = d_3$ for $|\psi_-\rangle$. The two eigenvectors (4.28) and (4.29) correspond to different gauge choices. If we cannot pick a gauge which is smooth and continuous in the entirety of the first Brillouin zone, the system must be topologically non-trivial. The reason is that the wave function must be well-defined in the entire Brillouin zone, which means we must patch over the singularities that prevent us from picking a global gauge by choosing different, well-defined gauges in different regions of the Brillouin zone. The gauge dependence of the Berry connection therefore guarantees a non-zero Chern number, per (4.23). The lack of a global, well-defined gauge will manifest itself in gauge invariant quantities such as the Berry curvature and Chern number. Let us now focus on the insulator's filled band $|\psi_-\rangle$. Since our model has only two bands, it follows from (4.21) that the Chern number of the unfilled band is simply the negative of that of the filled band. The components of the Berry connection are given by

$$A_{-,k_x} = i \langle \psi_- | \partial_{k_x} | \psi_- \rangle \quad \text{and} \quad A_{-,k_y} = i \langle \psi_- | \partial_{k_y} | \psi_- \rangle , \tag{4.30}$$

where the derivatives act on both components of the eigenvector. The Berry curvature, which in two dimensions has just one unique component, is

$$\Omega_-(\vec{k}) = \partial_{k_x} A_{-,k_y} - \partial_{k_y} A_{-,k_x}. \tag{4.31}$$

The Chern number is the integral of the Berry curvature over the two-dimensional Brillouin zone. The Brillouin zone is a closed manifold because it is topologically equivalent to a 2-torus; each point on the right and top edges of the square lattice Brillouin zone is equivalent to the corresponding points on the left and bottom edges, respectively, since they are separated by a reciprocal lattice vector. The Chern number of the filled band is therefore given by

$$C_- = \frac{1}{2\pi} \int_{-\pi/a}^{\pi/a} \int_{-\pi/a}^{\pi/a} dk_x \; dk_y \; \Omega_-(\vec{k}). \tag{4.32}$$

There are four potentially distinct topological phases: $M < -4$, $-4 < M < -2$, $-2 < M < 0$ and $M > 0$. Figure 3 shows the Berry curvature of the filled band in the first Brillouin zone for four different mass terms M, one in each of these regions. It is hard to draw conclusions about the topology of the system in these regions based on the Berry curvature, so we have to compute the corresponding Chern number to obtain the phase diagram. While the expression for the Berry curvature can be obtained through straightforward differentiation,[5] its functional form is not simple enough that we can analytically perform the integral (4.32). We therefore resort to numerical integration to identify the phase diagram, which is shown in Figure 4a. The Chern number is seen to be invariant, up to a small numerical error, in each of the aforementioned phases. The vertical lines at $M = -4$, $M = -2$, and $M = 0$ indicate the gap closing points that separate the four phases. For completeness, we show the Chern numbers of the unfilled band in each of the four phases in Figure 4b. This confirms that the Chern number of the unfilled band is the negative of that of the filled band.

[5]For simplicity, we chose to perform the differentiation symbolically on a computer instead of by hand here.

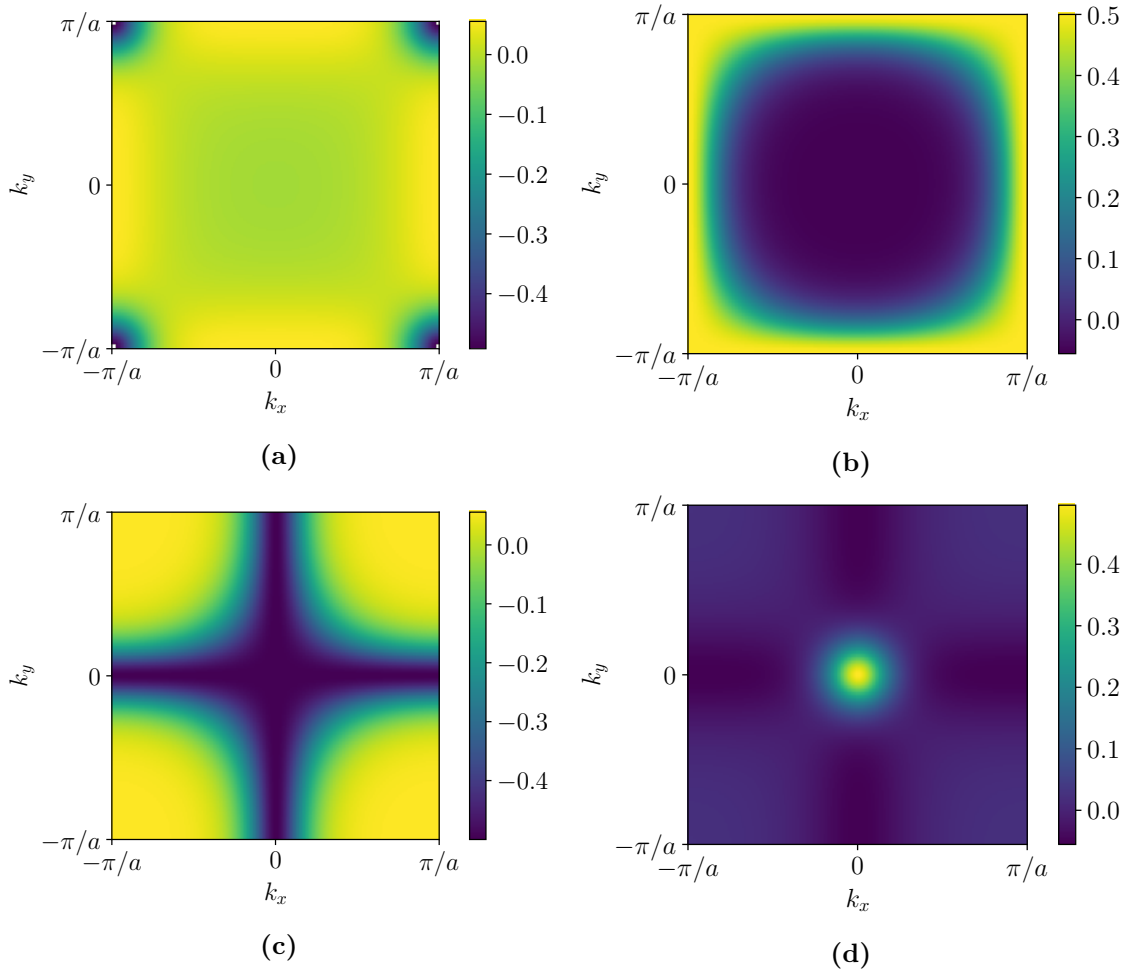


Figure 3: Berry curvatures of the filled band of the Chern insulator for mass terms a) $M = -5$, b) $M = -3$, c) $M = -1$ and d) $M = 1$.

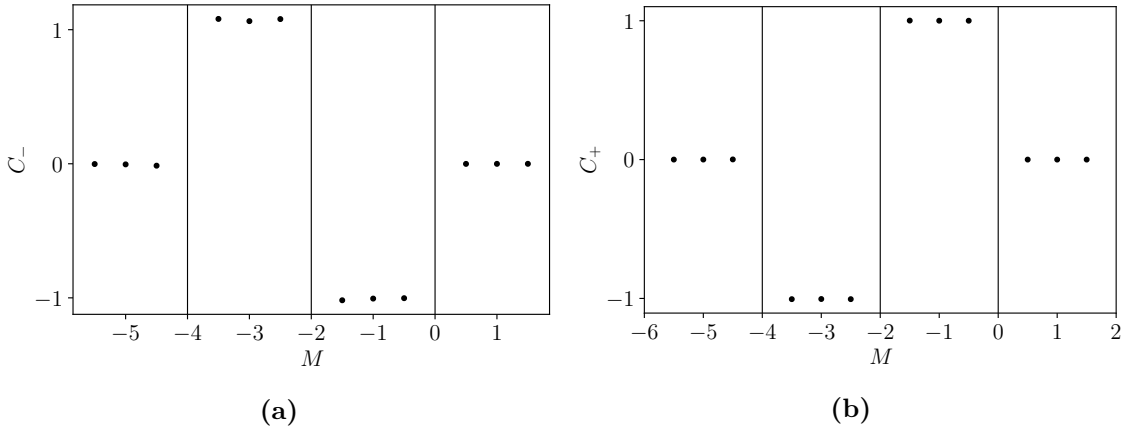


Figure 4: Chern numbers of the a) filled band and b) unfilled band in the topologically distinct phases of the Chern insulator.

5 Linear Response Theory

The diagonalization procedures described in Section 2 provide ways of obtaining the energy spectrum of a second quantized spin Hamiltonian. Once the spectrum is known, we can calculate how the spin system responds to external perturbations, which provides insight into its properties. In linear response theory, the response is considered to be proportional to the perturbation, and all we have to know to determine the system's response is the proportionality factor. In quantum mechanics, physical observables are associated with mathematical operators. Computing experimentally measurable quantities therefore corresponds to finding the expectation value of the associated operators. The observable response is therefore the difference between the expectation values of the operator in the perturbed state and the unperturbed state. Following Bruus and Flensberg [12, p. 95-98], we will now derive the so-called Kubo formula, which lets us compute this response to linear order. We will then introduce two types of Green's functions that will aid us in deriving a general expression for the current response to vector potentials, which is the main goal of this Section.

5.1 The General Kubo Formula

Consider a quantum mechanical system in thermodynamic equilibrium, described by a time-independent Hamiltonian $\hat{H}_0$. The expectation value of an operator $\hat{A}$ corresponding to a physical observable is given by (2.60):

$$\langle \hat{A} \rangle = \mathrm{Tr}\Big\{ \hat{\rho}\hat{A} \Big\} = \frac{1}{Z_0} \sum_n \langle n | \, e^{-\beta \hat{H}_0} \hat{A} \, | n \rangle = \frac{1}{Z_0} \sum_n \langle n | \, \hat{A} \, | n \rangle \, e^{-\beta E_n}, \tag{5.1}$$

where $Z_0 = \mathrm{Tr}\Big\{ e^{-\beta \hat{H}_0} \Big\}$ is the partition function and $\{|n\rangle\}$ are the eigenstates of the Hamiltonian $\hat{H}_0$ corresponding to energies E_n. Then, at time $t = t_0$, a weak perturbation $\hat{V}(t)$ is introduced which makes the Hamiltonian time-dependent:

$$\hat{H}(t) = \hat{H}_0 + \hat{V}(t)\theta(t - t_0). \tag{5.2}$$

Here, $\theta(t)$ denotes the unit step function. The perturbation makes the eigenstates $\{|n\rangle\}$ time-dependent, with the time evolution governed by the perturbed Hamiltonian (5.2). Since the perturbation is assumed weak, the states remain Boltzmann distributed. The expectation value of $\hat{A}$ for times $t > t_0$ then becomes

$$\langle \hat{A}(t) \rangle = \frac{1}{Z_0} \sum_n \langle n(t) | \, \hat{A} \, | n(t) \rangle \, e^{-\beta E_n}. \tag{5.3}$$

It is now convenient to introduce the interaction picture, where state vectors and operators are given by

$$|n_I(t)\rangle = e^{i\hat{H}_0 t/\hbar} \, |n(t)\rangle \quad \text{and} \quad \hat{A}_I(t) = e^{i\hat{H}_0 t/\hbar} \hat{A}(t) e^{-i\hat{H}_0 t/\hbar}, \tag{5.4}$$

respectively. Here, $\hat{A}(t)$ is given in the Heisenberg picture, where the state vectors are time-independent while the operators are given by

$$\hat{A}(t) = e^{i\hat{H}t/\hbar} \hat{A} e^{-i\hat{H}t/\hbar}. \tag{5.5}$$

In the interaction picture, the fast time dependence due to $\hat{H}_0$ is extracted from $|n(t)\rangle$ and $\hat{A}(t)$, leaving only the time dependence due to $\hat{V}(t)$.[6] This can be seen by inserting the state vector $|n_I(t)\rangle$ into the time-dependent Schrödinger equation:

$$i\hbar\frac{\partial}{\partial t}|n_I(t)\rangle = \hat{V}(t)\,|n_I(t)\rangle\,. \tag{5.6}$$

This suggests that the time evolution of $|n_I(t)\rangle$ from time t_0 to t is given by a unitary operator $\hat{U}(t,t_0)$ that depends on $\hat{V}(t)$:

$$|n_I(t)\rangle = \hat{U}(t,t_0)\,|n_I(t_0)\rangle\,. \tag{5.7}$$

Inserting (5.7) into (5.6) yields a differential equation for $\hat{U}(t,t_0)$, which when integrated becomes

$$\hat{U}(t,t_0) = I - \frac{i}{\hbar}\int_{t_0}^{t} dt'\ \hat{V}(t')\hat{U}(t',t_0), \tag{5.8}$$

where the first term on the right hand side comes from $\hat{U}(t_0,t_0) = I$. This equation may be solved iteratively for $\hat{U}(t,t_0)$, with the starting point being the boundary condition $\hat{U}(t',t_0) = I$. The state vector $|n(t)\rangle$ in (5.3) can then be expressed in terms of the interaction picture state vector as

$$|n(t)\rangle = e^{-i\hat{H}_0 t/\hbar}\,|n_I(t)\rangle = e^{-i\hat{H}_0 t/\hbar}\hat{U}(t,t_0)\,|n_I(t_0)\rangle = e^{-i\hat{H}_0 t/\hbar}\hat{U}(t,t_0)e^{i\hat{H}_0 t_0/\hbar}\,|n(t_0)\rangle\,. \tag{5.9}$$

Since we are interested in the system's linear response, we keep only first order terms in the perturbation $\hat{V}(t)$, which means that (5.8) reduces to

$$\hat{U}(t,t_0) = I - \frac{i}{\hbar}\int_{t_0}^{t} dt'\ \hat{V}(t'). \tag{5.10}$$

Inserting (5.9) with $\hat{U}(t,t_0)$ given by (5.10) into (5.3), we obtain

$$\begin{aligned}
\langle\hat{A}(t)\rangle &= \frac{1}{Z_0}\sum_n \langle n(t_0)|\, e^{-i\hat{H}_0 t_0/\hbar}\hat{U}^\dagger(t,t_0)e^{i\hat{H}_0 t/\hbar}\hat{A}e^{-i\hat{H}_0 t/\hbar}\hat{U}(t,t_0)e^{i\hat{H}_0 t_0/\hbar}\,|n(t_0)\rangle\, e^{-\beta E_n} \\
&= \frac{1}{Z_0}\sum_n \langle n|\left(I + \frac{i}{\hbar}\int_{t_0}^{t} dt'\ \hat{V}(t')\right)\hat{A}(t)\left(I - \frac{i}{\hbar}\int_{t_0}^{t} dt'\ \hat{V}(t')\right)|n\rangle\, e^{-\beta E_n} \\
&= \frac{1}{Z_0}\sum_n \langle n|\left(\hat{A}(t) + \frac{i}{\hbar}\int_{t_0}^{t} dt'\ \hat{V}(t')\hat{A}(t) - \frac{i}{\hbar}\int_{t_0}^{t} dt'\ \hat{A}(t)\hat{V}(t')\right)|n\rangle\, e^{-\beta E_n} \\
&= \frac{1}{Z_0}\sum_n \langle n|\left(\hat{A}(t) - \frac{i}{\hbar}\int_{t_0}^{t} dt'\ [\hat{A}(t),\hat{V}(t')]\right)|n\rangle\, e^{-\beta E_n} \\
&= \frac{1}{Z_0}\sum_n \langle n|\,\hat{A}(t)\,|n\rangle\, e^{-\beta E_n} - \frac{i}{\hbar}\int_{t_0}^{t} dt'\ \frac{1}{Z_0}\sum_n \langle n|\,[\hat{A}(t),\hat{V}(t')]\,|n\rangle\, e^{-\beta E_n} \\
&= \langle\hat{A}\rangle_0 - \frac{i}{\hbar}\int_{t_0}^{t} dt'\ \langle[\hat{A}(t),\hat{V}(t')]\rangle_0,
\end{aligned} \tag{5.11}$$

[6] Recall that for a time-independent Hamiltonian $\hat{H}_0$, the time evolution of the state vector $|n\rangle$ is given by $|n(t)\rangle = e^{-i\hat{H}_0 t/\hbar}\,|n\rangle$ in the Schrödinger picture.

where $\langle \cdot \rangle_0$ means that the average is taken with respect to the time-independent Hamiltonian $\hat{H}_0$. Equation (5.11) is the announced Kubo formula, which gives us an expression for the linear response of the observable $\langle \hat{A} \rangle$ to the external perturbation $\hat{V}(t)$. It is worth noting that the time dependence of $\hat{A}$ in the commutator is described by the Hamiltonian $\hat{H}_0$ of the unperturbed system.

5.2 Green's Functions and the Matsubara Formalism

With the Kubo formula (5.11), computing the linear response of an observable amounts to calculating the expectation value of the commutator of the observable and the perturbation. Calculating such correlation functions is not necessarily trivial, particularly when temperature is also taken into account. Fortunately, this is made easier by the introduction of Green's functions. We will now present the real time Green's functions and some of its properties, and then introduce the Matsubara formalism, which is conveniently used to calculate expectation values at finite temperatures. This Section is based on results derived by Ottesen [21] and Mahan [22, p. 109-142].

5.2.1 Real Time Green's Functions

The retarded and advanced real time two-point Green's functions are defined by

$$G^{r,a}_{ij}(t,t') = \mp \frac{i}{\hbar} \theta(\pm(t-t')) \langle [\hat{a}_i(t), \hat{a}^\dagger_j(t')]_\zeta \rangle, \tag{5.12}$$

where the upper sign defines the retarded Green's function $G^r_{ij}(t,t')$ and the lower sign defines the advanced Green's function $G^a_{ij}(t,t')$. As before, $\{\hat{a}_i\}$ are fermionic or bosonic operators. The Green's functions (5.12) depend only on relative time because of the unit step function, which makes it natural to introduce their Fourier transforms:

$$\begin{aligned} \tilde{G}^{r,a}_{ij}(\omega) &= \int_{-\infty}^{\infty} dt \; G^{r,a}_{ij}(t) e^{i\omega t}. \\ G^{r,a}_{ij}(t) &= \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \; \tilde{G}^{r,a}_{ij}(\omega) e^{-i\omega t}. \end{aligned} \tag{5.13}$$

It is easy to verify that $\tilde{G}^r_{ij}(\omega) = (\tilde{G}^a_{ij}(\omega))^*$, which means that once we know the retarded (advanced) Green's function, the advanced (retarded) one automatically follows. We define the spectral weight

$$C_{ij}(t,t') = \frac{\langle [\hat{a}_i(t), \hat{a}^\dagger_j(t')]_\zeta \rangle}{\hbar}, \tag{5.14}$$

which lets us write the Green's functions as

$$G^{r,a}_{ij}(t-t') = \mp i \theta(\pm(t-t')) C_{ij}(t,t'). \tag{5.15}$$

To see why we call $C_{ij}(t,t')$ the spectral weight we write out (5.14) explicitly:

$$\begin{aligned} C_{ij}(t,t') &= \frac{1}{\hbar Z}\,\mathrm{Tr}\Big\{e^{-\beta\hat{H}}[\hat{a}_i(t),\hat{a}_j^\dagger(t')]_\zeta\Big\} \\ &= \frac{1}{\hbar Z}\sum_n e^{-\beta E_n}\big[\,\langle n|\,\hat{a}_i(t)\hat{a}_j^\dagger(t')\,|n\rangle - \zeta\,\langle n|\,\hat{a}_j^\dagger(t')\hat{a}_i(t)\,|n\rangle\,\big] \\ &= \frac{1}{\hbar Z}\sum_{n,m} e^{-\beta E_n}\big[\,\langle n|\,\hat{a}_i(t)\,|m\rangle\,\langle m|\,\hat{a}_j^\dagger(t')\,|n\rangle - \zeta\,\langle n|\,\hat{a}_j^\dagger(t')\,|m\rangle\,\langle m|\,\hat{a}_i(t)\,|n\rangle\,\big] \\ &= \frac{1}{\hbar Z}\sum_{n,m}\big[e^{-\beta E_n} - \zeta e^{-\beta E_m}\big]\,\langle n|\,\hat{a}_i(t)\,|m\rangle\,\langle m|\,\hat{a}_j^\dagger(t')\,|n\rangle\,, \end{aligned} \tag{5.16}$$

where the trace was taken using the eigenstates $\{|n\rangle\}$ of the Hamiltonian $\hat{H}$ and a complete set of states $\sum_m |m\rangle\langle m| = I$ was inserted in the third equality. $Z = \mathrm{Tr}\Big\{e^{-\beta\hat{H}}\Big\}$ is the partition function of the system. In the last equality, the dummy indices n and m were interchanged in the second summand. Expressing the matrix elements in terms of (5.5), the spectral weight can be written as

$$C_{ij}(t,t') = \frac{1}{\hbar Z}\sum_{n,m}\big[e^{-\beta E_n} - \zeta e^{-\beta E_m}\big]\,\langle n|\,\hat{a}_i\,|m\rangle\,\langle m|\,\hat{a}_j^\dagger\,|n\rangle\,e^{i(E_n-E_m)(t-t')/\hbar} = C_{ij}(t-t'). \tag{5.17}$$

Since the time dependence of $C_{ij}(t)$ is present only in the exponential factor, its Fourier transform is a delta function. Defining $\omega_n = E_n/\hbar$ and $\omega_m = E_m/\hbar$, we obtain

$$\tilde{C}_{ij}(\omega) = \frac{1}{\hbar Z}\sum_{n,m}\big[e^{-\beta E_n} - \zeta e^{-\beta E_m}\big]\,\langle n|\,\hat{a}_i\,|m\rangle\,\langle m|\,\hat{a}_j^\dagger\,|n\rangle\,2\pi\delta(\omega+\omega_n-\omega_m). \tag{5.18}$$

The reason why this is interesting is because a Green's function can be expressed in its so-called spectral representation:

$$\tilde{G}(z) = \int_{-\infty}^{\infty}\frac{d\omega'}{2\pi}\frac{-2\,\mathrm{Im}\{\tilde{G}(\omega')\}}{z-\omega'}. \tag{5.19}$$

Consider then the retarded Green's function $G_{ij}^r(t)$ defined in (5.12). Inserting (5.17), its Fourier transform is found to be

$$\tilde{G}_{ij}^r(\omega) = \frac{1}{\hbar Z}\sum_{n,m}\frac{e^{-\beta E_n} - \zeta e^{-\beta E_m}}{\omega+\omega_n-\omega_m+i\delta}\,\langle n|\,\hat{a}_i\,|m\rangle\,\langle m|\,\hat{a}_j^\dagger\,|n\rangle\,, \tag{5.20}$$

where δ is an infinitesimal parameter introduced in the Fourier transform to ensure that the integral converges as $t\to\infty$.[7] Using the relation

$$\lim_{\delta\to 0^+}\frac{1}{\omega+\omega_n-\omega_m+i\delta} = P\frac{1}{\omega+\omega_n-\omega_m} - i\pi\delta(\omega+\omega_n-\omega_m),$$

[7]Physically, the integrand has to vanish to ensure that there is no correlation between events that are separated by an infinite period of time.

where P denotes Cauchy's principal value, it is easily confirmed that $-2\,\mathrm{Im}\{\tilde{G}^r_{ij}(\omega)\} = \tilde{C}_{ij}(\omega)$. We can then define the spectral representations of the Green's functions (5.12) as

$$\tilde{G}_{ij}(z) = \int_{-\infty}^{\infty} \frac{d\omega'}{2\pi}\, \frac{\tilde{C}_{ij}(\omega')}{z-\omega'}. \tag{5.21}$$

The real time two-point Green's functions are then obtained as

$$\tilde{G}^{r,a}_{ij}(\omega) = \lim_{\delta\to 0^+} \tilde{G}_{ij}(\omega \pm i\delta), \tag{5.22}$$

where the upper (lower) sign corresponds to the retarded (advanced) Green's function. It will soon become clear why the spectral representation is so useful.

5.2.2 The Matsubara Formalism

As we have seen, calculating average values of quantum mechanical operators at non-zero temperatures involves the density matrix (2.61). At the same time, the time evolution of such operators is described by exponential operators with imaginary arguments. It would be mathematically more convenient if the arguments of all exponential operators were real. To this end, we define $it/\hbar$ to be a real temperature factor, analogous to β, which implies that time t is treated as a complex quantity. The variable describing time evolution is then defined as imaginary time $\tau = it$. This is known as the Matsubara formalism, and it allows for simple calculation of finite temperature Green's functions.

5.2.3 Matsubara Green's Functions

The two-point Matsubara Green's function is given by

$$\begin{aligned}\mathcal{G}_{ij}(\tau_1,\tau_2) &= -\frac{1}{\hbar}\langle \hat{\mathrm{T}}\{\hat{a}_i(\tau_1)\hat{a}^\dagger_j(\tau_2)\}\rangle \\ &= -\frac{1}{\hbar}\theta(\tau_1-\tau_2)\langle \hat{a}_i(\tau_1)\hat{a}^\dagger_j(\tau_2)\rangle - \zeta\frac{1}{\hbar}\theta(\tau_2-\tau_1)\langle \hat{a}^\dagger_j(\tau_2)\hat{a}_i(\tau_1)\rangle = \mathcal{G}_{ij}(\tau_1-\tau_2),\end{aligned} \tag{5.23}$$

where $\hat{\mathrm{T}}$ is the time ordering operator for imaginary time. The time evolution of an operator $\hat{A}$ in the Matsubara formalism is found by substituting $\tau = it$ in (5.5):

$$\hat{A}(\tau) = e^{\hat{H}\tau/\hbar}\hat{A}e^{-\hat{H}\tau/\hbar}. \tag{5.24}$$

Like the real time Green's functions, the Matsubara Green's function is only dependent on relative imaginary time. Assuming that the system is in thermodynamic equilibrium, the Hamiltonian $\hat{H}$ and thus the density matrix $\hat{\rho}$ are time-independent. Taking the derivative of (5.23) with respect to τ, we obtain

$$\hbar\frac{\partial}{\partial\tau}\mathcal{G}_{ij}(\tau) = -\delta(\tau)\delta_{ij} - \frac{1}{\hbar}\langle\hat{\mathrm{T}}\{[\hat{H},\hat{a}_i(\tau)]\hat{a}^\dagger_j(0)\}\rangle, \tag{5.25}$$

where the commutator comes from

$$\frac{\partial}{\partial\tau}\hat{a}_i(\tau) = \frac{1}{\hbar}[\hat{H},\hat{a}_i(\tau)], \tag{5.26}$$

which is found by taking the derivative of (5.24). The delta function in the equation of motion (5.25) for the Matsubara Green's function shows that it is discontinuous at the origin. The unit step function in (5.23) lets us rewrite the Matsubara Green's function in the following way, which further elucidates the discontinuity:

$$\mathcal{G}_{ij}(\tau) = \begin{cases} -\langle \hat{a}_i(\tau)\hat{a}_j^\dagger(0)\rangle/\hbar & \text{for } \tau > 0, \\ -\zeta\langle \hat{a}_j^\dagger(0)\hat{a}_i(\tau)\rangle/\hbar & \text{for } \tau < 0. \end{cases} \tag{5.27}$$

Consider the interval $-\hbar\beta < \tau < 0$. Ignoring the partition function normalization of the density matrix, the expectation value in the Matsubara Green's function is given by

$$\begin{aligned} \langle \hat{a}_j^\dagger(0)\hat{a}_i(\tau)\rangle &= \mathrm{Tr}\Big\{ e^{-\beta\hat{H}}\hat{a}_j^\dagger e^{\hat{H}\tau/\hbar}\hat{a}_i e^{-\hat{H}\tau/\hbar}\Big\} = \mathrm{Tr}\Big\{ e^{\hat{H}\tau/\hbar}\hat{a}_i e^{-\hat{H}\tau/\hbar}e^{-\beta\hat{H}}\hat{a}_j^\dagger\Big\} \\ &= \mathrm{Tr}\Big\{ e^{-\beta\hat{H}}e^{\hat{H}(\tau+\hbar\beta)/\hbar}\hat{a}_i e^{-\hat{H}(\tau+\hbar\beta)/\hbar}\hat{a}_j^\dagger\Big\} \\ &= \mathrm{Tr}\Big\{ e^{-\beta\hat{H}}\hat{a}_i(\tau+\hbar\beta)\hat{a}_j^\dagger(0)\Big\} = \langle \hat{a}_i(\tau+\hbar\beta)\hat{a}_j^\dagger(0)\rangle, \end{aligned} \tag{5.28}$$

where we used the cyclic property of the trace in the second equality. This implies that the Matsubara Green's function satisfies $\mathcal{G}_{ij}(\tau+\hbar\beta) = \zeta\mathcal{G}_{ij}(\tau)$. Since $\zeta^2 = 1$, it follows that the Matsubara Green's function is periodic with $2\hbar\beta$:

$$\mathcal{G}_{ij}(\tau + 2n\hbar\beta) = \mathcal{G}_{ij}(\tau) \quad \forall \quad n \in \mathbb{Z}. \tag{5.29}$$

The Matsubara Green's function is then naturally expressed as a Fourier series:

$$\mathcal{G}_{ij}(\tau) = \frac{1}{\hbar\beta}\sum_{n=-\infty}^{\infty} \tilde{\mathcal{G}}_{ij}(n)e^{-\frac{in\pi}{\hbar\beta}\tau}. \tag{5.30}$$

The inverse transform is found by multiplying (5.30) by $\exp(im\pi\tau/\hbar\beta)$ and integrating both sides over the interval $\tau \in (-\hbar\beta, \hbar\beta)$. Using the orthogonality of the complex exponentials and the periodicity of the Matsubara Green's function, we obtain

$$\tilde{\mathcal{G}}_{ij}(n) = \frac{1}{2}(1+\zeta e^{-in\pi})\int_0^{\hbar\beta} d\tau\ \mathcal{G}_{ij}(\tau)e^{\frac{in\pi}{\hbar\beta}\tau}. \tag{5.31}$$

The prefactor in (5.31) vanishes for odd (even) n if the Matsubara Green's function is bosonic (fermionic). To this end, we define the Matsubara frequencies

$$\mu_n = \begin{cases} 2n\pi/\hbar\beta & \text{for bosons,} \\ (2n+1)\pi/\hbar\beta & \text{for fermions.} \end{cases} \tag{5.32}$$

The Fourier transforms may then be rewritten in terms of the Matsubara frequencies:

$$\begin{aligned} \tilde{\mathcal{G}}_{ij}(i\mu_n) &= \int_0^{\hbar\beta} d\tau\ \mathcal{G}_{ij}(\tau)e^{i\mu_n\tau}, \\ \mathcal{G}_{ij}(\tau) &= \frac{1}{\hbar\beta}\sum_{\mu_n} \tilde{\mathcal{G}}_{ij}(i\mu_n)e^{-i\mu_n\tau}. \end{aligned} \tag{5.33}$$

We are now ready to establish the connection between Matsubara Green's functions and real time Green's functions. Considering both time t and τ to be complex parameters, the definition $\tau = it$ means the two are related by $\mathrm{Re}\{\tau\} = -\mathrm{Im}\{t\}$ and $\mathrm{Im}\{\tau\} = \mathrm{Re}\{t\}$. A positive Matsubara frequency μ_n then means that $\exp(i\mu_n\tau) = \exp(\mu_n(-\mathrm{Re}\{t\} - i\mathrm{Im}\{t\}))$ vanishes as $\mathrm{Re}\{t\}$ approaches positive infinity. This allows us to write

$$\begin{aligned}\tilde{\mathcal{G}}_{ij}(i\mu_n) &= -\frac{1}{\hbar}\int_0^{\hbar\beta} d\tau\ \langle \hat{a}_i(\tau)\hat{a}_j^\dagger(0)\rangle e^{i\mu_n\tau} \\ &= -\frac{1}{\hbar}\int_0^{\infty} d\tau\ \langle \hat{a}_i(\tau)\hat{a}_j^\dagger(0)\rangle e^{i\mu_n\tau} + \frac{1}{\hbar}\int_{\hbar\beta}^{\infty} d\tau\ \langle \hat{a}_i(\tau)\hat{a}_j^\dagger(0)\rangle e^{i\mu_n\tau} \\ &= -\frac{1}{\hbar}\int_0^{\infty} d\tau\ \langle \hat{a}_i(\tau)\hat{a}_j^\dagger(0)\rangle e^{i\mu_n\tau} + \frac{1}{\hbar}e^{i\mu_n\hbar\beta}\int_0^{\infty} d\tau\ \langle \hat{a}_i(\tau+\hbar\beta)\hat{a}_j^\dagger(0)\rangle e^{i\mu_n\tau}.\end{aligned} \tag{5.34}$$

We now express the integrals in terms of time t. Inserting the τ-dependent operators in (5.34) into (5.24), we readily find $\hat{a}_i(\tau) = \hat{a}_i(t)$ and $\hat{a}_i(\tau+\hbar\beta) = \hat{a}_i(t-i\hbar\beta)$. Identifying the prefactor $\exp(i\mu_n\hbar\beta) = \zeta$, (5.34) can be expressed in terms of time t as

$$\begin{aligned}\tilde{\mathcal{G}}_{ij}(i\mu_n) &= -\frac{i}{\hbar}\int_0^{\infty} dt\ \langle \hat{a}_i(t)\hat{a}_j^\dagger(0)\rangle e^{i(i\mu_n)t} + \zeta\frac{i}{\hbar}\int_0^{\infty} dt\ \langle \hat{a}_i(t-i\hbar\beta)\hat{a}_j^\dagger(0)\rangle e^{i(i\mu_n)t} \\ &= -\frac{i}{\hbar}\int_0^{\infty} dt\ \langle \hat{a}_i(t)\hat{a}_j^\dagger(0)\rangle e^{i(i\mu_n)t} + \zeta\frac{i}{\hbar}\int_0^{\infty} dt\ \langle \hat{a}_j^\dagger(0)\hat{a}_i(t)\rangle e^{i(i\mu_n)t} \\ &= -\frac{i}{\hbar}\int_0^{\infty} dt\ \langle [\hat{a}_i(t), \hat{a}_j^\dagger(0)]_\zeta\rangle e^{i(i\mu_n)t} = -\frac{i}{\hbar}\int_{-\infty}^{\infty} dt\ \theta(t)\langle [\hat{a}_i(t), \hat{a}_j^\dagger(0)]_\zeta\rangle e^{i(i\mu_n)t},\end{aligned} \tag{5.35}$$

where we used a similar approach to the one shown in (5.28) to rewrite the second integral in the second equality. The expression for $\tilde{\mathcal{G}}_{ij}(i\mu_n)$ in (5.35) is the Fourier transform (5.13) of the retarded real time Green's function $G_{ij}^r(t)$ defined in (5.12), with the frequency ω replaced by the Matsubara frequency $i\mu_n$. We may then write

$$\tilde{\mathcal{G}}_{ij}(i\mu_n) = \int_{-\infty}^{\infty}\frac{d\omega'}{2\pi}\frac{\tilde{C}_{ij}(\omega')}{i\mu_n - \omega'} = \tilde{G}_{ij}(i\mu_n). \tag{5.36}$$

From (5.21) and (5.22) we see that the retarded real time Green's function is analytical in the upper half of the complex plane. This is also the case for the Matsubara Green's function. This allows us to connect the two by the analytical continuation

$$\tilde{G}_{ij}^r(\omega) = \lim_{i\mu_n\to\omega+i\delta}\tilde{\mathcal{G}}_{ij}(i\mu_n). \tag{5.37}$$

A similar procedure can be used to show that the advanced real time Green's function can be obtained from the Matsubara Green's function by the analytical continuation $i\mu_n \to \omega - i\delta$ for negative μ_n.

5.2.4 Non-Interacting Particle Ensembles

In Section 2, we saw how we could diagonalize Hamiltonians describing ensembles of non-interacting particles. We now study such a system in the Matsubara formalism. A general

non-interacting Hamiltonian may be written

$$\hat{H} = \sum_{i,j} \mathbb{H}_{ij} \hat{a}_i^\dagger \hat{a}_j. \tag{5.38}$$

The specific form of the Hamiltonian lets us find an explicit expression for the differential equation (5.25). To do so, we have to establish what the commutator $[\hat{H}, \hat{a}_i(\tau)]$ is. In the time-independent case, we have

$$\hat{H}\hat{a}_i = \sum_{j,l} \mathbb{H}_{jl} \hat{a}_j^\dagger \hat{a}_l \hat{a}_i = \sum_{j,l} \mathbb{H}_{jl} (\hat{a}_i \hat{a}_j^\dagger - \delta_{ij}) \hat{a}_l = \hat{a}_i \hat{H} - \sum_l \mathbb{H}_{il} \hat{a}_l. \tag{5.39}$$

Since the Hamiltonian commutes with itself, we have:

$$\begin{aligned} \hat{H}\hat{a}_i(\tau) &= e^{\hat{H}\tau/\hbar} \hat{H} \hat{a}_i e^{-\hat{H}\tau/\hbar}, \\ \hat{a}_i(\tau)\hat{H} &= e^{\hat{H}\tau/\hbar} \hat{a}_i \hat{H} e^{-\hat{H}\tau/\hbar}. \end{aligned} \tag{5.40}$$

The commutator is then given by

$$\begin{aligned} [\hat{H}, \hat{a}_i(\tau)] &= e^{\hat{H}\tau/\hbar} \hat{H} \hat{a}_i e^{-\hat{H}\tau/\hbar} - e^{\hat{H}\tau/\hbar} \hat{a}_i \hat{H} e^{-\hat{H}\tau/\hbar} = e^{\hat{H}\tau/\hbar} (\hat{H}\hat{a}_i - \hat{a}_i \hat{H}) e^{-\hat{H}\tau/\hbar} \\ &= e^{\hat{H}\tau/\hbar} \Big(- \sum_l \mathbb{H}_{il} \hat{a}_l \Big) e^{-\hat{H}\tau/\hbar} = - \sum_l \mathbb{H}_{il} \hat{a}_l(\tau). \end{aligned} \tag{5.41}$$

Inserting the commutator (5.41) into (5.25), we obtain

$$\hbar \frac{\partial}{\partial \tau} \mathcal{G}_{ij}(\tau) = \delta(\tau)\delta_{ij} - \sum_l \mathbb{H}_{il} \mathcal{G}_{lj}(\tau). \tag{5.42}$$

Defining $\mathbb{G}(\tau)$ as the matrix whose elements are the two-point Matsubara Green's functions $\mathcal{G}_{ij}(\tau)$, these coupled differential equations may be written as a matrix equation:

$$\Big(- \hbar I \frac{\partial}{\partial \tau} - \mathbb{H} \Big) \mathbb{G}(\tau) = \delta(\tau) I. \tag{5.43}$$

This equation is more conveniently expressed by its Fourier transform:

$$(i\hbar\mu_n I - \mathbb{H}) \tilde{\mathbb{G}}(i\mu_n) = I. \tag{5.44}$$

The Matsubara Green's functions for a system described by a general non-interacting Hamiltonian are identified by computing the inverse of the matrix $(i\hbar\mu_n I - \mathbb{H})$. Consider now the case where the Hamiltonian is diagonal, $\mathbb{H}_{ij} = \varepsilon_i \delta_{ij}$. The commutator (5.41) then reduces to $-\varepsilon_i \hat{a}_i(\tau)$, and we can solve (5.26) to obtain an explicit expression for $\hat{a}_i(\tau)$:

$$\hat{a}_i(\tau) = e^{-\varepsilon_i \tau/\hbar} \hat{a}_i(0). \tag{5.45}$$

Inserting this into (5.23) and using $\theta(-\tau) = 1 - \theta(\tau)$, we obtain

$$\begin{aligned} \mathcal{G}_{ij}(\tau) &= -\frac{1}{\hbar} \theta(\tau) e^{-\varepsilon_i \tau/\hbar} \langle \big[\hat{a}_i \hat{a}_j^\dagger - \zeta \hat{a}_j^\dagger \hat{a}_i \big] \rangle - \zeta \frac{1}{\hbar} e^{-\varepsilon_i \tau/\hbar} \langle \hat{a}_j^\dagger \hat{a}_i \rangle \\ &= -\frac{e^{-\varepsilon_i \tau/\hbar}}{\hbar} \big[\theta(\tau)\delta_{ij} + \zeta \langle \hat{a}_j^\dagger \hat{a}_i \rangle \big]. \end{aligned} \tag{5.46}$$

Using the commutation relation (2.12), the expectation value of the operator product is given by

$$\langle \hat{a}_j^\dagger \hat{a}_i \rangle = \frac{\mathrm{Tr}\Big\{ e^{-\beta \hat{H}} \hat{a}_j^\dagger \hat{a}_i \Big\}}{Z} = \frac{-\zeta \delta_{ij} \, \mathrm{Tr}\Big\{ e^{-\beta \hat{H}} \Big\} + \zeta \, \mathrm{Tr}\Big\{ e^{-\beta \hat{H}} \hat{a}_i \hat{a}_j^\dagger \Big\}}{Z}. \tag{5.47}$$

The second trace in the numerator may be rewritten using the cyclic property of the trace and (5.39):

$$\mathrm{Tr}\Big\{ e^{-\beta \hat{H}} \hat{a}_i \hat{a}_j^\dagger \Big\} = \mathrm{Tr}\Big\{ \hat{a}_j^\dagger e^{-\beta \hat{H}} \hat{a}_i \Big\} = \mathrm{Tr}\Big\{ \hat{a}_j^\dagger \hat{a}_i e^{-\beta \hat{H}} e^{\beta \varepsilon_i} \Big\} = e^{\beta \varepsilon_i} \, \mathrm{Tr}\Big\{ e^{-\beta \hat{H}} \hat{a}_j^\dagger \hat{a}_i \Big\}. \tag{5.48}$$

Solving (5.47) for $\mathrm{Tr}\Big\{ e^{-\beta \hat{H}} \hat{a}_j^\dagger \hat{a}_i \Big\}$, we obtain a simple expression for the expectation value of the operator product:

$$\langle \hat{a}_j^\dagger \hat{a}_i \rangle = \frac{\delta_{ij}}{e^{\beta \varepsilon_i} - \zeta} = \delta_{ij} n_\zeta(\varepsilon_i), \tag{5.49}$$

where $n_\zeta(\varepsilon_i)$ is the Bose-Einstein distribution function for $\zeta = 1$ and the Fermi-Dirac distribution function for $\zeta = -1$. It is now evident that the Matsubara Green's function (5.46) is non-zero only if $i = j$, which prompts us to write it as

$$\mathcal{G}_i(\tau) = -\frac{e^{-\varepsilon_i \tau / \hbar}}{\hbar} \big[\theta(\tau) + \zeta n_\zeta(\varepsilon_i) \big]. \tag{5.50}$$

This expression can easily be shown to satisfy the equation of motion (5.25). The Fourier transform (5.33) of (5.50) is

$$\tilde{\mathcal{G}}_i(i\mu_n) = \frac{1}{i\hbar\mu_n - \varepsilon_i}, \tag{5.51}$$

which agrees with (5.43). Through the analytical continuation (5.37), this equation provides simple expressions for the real time Green's functions of a non-interacting, diagonal Hamiltonian.

5.3 Current Response to Vector Potentials

We will now use the Kubo formula (5.11) and the Green's functions machinery introduced in the previous Sections to compute the response function for a particular type of perturbation. Consider the case where the perturbation $\hat{V}(t)$ is given by a vector potential $\vec{A}(\vec{r}, t)$ which couples to a current density $\vec{J}(\vec{r}, t)$:

$$\hat{V}(t) = -\int d^d r \; \hat{\vec{j}}(\vec{r}, t) \cdot \vec{A}(\vec{r}, t), \tag{5.52}$$

where the integral is performed in d dimensions. We distinguish between the current density $\vec{J}(\vec{r}, t)$, which is what is measured in an experiment, and the microscopic current density operator $\hat{\vec{j}}(\vec{r}, t)$, which appears in the Hamiltonian. A coupling of the form (5.52) describes the case where an electric current is induced by an applied electric field, or when a thermal current is induced by a temperature gradient. Since we are familiar with Ohm's law from electromagnetism, we use the associated notation to discuss general current response in what

follows. The current density $\vec{J}(\vec{r},t)$ induced by an applied field $\vec{E}(\vec{r},t)$, related to the vector potential $\vec{A}(\vec{r},t)$ by

$$\vec{E}(\vec{r},t) = -\frac{\partial}{\partial t}\vec{A}(\vec{r},t),^{8} \tag{5.53}$$

may be written as

$$J_i(\vec{r},t) = \int d^d r' \int dt' \sum_l \sigma_{il}(\vec{r},t;\vec{r}',t')E_l(\vec{r}',t'), \tag{5.54}$$

where i and l denote spatial directions and $\sigma_{il}(\vec{r},t;\vec{r}',t')$ is the response function. Since we assume that the system is in thermodynamic equilibrium before the perturbation is applied, the response function depends only on relative time: $\sigma_{il}(\vec{r},t;\vec{r}',t') = \sigma_{il}(\vec{r},\vec{r}',t-t')$. The temporal integral in (5.54) is then a convolution, which means the temporal Fourier transform of the current density is just the product of the Fourier transforms of the convoluted functions:

$$J_i(\vec{r},\omega) = \int d^d r' \sum_l \sigma_{il}(\vec{r},\vec{r}',\omega)E_j(\vec{r}',\omega). \tag{5.55}$$

In general, the response function does not depend only on relative spatial positions. However, if the system we consider has a crystal structure, the macroscopic response function can be regarded as an average of microscopic response functions for the constituent unit cells. The system effectively becomes translationally invariant, such that $\sigma_{il}(\vec{r},\vec{r}',\omega) = \sigma_{il}(\vec{r}-\vec{r}',\omega)$, and the remaining spatial integral in (5.55) is also a convolution. The Fourier transformed current density then takes on a very simple form compared to (5.54):

$$J_i(\vec{k},\omega) = \sum_l \sigma_{il}(\vec{k},\omega)E_l(\vec{k},\omega). \tag{5.56}$$

Fourier transforming (5.53) yields

$$\vec{E}(\vec{k},\omega) = i\omega\vec{A}(\vec{k},\omega), \tag{5.57}$$

and by defining the vector potential response function

$$K_{il}(\vec{k},\omega) = -i\omega\sigma_{il}(\vec{k},\omega), \tag{5.58}$$

we may write the relation between the current density and the vector potential as

$$J_i(\vec{k},\omega) = -\sum_l K_{il}(\vec{k},\omega)A_l(\vec{k},\omega). \tag{5.59}$$

We now use the Kubo formula to find an expression for the current density, which we will compare to (5.59) to identify the response function. The current response to the perturbation (5.52) is given by:

$$\begin{aligned} J_i(\vec{r},t) &= \langle\hat{j}_i(\vec{r},t)\rangle_0 - \frac{i}{\hbar}\int_{t_0}^{t} dt' \ \langle[\hat{j}_i(\vec{r},t),\hat{V}(t')]\rangle_0 \\ &= \langle\hat{j}_i(\vec{r},t)\rangle_0 + \frac{i}{\hbar}\int_{t_0}^{t} dt' \int d^d r' \sum_l \langle[\hat{j}_i(\vec{r},t),\hat{j}_l(\vec{r}',t')]\rangle_0 A_l(\vec{r}',t'). \end{aligned} \tag{5.60}$$

[8] This can be accomplished by a suitable choice of gauge.

We introduce the current-current correlation function

$$\Pi_{il}(\vec{r},t;\vec{r}',t') = -\frac{i}{\hbar}\theta(t-t')\langle[\hat{j}_i(\vec{r},t),\hat{j}_l(\vec{r}',t')]\rangle_0, \tag{5.61}$$

which when setting $t_0 = -\infty$ lets us write (5.60) as

$$J_i(\vec{r},t) = \langle\hat{j}_i(\vec{r},t)\rangle_0 - \int_{-\infty}^{\infty} dt' \int d^d r' \sum_l \Pi_{il}(\vec{r},t;\vec{r}',t')A_l(\vec{r}',t'). \tag{5.62}$$

The correlation function (5.61) has the exact same form as the retarded real time Green's function (5.12), with the creation and annihilation operators replaced by current density operators. Invoking translational and temporal invariance like before, the Fourier transform of (5.62) is given by

$$J_i(\vec{k},\omega) = \langle\hat{j}_i(\vec{k},\omega)\rangle_0 - \sum_l \Pi_{il}(\vec{k},\omega)A_l(\vec{k},\omega). \tag{5.63}$$

A current density associated with a vector potential may be decomposed into a diamagnetic and a paramagnetic part, where the former is proportional to the vector potential and the latter is independent of it. Therefore, the first term on the right hand side of (5.63) is strictly diamagnetic; since there is no current in the equilibrium state in the absence of the perturbation, the expectation value of the paramagnetic contribution must be zero. On the other hand, the current-current correlation in the second term is strictly paramagnetic since we are considering the response to linear order in the perturbation. This analysis suggests that the vector potential response function (5.58) can also be decomposed into a diamagnetic and a paramagnetic part:

$$K_{il} = K_{il}^D + K_{il}^P. \tag{5.64}$$

It can be shown that the diamagnetic current depends on the frequency ω only through the vector potential that drives it. We then see from (5.59) and (5.63) that the diamagnetic response $K_{il}^D(\vec{k},\omega)$ is independent of the frequency ω, while the paramagnetic response $K_{il}^P(\vec{k},\omega) = \Pi_{il}(\vec{k},\omega)$. At first glance, this seems highly problematic, since the response function

$$\sigma_{il}(\vec{k},\omega) = \frac{i}{\omega}K_{il}(\vec{k},\omega) \tag{5.65}$$

diverges as $\omega \to 0$ due to the diamagnetic contribution. To avoid this divergence, we impose[9] that $K_{il}^D(\vec{k}) = -K_{il}^P(\vec{k},0)$, which lets us write the response function as

$$\sigma_{il}(\vec{k},\omega) = \frac{i}{\omega}\big[\Pi_{il}(\vec{k},\omega) - \Pi_{il}(\vec{k},0)\big]. \tag{5.66}$$

The response function is now expressed entirely in terms of retarded Green's functions! In order to compute it explicitly, we must first obtain an expression for the correlation function $\Pi_{il}(\vec{k},\omega)$. We start off by writing the applied field in terms of a complex exponential:

$$\vec{E}(\vec{r},t) = \vec{E}_0 e^{i(\vec{k}\cdot\vec{r}-\omega t)}, \tag{5.67}$$

[9]This imposition is not actually made by hand to avoid the divergence, but appears rigorously in the derivation if we go to second order in the the perturbation.

where $\vec{E}_0$ describes the amplitude and direction of the field. The corresponding vector potential is found by solving (5.53):

$$\vec{A}(\vec{r},t) = -\frac{i}{\omega}\vec{E}(\vec{r},t). \tag{5.68}$$

The perturbation $\hat{V}(t)$ then becomes

$$\begin{aligned}\hat{V}(t) &= -\int d^d r\ \hat{\vec{j}}(\vec{r},t)e^{i\vec{k}\cdot\vec{r}}\Big(-\frac{i}{\omega}\vec{E}_0 e^{-i\omega t}\Big) = -\hat{\vec{j}}(-\vec{k},t)\Big(-\frac{i}{\omega}\vec{E}_0 e^{-i\omega t}\Big) \\ &= -\hat{\vec{j}}(-\vec{k},t)e^{-i\vec{k}\cdot\vec{r}}\vec{A}(\vec{r},t),\end{aligned} \tag{5.69}$$

where we used the definition of the Fourier transform in the second equality. Inserting (5.69) into the Kubo formula (5.60), the current-current correlation function is found to be

$$\Pi_{il}(\vec{r},t;\vec{k},t') = -\frac{i}{\hbar}\theta(t-t')e^{-i\vec{k}\cdot\vec{r}}\langle[\hat{j}_i(\vec{r},t),\hat{j}_l(-\vec{k},t')]\rangle_0. \tag{5.70}$$

The correlation function depends only on relative time, so we set $t' = 0$. In order to get rid of the microscopic fluctuations in the response function and thereby justify our assumption of translational invariance, we perform a spatial average of the correlation function (5.70) over the entire volume V of the system:

$$\begin{aligned}\Pi_{il}(\vec{k},t) &= -\frac{i}{\hbar V}\theta(t)\int d^d r\ e^{-i\vec{k}\cdot\vec{r}}\langle[\hat{j}_i(\vec{r},t),\hat{j}_l(-\vec{k},0)]\rangle_0 \\ &= -\frac{i}{\hbar V}\theta(t)\langle[\hat{j}_i(\vec{k},t),\hat{j}_l(-\vec{k},0)]\rangle_0.\end{aligned} \tag{5.71}$$

This prompts the definition of the Matsubara current-current correlation function

$$\pi_{il}(\vec{k},i\omega_n) = -\frac{1}{\hbar V}\int_0^{\hbar\beta} d\tau\ e^{i\omega_n\tau}\langle\hat{\mathrm{T}}\{\hat{j}_i(\vec{k},\tau)\hat{j}_l(-\vec{k},0)\}\rangle_0, \tag{5.72}$$

where ω_n is a bosonic Matsubara frequency.[10] The response function (5.66) can readily be calculated in the more convenient Matsubara formalism, as the two correlation functions are related by the analytical continuation

$$\Pi_{il}(\vec{k},\omega) = \lim_{i\omega_n\to\omega+i\delta}\pi_{il}(\vec{k},i\omega_n). \tag{5.73}$$

[10]Regardless of whether a current density operator is related to fermions or bosons, the product of two current density operators is periodic with $\hbar\beta$.

6 Theory of Hall Conductivities

With the linear response theory established in detail in Section 5, we are now well equipped to compute response functions for a range of perturbations in both fermionic and bosonic systems. As foreshadowed by Section 5.3, we are particularly interested in obtaining conductivity formulas describing current responses to vector potentials. As an example of this, we derive the Hall conductivity in an electronic insulator in great detail. We then present transverse conductivity formulas describing spin and thermal current responses to a temperature gradient in bosonic spin systems, which we will eventually use to study the Kane-Mele-Hubbard model. As we shall see, all of these conductivity formulas are directly related to the topological properties of the system.

6.1 The Hall Conductivity

The Hall conductivity σ_{xy}, which is the transverse current response to an applied electric field, is perhaps the most famous example of the link between transport and topology in a quantum mechanical system. It describes the quantum Hall effect, which is the quantized version of the well known classical Hall effect. In fact, the Hall conductivity is proportional to the Chern number, making it a topological invariant. We now show this by computing σ_{xy} in the so-called flat-band limit of an insulator using linear response theory.

6.1.1 The Current Density Operator

To compute the Hall conductivity σ_{xy} using (5.66), we must first derive the expression for the current density operator $\hat{\vec{j}}$. We do this using the continuity equation

$$-e\frac{\partial}{\partial t}\hat{n}(\vec{r}) + \nabla \cdot \hat{\vec{j}}(\vec{r}) = 0, \tag{6.1}$$

where $-e$ is the charge of an electron and $\hat{n}(\vec{r}) = \sum_\alpha n_\alpha(\vec{r})$ is the number operator for electrons with orbital indices α at the position $\vec{r}$. Taking the Fourier transform of (6.1), we obtain

$$-e\frac{\partial}{\partial t}\hat{n}(\vec{k}) - i\vec{k} \cdot \hat{\vec{j}}(\vec{k}) = 0. \tag{6.2}$$

The Fourier transform of the number operator is given by [12, p. 16-22]

$$\hat{n}(\vec{k}) = \sum_{\vec{q}} \sum_{\alpha} \hat{c}^\dagger_{\vec{q}\alpha} \hat{c}_{\vec{q}+\vec{k}\alpha}, \tag{6.3}$$

where the sum over $\vec{q}$ is taken over the first Brillouin zone. The time evolution of the number operator is found by taking the derivative of (5.5):

$$\frac{\partial}{\partial t}\hat{n}(\vec{k}) = \frac{i}{\hbar}[\hat{H}, \hat{n}(\vec{k})]. \tag{6.4}$$

We write the Hamiltonian of the system

$$\hat{H} = \sum_{\vec{k}} \sum_{\alpha,\beta} \hat{c}^\dagger_{\vec{k}\alpha} \mathbb{H}_{\alpha\beta}(\vec{k}) \hat{c}_{\vec{k}\beta}, \tag{6.5}$$

and use the fermionic commutation relations to obtain

$$\begin{aligned}[\hat{H}, \hat{n}(\vec{k})] &= \sum_{\vec{p}} \sum_{\vec{q}} \sum_{\alpha,\beta,\gamma} \mathbb{H}_{\alpha\beta}(\vec{p}) [\hat{c}^\dagger_{\vec{p}\alpha} \hat{c}_{\vec{p}\beta}, \hat{c}^\dagger_{\vec{q}\gamma} \hat{c}_{\vec{q}+\vec{k}\gamma}] \\ &= \sum_{\vec{q}} \sum_{\alpha,\beta} [\mathbb{H}_{\alpha\beta}(\vec{q}) - \mathbb{H}_{\alpha\beta}(\vec{q}+\vec{k})] \hat{c}^\dagger_{\vec{q}\alpha} \hat{c}_{\vec{q}+\vec{k}\beta}.\end{aligned} \tag{6.6}$$

We focus on the low energy, long wavelength case where $\vec{k}$ is small, which means the applied external field varies slowly in space. This is often the case in practice. To first order in $\vec{k}$, we have $\mathbb{H}_{\alpha\beta}(\vec{q}+\vec{k}) - \mathbb{H}_{\alpha\beta}(\vec{q}) \approx \vec{k} \cdot \nabla_{\vec{q}} \mathbb{H}_{\alpha\beta}(\vec{q})$, and the continuity equation becomes

$$i\vec{k} \cdot \hat{\vec{j}}(\vec{k}) = i\vec{k} \cdot \frac{e}{\hbar} \sum_{\vec{q}} \sum_{\alpha,\beta} \nabla_{\vec{q}} \mathbb{H}_{\alpha\beta}(\vec{q}) \hat{c}^\dagger_{\vec{q}\alpha} \hat{c}_{\vec{q}+\vec{k}\beta}. \tag{6.7}$$

We have thus obtained an expression for the current density operator:

$$\hat{\vec{j}}(\vec{k}) = \frac{e}{\hbar} \sum_{\vec{q}} \sum_{\alpha,\beta} \nabla_{\vec{q}} \mathbb{H}_{\alpha\beta}(\vec{q}) \hat{c}^\dagger_{\vec{q}\alpha} \hat{c}_{\vec{q}+\vec{k}\beta}. \tag{6.8}$$

6.1.2 The Flat-Band Limit of an Insulator

As we see from (6.8), the current density operator generally depends on the details of the Hamiltonian. If two systems are described by Hamiltonians $\hat{H}_1$ and $\hat{H}_2$, with different eigenvalues and eigenstates, their response to an external perturbation will generally be different. Suppose then that the response function depends only on the topological properties of the system. If the two Hamiltonians are adiabatically connected, that is, they have the same eigenstates, we can use the current density operator of either to compute the response function. This motivates us to find an insulating Hamiltonian for which the current density operator takes a particularly simple form. Consider therefore the Hamiltonian

$$\hat{H} = \sum_{\vec{k}} \hat{\mathbf{c}}^\dagger_{\vec{k}} \mathbb{H}(\vec{k}) \hat{\mathbf{c}}_{\vec{k}}, \tag{6.9}$$

where $\hat{\mathbf{c}}_{\vec{k}}$ is a vector containing fermionic annihilation operators and $\mathbb{H}(\vec{k})$ is an $N \times N$ matrix. Suppose that p out of the Hamiltonian's N bands are filled. If we take the Fermi level, which lies in the gap between the filled and unfilled states, to be zero, the bands $E_n(\vec{k})$ satisfy

$$E_1(\vec{k}) \leq \cdots \leq E_p(\vec{k}) < 0 < E_{p+1}(\vec{k}) \leq \cdots \leq E_N(\vec{k}). \tag{6.10}$$

We can then pick two constant energies $E^- < 0$ and $E^+ > 0$, which the filled and unfilled bands can be adiabatically connected to through the interpolations

$$\begin{aligned}\tilde{E}_n(\vec{k}, t) &= E_n(\vec{k})(1-t) + E^- t; \quad 1 \leq n \leq p, \\ \tilde{E}_n(\vec{k}, t) &= E_n(\vec{k})(1-t) + E^+ t; \quad p+1 \leq n \leq N,\end{aligned} \tag{6.11}$$

where $t \in [0, 1]$. These adiabatic interpolations do not alter the eigenstates of the Hamiltonian. We can therefore define the interpolated Hamiltonian matrix $\mathbb{H}(\vec{k}, t)$, which can

be diagonalized by the unitary[11] matrix $\mathbb{U}_{\vec{k}}$ whose columns are the eigenvectors $\{|n(\vec{k})\rangle\}$ of $\mathbb{H}(\vec{k}) = \mathbb{H}(\vec{k}, 0)$. The interpolated Hamiltonian may therefore be written

$$\mathbb{H}(\vec{k}, t) = \mathbb{U}_{\vec{k}} \operatorname{diag}(\tilde{E}_1(\vec{k}, t), \ldots, \tilde{E}_p(\vec{k}, t), \tilde{E}_{p+1}(\vec{k}, t), \ldots \tilde{E}_N(\vec{k}, t))\mathbb{U}_{\vec{k}}^\dagger, \tag{6.12}$$

which for $t = 1$ takes the simple form

$$\mathbb{H}(\vec{k}, 1) = E^- \sum_{n=1}^{p} |n(\vec{k})\rangle \langle n(\vec{k})| + E^+ \sum_{n=p+1}^{N} |n(\vec{k})\rangle \langle n(\vec{k})| . \tag{6.13}$$

The Hamiltonian is the sum of the projectors $P_n(\vec{k}) = |n(\vec{k})\rangle \langle n(\vec{k})|$ of the filled and unfilled bands onto the two flat bands E^- and E^+. This is known as the flat-band limit. The explicit expression for the Hall conductivity is easily obtained in the flat-band limit of an insulator.

6.1.3 The Matsubara Current-Current Correlation Function

We now compute the Matsubara current-current correlation function (5.72) in the flat-band limit of an insulator. For simplicity, we consider a two-band Hamiltonian in the following. In the flat-band limit, the two-band Hamiltonian is given by

$$\mathbb{H}(\vec{k}) = E^- |\psi_-\rangle \langle \psi_-| + E^+ |\psi_+\rangle \langle \psi_+| = E^- P_-(\vec{k}) + E^+ P_+(\vec{k}), \tag{6.14}$$

where $|\psi_\pm\rangle = |\psi_\pm(\vec{k})\rangle$ are the eigenstates corresponding to the filled and unfilled bands. The projectors satisfy

$$P_- + P_+ = I; \quad P_\mp P_\pm = 0; \quad (P_\pm)^2 = P_\pm, \tag{6.15}$$

which will be useful in the computation that follows. We begin by inserting the current density operator (6.8) into the Matsubara correlation function (5.72):

$$\begin{aligned} \pi_{xy}(\vec{k}, i\omega_n) = &- \frac{e^2}{\hbar^3 V} \sum_{\vec{q},\vec{p}} \sum_{\alpha,\beta,\gamma,\delta} \frac{\partial \mathbb{H}_{\alpha\beta}(\vec{q})}{\partial q_x} \frac{\partial \mathbb{H}_{\gamma\delta}(\vec{p})}{\partial p_y} \\ &\times \int_0^{\hbar\beta} d\tau \; e^{i\omega_n \tau} \langle \hat{\mathrm{T}}\{\hat{c}^\dagger_{\vec{q}\alpha}(\tau) \hat{c}_{\vec{q}+\vec{k}\beta}(\tau) \hat{c}^\dagger_{\vec{p}\gamma} \hat{c}_{\vec{p}-\vec{k}\delta}\}\rangle_0. \end{aligned} \tag{6.16}$$

The correlation function involves a four-point Matsubara Green's function. By Wick's theorem, an n-point Matsubara Green's function is equal to the sum of products of inequivalent two-point Matsubara Green's functions. The sign of each term in the sum is determined by the number of times two operators are exchanged in the n-point Green's function; exchanging two bosonic operators does not produce a sign change, whereas exchanging two fermionic operators produces a negative sign. The four-point Matsubara Green's function can therefore be written

$$\begin{aligned} \frac{1}{\hbar^2} \langle \hat{\mathrm{T}}\{\hat{c}^\dagger_{\vec{q}\alpha}(\tau) \hat{c}_{\vec{q}+\vec{k}\beta}(\tau) \hat{c}^\dagger_{\vec{p}\gamma} \hat{c}_{\vec{p}-\vec{k}\delta}\}\rangle_0 = &\frac{1}{\hbar^2} \langle \hat{\mathrm{T}}\{\hat{c}_{\vec{q}+\vec{k}\beta}(\tau) \hat{c}^\dagger_{\vec{q}\alpha}(\tau)\}\rangle_0 \langle \hat{\mathrm{T}}\{\hat{c}_{\vec{p}-\vec{k}\delta} \hat{c}^\dagger_{\vec{p}\gamma}\}\rangle_0 \\ &- \frac{1}{\hbar^2} \langle \hat{\mathrm{T}}\{\hat{c}_{\vec{q}+\vec{k}\beta}(\tau) \hat{c}^\dagger_{\vec{p}\gamma}\}\rangle_0 \langle \hat{\mathrm{T}}\{\hat{c}_{\vec{p}-\vec{k}\delta} \hat{c}^\dagger_{\vec{q}\alpha}(\tau)\}\rangle_0. \end{aligned} \tag{6.17}$$

[11] Recall from Section 2.2 that the operator vector must be canonically transformed when the Hamiltonian is diagonalized. This is accomplished by a unitary matrix when the operators are fermionic.

The first term involves the average of the current density operator in the equilibrium state, which is zero. Because we assume that our system is translationally invariant, the two-point Matsubara Green's function $\mathcal{G}_{ij}(\vec{r},\vec{r}')$ depends only on the relative coordinate $\vec{r}-\vec{r}'$. Fourier transforming its constituent operators, we obtain

$$\mathcal{G}_{ij}(\vec{r}-\vec{r}') = \frac{1}{N}\sum_{\vec{k},\vec{k}'} e^{i\vec{k}\cdot\vec{r}}\mathcal{G}_{ij}(\vec{k},\vec{k}')e^{-i\vec{k}'\cdot\vec{r}'} = \frac{1}{N}\sum_{\vec{k},\vec{k}'} e^{i\vec{k}\cdot(\vec{r}-\vec{r}')}\mathcal{G}_{ij}(\vec{k},\vec{k}')e^{i(\vec{k}-\vec{k}')\cdot\vec{r}'}, \tag{6.18}$$

and since there is no explicit dependence on $\vec{r}'$, we must have $\mathcal{G}_{ij}(\vec{k},\vec{k}') = \mathcal{G}_{ij}(\vec{k})\delta_{\vec{k},\vec{k}'}$. We therefore have

$$\frac{1}{\hbar^2}\langle\hat{\mathrm{T}}\{\hat{c}^\dagger_{\vec{q}\alpha}(\tau)\hat{c}_{\vec{q}+\vec{k}\beta}(\tau)\hat{c}^\dagger_{\vec{p}\gamma}\hat{c}_{\vec{p}-\vec{k}\delta}\}\rangle_0 = -\mathcal{G}_{\beta\gamma}(\vec{q}+\vec{k},\tau)\mathcal{G}_{\delta\alpha}(\vec{p}-\vec{k},-\tau)\delta_{\vec{p},\vec{q}+\vec{k}}, \tag{6.19}$$

and the Wick-decomposed correlation function is

$$\begin{aligned}\pi_{xy}(\vec{k},i\omega_n) =& \frac{e^2}{\hbar V}\sum_{\vec{q}}\sum_{\alpha,\beta,\gamma,\delta}\int_0^{\hbar\beta} d\tau\ e^{i\omega_n\tau} \\ &\times\left[\frac{\partial\mathbb{H}_{\alpha\beta}(\vec{q})}{\partial q_x}\mathcal{G}_{\beta\gamma}(\vec{q}+\vec{k},\tau)\frac{\partial\mathbb{H}_{\gamma\delta}(\vec{q}+\vec{k})}{\partial q_y}\mathcal{G}_{\delta\alpha}(\vec{q},-\tau)\right].\end{aligned} \tag{6.20}$$

The integral is easily performed by expressing the Matsubara Green's functions in terms of their Fourier transforms (5.33):

$$\begin{aligned}&\int_0^{\hbar\beta} d\tau\ e^{i\omega_n\tau}\mathcal{G}_{\beta\gamma}(\vec{q}+\vec{k},\tau)\mathcal{G}_{\delta\alpha}(\vec{q},-\tau) \\ &= \frac{1}{(\hbar\beta)^2}\sum_{\nu_m}\sum_{\nu_n}\int_0^{\hbar\beta} d\tau\ e^{i(\omega_n-\nu_m+\nu_n)\tau}\tilde{\mathcal{G}}_{\beta\gamma}(\vec{q}+\vec{k},i\nu_m)\tilde{\mathcal{G}}_{\delta\alpha}(\vec{q},i\nu_n) \\ &= \frac{1}{\hbar\beta}\sum_{\nu_m}\sum_{\nu_n}\delta_{\nu_m,\omega_n+\nu_n}\tilde{\mathcal{G}}_{\beta\gamma}(\vec{q}+\vec{k},i\nu_m)\tilde{\mathcal{G}}_{\delta\alpha}(\vec{q},i\nu_n) \\ &= \frac{1}{\hbar\beta}\sum_{\nu_n}\tilde{\mathcal{G}}_{\beta\gamma}(\vec{q}+\vec{k},i\omega_n+i\nu_n)\tilde{\mathcal{G}}_{\delta\alpha}(\vec{q},i\nu_n),\end{aligned} \tag{6.21}$$

where ν_m and ν_n are fermionic Matsubara frequencies and we used the orthogonality of the complex exponentials in the second equality. The correlation function then becomes

$$\begin{aligned}\pi(\vec{k},i\omega_n) &= \frac{e^2}{\hbar^2 V\beta}\sum_{\vec{q}}\sum_{\nu_n}\sum_{\alpha,\beta,\gamma,\delta}\left[\frac{\partial\mathbb{H}_{\alpha\beta}(\vec{q})}{\partial q_x}\tilde{\mathcal{G}}_{\beta\gamma}(\vec{q}+\vec{k},i\omega_n+i\nu_n)\frac{\partial\mathbb{H}_{\gamma\delta}(\vec{q}+\vec{k})}{\partial q_y}\tilde{\mathcal{G}}_{\delta\alpha}(\vec{q},i\nu_n)\right] \\ &= \frac{e^2}{\hbar^2 V\beta}\sum_{\vec{q}}\sum_{\nu_n}\mathrm{Tr}\left\{\frac{\partial\mathbb{H}(\vec{q})}{\partial q_x}\tilde{\mathbb{G}}(\vec{q}+\vec{k},i\omega_n+i\nu_n)\frac{\partial\mathbb{H}(\vec{q}+\vec{k})}{\partial q_y}\tilde{\mathbb{G}}(\vec{q},i\nu_n)\right\}.\end{aligned} \tag{6.22}$$

Since the Hamiltonian matrix is diagonal in the projectors, the Matsubara Green's function matrix takes the form

$$\tilde{\mathbb{G}}(\vec{k},i\omega_n) = \frac{P_-(\vec{k})}{i\hbar\omega_n - E_-} + \frac{P_+(\vec{k})}{i\hbar\omega_n - E_+}, \tag{6.23}$$

which is easily confirmed to satisfy (5.44) using the projector relations (6.15). The first of these relations also allows us to rewrite the Hamiltonian matrix as $\mathbb{H}(\vec{k}) = E_+ I + (E_- - E_+)P_-(\vec{k})$, which means the derivatives in (6.22) can be expressed in terms of the projector of the filled band. We then get

$$\begin{aligned}\pi_{xy}(\vec{k}, i\omega_n) =& \frac{e^2(E_- - E_+)^2}{\hbar^2 V\beta} \sum_{\vec{q}} \sum_{\nu_n} \mathrm{Tr}\left\{ \frac{\partial P_-(\vec{q})}{\partial q_x} \left(\frac{P_-(\vec{q}+\vec{k})}{i\hbar\omega_n + i\hbar\nu_n - E_-} + \frac{P_+(\vec{q}+\vec{k})}{i\hbar\omega_n + i\hbar\nu_n - E_+} \right) \right. \\ & \left. \times \frac{\partial P_-(\vec{q}+\vec{k})}{\partial q_y} \left(\frac{P_-(\vec{q})}{i\hbar\nu_n - E_-} + \frac{P_+(\vec{q})}{i\hbar\nu_n - E_+} \right) \right\}.\end{aligned} \tag{6.24}$$

When the spatial variation of the applied electric field is small, the Hall conductivity is well described by the uniform conductivity tensor $\sigma_{xy}(\omega) = \sigma_{xy}(0,\omega)$. We therefore set $\vec{k} = 0$ in the following. To continue the derivation, we write out the products of projectors in (6.24) explicitly:

$$\begin{aligned}(\partial_i P_-)(P_- + P_+)(\partial_j P_-)(P_- + P_+) =& (\partial_i P_-)P_+(\partial_j P_-)P_+ + (\partial_i P_-)P_+(\partial_j P_-)P_- \\ & + (\partial_i P_-)P_-(\partial_j P_-)P_+ + (\partial_i P_-)P_-(\partial_j P_-)P_-.\end{aligned} \tag{6.25}$$

Applying the differential operator to (6.15) yields

$$\partial_i P_\pm = -\partial_i P_\mp; \quad (\partial_i P_\pm)P_\mp = -P_\pm(\partial_i P_\mp); \quad P_\pm(\partial_i P_\pm) = \partial_i P_\pm - (\partial_i P_\pm)P_\pm, \tag{6.26}$$

which lets us simplify the four terms in (6.25):

$$(\partial_i P_-)P_+(\partial_j P_-)P_+ = -(\partial_i P_-)P_+P_-(\partial_j P_+) = 0, \tag{6.27}$$

$$\begin{aligned}(\partial_i P_-)P_+(\partial_j P_-)P_- &= (\partial_i P_-)P_+[\partial_j P_- - \partial_j P_-(\partial_j P_-)] \\ &= (\partial_i P_-)P_+(\partial_j P_-) = -(\partial_i P_-)(\partial_j P_+)P_-,\end{aligned} \tag{6.28}$$

$$\begin{aligned}(\partial_i P_-)P_-(\partial_j P_-)P_+ &= (\partial_i P_-)[(\partial_j P_-) - (\partial_j P_-)P_-]P_+ \\ &= (\partial_i P_-)(\partial_j P_-)P_+ = -(\partial_i P_-)(\partial_j P_+)P_+,\end{aligned} \tag{6.29}$$

$$\begin{aligned}(\partial_i P_-)P_-(\partial_j P_-)P_- &= (\partial_i P_-)[(\partial_j P_-) - (\partial_j P_-)P_-]P_- \\ &= (\partial_i P_-)(\partial_j P_-)P_- - (\partial_i P_-)(\partial_j P_-)P_- = 0.\end{aligned} \tag{6.30}$$

The correlation function is therefore

$$\begin{aligned}\pi_{xy}(i\omega_n) =& -\frac{e^2(E_- - E_+)^2}{\hbar^2 V\beta} \sum_{\vec{q}} \sum_{\nu_n} \mathrm{Tr}\left\{ \frac{(\partial_{q_x} P_-(\vec{q}))(\partial_{q_y} P_+(\vec{q}))P_-(\vec{q})}{(i\hbar\omega_n + i\hbar\nu_n - E_+)(i\hbar\nu_n - E_-)} \right. \\ & \left. + \frac{(\partial_{q_x} P_-(\vec{q}))(\partial_{q_y} P_+(\vec{q}))P_+(\vec{q})}{(i\hbar\omega_n + i\hbar\nu_n - E_-)(i\hbar\nu_n - E_+)} \right\}.\end{aligned} \tag{6.31}$$

The next step is to perform the summation over the fermionic Matsubara frequencies ν_n. To do this, we note that $i\hbar\nu_n$ are the simple poles of the Fermi-Dirac distribution function

$n(z) = 1/(1+\exp(\beta z))$. If a function $f(z)$ has simple poles at $z = z_i$, the residue theorem states that

$$\oint_{\mathcal{C}} dz\ f(z) = 2\pi i \sum_i \mathrm{Res}(f(z_i)), \tag{6.32}$$

where the contour $\mathcal{C}$ encloses the simple poles and

$$\mathrm{Res}(f(z_i)) = \lim_{z\to z_i}(z - z_i)f(z). \tag{6.33}$$

In particular, the residue of $n(z)$ at the simple pole $i\hbar\nu_n$ is

$$\begin{aligned}\mathrm{Res}(n(i\hbar\nu_n)) &= \lim_{z\to i\hbar\nu_n} \frac{z - i\hbar\nu_n}{1+e^{\beta z}} = \lim_{z\to i\hbar\nu_n} \frac{z - i\hbar\nu_n}{1+e^{\beta(z-i\hbar\nu_n+i\hbar\nu_n)}} = \lim_{z\to i\hbar\nu_n} \frac{z - i\hbar\nu_n}{1-e^{\beta(z-i\hbar\nu_n)}} \\ &= \lim_{z\to i\hbar\nu_n} \frac{z - i\hbar\nu_n}{1-1-\beta(z-i\hbar\nu_n)-\ldots} = \lim_{z\to i\hbar\nu_n} -\frac{1}{\beta}\frac{z-i\hbar\nu_n}{z-i\hbar\nu_n} = -\frac{1}{\beta}.\end{aligned} \tag{6.34}$$

Consider now a function $f(z) = g(z)n(z)$. Let $\mathcal{C}_1$ be the contour running vertically to the right and left of the imaginary axis, where the poles of $n(z)$ lie, in a counterclockwise fashion, and let $\mathcal{C}_2$ be the clockwise contour enclosing the poles z_i of $g(z)$. By the residue theorem, we then have

$$\sum_{\nu_n} g(i\hbar\nu_n) = \frac{-\beta}{2\pi i}\oint_{\mathcal{C}_1} dz\ g(z)n(z) = \frac{\beta}{2\pi i}\oint_{\mathcal{C}_2} dz\ g(z)n(z) = \beta\sum_i n(z_i)\mathrm{Res}(g(z_i)), \tag{6.35}$$

where we deformed $\mathcal{C}_1$ into $\mathcal{C}_2$ in the second equality. Let us consider the function

$$g_1(z) = \frac{1}{(z+i\hbar\omega_n - E_+)(z-E_-)} = \frac{1}{i\hbar\omega_n + E_- - E_+}\left(\frac{1}{z-E_-} - \frac{1}{z+i\hbar\omega_n - E_+}\right), \tag{6.36}$$

which for $z = i\hbar\nu_n$ is the first frequency summand in (6.31). This function has simple poles at $z_1 = E_-$ and $z_2 = E_+ - i\hbar\omega_n$, and the corresponding residues are $\mathrm{Res}(g_1(z_1)) = 1/(i\hbar\omega_n + E_- - E_+)$ and $\mathrm{Res}(g_1(z_2)) = -1/(i\hbar\omega_n + E_- - E_+)$. By (6.35), the Matsubara frequency summation of the first term in (6.31) then yields

$$\begin{aligned}\sum_{\nu_n} g_1(i\hbar\nu_n) &= \frac{\beta}{i\hbar\omega_n + E_- - E_+}\Big(n(E_-) - n(E_+ - i\hbar\omega_n)\Big) \\ &= \frac{\beta}{i\hbar\omega_n + E_- - E_+}\Big(n(E_-) - n(E_+)\Big),\end{aligned} \tag{6.37}$$

where the last equality comes from ω_n being a bosonic Matsubara frequency. Similarly, if $g_2(z = i\hbar\nu_n)$ denotes the second frequency summand in (6.31), the Matsubara frequency summation of the second term in (6.31) yields

$$\sum_{\nu_n} g_2(i\hbar\nu_n) = \frac{\beta}{i\hbar\omega_n + E_+ - E_-}\Big(n(E_+) - n(E_-)\Big). \tag{6.38}$$

Since only the lower band is occupied, we have $n(E_-) = 1$ and $n(E_+) = 0$, and the correlation function becomes

$$\begin{aligned}\pi_{xy}(i\omega_n) = &- \frac{e^2(E_- - E_+)^2}{\hbar^2 V} \sum_{\vec{q}} \mathrm{Tr}\left\{ \frac{(\partial_{q_x} P_-(\vec{q}))(\partial_{q_y} P_+(\vec{q}))P_-(\vec{q})}{E_- - E_+ + i\hbar\omega_n} \right. \\ &\left. + \frac{(\partial_{q_x} P_-(\vec{q}))(\partial_{q_y} P_+(\vec{q}))P_+(\vec{q})}{E_- - E_+ - i\hbar\omega_n} \right\}.\end{aligned} \tag{6.39}$$

Consider now

$$\begin{aligned}\mathrm{Tr}\Big\{(\partial_{q_x} P_-(\vec{q}))(\partial_{q_y} P_+(\vec{q}))P_+(\vec{q})\Big\} = &\,\mathrm{Tr}\Big\{(\partial_{q_x} P_-(\vec{q}))(\partial_{q_y} P_+(\vec{q}))\Big\} \\ &- \mathrm{Tr}\Big\{(\partial_{q_x} P_-(\vec{q}))(\partial_{q_y} P_+(\vec{q}))P_-(\vec{q})\Big\}.\end{aligned} \tag{6.40}$$

The Hall conductivity is antisymmetric, $\sigma_{yx} = -\sigma_{xy}$. However, if we interchange x and y in the first trace in (6.40), we obtain

$$\begin{aligned}\mathrm{Tr}\Big\{(\partial_{q_y} P_-(\vec{q}))(\partial_{q_x} P_+(\vec{q}))\Big\} &= \mathrm{Tr}\Big\{(-\partial_{q_y} P_+(\vec{q}))(-\partial_{q_x} P_-(\vec{q}))\Big\} \\ &= \mathrm{Tr}\Big\{(\partial_{q_x} P_-(\vec{q}))(\partial_{q_y} P_+(\vec{q}))\Big\},\end{aligned} \tag{6.41}$$

where we used the cyclic property of the trace in the last equality. This term is symmetric in x and y, and therefore does not contribute to the Hall conductivity. Inserting (6.40) into (6.39) and using the first relation in (6.26), the correlation function is expressed only in terms of the projector of the filled band:

$$\begin{aligned}\pi_{xy}(i\omega_n) = &- \frac{e^2(E_- - E_+)^2}{\hbar^2 V} \frac{2i\hbar\omega_n}{(E_- - E_+)^2 - (i\hbar\omega_n)^2} \\ &\times \sum_{\vec{q}} \mathrm{Tr}\Big\{(\partial_{q_x} P_-(\vec{q}))(\partial_{q_y} P_-(\vec{q}))P_-(\vec{q})\Big\}.\end{aligned} \tag{6.42}$$

The analytical continuation (5.73) lets us write (6.42) in terms of a real frequency:

$$\begin{aligned}\Pi_{xy}(\omega) = &- \frac{e^2(E_- - E_+)^2}{\hbar^2 V} \frac{2\hbar(\omega + i\delta)}{(E_- - E_+)^2 - \hbar^2(\omega + i\delta)^2} \\ &\times \sum_{\vec{q}} \mathrm{Tr}\Big\{(\partial_{q_x} P_-(\vec{q}))(\partial_{q_y} P_-(\vec{q}))P_-(\vec{q})\Big\}.\end{aligned} \tag{6.43}$$

The Hall conductivity σ_{xy} is obtained by taking the DC limit $\omega \to 0$ of (5.66):

$$\begin{aligned}\sigma_{xy} &= \lim_{\omega\to 0} \frac{i}{\omega}\big[\Pi_{xy}(\omega) - \Pi_{xy}(0)\big] = i\frac{\partial \Pi_{xy}(\omega)}{\partial\omega}\bigg|_{\omega=0} \\ &= -\frac{ie^2(E_- - E_+)^2}{\hbar^2 V} \frac{2\hbar}{(E_- - E_+)^2} \sum_{\vec{q}} \mathrm{Tr}\Big\{(\partial_{q_x} P_-(\vec{q}))(\partial_{q_y} P_-(\vec{q}))P_-(\vec{q})\Big\} \\ &= -\frac{2ie^2}{\hbar V} \sum_{\vec{q}} \mathrm{Tr}\Big\{(\partial_{q_x} P_-(\vec{q}))(\partial_{q_y} P_-(\vec{q}))P_-(\vec{q})\Big\}.\end{aligned} \tag{6.44}$$

Expressing the projector of the filled band in terms of the eigenstate $|\psi_-\rangle$, we find

$$\begin{aligned}\mathrm{Tr}\Big\{(\partial_{q_x}P_-(\vec{q}))(\partial_{q_y}P_-(\vec{q}))P_-(\vec{q})\Big\} =&\big(\,\langle\psi_-|\partial_{q_y}\psi_-\rangle + \langle\partial_{q_y}\psi_-|\psi_-\rangle\,\big)\,\langle\psi_-|\partial_{q_x}\psi_-\rangle \\ &+ \langle\partial_{q_x}\psi_-|\partial_{q_y}\psi_-\rangle + \langle\partial_{q_x}\psi_-|\psi_-\rangle\,\langle\partial_{q_y}\psi_-|\psi_-\rangle\,,\end{aligned} \tag{6.45}$$

where we used the orthonormality of the eigenstates when taking the trace. The first term in (6.45) is zero, since it is proportional to the derivative of $\langle\psi_-|\psi_-\rangle = 1$. The last term is invariant under $x \leftrightarrow y$ and does not contribute to the Hall conductivity, which becomes

$$\sigma_{xy} = \frac{\sigma_{xy}-\sigma_{yx}}{2} = -\frac{ie^2}{\hbar V}\sum_{\vec{q}}\big[\,\langle\partial_{q_x}\psi_-|\partial_{q_y}\psi_-\rangle - \langle\partial_{q_y}\psi_-|\partial_{q_x}\psi_-\rangle\,\big] = -\frac{e^2}{\hbar V}\sum_{\vec{q}}\Omega_-(\vec{q}), \tag{6.46}$$

where we identified the Berry curvature $\Omega_-(\vec{q})$ from (4.16) in the last equality. Replacing the discrete sum with the integral over the first Brillouin zone, $\sum_{\vec{q}} = V/(2\pi)^2\iint_{\mathrm{BZ}} d^2k$, the Hall conductivity becomes

$$\sigma_{xy} = -\frac{e^2}{\hbar}\frac{1}{(2\pi)^2}\iint_{\mathrm{BZ}} d^2k\;\Omega_-(\vec{q}) = -\frac{e^2}{h}\frac{1}{2\pi}\iint_{\mathrm{BZ}} d^2k\;\Omega_-(\vec{q}) = -\frac{e^2}{h}C_-. \tag{6.47}$$

The Hall conductivity is quantized in units of e^2/h, and its magnitude is determined by the Chern number of the filled band.[12] This gives physical meaning to the phases of the Chern insulator considered in Section 4.3; its topologically non-trivial phases exhibit the quantum Hall effect. The generalization of (6.47) to the multi-band Hamiltonian is straightforward, and one finds that the Hall conductivity is proportional to the sum of the Chern numbers of all filled bands.

6.2 Transverse Transport Phenomena in Bosonic Spin Systems

In the quantum Hall effect, the induced electric current is carried by electrons. Suppose instead that our system is described by charge-neutral spin excitations. These do not exhibit the quantum Hall effect in response to an external electric field, but we can induce different types of currents by subjecting the system to a temperature gradient ∇T. Here, we will consider two such phenomena: the spin Nernst effect and the thermal Hall effect. In the former, the temperature gradient induces a transverse spin current caused by the separation of oppositely polarized spins, whereas the latter describes the emergence of a transverse spin-carried heat current. We wish to compute the corresponding conductivities, as they can provide essential information about the nature of the spin system. Once we know the expression for the current density operators, these can be obtained using (5.66), (5.72) and (5.73). However, obtaining these expressions is not straightforward; since thermal forces due to temperature gradients are of a statistical nature, it is not immediately clear how they can be represented in the Hamiltonian. The problem is solved using the approach of Luttinger [23]. Suppose we have an operator $\hat{O}$ whose density

$$\hat{o}(\vec{r}) = \frac{1}{2}\hat{\mathbf{b}}^\dagger(\vec{r})\,\hat{O}\,\hat{\mathbf{b}}(\vec{r}) \tag{6.48}$$

[12] The sign in (6.47) may be absorbed into the definition of the Chern number.

acts as a source for a temperature induced current [24]. Here,

$$\hat{\mathbf{b}}(\vec{r}) = \begin{bmatrix} \hat{b}_1(\vec{r}) & \dots & \hat{b}_N(\vec{r}) & \hat{b}_1^\dagger(\vec{r}) & \dots & \hat{b}_N^\dagger(\vec{r}) \end{bmatrix}^\mathsf{T} \tag{6.49}$$

is a bosonic operator vector. We introduce a pseudo-gravitational potential $\Phi(\vec{r})$ which couples to the density operator through

$$\hat{V}(t) = \int d^d r \; \Phi(\vec{r})\hat{o}(\vec{r}). \tag{6.50}$$

The pseudo-gravitational potential attempts to balance the temperature gradient, and in equilibrium the two are related by $\nabla\Phi = \nabla T/T$. The perturbation (6.50) may then be combined with the unperturbed Hamiltonian

$$\hat{H}_0 = \frac{1}{2}\int d^d r \; \hat{\mathbf{b}}^\dagger(\vec{r})\, \mathbb{H}(\vec{r})\, \hat{\mathbf{b}}(\vec{r}), \tag{6.51}$$

where $\mathbb{H}(\vec{r})$ is a $2N \times 2N$ matrix, to obtain

$$\hat{H} = \frac{1}{2}\int d^d r \; \left(1 + \frac{\Phi}{2}\right)\hat{\mathbf{b}}^\dagger(\vec{r})\, \mathbb{H}(\vec{r})\left(1 + \frac{\Phi}{2}\right)\hat{\mathbf{b}}(\vec{r}). \tag{6.52}$$

We now have a Hamiltonian that accounts for the thermal forces. The continuity equation

$$\frac{\partial}{\partial t}\hat{o}(\vec{r}) + \nabla \cdot \hat{\vec{j}}(\vec{r}) = 0 \tag{6.53}$$

may be used to derive the current density operator associated with the density operator $\hat{o}(\vec{r})$. In the spin Nernst effect, the density operator is the local spin operator in the z-direction, $\hat{o}(\vec{r}) = \hat{S}^z(\vec{r})$; in the thermal Hall effect, it is the energy density, $\hat{o}(\vec{r}) = 1/2\big(1 + \Phi/2\big)\hat{\mathbf{b}}^\dagger(\vec{r})\mathbb{H}(\vec{r})\big(1+\Phi/2\big)\hat{\mathbf{b}}(\vec{r})$. Before we present the corresponding conductivity formulas, we note that the modification of the Hamiltonian gives rise to a magnetization current caused by the rotational motion of the spin excitations [25, 26]. The contribution of the magnetization current is evaluated in the system's equilibrium state, as opposed to the Kubo current appearing in linear response theory. Both of these current contributions must be taken into account to get the correct conductivity formulas.

6.2.1 The Spin Nernst Effect

The derivation of the spin Nernst conductivity is mathematically more involved than the approach used to derive the Hall conductivity in Section 6.1. We will therefore only present the current density operators, to see what the starting point of the derivation is, and the final result for the spin Nernst conductivity. The full derivation is due to Zyuzin and Kovalev [24]. As mentioned in the previous Section, there are two current contributions:

$$\begin{aligned} \hat{\vec{j}}^{\,(1)}(\vec{r}) &= \hat{\mathbf{b}}^\dagger(\vec{r})\, \hat{O}\sigma_3\hat{\vec{v}}\, \hat{\mathbf{b}}(\vec{r}), \\ \hat{\vec{j}}^{\,(2)}(\vec{r}) &= \frac{1}{2}\sum_j \hat{\mathbf{b}}^\dagger(\vec{r})\, \hat{O}\sigma_3\big(\hat{r}_j\hat{\vec{v}} + \hat{\vec{v}}\hat{r}_j\big)\, \hat{\mathbf{b}}(\vec{r})\partial_j\Phi(\vec{r}), \end{aligned} \tag{6.54}$$

where

$$\hat{O} = \hbar \begin{bmatrix} (\sigma_3)_{N\times N} & 0 \\ 0 & (\sigma_3)_{N\times N} \end{bmatrix}, \tag{6.55}$$

$\hat{\vec{r}}$ is the position operator and $\hat{\vec{v}} = i/\hbar[\hat{H}, \hat{\vec{r}}]$. The first current density operator $\hat{\vec{j}}^{(1)}$ is treated using linear response theory. While the spin Nernst conductivity does depend on the topological properties of the bands, it is not a topological invariant. Therefore, we cannot compute the Matsubara current-current correlation function in the flat-band limit, which makes the computation more cumbersome than for the Hall conductivity. The second current density operator $\hat{\vec{j}}^{(2)}$ represents the magnetization current, and its contribution to the conductivity is described in Reference [27]. Combining the two contributions yields the expression for the spin Nernst conductivity [26, 28]:

$$\alpha_{xy} = \frac{k_B}{\hbar V} \sum_{\vec{k}} \sum_{n=1}^{N} c_1\big[g\big(E_{n\vec{k}}\big)\big]\Omega_n(\vec{k}), \tag{6.56}$$

where $g\big(E_{n\vec{k}}\big)$ is the Bose-Einstein distribution of the nth energy band $E_{n\vec{k}}$ and $\Omega_n(\vec{k})$ is the corresponding Berry curvature. For non-interacting bosonic Hamiltonians, the Berry curvature may be expressed as

$$\Omega_n(\vec{k}) = i\epsilon_{xy}\left[\sigma_3 \frac{\partial \mathbb{T}^\dagger_{\vec{k}}}{\partial k_x} \sigma_3 \frac{\partial \mathbb{T}_{\vec{k}}}{\partial k_y}\right]_{nn}, \tag{6.57}$$

where ϵ_{xy} is the Levi-Civita symbol and $\mathbb{T}$ is the canonical transformation matrix defined in (2.27). The function $c_1(x)$ that weights the sum over the Berry curvature is given by

$$c_1(x) = (1+x)\ln(1+x) - x\ln x. \tag{6.58}$$

In its current form, the Berry curvature (6.57) is not well suited for numerical calculations. Using the pseudo-unitarity (2.42) and eigenvalue equation (2.45), we can obtain the bosonic matrix analogue of (4.20), which is manifestly gauge invariant. The Berry curvature may then be written [29]

$$\Omega_n(\vec{k}) = i\epsilon_{xy} \sum_{m\neq n} \frac{\big(\mathbb{T}^{\dagger(n)}_{\vec{k}} \sigma_3 V_{k_x}(\vec{k}) \mathbb{T}^{(m)}_{\vec{k}}\big)\big(\mathbb{T}^{\dagger(m)}_{\vec{k}} \sigma_3 V_{k_y}(\vec{k}) \mathbb{T}^{(n)}_{\vec{k}}\big)}{\big(E_{n\vec{k}} - E_{m\vec{k}}\big)^2}, \tag{6.59}$$

where $\mathbb{T}^{(n)}_{\vec{k}}$ is the eigenvector corresponding to $E_{n\vec{k}}$ and $V_i(\vec{k}) = \mathbb{T}^\dagger_{\vec{k}} \partial_i \mathbb{H}(\vec{k}) \mathbb{T}_{\vec{k}}$ for $i \in \{k_x, k_y\}$.

6.2.2 The Thermal Hall Effect

The derivation of the thermal Hall conductivity is more cumbersome than that of the spin Nernst conductivity, and a full derivation is given in Reference [25]. The Kubo and magnetic

current density operators are in this case given by

$$\begin{aligned}\hat{\vec{j}}^{(1)}(\vec{r}) &= \frac{1}{4}\hat{\mathbf{b}}^{\dagger}(\vec{r})\big(\hat{\vec{v}}\sigma_3\mathbb{H}(\vec{r}) + \mathbb{H}(\vec{r})\sigma_3\hat{\vec{v}}\big)\hat{\mathbf{b}}(\vec{r}),\\ \hat{\vec{j}}^{(2)}(\vec{r}) &= \frac{1}{8}\sum_j \Big[-i\hat{\mathbf{b}}^{\dagger}(\vec{r})\big(\hat{\vec{v}}\sigma_3\hat{v}_j - \hat{v}_j\sigma_3\hat{\vec{v}}\big)\hat{\mathbf{b}}(\vec{r}) + \hat{\mathbf{b}}^{\dagger}(\vec{r})\big(\hat{r}_j\hat{\vec{v}}\sigma_3 + 3\hat{\vec{v}}\sigma_3\hat{r}_j\big)\,\mathbb{H}(\vec{r})\,\hat{\mathbf{b}}(\vec{r})\\ &\qquad + \hat{\mathbf{b}}^{\dagger}(\vec{r})\,\mathbb{H}(\vec{r})\,\big(3\hat{r}_j\sigma_3\hat{\vec{v}} + \sigma_3\hat{\vec{v}}\hat{r}_j\big)\hat{\mathbf{b}}(\vec{r})\Big]\partial_j\Phi(\vec{r}),\end{aligned} \tag{6.60}$$

whose net response to the applied temperature gradient is the thermal Hall conductivity

$$\kappa_{xy} = -\frac{k_B^2 T}{\hbar V}\sum_{\vec{k}}\sum_{n=1}^{N}\left[c_2[g\big(E_{n\vec{k}}\big)] - \frac{\pi^2}{3}\right]\Omega_n(\vec{k}), \tag{6.61}$$

with the weighting function $c_2(x)$ given by

$$c_2(x) = (1+x)\left(\ln\frac{1+x}{x}\right)^2 - (\ln x)^2 - 2\mathrm{Li}_2(-x), \tag{6.62}$$

where $\mathrm{Li}_2(x)$ is the dilogarithm:

$$\mathrm{Li}_2(x) = -\int_0^x dt\ \frac{\ln(1-t)}{t}. \tag{6.63}$$

By (4.21), the term proportional to $\pi^2/3$ in (6.61) vanishes when we sum over the bands, and we have

$$\kappa_{xy} = -\frac{k_B^2 T}{\hbar V}\sum_{\vec{k}}\sum_{n=1}^{N} c_2[g\big(E_{n\vec{k}}\big)]\Omega_n(\vec{k}). \tag{6.64}$$

In the derivations of (6.56) and (6.61), it is assumed that the Hamiltonian matrix satisfies the particle-hole symmetry

$$\mathbb{H}(\vec{k}) = \sigma_1\mathbb{H}^{\mathsf{T}}(-\vec{k})\sigma_1, \tag{6.65}$$

where σ_1 is the $2N \times 2N$ Pauli matrix. If the Hamiltonian matrix does not have this symmetry, the sum over the N bands in (6.56) and (6.64) must be replaced by the sum over the $2N$ eigenvectors of the dynamical matrix $\mathbb{K} = \sigma_3\mathbb{H}(\vec{k})$ that make up the columns of $\mathbb{T}_{\vec{k}}$ [30]. The expressions are otherwise unchanged. In either case, the spin Nernst conductivity α_{xy} and the thermal Hall conductivity κ_{xy} are proportional to the Berry curvature, and we therefore expect enhanced responses to the temperature gradient in phases with non-trivial topology.

6.2.3 The Thermal Vector Potential

Before moving on, we note that the perturbation (6.50) does not have the same form as (5.52), which was used to derive (5.66) in a very straightforward manner. However, the latter equation is used to obtain the Kubo contribution to both the spin Nernst conductivity (6.56)

and the thermal Hall conductivity (6.61) from a perturbation of the form (6.50). It may seem coincidental that (5.66) applies to both cases, but it can be shown that the two perturbation terms (5.52) and (6.50) are equivalent, with the thermal vector potential given by [31]

$$\vec{A}(\vec{r},t) = \int_{-\infty}^{t} dt' \ \nabla\Phi(\vec{r},t'). \tag{6.66}$$

The thermal vector potential satisfies

$$\frac{\partial}{\partial t}\vec{A}(\vec{r},t) = \nabla\Phi(\vec{r},t) = \frac{\nabla T}{T}, \tag{6.67}$$

which means the applied perturbation in (5.54) is given by

$$\vec{E}(\vec{r},t) = -\frac{\nabla T}{T}. \tag{6.68}$$

A generic current density induced by a temperature gradient may be written very generally as

$$J_i(\vec{r},t) = -\int d^d r' \int dt' \sum_l \chi_{il}(\vec{r},t;\vec{r}',t')\partial_l T(\vec{r}',t'), \tag{6.69}$$

where χ is the response function. For (5.54) and (6.69) to describe the same phenomenon, the response function χ must be related to the response function σ by

$$\chi_{il} = \frac{\sigma_{il}}{T}. \tag{6.70}$$

This is precisely the relation used in the derivations of (6.56) and (6.61), so we conclude that our derivation of (5.66) is valid.

7 Schwinger Boson Mean Field Theory of the Extended Kane-Mele-Hubbard Model

We will now rederive the results obtained by Ushakov [6] for the ground state energy of an extended Kane-Mele-Hubbard model using Schwinger boson mean field theory. The Schwinger boson approach is well suited for studying the physical phases that can exist in magnetically frustrated honeycomb lattice materials; of particular interest is the potential quantum spin liquid phase. Before we get into the treatment of the model, it is useful to present some of the properties of the honeycomb lattice and the corresponding reciprocal lattice.

7.1 The Honeycomb Lattice

The honeycomb lattice is shown in Figure 5. It consists of two superimposed hexagonal Bravais lattices, which we will refer to as sublattices A and B. A lattice site on sublattice X has three nearest neighbor lattice sites whose positions are given by $\vec{\delta}_1^{(X)}$, $\vec{\delta}_2^{(X)}$ and $\vec{\delta}_3^{(X)}$. The nearest neighbor vectors on sublattice A are given by:

$$\begin{aligned}\vec{\delta}_1^{(A)} &= \frac{a}{2}(\sqrt{3}\hat{e}_1 + \hat{e}_2),\\ \vec{\delta}_2^{(A)} &= \frac{a}{2}(-\sqrt{3}\hat{e}_1 + \hat{e}_2),\\ \vec{\delta}_3^{(A)} &= -a\hat{e}_2,\end{aligned} \tag{7.1}$$

where $\hat{e}_1$ and $\hat{e}_2$ are the unit vectors in the x- and y-directions, respectively. The length of these vectors is a, which is the lattice constant of the honeycomb lattice. From Figure 5 it is clear that $\vec{\delta}_1^{(B)} = -\vec{\delta}_1^{(A)}$, $\vec{\delta}_2^{(B)} = -\vec{\delta}_2^{(A)}$ and $\vec{\delta}_3^{(B)} = -\vec{\delta}_3^{(A)}$. Each lattice site also has six next nearest neighbor sites located at $\vec{\epsilon}_1$, $\vec{\epsilon}_2$, $\vec{\epsilon}_3$, $\vec{\epsilon}_4$, $\vec{\epsilon}_5$ and $\vec{\epsilon}_6$; due to the symmetry of the

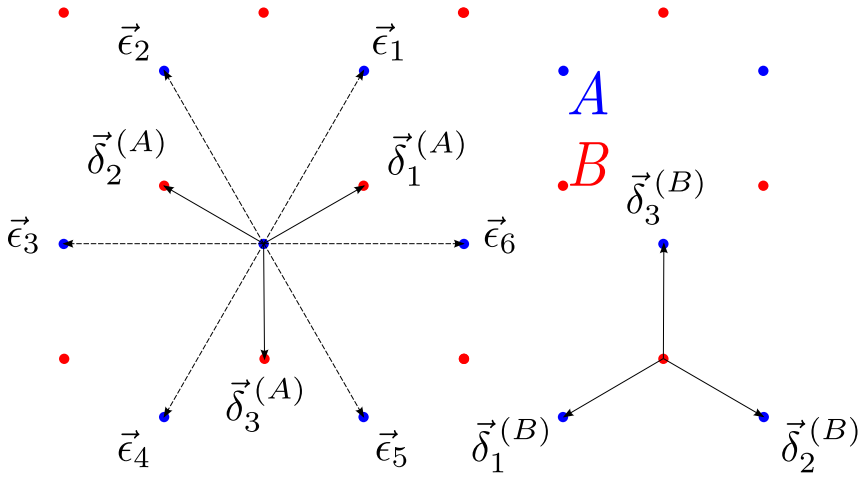


Figure 5: The honeycomb lattice consists of two superimposed hexagonal sublattices A and B. Each lattice site has three nearest neighbors and six next nearest neighbors.

honeycomb lattice these vectors are the same for both sublattices. They are given by:

$$\begin{aligned}
\vec{\epsilon}_1 &= \frac{a}{2}(\sqrt{3}\hat{e}_1 + 3\hat{e}_2), \\
\vec{\epsilon}_2 &= \frac{a}{2}(-\sqrt{3}\hat{e}_1 + 3\hat{e}_2), \\
\vec{\epsilon}_3 &= -\sqrt{3}a\hat{e}_1, \\
\vec{\epsilon}_4 &= -\vec{\epsilon}_1, \\
\vec{\epsilon}_5 &= -\vec{\epsilon}_2, \\
\vec{\epsilon}_6 &= -\vec{\epsilon}_3.
\end{aligned} \tag{7.2}$$

We will not consider further neighbors in the treatment of the Kane-Mele-Hubbard model. This allows electrons on the lattice to hop between nearest and next nearest neighbor sites only. The nearest neighbor hopping is trivial, but the next nearest neighbor hopping requires closer consideration. Suppose that an electron located at site i hops to a next nearest neighbor site j through an intermediate site k. The electron can then end up in two possible next nearest neighbor sites j_1 and j_2, depending on whether it makes a left or right turn from the intermediate site. If we denote the vector from site i to k as $\vec{d}_{ik}$ and the vector from site k to j as $\vec{d}_{kj}$, we can define the vector

$$\vec{\nu}_{ij} = \frac{\vec{d}_{ik} \times \vec{d}_{kj}}{|\vec{d}_{ik} \times \vec{d}_{kj}|}, \tag{7.3}$$

which points out of the plane for a left turn and into the plane for a right turn. Defining the out-of-plane axis to be the z-axis, the two possibilities can be distinguished by the scalar parameter

$$\nu_{ij} = \vec{\nu}_{ij} \cdot \hat{e}_3 = \begin{cases} +1 & \text{for a left turn,} \\ -1 & \text{for a right turn.} \end{cases} \tag{7.4}$$

Denoting the next nearest neighbor vectors by $\vec{\epsilon}_l$, we can associate the parameter $\nu_{ij} = \nu_l^{(X)}$ on sublattice X with these vectors in the following way:

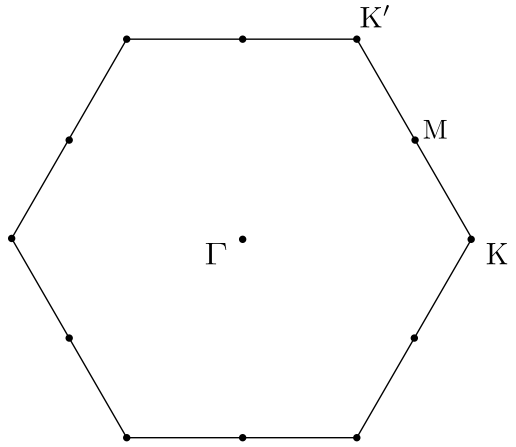
$$\nu_l^{(X)} = \begin{cases} (-1)^{l+1} & \text{if } X = A, \\ (-1)^l & \text{if } X = B. \end{cases} \tag{7.5}$$

From (7.5) it is clear that $\nu_l^{(A)} = -\nu_l^{(B)}$.

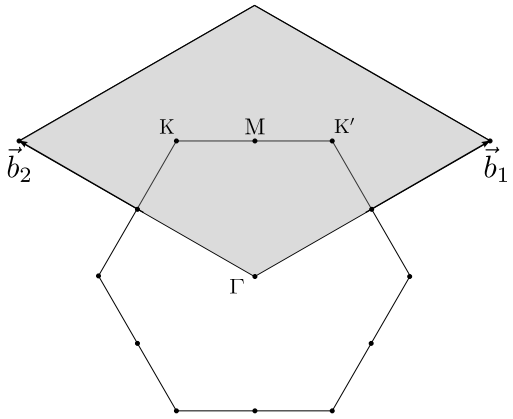
7.2 The Reciprocal Lattice

We will impose periodic boundary conditions on the honeycomb lattice in what follows, which makes it useful to express operators in terms of their Fourier transforms. It is therefore helpful to define the reciprocal lattice of the honeycomb lattice. To do so, we first choose two primitive lattice vectors $\vec{a}_1 = \vec{\epsilon}_1$ and $\vec{a}_2 = \vec{\epsilon}_2$ for the honeycomb lattice. The primitive vectors of the reciprocal lattice satisfy [32, p. 29]

$$\vec{a}_i \cdot \vec{b}_j = 2\pi\delta_{ij}, \tag{7.6}$$



(a) The first Brillouin zone of the hexagonal lattice is the Wigner-Seitz cell in reciprocal space. Γ, K, M and K′ denote high symmetry points.



(b) The numerical Brillouin zone is the gray-shaded area spanned by the primitive reciprocal lattice vectors $\vec{b}_1$ and $\vec{b}_2$.

Figure 6: Two equivalent choices for the primitive cell of the hexagonal reciprocal lattice. Figure 6a shows the conventional first Brillouin zone, while Figure 6b shows the numerical Brillouin zone, which is more suited for evaluation of sums in reciprocal space.

from which they can be constructed as

$$\begin{aligned}\vec{b}_1 &= 2\pi\frac{\vec{a}_2\times\hat{e}_3}{\vec{a}_1\cdot\vec{a}_2\times\hat{e}_3} = \frac{2\pi}{\sqrt{3}a}\left(\hat{e}_1+\frac{1}{\sqrt{3}}\hat{e}_2\right),\\ \vec{b}_2 &= 2\pi\frac{\hat{e}_3\times\vec{a}_1}{\vec{a}_1\cdot\vec{a}_2\times\hat{e}_3} = \frac{2\pi}{\sqrt{3}a}\left(-\hat{e}_1+\frac{1}{\sqrt{3}}\hat{e}_2\right).\end{aligned} \tag{7.7}$$

This shows that the reciprocal lattice is a hexagonal lattice. The first Brillouin zone of the reciprocal lattice is shown in Figure 6a. As we know, the Fourier transform is given by a sum over the first Brillouin zone, which is a primitive cell of the reciprocal lattice. However, if we wish to evaluate these sums numerically, it is more convenient to sum over the equivalent primitive cell shown in Figure 6b. For this reason, we will refer to this primitive cell as the numerical Brillouin zone. A generic point $\vec{k}$ in the numerical Brillouin zone may be written as a linear combination of the reciprocal lattice vectors:

$$\vec{k} = k_x\hat{e}_1 + k_y\hat{e}_2 = m_1\vec{b}_1 + m_2\vec{b}_2. \tag{7.8}$$

Suppose that each of the sublattices A and B have $N = N_1N_2$ lattice sites, where N_i is the number of lattice sites in the direction of the primitive lattice vector $\vec{a}_i$. For an operator $\hat{a}_{\vec{r}}$, periodic boundary conditions mean

$$\hat{a}_{\vec{r}+N_i\vec{a}_i} = \hat{a}_{\vec{r}} \implies \sum_{\vec{k}} e^{iN_i\vec{k}\cdot\vec{a}_i}e^{i\vec{k}\cdot\vec{r}}\hat{a}_{\vec{k}} = \sum_{\vec{k}} e^{i\vec{k}\cdot\vec{r}}\hat{a}_{\vec{k}} \implies N_i\vec{k}\cdot\vec{a}_i = 2\pi n_i; \quad n_i\in\mathbb{Z}. \tag{7.9}$$

Inserting (7.8) into the scalar product and using (7.6), we obtain

$$m_i = \frac{n_i}{N_i}; \quad n_i\in\{0,\ldots,N_i\}, \quad i\in\{1,2\}, \tag{7.10}$$

where the upper bound on n_i ensures that $\vec{k}$ remains in the numerical Brillouin zone. Taking $N_1 = N_2 = \mathcal{N}$ and inserting (7.7) and (7.10) into (7.8), we find that the numerical Brillouin zone is the set

$$\text{NBZ} = \left\{ \vec{k} \mid k_x = \frac{2\pi}{\sqrt{3}\mathcal{N}a}(n_1 - n_2),\ k_y = \frac{2\pi}{3\mathcal{N}a}(n_1 + n_2);\ n_1, n_2 \in \{0, \dots, \mathcal{N}\} \right\}. \tag{7.11}$$

This parametrization makes it easy to construct the grid of reciprocal lattice vectors needed for numerical computations. With the definitions of the honeycomb lattice and reciprocal lattice in mind, we now present the extended Kane-Mele-Hubbard model.

7.3 The Extended Kane-Mele-Hubbard Model

The model we will be considering is given by

$$\begin{aligned} \hat{H} = &- t_1 \sum_{\langle i,j \rangle} \sum_{\alpha} \hat{c}^\dagger_{i\alpha} \hat{c}_{j\alpha} + U \sum_i \hat{n}_{i\uparrow} \hat{n}_{i\downarrow} \\ &+ i\lambda_R \sum_{\langle i,j \rangle} \sum_{\alpha} (\vec{\sigma} \times \hat{d}_{ij})_3 \hat{c}^\dagger_{i\alpha} \hat{c}_{j\alpha} - t_2 \sum_{\langle\langle i,j \rangle\rangle} \sum_{\alpha} \hat{c}^\dagger_{i\alpha} \hat{c}_{j\alpha} + i\lambda_I \sum_{\langle\langle i,j \rangle\rangle} \sum_{\alpha,\beta} \nu_{ij} (\sigma_3)_{\alpha\beta} \hat{c}^\dagger_{i\alpha} \hat{c}_{j\beta}. \end{aligned} \tag{7.12}$$

Here, $\{\hat{c}^\dagger_{i\alpha}\}$ and $\{\hat{c}_{i\alpha}\}$ are creation and annihilation operators for electrons with spin α at lattice sites i. The first term describes nearest neighbor hopping, with $t_1 > 0$ being the nearest neighbor hopping amplitude. The second term is the Hubbard term, which for $U > 0$ describes the repulsion of two electrons with opposite spins located at the same site i. The third term is a nearest neighbor spin-orbit coupling known as the Rashba coupling. λ_R is the coupling amplitude and $\hat{d}_{ij}$ is the unit vector pointing from site i to site j. The fourth term is the next nearest neighbor hopping term, where the hopping amplitude $t_2 > 0$. The final term describes the intrinsic next nearest neighbor spin-orbit coupling, with λ_I being the coupling amplitude. If the system is invariant under the inversion $z \to -z$, the Rashba coupling vanishes. Such a symmetry may be broken by e.g. strain from a substrate or by applying an electric field in the z-direction. In the following, we will consider the case where the inversion symmetry is preserved. The Schwinger boson transformation of the Rashba coupling is presented in Appendix B. We do not study it in the mean field approximation because of how complex the Hamiltonian becomes when all of the above terms are included in the model.

We will assume here that $t_1 << U$, $t_2 << U$. The Hamiltonian may then be expanded to lowest order in t_1/U and t_2/U, which lets us write the Hamiltonian (7.12), without the Rashba coupling, in terms of spin operators as:

$$\begin{aligned} \hat{H} = &J_1 \sum_{\langle i,j \rangle} \hat{\vec{S}}_i \cdot \hat{\vec{S}}_j + J_2 \sum_{\langle\langle i,j \rangle\rangle} \hat{\vec{S}}_i \cdot \hat{\vec{S}}_j \\ &+ \Gamma \sum_{\langle\langle i,j \rangle\rangle} \hat{\vec{S}}_i \cdot [\text{diag}(-1,-1,1) \hat{\vec{S}}_j] + D_\perp \sum_{\langle\langle i,j \rangle\rangle} \nu_{ij} \hat{e}_3 \cdot (\hat{\vec{S}}_i \times \hat{\vec{S}}_j). \end{aligned} \tag{7.13}$$

Here,

$$J_1 = \frac{2t_1^2}{U}; \quad J_2 = \frac{2t_2^2}{U}; \quad \Gamma = \frac{2\lambda_I^2}{U}; \quad D_\perp = \frac{4t_2\lambda_I}{U}. \tag{7.14}$$

The first two terms in (7.13) are isotropic Heisenberg terms for nearest and next nearest neighbor spins, respectively. The third term is a next nearest neighbor anisotropic Heisenberg term, while the fourth term is the out-of-plane next nearest neighbor Dzyaloshinskii-Moriya interaction (DMI). We will refer to these terms as $\hat{H}_1$, $\hat{H}_2$, $\hat{H}_\Gamma$ and $\hat{H}_{D_\perp}$, respectively. We will also study the effect of an external magnetic field $\vec{B}_{\text{ext}} = B_{\text{ext}}\hat{e}_3$ through the Zeeman coupling

$$\hat{H}_Z = -\frac{\gamma\mu_B}{\hbar}B_{\text{ext}}\sum_i \hat{S}_i^z = -B\sum_i \hat{S}_i^z, \tag{7.15}$$

where we absorbed the gyromagnetic ratio γ, the Bohr magneton μ_B and a factor $1/\hbar$ into the definition of the effective magnetic field B.

7.4 The Schwinger Boson Transformation of the Spin Hamiltonian

We wish to express the spin operators in (7.13) in terms of Schwinger bosons:

$$\hat{\vec{S}}_i = \frac{\hbar}{2}\sum_{\alpha,\beta}\hat{b}_{i\alpha}^\dagger\vec{\sigma}_{\alpha\beta}\hat{b}_{i\beta}. \tag{7.16}$$

The Schwinger bosons are constrained at each site i by

$$\hat{n}_i = \sum_\alpha \hat{b}_{i\alpha}^\dagger\hat{b}_{i\alpha} = 2S = 1, \tag{7.17}$$

since we are dealing with electrons with spin $S = 1/2$. Before we insert (7.16) into the Hamiltonian, we establish some useful mathematical relations. Consider first the sum $\sum_{\alpha,\beta}\hat{b}_{j\alpha}^\dagger\hat{b}_{i\beta}\hat{b}_{i\beta}^\dagger\hat{b}_{j\alpha}$. Making use of the bosonic commutation relations and the constraint (7.17), we get

$$\sum_{\alpha,\beta}\hat{b}_{j\alpha}^\dagger\hat{b}_{i\beta}\hat{b}_{i\beta}^\dagger\hat{b}_{j\alpha} = \sum_{\alpha,\beta}\hat{b}_{j\alpha}^\dagger(1+\hat{b}_{i\beta}^\dagger\hat{b}_{i\beta})\hat{b}_{j\alpha} = 2\sum_\alpha\hat{b}_{j\alpha}^\dagger\hat{b}_{j\alpha} + \sum_{\alpha,\beta}\hat{b}_{j\alpha}^\dagger\hat{b}_{i\beta}^\dagger\hat{b}_{i\beta}\hat{b}_{j\alpha} = 3. \tag{7.18}$$

Consider also the tensor

$$g_{\alpha\beta\gamma\delta} = \vec{\sigma}_{\alpha\beta}\cdot(\mathbf{A}\vec{\sigma}_{\gamma\delta}) = \sum_{i,j=1}^{3}(\sigma_i)_{\alpha\beta}A_{ij}(\sigma_j)_{\gamma\delta}. \tag{7.19}$$

Using the properties of the Pauli matrices given in Appendix A, this can be written in matrix form as

$$\begin{aligned} G(\mathbf{A}) &= \begin{bmatrix} g_{1111} & g_{1112} & g_{1121} & g_{1122} \\ g_{1211} & g_{1212} & g_{1221} & g_{1222} \\ g_{2111} & g_{2112} & g_{2121} & g_{2122} \\ g_{2211} & g_{2212} & g_{2221} & g_{2222} \end{bmatrix} \\ &= \begin{bmatrix} A_{33} & A_{31}-iA_{32} & A_{31}+iA_{32} & -A_{33} \\ A_{13}-iA_{23} & A_{11}-A_{22}-i(A_{12}+A_{21}) & A_{11}+A_{22}+i(A_{12}-A_{21}) & -A_{13}+iA_{23} \\ A_{13}+iA_{23} & A_{11}+A_{22}-i(A_{12}-A_{21}) & A_{11}-A_{22}+i(A_{12}+A_{21}) & -A_{13}-iA_{23} \\ -A_{33} & -A_{31}+iA_{32} & -A_{31}-iA_{32} & A_{33} \end{bmatrix}. \end{aligned} \tag{7.20}$$

It is particularly interesting to establish what this tensor is for a diagonal matrix $\mathbf{D} = \mathrm{diag}(D_1, D_1, D_3)$. Inserting $\mathbf{D}$ into (7.20) we get

$$G(\mathbf{D}) = \begin{bmatrix} D_3 & 0 & 0 & -D_3 \\ 0 & 0 & 2D_1 & 0 \\ 0 & 2D_1 & 0 & 0 \\ -D_3 & 0 & 0 & D_3 \end{bmatrix}. \tag{7.21}$$

We are now ready to study the Hamiltonian using the Schwinger boson transformation. We start with the isotropic interactions present in $\hat{H}_1$ and $\hat{H}_2$. We have

$$\begin{aligned} \hat{\vec{S}}_i \cdot \hat{\vec{S}}_j &= \frac{\hbar^2}{4} \sum_{\alpha,\beta,\gamma,\delta} \hat{b}^\dagger_{i\alpha} \hat{b}_{i\beta} \hat{b}^\dagger_{j\gamma} \hat{b}_{j\delta} (\vec{\sigma}_{\alpha\beta} \cdot \vec{\sigma}_{\gamma\delta}) = \frac{\hbar^2}{4} \sum_{\alpha,\beta,\gamma,\delta} \hat{b}^\dagger_{j\gamma} (\hat{b}_{i\beta} \hat{b}^\dagger_{i\alpha} - \delta_{\alpha\beta}) \hat{b}_{j\delta} (\vec{\sigma}_{\alpha\beta} \cdot \vec{\sigma}_{\gamma\delta}) \\ &= \frac{\hbar^2}{4} \sum_{\alpha,\beta,\gamma,\delta} \hat{b}^\dagger_{j\gamma} \hat{b}_{i\beta} \hat{b}^\dagger_{i\alpha} \hat{b}_{j\delta} (\vec{\sigma}_{\alpha\beta} \cdot \vec{\sigma}_{\gamma\delta}) - \frac{\hbar^2}{4} \sum_{\alpha,\gamma,\delta} \hat{b}^\dagger_{j\gamma} \hat{b}_{j\delta} (\vec{\sigma}_{\alpha\alpha} \cdot \vec{\sigma}_{\gamma\delta}). \end{aligned} \tag{7.22}$$

Using (7.21) with $\mathbf{D}$ as the identity matrix we obtain

$$\begin{aligned} \hat{\vec{S}}_i \cdot \hat{\vec{S}}_j &= \frac{\hbar^2}{4} \Big[\hat{b}^\dagger_{j\uparrow} \hat{b}_{i\uparrow} \hat{b}^\dagger_{i\uparrow} \hat{b}_{j\uparrow} - \hat{b}^\dagger_{j\downarrow} \hat{b}_{i\uparrow} \hat{b}^\dagger_{i\uparrow} \hat{b}_{j\downarrow} + 2\hat{b}^\dagger_{j\downarrow} \hat{b}_{i\downarrow} \hat{b}^\dagger_{i\uparrow} \hat{b}_{j\uparrow} + 2\hat{b}^\dagger_{j\uparrow} \hat{b}_{i\uparrow} \hat{b}^\dagger_{i\downarrow} \hat{b}_{j\downarrow} \\ &\qquad - \hat{b}^\dagger_{j\uparrow} \hat{b}_{i\downarrow} \hat{b}^\dagger_{i\downarrow} \hat{b}_{j\uparrow} + \hat{b}^\dagger_{j\downarrow} \hat{b}_{i\downarrow} \hat{b}^\dagger_{i\downarrow} \hat{b}_{j\downarrow} \Big] - \frac{\hbar^2}{4} \Big[\hat{b}^\dagger_{j\uparrow} \hat{b}_{j\uparrow} - \hat{b}^\dagger_{j\downarrow} \hat{b}_{j\downarrow} - \hat{b}^\dagger_{j\uparrow} \hat{b}_{j\uparrow} + \hat{b}^\dagger_{j\downarrow} \hat{b}_{j\downarrow} \Big] \\ &= \frac{\hbar^2}{4} \Big[2\hat{b}^\dagger_{j\uparrow} \hat{b}_{i\uparrow} \hat{b}^\dagger_{i\uparrow} \hat{b}_{j\uparrow} + 2\hat{b}^\dagger_{j\downarrow} \hat{b}_{i\downarrow} \hat{b}^\dagger_{i\downarrow} \hat{b}_{j\downarrow} + 2\hat{b}^\dagger_{j\downarrow} \hat{b}_{i\downarrow} \hat{b}^\dagger_{i\uparrow} \hat{b}_{j\uparrow} + 2\hat{b}^\dagger_{j\uparrow} \hat{b}_{i\uparrow} \hat{b}^\dagger_{i\downarrow} \hat{b}_{j\downarrow} \\ &\qquad - \sum_{\alpha,\beta} \hat{b}^\dagger_{j\alpha} \hat{b}_{i\beta} \hat{b}^\dagger_{i\beta} \hat{b}_{j\alpha} \Big] = \frac{\hbar^2}{2} \sum_{\alpha,\beta} \hat{b}^\dagger_{j\alpha} \hat{b}_{i\alpha} \hat{b}^\dagger_{i\beta} \hat{b}_{j\beta} - \frac{3\hbar^2}{4}, \end{aligned} \tag{7.23}$$

where we used (7.18) in the last equality. The terms in the sum may be decoupled using the operators

$$\hat{B}_{ij} = \hat{b}^\dagger_{i\uparrow} \hat{b}_{j\uparrow} + \hat{b}^\dagger_{i\downarrow} \hat{b}_{j\downarrow}; \quad \hat{A}_{ij} = \hat{b}_{i\uparrow} \hat{b}_{j\downarrow} - \hat{b}_{i\downarrow} \hat{b}_{j\uparrow}. \tag{7.24}$$

$\hat{B}_{ij}$ is a hopping term, whereas $\hat{A}^\dagger_{ij}$ creates a spin singlet on the oriented bond between sites i and j. The former is characteristic of ordered systems, while the latter is most prominent in disordered systems [33]. These operators allow us to write (7.23) as

$$\hat{\vec{S}}_i \cdot \hat{\vec{S}}_j = \frac{\hbar^2}{4} \Big[\hat{B}^\dagger_{ij} \hat{B}_{ij} - \hat{A}^\dagger_{ij} \hat{A}_{ij} - 1 \Big]. \tag{7.25}$$

Next, we consider the anisotropic term in $\hat{H}_\gamma$:

$$\begin{aligned}
\hat{\vec{S}}_i \cdot [\mathrm{diag}(-1,-1,1)\hat{\vec{S}}_j] &= \frac{\hbar^2}{4}\sum_{\alpha,\beta,\gamma,\delta} \hat{b}^\dagger_{i\alpha}\hat{b}_{i\beta}\hat{b}^\dagger_{j\gamma}\hat{b}_{j\delta}(\vec{\sigma}_{\alpha\beta}\cdot[\mathrm{diag}(-1,-1,1)\vec{\sigma}_{\gamma\delta}]) \\
&= \frac{\hbar^2}{4}\sum_{\alpha,\beta,\gamma,\delta} \hat{b}^\dagger_{j\gamma}(\hat{b}_{i\beta}\hat{b}^\dagger_{i\alpha}-\delta_{\alpha\beta})\hat{b}_{j\delta}(\vec{\sigma}_{\alpha\beta}\cdot[\mathrm{diag}(-1,-1,1)\vec{\sigma}_{\gamma\delta}]) \\
&= \frac{\hbar^2}{4}\sum_{\alpha,\beta,\gamma,\delta} \hat{b}^\dagger_{j\gamma}\hat{b}_{i\beta}\hat{b}^\dagger_{i\alpha}\hat{b}_{j\delta}(\vec{\sigma}_{\alpha\beta}\cdot[\mathrm{diag}(-1,-1,1)\vec{\sigma}_{\gamma\delta}]) \\
&\quad - \frac{\hbar^2}{4}\sum_{\alpha,\gamma,\delta} \hat{b}^\dagger_{j\gamma}\hat{b}_{j\delta}(\vec{\sigma}_{\alpha\alpha}\cdot[\mathrm{diag}(-1,-1,1)\vec{\sigma}_{\gamma\delta}]).
\end{aligned} \tag{7.26}$$

The last term is zero by (7.21), so we are left with

$$\begin{aligned}
\hat{\vec{S}}_i \cdot [\mathrm{diag}(-1,-1,1)\hat{\vec{S}}_j] &= \frac{\hbar^2}{4}\Big[\hat{b}^\dagger_{j\uparrow}\hat{b}_{i\uparrow}\hat{b}^\dagger_{i\uparrow}\hat{b}_{j\uparrow} - \hat{b}^\dagger_{j\downarrow}\hat{b}_{i\uparrow}\hat{b}^\dagger_{i\uparrow}\hat{b}_{j\downarrow} - 2\hat{b}^\dagger_{j\downarrow}\hat{b}_{i\downarrow}\hat{b}^\dagger_{i\uparrow}\hat{b}_{j\uparrow} - 2\hat{b}^\dagger_{j\uparrow}\hat{b}_{i\uparrow}\hat{b}^\dagger_{i\downarrow}\hat{b}_{j\downarrow} \\
&\quad - \hat{b}^\dagger_{j\uparrow}\hat{b}_{i\downarrow}\hat{b}^\dagger_{i\downarrow}\hat{b}_{j\uparrow} + \hat{b}^\dagger_{j\downarrow}\hat{b}_{i\downarrow}\hat{b}^\dagger_{i\downarrow}\hat{b}_{j\downarrow}\Big] \\
&= \frac{\hbar^2}{4}\Big[2\hat{b}^\dagger_{j\uparrow}\hat{b}_{i\uparrow}\hat{b}^\dagger_{i\uparrow}\hat{b}_{j\uparrow} + 2\hat{b}^\dagger_{j\downarrow}\hat{b}_{i\downarrow}\hat{b}^\dagger_{i\downarrow}\hat{b}_{j\downarrow} - 2\hat{b}^\dagger_{j\downarrow}\hat{b}_{i\downarrow}\hat{b}^\dagger_{i\uparrow}\hat{b}_{j\uparrow} - 2\hat{b}^\dagger_{j\uparrow}\hat{b}_{i\uparrow}\hat{b}^\dagger_{i\downarrow}\hat{b}_{j\downarrow} - \sum_{\alpha,\beta}\hat{b}^\dagger_{j\alpha}\hat{b}_{i\beta}\hat{b}^\dagger_{i\beta}\hat{b}_{j\alpha}\Big] \\
&= \frac{\hbar^2}{2}\sum_{\alpha,\beta}(2\delta_{\alpha\beta}-1)\hat{b}^\dagger_{j\alpha}\hat{b}_{i\alpha}\hat{b}^\dagger_{i\beta}\hat{b}_{j\beta} - \frac{3\hbar^2}{4}.
\end{aligned} \tag{7.27}$$

Defining

$$\hat{D}_{ij} = \hat{b}^\dagger_{i\uparrow}\hat{b}_{j\uparrow} - \hat{b}^\dagger_{i\downarrow}\hat{b}_{j\downarrow}; \quad \hat{E}_{ij} = \hat{b}_{i\uparrow}\hat{b}_{j\downarrow} + \hat{b}_{i\downarrow}\hat{b}_{j\uparrow}, \tag{7.28}$$

which are variants of $\hat{B}_{ij}$ and $\hat{A}_{ij}$, respectively, we can write the anisotropic term as

$$\hat{\vec{S}}_i \cdot [\mathrm{diag}(-1,-1,1)\hat{\vec{S}}_j] = \frac{\hbar^2}{4}\Big[\hat{D}^\dagger_{ij}\hat{D}_{ij} - \hat{E}^\dagger_{ij}\hat{E}_{ij} - 1\Big]. \tag{7.29}$$

Finally, we consider the DMI term

$$\begin{aligned}
\hat{e}_3 \cdot (\hat{\vec{S}}_i \times \hat{\vec{S}}_j) &= \hat{S}^x_i\hat{S}^y_j - \hat{S}^y_i\hat{S}^x_j \\
&= \frac{i\hbar^2}{4}\Big[(\hat{b}^\dagger_{i\uparrow}\hat{b}_{i\downarrow} + \hat{b}^\dagger_{i\downarrow}\hat{b}_{i\uparrow})(\hat{b}^\dagger_{j\downarrow}\hat{b}_{j\uparrow} - \hat{b}^\dagger_{j\uparrow}\hat{b}_{j\downarrow}) - (\hat{b}^\dagger_{i\downarrow}\hat{b}_{i\uparrow} - \hat{b}^\dagger_{i\uparrow}\hat{b}_{i\downarrow})(\hat{b}^\dagger_{j\uparrow}\hat{b}_{j\downarrow} + \hat{b}^\dagger_{j\downarrow}\hat{b}_{j\uparrow})\Big] \\
&= \frac{i\hbar^2}{4}\Big[2\hat{b}^\dagger_{i\uparrow}\hat{b}_{i\downarrow}\hat{b}^\dagger_{j\downarrow}\hat{b}_{j\uparrow} - 2\hat{b}^\dagger_{i\downarrow}\hat{b}_{i\uparrow}\hat{b}^\dagger_{j\uparrow}\hat{b}_{j\downarrow}\Big] = \frac{i\hbar^2}{4}\Big[2\hat{b}^\dagger_{j\downarrow}\hat{b}_{i\downarrow}\hat{b}^\dagger_{i\uparrow}\hat{b}_{j\uparrow} - 2\hat{b}^\dagger_{j\uparrow}\hat{b}_{i\uparrow}\hat{b}^\dagger_{i\downarrow}\hat{b}_{j\downarrow}\Big].
\end{aligned} \tag{7.30}$$

Using (7.24) and (7.28), this term can be written as

$$\hat{e}_3 \cdot (\hat{\vec{S}}_i \times \hat{\vec{S}}_j) = \frac{i\hbar^2}{4}\Big[\hat{B}^\dagger_{ij}\hat{D}_{ij} - \hat{D}^\dagger_{ij}\hat{B}_{ij}\Big] \tag{7.31}$$

or

$$\hat{e}_3 \cdot (\hat{\vec{S}}_i \times \hat{\vec{S}}_j) = \frac{i\hbar^2}{4}\left[\hat{A}^\dagger_{ij}\hat{E}_{ij} - \hat{E}^\dagger_{ij}\hat{A}_{ij}\right]. \tag{7.32}$$

To include all decoupling operators we add (7.31) and (7.32) and divide by two, such that the decoupled DMI term is given by

$$\hat{e}_3 \cdot (\hat{\vec{S}}_i \times \hat{\vec{S}}_j) = \frac{i\hbar^2}{8}\left[\hat{A}^\dagger_{ij}\hat{E}_{ij} + \hat{B}^\dagger_{ij}\hat{D}_{ij} - \hat{D}^\dagger_{ij}\hat{B}_{ij} - \hat{E}^\dagger_{ij}\hat{A}_{ij}\right]. \tag{7.33}$$

7.5 The Zero-Flux Mean Field Ansatz

With the interactions written in terms of the decoupling operators (7.24) and (7.28), we can use the mean field approximation (2.57) to make the Hamiltonian quadratic in the bosonic operators. Dropping the constant terms in (7.25) and (7.29), we obtain

$$\begin{aligned}\hat{\vec{S}}_i \cdot \hat{\vec{S}}_j \approx &\frac{\hbar^2}{4}\left[\langle\hat{B}_{ij}\rangle^*\hat{B}_{ij} - \langle\hat{A}_{ij}\rangle^*\hat{A}_{ij} + \langle\hat{B}_{ij}\rangle\hat{B}^\dagger_{ij} - \langle\hat{A}_{ij}\rangle\hat{A}^\dagger_{ij}\right] \\ &- \frac{\hbar^2}{4}|\langle\hat{B}_{ij}\rangle|^2 + \frac{\hbar^2}{4}|\langle\hat{A}_{ij}\rangle|^2,\end{aligned} \tag{7.34}$$

$$\begin{aligned}\hat{\vec{S}}_i \cdot [\mathrm{diag}(-1,-1,1)\hat{\vec{S}}_j] \approx &\frac{\hbar^2}{4}\left[\langle\hat{D}_{ij}\rangle^*\hat{D}_{ij} - \langle\hat{E}_{ij}\rangle^*\hat{E}_{ij} + \langle\hat{D}_{ij}\rangle\hat{D}^\dagger_{ij} - \langle\hat{E}_{ij}\rangle\hat{E}^\dagger_{ij}\right] \\ &- \frac{\hbar^2}{4}|\langle\hat{D}_{ij}\rangle|^2 + \frac{\hbar^2}{4}|\langle\hat{E}_{ij}\rangle|^2,\end{aligned} \tag{7.35}$$

$$\begin{aligned}\hat{e}_3 \cdot (\hat{\vec{S}}_i \times \hat{\vec{S}}_j) \approx &\frac{i\hbar^2}{8}\Big[\langle\hat{A}_{ij}\rangle^*\hat{E}_{ij} + \langle\hat{E}_{ij}\rangle\hat{A}^\dagger_{ij} + \langle\hat{B}_{ij}\rangle^*\hat{D}_{ij} \\ &+ \langle\hat{D}_{ij}\rangle\hat{B}^\dagger_{ij} - \langle\hat{A}_{ij}\rangle^*\langle\hat{E}_{ij}\rangle - \langle\hat{B}_{ij}\rangle^*\langle\hat{D}_{ij}\rangle\Big] + \mathrm{H.c.},\end{aligned} \tag{7.36}$$

where H.c. denotes the Hermitian conjugate. To enforce the constraint (7.17) on average, we add the Lagrange term

$$\hat{H}_L = \hbar^2 \sum_i \mu_i(\hat{n}_i - \kappa) \tag{7.37}$$

to the mean field Hamiltonian. The number of mean field parameters may be reduced by a particular mean field Ansatz. We will employ the same Ansatz as Ushakov [6] and Vaezi et al. [7], which is based on the projective symmetry group analysis conducted by Wang [34]. We briefly outline the motivation behind this analysis. The Schwinger boson transformed spin Hamiltonian is invariant under local gauge transformations

$$\hat{b}_{i\alpha} \to e^{i\phi(i)}\hat{b}_{i\alpha}. \tag{7.38}$$

However, the mean field parameters are not, and transform as

$$\langle\hat{A}_{ij}\rangle \to e^{i[\phi(i)+\phi(j)]}\langle\hat{A}_{ij}\rangle; \quad \langle\hat{B}_{ij}\rangle \to e^{i[-\phi(i)+\phi(j)]}\langle\hat{B}_{ij}\rangle. \tag{7.39}$$

Table 1: Mean field Ansatz of the Schwinger boson zero-flux mean field theory on the two sublattices A and B of the honeycomb lattice. Only non-zero parameters are shown.

$$\begin{array}{llllll} \text{A:} & \langle \hat{A}_{\langle i,j\rangle} \rangle = -\Delta_1 & \langle \hat{A}_{\langle\langle i,j\rangle\rangle} \rangle = +\nu_{ij}\Delta_2 & \langle \hat{E}_{\langle\langle i,j\rangle\rangle} \rangle = \Delta_3 & \mu_i = \mu \\ \text{B:} & \langle \hat{A}_{\langle i,j\rangle} \rangle = +\Delta_1 & \langle \hat{A}_{\langle\langle i,j\rangle\rangle} \rangle = -\nu_{ij}\Delta_2 & \langle \hat{E}_{\langle\langle i,j\rangle\rangle} \rangle = \Delta_3 & \mu_i = \mu \end{array}$$

This seems problematic, since the physical spin state is gauge invariant. The gauge invariance of the spin state is satisfied if the constraint (7.17) is implemented exactly, and not just on mean field level through the Lagrange term (7.37). We must therefore project the mean field spin states onto the physical space of spin states. This is accomplished by demanding that the mean field parameters be invariant under a combined physical and gauge transformation. We associate a local gauge group element $e^{i\phi_X(i)}$ to each generator X of the physical symmetry group of the lattice, and demand that the symmetry operation

$$\hat{b}_{i\alpha} \to e^{i\phi_X[X(i)]}\hat{b}_{X(i)\alpha} \tag{7.40}$$

leaves the mean field parameters unchanged. This places constraints on the allowed mean field parameters. Analyzing these constraints, Wang [34] finds that the so-called zero-flux mean field Ansatz, which is the one used by Ushakov [6] and Vaezi et al. [7], is a promising candidate for potential quantum spin liquid states on the honeycomb lattice. In the zero-flux state, the gauge flux through each plaquette of the honeycomb lattice is zero. The non-zero mean field parameters of our zero-flux Ansatz is shown in Table 1. Δ_1, Δ_2, Δ_3 and μ will be taken to be real in the following.

7.6 The Fourier Transform of the Mean Field Hamiltonian

The next step is to Fourier transform the mean field Hamiltonian to decouple the operators acting on different lattice sites. We start off with the Lagrange term $\hat{H}_L$. We have

$$\begin{aligned} \sum_i \hat{n}_i &= \sum_i \sum_\alpha \hat{b}^\dagger_{i\alpha}\hat{b}_{i\alpha} = \sum_{\vec{r}} \sum_X \sum_\alpha \hat{b}^\dagger_{\vec{r}\alpha X}\hat{b}_{\vec{r}\alpha X} = \frac{1}{N}\sum_{\vec{r}}\sum_X\sum_\alpha\sum_{\vec{k},\vec{k}'} e^{i(\vec{k}'-\vec{k})\cdot\vec{r}}\hat{b}^\dagger_{\vec{k}\alpha X}\hat{b}_{\vec{k}'\alpha X} \\ &= \sum_X\sum_\alpha\sum_{\vec{k},\vec{k}'}\delta_{\vec{k}\vec{k}'}\hat{b}^\dagger_{\vec{k}\alpha X}\hat{b}_{\vec{k}'\alpha X} = \sum_X\sum_\alpha\sum_{\vec{k}}\hat{b}^\dagger_{\vec{k}\alpha X}\hat{b}_{\vec{k}\alpha X} \\ &= \sum_X\sum_{\vec{k}}\left[\hat{b}^\dagger_{\vec{k}\uparrow X}\hat{b}_{\vec{k}\uparrow X} + \hat{b}^\dagger_{\vec{k}\downarrow X}\hat{b}_{\vec{k}\downarrow X}\right] = \sum_X\sum_{\vec{k}}\left[\hat{b}^\dagger_{\vec{k}\uparrow X}\hat{b}_{\vec{k}\uparrow X} + \hat{b}^\dagger{}_{-\vec{k}\downarrow X}\hat{b}_{-\vec{k}\downarrow X}\right] \\ &= \sum_X\sum_{\vec{k}}\left[\hat{b}^\dagger_{\vec{k}\uparrow X}\hat{b}_{\vec{k}\uparrow X} + \hat{b}_{-\vec{k}\downarrow X}\hat{b}^\dagger{}_{-\vec{k}\downarrow X} - 1\right] \\ &= \sum_{\vec{k}}\left[\hat{b}^\dagger_{\vec{k}\uparrow A}\hat{b}_{\vec{k}\uparrow A} + \hat{b}^\dagger_{\vec{k}\uparrow B}\hat{b}_{\vec{k}\uparrow B} + \hat{b}_{-\vec{k}\downarrow A}\hat{b}^\dagger{}_{-\vec{k}\downarrow A} + \hat{b}_{-\vec{k}\downarrow B}\hat{b}^\dagger{}_{-\vec{k}\downarrow B} - 2\right], \end{aligned} \tag{7.41}$$

which means that

$$\begin{aligned}\hat{H}_L &= \hbar^2\mu\sum_i(\hat{n}_i-\kappa) = \hbar^2\mu\Big(\sum_i\hat{n}_i-\sum_i\kappa\Big) = \hbar^2\mu\Big(\sum_i\hat{n}_i-\sum_{\vec{r}}\sum_X\kappa\Big)\\ &= \hbar^2\mu\Big(\sum_i\hat{n}_i-2\sum_{\vec{r}}\kappa\Big) = \hbar^2\mu\Big(\sum_i\hat{n}_i-2\sum_{\vec{k}}\kappa\Big)\\ &= \hbar^2\mu\sum_{\vec{k}}\Big[\hat{b}^\dagger_{\vec{k}\uparrow A}\hat{b}_{\vec{k}\uparrow A}+\hat{b}^\dagger_{\vec{k}\uparrow B}\hat{b}_{\vec{k}\uparrow B}+\hat{b}_{-\vec{k}\downarrow A}\hat{b}^\dagger{}_{-\vec{k}\downarrow A}+\hat{b}_{-\vec{k}\downarrow B}\hat{b}^\dagger{}_{-\vec{k}\downarrow B}-2-2\kappa\Big] \qquad (7.42)\\ &= \hbar^2\sum_{\vec{k}}\Big[\hat{\mathbf{b}}^\dagger_{\vec{k}}\begin{bmatrix}\mu&0&0&0\\0&\mu&0&0\\0&0&\mu&0\\0&0&0&\mu\end{bmatrix}\hat{\mathbf{b}}_{\vec{k}}-(2+2\kappa)\mu\Big],\end{aligned}$$

where

$$\hat{\mathbf{b}}_{\vec{k}} = \begin{bmatrix}\hat{b}_{\vec{k}\uparrow A} & \hat{b}_{\vec{k}\uparrow B} & \hat{b}^\dagger_{-\vec{k}\downarrow A} & \hat{b}^\dagger_{-\vec{k}\downarrow B}\end{bmatrix}^\mathsf{T}. \qquad (7.43)$$

The Zeeman term $\hat{H}_Z$ is Fourier transformed in much the same way as the Lagrange term. In the Schwinger boson representation, it takes the form

$$\hat{H}_Z = -\frac{\hbar B}{2}\sum_i\left(\hat{b}^\dagger_{i\uparrow}\hat{b}_{i\uparrow}-\hat{b}^\dagger_{i\downarrow}\hat{b}_{i\downarrow}\right) = -\frac{\hbar B}{2}\sum_i\sum_\alpha\alpha\hat{b}^\dagger_{i\alpha}\hat{b}_{i\alpha}, \qquad (7.44)$$

and by comparison to (7.41) we immediately obtain

$$\begin{aligned}\hat{H}_Z &= \hbar^2\Big(-\frac{B}{2\hbar}\Big)\sum_{\vec{k}}\Big[\hat{b}^\dagger_{\vec{k}\uparrow A}\hat{b}_{\vec{k}\uparrow A}+\hat{b}^\dagger_{\vec{k}\uparrow B}\hat{b}_{\vec{k}\uparrow B}-\hat{b}_{-\vec{k}\downarrow A}\hat{b}^\dagger{}_{-\vec{k}\downarrow A}-\hat{b}_{-\vec{k}\downarrow B}\hat{b}^\dagger{}_{-\vec{k}\downarrow B}+2\Big]\\ &= \hbar^2\sum_{\vec{k}}\Big[\hat{\mathbf{b}}^\dagger_{\vec{k}}\begin{bmatrix}-B/2\hbar&0&0&0\\0&-B/2\hbar&0&0\\0&0&B/2\hbar&0\\0&0&0&B/2\hbar\end{bmatrix}\hat{\mathbf{b}}_{\vec{k}}-\frac{B}{\hbar}\Big].\end{aligned} \qquad (7.45)$$

Since the constant term does not depend on any mean field parameters, we drop it in the following.

We now move on to the nearest neighbor isotropic mean field Hamiltonian:

$$\hat{H}_{1(\mathrm{MF})} = -\frac{\hbar^2J_1}{4}\sum_{\langle i,j\rangle}\Big[\langle\hat{A}_{ij}\rangle^*\hat{A}_{ij}+\langle\hat{A}_{ij}\rangle\hat{A}^\dagger_{ij}\Big]+\frac{\hbar^2J_1}{4}\sum_{\langle i,j\rangle}|\langle\hat{A}_{ij}\rangle|^2. \qquad (7.46)$$

The first operator term is

$$\sum_{\langle i,j\rangle}\langle\hat{A}_{ij}\rangle^*\hat{A}_{ij} = \sum_{\langle i,j\rangle}\langle\hat{A}_{ij}\rangle^*(\hat{b}_{i\uparrow}\hat{b}_{j\downarrow}-\hat{b}_{i\downarrow}\hat{b}_{j\uparrow}). \qquad (7.47)$$

Consider the first part of the right hand side of (7.47):

$$\begin{aligned}\sum_{\langle i,j\rangle}\langle\hat{A}_{ij}\rangle^*\hat{b}_{i\uparrow}\hat{b}_{j\downarrow} &= \sum_{\vec{r}}\sum_{l=1}^{3}\left[\langle\hat{A}_{\vec{r},\vec{r}+\vec{\delta}_l^{(A)}}\rangle^*\hat{b}_{\vec{r}\uparrow A}\hat{b}_{\vec{r}+\vec{\delta}_l^{(A)}\downarrow B}+\langle\hat{A}_{\vec{r},\vec{r}+\vec{\delta}_l^{(B)}}\rangle^*\hat{b}_{\vec{r}\uparrow B}\hat{b}_{\vec{r}+\vec{\delta}_l^{(B)}\downarrow A}\right]\\ &= \Delta_1\sum_{\vec{r}}\sum_{l=1}^{3}\left[\hat{b}_{\vec{r}\uparrow B}\hat{b}_{\vec{r}+\vec{\delta}_l^{(B)}\downarrow A}-\hat{b}_{\vec{r}\uparrow A}\hat{b}_{\vec{r}+\vec{\delta}_l^{(A)}\downarrow B}\right]\\ &= \Delta_1\sum_{\vec{k}}\sum_{l=1}^{3}\left[e^{-i\vec{k}\cdot\vec{\delta}_l^{(B)}}\hat{b}_{\vec{k}\uparrow B}\hat{b}_{-\vec{k}\downarrow A}-e^{-i\vec{k}\cdot\vec{\delta}_l^{(A)}}\hat{b}_{\vec{k}\uparrow A}\hat{b}_{-\vec{k}\downarrow B}\right].\end{aligned} \tag{7.48}$$

Since the second part of the right hand side of (7.47) is the same as the first, but with the spin indices interchanged, we get

$$\begin{aligned}\sum_{\langle i,j\rangle}\langle\hat{A}_{ij}\rangle^*\hat{A}_{ij} &= \Delta_1\sum_{\vec{k}}\sum_{l=1}^{3}\Big[e^{-i\vec{k}\cdot\vec{\delta}_l^{(B)}}\hat{b}_{-\vec{k}\downarrow A}\hat{b}_{\vec{k}\uparrow B}+e^{i\vec{k}\cdot\vec{\delta}_l^{(A)}}\hat{b}_{-\vec{k}\downarrow A}\hat{b}_{\vec{k}\uparrow B}\\ &\qquad -e^{i\vec{k}\cdot\vec{\delta}_l^{(B)}}\hat{b}_{-\vec{k}\downarrow B}\hat{b}_{\vec{k}\uparrow A}-e^{-i\vec{k}\cdot\vec{\delta}_l^{(A)}}\hat{b}_{-\vec{k}\downarrow B}\hat{b}_{\vec{k}\uparrow A}\Big]\\ &= 2\Delta_1\sum_{\vec{k}}\sum_{l=1}^{3}\left[e^{i\vec{k}\cdot\vec{\delta}_l^{(A)}}\hat{b}_{-\vec{k}\downarrow A}\hat{b}_{\vec{k}\uparrow B}-e^{-i\vec{k}\cdot\vec{\delta}_l^{(A)}}\hat{b}_{-\vec{k}\downarrow B}\hat{b}_{\vec{k}\uparrow A}\right]\\ &= 2\Delta_1\sum_{\vec{k}}\left[\eta_{\vec{k}}\hat{b}_{-\vec{k}\downarrow A}\hat{b}_{\vec{k}\uparrow B}-\eta_{\vec{k}}^*\hat{b}_{-\vec{k}\downarrow B}\hat{b}_{\vec{k}\uparrow A}\right],\end{aligned} \tag{7.49}$$

where we defined

$$\eta_{\vec{k}}=\sum_{l=1}^{3}e^{i\vec{k}\cdot\vec{\delta}_l^{(A)}}=e^{-ik_y a}+2e^{ik_y a/2}\cos\left(\frac{\sqrt{3}a}{2}k_x\right). \tag{7.50}$$

The second operator term in (7.46) is just the conjugate of (7.49), so we have

$$\begin{aligned}\hat{H}_{1(\mathrm{MF})} &= \frac{\hbar^2 J_1\Delta_1}{2}\sum_{\vec{k}}\left[\eta_{\vec{k}}\big(\hat{b}^\dagger_{\vec{k}\uparrow A}\hat{b}^\dagger_{-\vec{k}\downarrow B}-\hat{b}_{-\vec{k}\downarrow A}\hat{b}_{\vec{k}\uparrow B}\big)+\eta_{\vec{k}}^*\big(\hat{b}_{-\vec{k}\downarrow B}\hat{b}_{\vec{k}\uparrow A}-\hat{b}^\dagger_{\vec{k}\uparrow B}\hat{b}^\dagger_{-\vec{k}\downarrow A}\big)\right]\\ &\quad+\frac{\hbar^2 J_1}{4}\sum_{\vec{k}}6\Delta_1^2\\ &= \hbar^2\sum_{\vec{k}}\left[\frac{J_1\Delta_1}{2}\hat{\mathbf{b}}^\dagger_{\vec{k}}\begin{bmatrix}0&0&0&\eta_{\vec{k}}\\0&0&-\eta_{\vec{k}}^*&0\\0&-\eta_{\vec{k}}&0&0\\\eta_{\vec{k}}^*&0&0&0\end{bmatrix}\hat{\mathbf{b}}_{\vec{k}}+\frac{3}{2}J_1\Delta_1^2\right],\end{aligned} \tag{7.51}$$

where we used that there are two sublattices and three nearest neighbors for each lattice site for the constant term.

The next term is the next nearest neighbor isotropic mean field Hamiltonian:

$$\hat{H}_{2(\mathrm{MF})} = -\frac{\hbar^2 J_2}{4} \sum_{\langle\langle i,j\rangle\rangle} \left[\langle \hat{A}_{ij} \rangle^* \hat{A}_{ij} + \langle \hat{A}_{ij} \rangle \hat{A}_{ij}^\dagger \right] + \frac{\hbar^2 J_2}{4} \sum_{\langle\langle i,j\rangle\rangle} |\langle \hat{A}_{ij} \rangle|^2. \tag{7.52}$$

We follow the steps taken in the treatment of the nearest neighbor isotropic mean field Hamiltonian. The first operator term is

$$\sum_{\langle\langle i,j\rangle\rangle} \langle \hat{A}_{ij} \rangle^* \hat{A}_{ij} = \sum_{\langle\langle i,j\rangle\rangle} \langle \hat{A}_{ij} \rangle^* (\hat{b}_{i\uparrow} \hat{b}_{j\downarrow} - \hat{b}_{i\downarrow} \hat{b}_{j\uparrow}). \tag{7.53}$$

and the first part of the right hand side of (7.53) is

$$\begin{aligned} \sum_{\langle\langle i,j\rangle\rangle} \langle \hat{A}_{ij} \rangle^* \hat{b}_{i\uparrow} \hat{b}_{j\downarrow} &= \sum_{\vec{r}} \sum_{l=1}^{6} \left[\langle \hat{A}_{\vec{r},\vec{r}+\vec{\epsilon}_l} \rangle^* \hat{b}_{\vec{r}\uparrow A} \hat{b}_{\vec{r}+\vec{\epsilon}_l\downarrow A} + \langle \hat{A}_{\vec{r},\vec{r}+\vec{\epsilon}_l} \rangle^* \hat{b}_{\vec{r}\uparrow B} \hat{b}_{\vec{r}+\vec{\epsilon}_l\downarrow B} \right] \\ &= \Delta_2 \sum_{\vec{r}} \sum_{l=1}^{6} \left[\nu_l^{(A)} \hat{b}_{\vec{r}\uparrow A} \hat{b}_{\vec{r}+\vec{\epsilon}_l\downarrow A} - \nu_l^{(B)} \hat{b}_{\vec{r}\uparrow B} \hat{b}_{\vec{r}+\vec{\epsilon}_l\downarrow B} \right] \\ &= \Delta_2 \sum_{\vec{k}} \sum_{l=1}^{6} \left[\nu_l^{(A)} e^{-i\vec{k}\cdot\vec{\epsilon}_l} \left(\hat{b}_{\vec{k}\uparrow A} \hat{b}_{-\vec{k}\downarrow A} + \hat{b}_{\vec{k}\uparrow B} \hat{b}_{-\vec{k}\downarrow B} \right) \right]. \end{aligned} \tag{7.54}$$

The second part of the right hand side of (7.53) is

$$\begin{aligned} \sum_{\langle\langle i,j\rangle\rangle} \langle \hat{A}_{ij} \rangle^* \hat{b}_{i\downarrow} \hat{b}_{j\uparrow} &= \Delta_2 \sum_{\vec{k}} \sum_{l=1}^{6} \left[\nu_l^{(A)} e^{-i\vec{k}\cdot\vec{\epsilon}_l} \left(\hat{b}_{\vec{k}\downarrow A} \hat{b}_{-\vec{k}\uparrow A} + \hat{b}_{\vec{k}\downarrow B} \hat{b}_{-\vec{k}\uparrow B} \right) \right] \\ &= \Delta_2 \sum_{\vec{k}} \sum_{l=1}^{6} \left[\nu_l^{(A)} e^{i\vec{k}\cdot\vec{\epsilon}_l} \left(\hat{b}_{\vec{k}\uparrow A} \hat{b}_{-\vec{k}\downarrow A} + \hat{b}_{\vec{k}\uparrow B} \hat{b}_{-\vec{k}\downarrow B} \right) \right]. \end{aligned} \tag{7.55}$$

We therefore have

$$\begin{aligned} \sum_{\langle\langle i,j\rangle\rangle} \langle \hat{A}_{ij} \rangle^* \hat{A}_{ij} &= \Delta_2 \sum_{\vec{k}} \sum_{l=1}^{6} \Big[\nu_l^{(A)} e^{-i\vec{k}\cdot\vec{\epsilon}_l} \left(\hat{b}_{\vec{k}\uparrow A} \hat{b}_{-\vec{k}\downarrow A} + \hat{b}_{\vec{k}\uparrow B} \hat{b}_{-\vec{k}\downarrow B} \right) \\ &\qquad - \nu_l^{(A)} e^{i\vec{k}\cdot\vec{\epsilon}_l} \left(\hat{b}_{\vec{k}\uparrow A} \hat{b}_{-\vec{k}\downarrow A} + \hat{b}_{\vec{k}\uparrow B} \hat{b}_{-\vec{k}\downarrow B} \right) \Big] \\ &= \Delta_2 \sum_{\vec{k}} \sum_{l=1}^{6} \left[\nu_l^{(A)} \left(e^{-i\vec{k}\cdot\vec{\epsilon}_l} - e^{i\vec{k}\cdot\vec{\epsilon}_l} \right) \left(\hat{b}_{\vec{k}\uparrow A} \hat{b}_{-\vec{k}\downarrow A} + \hat{b}_{\vec{k}\uparrow B} \hat{b}_{-\vec{k}\downarrow B} \right) \right] \\ &= -2i\Delta_2 \sum_{\vec{k}} \sum_{l=1}^{6} \nu_l^{(A)} \sin\left(\vec{k}\cdot\vec{\epsilon}_l\right) \left[\hat{b}_{\vec{k}\uparrow A} \hat{b}_{-\vec{k}\downarrow A} + \hat{b}_{\vec{k}\uparrow B} \hat{b}_{-\vec{k}\downarrow B} \right] \\ &= -2i\Delta_2 \sum_{\vec{k}} \xi_{\vec{k}} \left[\hat{b}_{\vec{k}\uparrow A} \hat{b}_{-\vec{k}\downarrow A} + \hat{b}_{\vec{k}\uparrow B} \hat{b}_{-\vec{k}\downarrow B} \right], \end{aligned} \tag{7.56}$$

where we defined

$$\xi_{\vec{k}} = \sum_{l=1}^{6} \nu_l^{(A)} \sin\left(\vec{k}\cdot\vec{\epsilon}_l\right) = 4\sin\left(\frac{\sqrt{3}a}{2}k_x\right)\left[\cos\left(\frac{3a}{2}k_y\right) - \cos\left(\frac{\sqrt{3}a}{2}k_x\right)\right]. \tag{7.57}$$

Inserting (7.56) and its conjugate into (7.52), we obtain:

$$\begin{aligned}\hat{H}_{2(\mathrm{MF})} &= \frac{\hbar^2 J_2 \Delta_2}{2}\sum_{\vec{k}}\left[i\xi_{\vec{k}}\left(\hat{b}_{\vec{k}\uparrow A}\hat{b}_{-\vec{k}\downarrow A} + \hat{b}_{\vec{k}\uparrow B}\hat{b}_{-\vec{k}\downarrow B}\right) - i\xi_{\vec{k}}\left(\hat{b}^\dagger_{\vec{k}\uparrow A}\hat{b}^\dagger_{-\vec{k}\downarrow A} + \hat{b}^\dagger_{\vec{k}\uparrow B}\hat{b}^\dagger_{-\vec{k}\downarrow B}\right)\right] \\ &\quad + \hbar^2\sum_{\vec{k}} 3J_2\Delta_2^2 \\ &= \hbar^2\sum_{\vec{k}}\left[\frac{J_2\Delta_2}{2}\hat{\mathbf{b}}^\dagger_{\vec{k}}\begin{bmatrix} 0 & 0 & -i\xi_{\vec{k}} & 0 \\ 0 & 0 & 0 & -i\xi_{\vec{k}} \\ i\xi_{\vec{k}} & 0 & 0 & 0 \\ 0 & i\xi_{\vec{k}} & 0 & 0 \end{bmatrix}\hat{\mathbf{b}}_{\vec{k}} + 3J_2\Delta_2^2\right],\end{aligned} \tag{7.58}$$

where we used that there are two sublattices and six next nearest neighbors for each lattice site for the constant term.

Next up is the next nearest neighbor anisotropic mean field Hamiltonian:

$$\hat{H}_{\Gamma(\mathrm{MF})} = -\frac{\hbar^2\Gamma}{4}\sum_{\langle\langle i,j\rangle\rangle}\left[\langle\hat{E}_{ij}\rangle^*\hat{E}_{ij} + \langle\hat{E}_{ij}\rangle\hat{E}^\dagger_{ij}\right] + \frac{\hbar^2\Gamma}{4}\sum_{\langle\langle i,j\rangle\rangle}|\langle\hat{E}_{ij}\rangle|^2. \tag{7.59}$$

Since the only difference between the operators $\hat{A}_{ij}$ and $\hat{E}_{ij}$ is a sign, the treatment of this Hamiltonian is very similar to the next nearest neighbor isotropic Hamiltonian (7.52). We have

$$\sum_{\langle\langle i,j\rangle\rangle}\langle\hat{E}_{ij}\rangle^*\hat{E}_{ij} = \Delta_3\sum_{\langle\langle i,j\rangle\rangle}\left[\hat{b}_{i\uparrow}\hat{b}_{j\downarrow} + \hat{b}_{i\downarrow}\hat{b}_{j\uparrow}\right] = 2\Delta_3\sum_{\vec{k}}\zeta_{\vec{k}}\left[\hat{b}_{\vec{k}\uparrow A}\hat{b}_{-\vec{k}\downarrow A} + \hat{b}_{\vec{k}\uparrow B}\hat{b}_{-\vec{k}\downarrow B}\right], \tag{7.60}$$

where

$$\zeta_{\vec{k}} = \sum_{l=1}^{6}\cos\left(\vec{k}\cdot\vec{\epsilon}_l\right) = 2\left[2\cos\left(\frac{\sqrt{3}a}{2}k_x\right)\cos\left(\frac{3a}{2}k_y\right) + \cos\left(\sqrt{3}ak_x\right)\right]. \tag{7.61}$$

The Fourier transformed next nearest neighbor anisotropic mean field Hamiltonian then reads:

$$\hat{H}_{\Gamma(\mathrm{MF})} = \hbar^2\sum_{\vec{k}}\left[\frac{\Gamma\Delta_3}{2}\hat{\mathbf{b}}^\dagger_{\vec{k}}\begin{bmatrix} 0 & 0 & -\zeta_{\vec{k}} & 0 \\ 0 & 0 & 0 & -\zeta_{\vec{k}} \\ -\zeta_{\vec{k}} & 0 & 0 & 0 \\ 0 & -\zeta_{\vec{k}} & 0 & 0 \end{bmatrix}\hat{\mathbf{b}}_{\vec{k}} + 3\Gamma\Delta_3^2\right]. \tag{7.62}$$

The remaining term is the next nearest neighbor mean field DMI Hamiltonian:

$$\begin{aligned}\hat{H}_{D_\perp(\mathrm{MF})} = \frac{i\hbar^2 D_\perp}{8}\sum_{\langle\langle i,j\rangle\rangle}\nu_{ij}\Big[\langle\hat{A}_{ij}\rangle^*\hat{E}_{ij} + \langle\hat{E}_{ij}\rangle\hat{A}_{ij}^\dagger - \langle\hat{E}_{ij}\rangle^*\hat{A}_{ij} - \langle\hat{A}_{ij}\rangle\hat{E}_{ij}^\dagger\Big] \\ + \frac{i\hbar^2 D_\perp}{8}\sum_{\langle\langle i,j\rangle\rangle}\nu_{ij}\Big[\langle\hat{E}_{ij}\rangle^*\langle\hat{A}_{ij}\rangle - \langle\hat{A}_{ij}\rangle^*\langle\hat{E}_{ij}\rangle\Big].\end{aligned} \tag{7.63}$$

Consider the first operator term:

$$\sum_{\langle\langle i,j\rangle\rangle}\nu_{ij}\langle\hat{A}_{ij}\rangle^*\hat{E}_{ij} = \sum_{\langle\langle i,j\rangle\rangle}\nu_{ij}\langle\hat{A}_{ij}\rangle^*\big(\hat{b}_{i\uparrow}\hat{b}_{j\downarrow} + \hat{b}_{i\downarrow}\hat{b}_{j\uparrow}\big). \tag{7.64}$$

The first part of the right hand side of (7.64) is

$$\begin{aligned}\sum_{\langle\langle i,j\rangle\rangle}\nu_{ij}\langle\hat{A}_{ij}\rangle^*\hat{b}_{i\uparrow}\hat{b}_{j\downarrow} &= \sum_{\vec{r}}\sum_{l=1}^{6}\Big[\nu_l^{(A)}\langle\hat{A}_{\vec{r},\vec{r}+\vec{\epsilon}_l}\rangle^*\hat{b}_{\vec{r}\uparrow A}\hat{b}_{\vec{r}+\vec{\epsilon}_l\downarrow A} + \nu_l^{(B)}\langle\hat{A}_{\vec{r},\vec{r}+\vec{\epsilon}_l}\rangle^*\hat{b}_{\vec{r}\uparrow B}\hat{b}_{\vec{r}+\vec{\epsilon}_l\downarrow B}\Big] \\ &= \Delta_2\sum_{\vec{r}}\sum_{l=1}^{6}\Big[\nu_l^{(A)}\nu_l^{(A)}\hat{b}_{\vec{r}\uparrow A}\hat{b}_{\vec{r}+\vec{\epsilon}_l\downarrow A} - \nu_l^{(B)}\nu_l^{(B)}\hat{b}_{\vec{r}\uparrow B}\hat{b}_{\vec{r}+\vec{\epsilon}_l\downarrow B}\Big] \\ &= \Delta_2\sum_{\vec{r}}\sum_{l=1}^{6}\Big[\hat{b}_{\vec{r}\uparrow A}\hat{b}_{\vec{r}+\vec{\epsilon}_l\downarrow A} - \hat{b}_{\vec{r}\uparrow B}\hat{b}_{\vec{r}+\vec{\epsilon}_l\downarrow B}\Big] \\ &= \Delta_2\sum_{\vec{k}}\sum_{l=1}^{6}\Big[e^{-i\vec{k}\cdot\vec{\epsilon}_l}\hat{b}_{\vec{k}\uparrow A}\hat{b}_{-\vec{k}\downarrow A} - e^{-i\vec{k}\cdot\vec{\epsilon}_l}\hat{b}_{\vec{k}\uparrow B}\hat{b}_{-\vec{k}\downarrow B}\Big] \\ &= \Delta_2\sum_{\vec{k}}\sum_{l=1}^{6}e^{-i\vec{k}\cdot\vec{\epsilon}_l}\Big[\hat{b}_{\vec{k}\uparrow A}\hat{b}_{-\vec{k}\downarrow A} - \hat{b}_{\vec{k}\uparrow B}\hat{b}_{-\vec{k}\downarrow B}\Big].\end{aligned} \tag{7.65}$$

The second part of the right hand side of (7.64) is found by interchanging the spin indices of the operators in (7.65):

$$\begin{aligned}\sum_{\langle\langle i,j\rangle\rangle}\nu_{ij}\langle\hat{A}_{ij}\rangle^*\hat{b}_{i\downarrow}\hat{b}_{j\uparrow} &= \Delta_2\sum_{\vec{k}}\sum_{l=1}^{6}e^{-i\vec{k}\cdot\vec{\epsilon}_l}\Big[\hat{b}_{\vec{k}\downarrow A}\hat{b}_{-\vec{k}\uparrow A} - \hat{b}_{\vec{k}\downarrow B}\hat{b}_{-\vec{k}\uparrow B}\Big] \\ &= \Delta_2\sum_{\vec{k}}\sum_{l=1}^{6}e^{i\vec{k}\cdot\vec{\epsilon}_l}\Big[\hat{b}_{\vec{k}\uparrow A}\hat{b}_{-\vec{k}\downarrow A} - \hat{b}_{\vec{k}\uparrow B}\hat{b}_{-\vec{k}\downarrow B}\Big],\end{aligned} \tag{7.66}$$

which means that (7.64) may be written as

$$\sum_{\langle\langle i,j\rangle\rangle}\nu_{ij}\langle\hat{A}_{ij}\rangle^*\hat{E}_{ij} = 2\Delta_2\sum_{\vec{k}}\zeta_{\vec{k}}\Big[\hat{b}_{\vec{k}\uparrow A}\hat{b}_{-\vec{k}\downarrow A} - \hat{b}_{\vec{k}\uparrow B}\hat{b}_{-\vec{k}\downarrow B}\Big], \tag{7.67}$$

with $\zeta_{\vec{k}}$ defined in (7.61). The transformation of the third operator term in (7.63) can immediately be identified by comparison to (7.56) and the definition of the mean field parameters

in Table 1:

$$\begin{aligned}\sum_{\langle\langle i,j\rangle\rangle}\nu_{ij}\langle\hat{E}_{ij}\rangle^*\hat{A}_{ij} &= \Delta_3\sum_{\langle\langle i,j\rangle\rangle}\nu_{ij}\Big[\hat{b}_{i\uparrow}\hat{b}_{j\downarrow}-\hat{b}_{i\downarrow}\hat{b}_{j\uparrow}\Big]\\ &= -2i\Delta_3\sum_{\vec{k}}\xi_{\vec{k}}\Big[\hat{b}_{\vec{k}\uparrow A}\hat{b}_{-\vec{k}\downarrow A}-\hat{b}_{\vec{k}\uparrow B}\hat{b}_{-\vec{k}\downarrow B}\Big],\end{aligned} \tag{7.68}$$

with $\xi_{\vec{k}}$ defined in (7.57). We also consider the constant term in the mean field DMI Hamiltonian:

$$\begin{aligned}&\sum_{\langle\langle i,j\rangle\rangle}\nu_{ij}\Big[\langle\hat{E}_{ij}\rangle^*\langle\hat{A}_{ij}\rangle-\langle\hat{A}_{ij}\rangle^*\langle\hat{E}_{ij}\rangle\Big]\\ &=\sum_{\vec{r}}\sum_{l=1}^{6}\sum_{X}\nu_l^{(X)}\Big[\langle\hat{E}_{\vec{r},\vec{r}+\vec{\epsilon}_l}\rangle^*\langle\hat{A}_{\vec{r},\vec{r}+\vec{\epsilon}_l}\rangle-\langle\hat{A}_{\vec{r},\vec{r}+\vec{\epsilon}_l}\rangle^*\hat{E}_{\vec{r},\vec{r}+\vec{\epsilon}_l}\rangle\Big]\\ &=\sum_{\vec{r}}\sum_{l=1}^{6}\Big[\nu_l^{(A)}(\nu_l^{(A)}\Delta_2\Delta_3-\nu_l^{(A)}\Delta_2\Delta_3)+\nu_l^{(B)}(\nu_l^{(B)}\Delta_2\Delta_3-\nu_l^{(B)}\Delta_2\Delta_3)\Big]\\ &=0.\end{aligned} \tag{7.69}$$

Since the fourth and second operator terms in (7.63) are the conjugates of (7.67) and (7.68), respectively, the next nearest neighbor mean field DMI Hamiltonian may be written

$$\begin{aligned}\hat{H}_{D_\perp(\mathrm{MF})} &= \hbar^2\sum_{\vec{k}}\Big[\big(-\frac{D_\perp\Delta_3}{4}\xi_{\vec{k}}+i\frac{D_\perp\Delta_2}{4}\zeta_{\vec{k}}\big)\big(\hat{b}_{\vec{k}\uparrow A}\hat{b}_{-\vec{k}\downarrow A}-\hat{b}_{\vec{k}\uparrow B}\hat{b}_{-\vec{k}\downarrow B}\big)\\ &\qquad+\big(-\frac{D_\perp\Delta_3}{4}\xi_{\vec{k}}-i\frac{D_\perp\Delta_2}{4}\zeta_{\vec{k}}\big)\big(\hat{b}^\dagger_{\vec{k}\uparrow A}\hat{b}^\dagger_{-\vec{k}\downarrow A}-\hat{b}^\dagger_{\vec{k}\uparrow B}\hat{b}^\dagger_{-\vec{k}\downarrow B}\big)\Big]\\ &= -\hbar^2\sum_{\vec{k}}\Big[\tau^*_{\vec{k}}\big(\hat{b}_{\vec{k}\uparrow A}\hat{b}_{-\vec{k}\downarrow A}-\hat{b}_{\vec{k}\uparrow B}\hat{b}_{-\vec{k}\downarrow B}\big)+\tau_{\vec{k}}\big(\hat{b}^\dagger_{\vec{k}\uparrow A}\hat{b}^\dagger_{-\vec{k}\downarrow A}-\hat{b}^\dagger_{\vec{k}\uparrow B}\hat{b}^\dagger_{-\vec{k}\downarrow B}\big)\Big]\\ &= \hbar^2\sum_{\vec{k}}\left[\hat{\mathbf{b}}^\dagger_{\vec{k}}\begin{bmatrix}0&0&-\tau_{\vec{k}}&0\\0&0&0&\tau_{\vec{k}}\\-\tau^*_{\vec{k}}&0&0&0\\0&\tau^*_{\vec{k}}&0&0\end{bmatrix}\hat{\mathbf{b}}_{\vec{k}}\right],\end{aligned} \tag{7.70}$$

where we defined

$$\tau_{\vec{k}}=\frac{D_\perp\Delta_3}{4}\xi_{\vec{k}}+i\frac{D_\perp\Delta_2}{4}\zeta_{\vec{k}}. \tag{7.71}$$

We summarize the findings of this Section in the following Proposition:

Proposition 7.1 The Kane-Mele-Hubbard Hamiltonian $\hat{H}$ treated with Schwinger bo-

son zero-flux mean field theory is given in matrix form as

$$\hat{H}_{(\mathrm{MF})} = \hbar^2 \sum_{\vec{k}} \Bigg[\hat{\mathbf{b}}^\dagger_{\vec{k}} \begin{bmatrix} \mu - B/2\hbar & 0 & -\psi_{\vec{k}} - \tau_{\vec{k}} & \tilde{\eta}_{\vec{k}} \\ 0 & \mu - B/2\hbar & -\tilde{\eta}^*_{\vec{k}} & -\psi_{\vec{k}} + \tau_{\vec{k}} \\ -\psi^*_{\vec{k}} - \tau^*_{\vec{k}} & -\tilde{\eta}_{\vec{k}} & \mu + B/2\hbar & 0 \\ \tilde{\eta}^*_{\vec{k}} & -\psi^*_{\vec{k}} + \tau^*_{\vec{k}} & 0 & \mu + B/2\hbar \end{bmatrix} \hat{\mathbf{b}}_{\vec{k}} + \frac{3}{2} J_1 \Delta_1^2 + 3 J_2 \Delta_2^2 + 3\Gamma \Delta_3^2 - (2 + 2\kappa)\mu \Bigg], \tag{7.72}$$

where

$$\hat{\mathbf{b}}_{\vec{k}} = \begin{bmatrix} \hat{b}_{\vec{k}\uparrow A} & \hat{b}_{\vec{k}\uparrow B} & \hat{b}^\dagger_{-\vec{k}\downarrow A} & \hat{b}^\dagger_{-\vec{k}\downarrow B} \end{bmatrix}^{\mathsf{T}} \tag{7.73}$$

is a bosonic operator vector. The matrix elements are given by

$$\tilde{\eta}_{\vec{k}} = \frac{J_1 \Delta_1}{2} \eta_{\vec{k}}; \quad \psi_{\vec{k}} = \frac{\Gamma \Delta_3}{2} \zeta_{\vec{k}} + i \frac{J_2 \Delta_2}{2} \xi_{\vec{k}}; \quad \tau_{\vec{k}} = \frac{D_\perp \Delta_3}{4} \xi_{\vec{k}} + i \frac{D_\perp \Delta_2}{4} \zeta_{\vec{k}}, \tag{7.74}$$

and

$$\eta_{\vec{k}} = e^{-ik_y a} + 2 e^{ik_y a/2} \cos\left(\frac{\sqrt{3} a}{2} k_x \right), \tag{7.75}$$

$$\xi_{\vec{k}} = 4 \sin\left(\frac{\sqrt{3} a}{2} k_x \right) \left[\cos\left(\frac{3a}{2} k_y \right) - \cos\left(\frac{\sqrt{3} a}{2} k_x \right) \right], \tag{7.76}$$

$$\zeta_{\vec{k}} = 2 \left[2 \cos\left(\frac{\sqrt{3} a}{2} k_x \right) \cos\left(\frac{3a}{2} k_y \right) + \cos\left(\sqrt{3} a k_x \right) \right]. \tag{7.77}$$

Before we diagonalize the Hamiltonian matrix

$$\mathbb{H}(\vec{k}) = \begin{bmatrix} \mu - B/2\hbar & 0 & -\psi_{\vec{k}} - \tau_{\vec{k}} & \tilde{\eta}_{\vec{k}} \\ 0 & \mu - B/2\hbar & -\tilde{\eta}^*_{\vec{k}} & -\psi_{\vec{k}} + \tau_{\vec{k}} \\ -\psi^*_{\vec{k}} - \tau^*_{\vec{k}} & -\tilde{\eta}_{\vec{k}} & \mu + B/2\hbar & 0 \\ \tilde{\eta}^*_{\vec{k}} & -\psi^*_{\vec{k}} + \tau^*_{\vec{k}} & 0 & \mu + B/2\hbar \end{bmatrix}, \tag{7.78}$$

we investigate whether it satisfies the particle-hole symmetry (6.65). From (7.50), (7.57) and (7.61), it is clear that

$$\tilde{\eta}_{-\vec{k}} = \tilde{\eta}^*_{\vec{k}}; \quad \psi_{-\vec{k}} = \psi^*_{\vec{k}}; \quad \tau_{-\vec{k}} = -\tau^*_{\vec{k}}. \tag{7.79}$$

It is easily shown that when σ_1 multiplies a 4×4 matrix from the left, it exchanges rows $1 \leftrightarrow 3$ and $2 \leftrightarrow 4$; when σ_1 multiplies the matrix from the right, it exchanges columns $1 \leftrightarrow 3$

and $2 \leftrightarrow 4$. Using this along with (7.79), we find

$$\sigma_1 \mathbb{H}^{\mathsf{T}}(-\vec{k})\sigma_1 = \begin{bmatrix} \mu + B/2\hbar & 0 & -\psi_{\vec{k}}^* + \tau_{\vec{k}}^* & -\tilde{\eta}_{\vec{k}} \\ 0 & \mu + B/2\hbar & \tilde{\eta}_{\vec{k}}^* & -\psi_{\vec{k}}^* - \tau_{\vec{k}}^* \\ -\psi_{\vec{k}} + \tau_{\vec{k}} & \tilde{\eta}_{\vec{k}} & \mu - B/2\hbar & 0 \\ -\tilde{\eta}_{\vec{k}}^* & -\psi_{\vec{k}} - \tau_{\vec{k}} & 0 & \mu - B/2\hbar \end{bmatrix} \neq \mathbb{H}(\vec{k}), \tag{7.80}$$

so the Hamiltonian does not have the particle-hole symmetry (6.65).

7.7 Diagonalization of the Hamiltonian Matrix

We can now use Theorem 2.2 to diagonalize the Hamiltonian matrix (7.78). First, we have to calculate its eigenvalues to ensure that it is positive definite. Using symbolic computation we find that the four eigenvalues are

$$\tilde{\lambda}_{\vec{k}}^{(i)} = \mu(\pm)_1 \sqrt{\left(\frac{B}{2\hbar}\right)^2 + |\tilde{\eta}_{\vec{k}}|^2 + |\psi_{\vec{k}}|^2 + |\tau_{\vec{k}}|^2 (\pm)_2 \sqrt{\Xi_{\vec{k}}}}, \tag{7.81}$$

where

$$\begin{aligned} \Xi_{\vec{k}} &= |\tilde{\eta}_{\vec{k}}|^2 \big[\tau_{\vec{k}}^2 + 2|\tau_{\vec{k}}|^2 + (\tau_{\vec{k}}^*)^2 - \psi_{\vec{k}}^2 + 2|\psi_{\vec{k}}|^2 - (\psi_{\vec{k}}^*)^2\big] + \big[\psi_{\vec{k}}\tau_{\vec{k}}^* + \psi_{\vec{k}}^*\tau_{\vec{k}}\big]^2 \\ &= |\tilde{\eta}_{\vec{k}}|^2 \big[(\tau_{\vec{k}} + \tau_{\vec{k}}^*)^2 - (\psi_{\vec{k}} - \psi_{\vec{k}}^*)^2\big] + 4\mathrm{Re}\{\psi_{\vec{k}}\tau_{\vec{k}}^*\}^2 \\ &= |\tilde{\eta}_{\vec{k}}|^2 \big[4\mathrm{Re}\{\tau_{\vec{k}}\}^2 + 4\mathrm{Im}\{\psi_{\vec{k}}\}^2\big] + 4\mathrm{Re}\{\psi_{\vec{k}}\tau_{\vec{k}}^*\}^2. \end{aligned} \tag{7.82}$$

Since

$$\begin{aligned} \psi_{\vec{k}}\tau_{\vec{k}}^* &= \big(\mathrm{Re}\{\psi_{\vec{k}}\} + i\mathrm{Im}\{\psi_{\vec{k}}\}\big)\big(\mathrm{Re}\{\tau_{\vec{k}}\} - i\mathrm{Im}\{\tau_{\vec{k}}\}\big) \\ &= \mathrm{Re}\{\psi_{\vec{k}}\}\mathrm{Re}\{\tau_{\vec{k}}\} - i\mathrm{Re}\{\psi_{\vec{k}}\}\mathrm{Im}\{\tau_{\vec{k}}\} + i\mathrm{Im}\{\psi_{\vec{k}}\}\mathrm{Re}\{\tau_{\vec{k}}\} + \mathrm{Im}\{\psi_{\vec{k}}\}\mathrm{Im}\{\tau_{\vec{k}}\}, \end{aligned} \tag{7.83}$$

we have

$$\Xi_{\vec{k}} = 4|\tilde{\eta}_{\vec{k}}|^2 \big[\mathrm{Re}\{\tau_{\vec{k}}\}^2 + \mathrm{Im}\{\psi_{\vec{k}}\}^2\big] + 4\big[\mathrm{Re}\{\psi_{\vec{k}}\}\mathrm{Re}\{\tau_{\vec{k}}\} + \mathrm{Im}\{\psi_{\vec{k}}\}\mathrm{Im}\{\tau_{\vec{k}}\}\big]^2. \tag{7.84}$$

Since $\Xi_{\vec{k}}$ is real and cannot be negative, we must impose the following constraints to ensure the positive definiteness of $\mathbb{H}(\vec{k})$:

$$\begin{aligned} \mu > &\sqrt{\left(\frac{B}{2\hbar}\right)^2 + |\tilde{\eta}_{\vec{k}}|^2 + |\psi_{\vec{k}}|^2 + |\tau_{\vec{k}}|^2 + \sqrt{\Xi_{\vec{k}}}}, \\ &\left(\frac{B}{2\hbar}\right)^2 + |\tilde{\eta}_{\vec{k}}|^2 + |\psi_{\vec{k}}|^2 + |\tau_{\vec{k}}|^2 \geq \sqrt{\Xi_{\vec{k}}}. \end{aligned} \tag{7.85}$$

Assuming that these constraints are satisfied, we can use symbolic computation to find the eigenvalues of the dynamical matrix $\mathbb{K} = \sigma_3 \mathbb{H}(\vec{k})$, which give us the energy dispersions:

$$\lambda_{\vec{k}}^{(i)} = (\pm)_1 \sqrt{\mu^2 - |\tilde{\eta}_{\vec{k}}|^2 - |\psi_{\vec{k}}|^2 - |\tau_{\vec{k}}|^2 (\pm)_2 \sqrt{\Xi_{\vec{k}}}} - \frac{B}{2\hbar}. \tag{7.86}$$

It is convenient to define

$$E_{\vec{k}}^{\pm} = \sqrt{\mu^2 - |\tilde{\eta}_{\vec{k}}|^2 - |\psi_{\vec{k}}|^2 - |\tau_{\vec{k}}|^2 \pm \sqrt{\Xi_{\vec{k}}}}. \tag{7.87}$$

By Theorem 2.2, the energy dispersions are positive, which imposes the additional constraint

$$E_{\vec{k}}^{-} > \frac{B}{2\hbar} \tag{7.88}$$

on the system. The diagonalized mean field Hamiltonian is given by

$$\begin{aligned}\hat{H}_{(\mathrm{MF})} = \hbar^2 \sum_{\vec{k}} \Big[& E_{\vec{k}}^{(1)} \hat{\beta}_{\vec{k}\uparrow A}^{\dagger} \hat{\beta}_{\vec{k}\uparrow A} + E_{\vec{k}}^{(2)} \hat{\beta}_{\vec{k}\uparrow B}^{\dagger} \hat{\beta}_{\vec{k}\uparrow B} + E_{\vec{k}}^{(3)} \hat{\beta}_{-\vec{k}\downarrow A}^{\dagger} \hat{\beta}_{-\vec{k}\downarrow A} + E_{\vec{k}}^{(4)} \hat{\beta}_{-\vec{k}\downarrow B}^{\dagger} \hat{\beta}_{-\vec{k}\downarrow B} \\ & + E_{\vec{k}}^{+} + E_{\vec{k}}^{-} + \frac{3}{2} J_1 \Delta_1^2 + 3 J_2 \Delta_2^2 + 3\Gamma \Delta_3^2 - (2 + 2\kappa)\mu \Big],\end{aligned} \tag{7.89}$$

where

$$E_{\vec{k}}^{(1)} = E_{\vec{k}}^{-} - \frac{B}{2\hbar}; \quad E_{\vec{k}}^{(2)} = E_{\vec{k}}^{+} - \frac{B}{2\hbar}; \quad E_{\vec{k}}^{(3)} = E_{\vec{k}}^{-} + \frac{B}{2\hbar}; \quad E_{\vec{k}}^{(4)} = E_{\vec{k}}^{+} + \frac{B}{2\hbar}. \tag{7.90}$$

We have again dropped a constant term $B/\hbar$, which stems from $E_{\vec{k}}^{(3)} + E_{\vec{k}}^{(4)}$, in (7.89). Evidently, there are no β-excitations in the ground state of the system, which means the ground state energy is given by

$$E_{0(\mathrm{MF})} = \hbar^2 \sum_{\vec{k}} \Big[E_{\vec{k}}^{+} + E_{\vec{k}}^{-} + \frac{3}{2} J_1 \Delta_1^2 + 3 J_2 \Delta_2^2 + 3\Gamma \Delta_3^2 - (2 + 2\kappa)\mu \Big]. \tag{7.91}$$

We present the findings of this Section in the following Proposition:

Proposition 7.2 The ground state energy of the Kane-Mele-Hubbard Hamiltonian $\hat{H}$ treated with Schwinger boson zero-flux mean field theory is given by

$$E_{0(\mathrm{MF})} = \hbar^2 \sum_{\vec{k}} \Big[E_{\vec{k}}^{+} + E_{\vec{k}}^{-} + \frac{3}{2} J_1 \Delta_1^2 + 3 J_2 \Delta_2^2 + 3\Gamma \Delta_3^2 - (2 + 2\kappa)\mu \Big], \tag{7.92}$$

where

$$E_{\vec{k}}^{\pm} = \sqrt{\mu^2 - |\tilde{\eta}_{\vec{k}}|^2 - |\psi_{\vec{k}}|^2 - |\tau_{\vec{k}}|^2 \pm \sqrt{\Xi_{\vec{k}}}} \tag{7.93}$$

and

$$\Xi_{\vec{k}} = 4|\tilde{\eta}_{\vec{k}}|^2 \big[\mathrm{Re}\{\tau_{\vec{k}}\}^2 + \mathrm{Im}\{\psi_{\vec{k}}\}^2 \big] + 4 \big[\mathrm{Re}\{\psi_{\vec{k}}\}\mathrm{Re}\{\tau_{\vec{k}}\} + \mathrm{Im}\{\psi_{\vec{k}}\}\mathrm{Im}\{\tau_{\vec{k}}\} \big]^2. \tag{7.94}$$

7.8 The Canonical Transformation Matrix

Having obtained the spectrum of the extended Kane-Mele-Hubbard Hamiltonian in the previous Section, we can readily identify the canonical transformation matrix $\mathbb{T}_{\vec{k}}$ that diagonalizes it. The column vectors comprising $\mathbb{T}_{\vec{k}}$ are found by solving (2.47) for the eigenvectors

of the dynamical matrix $\mathbb{K}$, whose eigenvalues are $E^{(1)}_{\vec{k}}$, $E^{(2)}_{\vec{k}}$, $-E^{(3)}_{\vec{k}}$ and $-E^{(4)}_{\vec{k}}$. This is most conveniently done using symbolic computation. Defining

$$\Lambda_{\vec{k}} = |\tau_{\vec{k}}|^2 + |\psi_{\vec{k}}|^2 + (\tau^*_{\vec{k}})^2 - (\psi^*_{\vec{k}})^2, \tag{7.95}$$

$$\Sigma^{\pm}_{\vec{k}} = 2|\tilde{\eta}_{\vec{k}}|^2(\mathrm{Re}\{\tau_{\vec{k}}\} - i\mathrm{Im}\{\psi_{\vec{k}}\}) + (\tau^*_{\vec{k}} + \psi^*_{\vec{k}})(2\mathrm{Re}\{\psi_{\vec{k}}\tau^*_{\vec{k}}\} \pm \sqrt{\Xi_{\vec{k}}}), \tag{7.96}$$

for brevity of notation, the unnormalized eigenvectors of $\mathbb{K}$ are given by

$$\tilde{\mathbb{T}}^{(1)}_{\vec{k}} = \begin{bmatrix} -\dfrac{(\mu + E^-_{\vec{k}})(2\mathrm{Re}\{\psi_{\vec{k}}\tau^*_{\vec{k}}\} + \sqrt{\Xi_{\vec{k}}})}{\tilde{\eta}^*_{\vec{k}}(\sqrt{\Xi_{\vec{k}}} + \Lambda_{\vec{k}})} \\ -\dfrac{2(\mu + E^-_{\vec{k}})(\mathrm{Re}\{\tau_{\vec{k}}\} - i\mathrm{Im}\{\psi_{\vec{k}}\})}{\sqrt{\Xi_{\vec{k}}} + \Lambda_{\vec{k}}} \\ -\dfrac{\Sigma^+_{\vec{k}}}{\tilde{\eta}^*_{\vec{k}}(\sqrt{\Xi_{\vec{k}}} + \Lambda_{\vec{k}})} \\ 1 \end{bmatrix}; \quad \tilde{\mathbb{T}}^{(2)}_{\vec{k}} = \begin{bmatrix} \dfrac{(\mu + E^+_{\vec{k}})(2\mathrm{Re}\{\psi_{\vec{k}}\tau^*_{\vec{k}}\} - \sqrt{\Xi_{\vec{k}}})}{\tilde{\eta}^*_{\vec{k}}(\sqrt{\Xi_{\vec{k}}} - \Lambda_{\vec{k}})} \\ \dfrac{2(\mu + E^+_{\vec{k}})(\mathrm{Re}\{\tau_{\vec{k}}\} - i\mathrm{Im}\{\psi_{\vec{k}}\})}{\sqrt{\Xi_{\vec{k}}} - \Lambda_{\vec{k}}} \\ \dfrac{\Sigma^-_{\vec{k}}}{\tilde{\eta}^*_{\vec{k}}(\sqrt{\Xi_{\vec{k}}} - \Lambda_{\vec{k}})} \\ 1 \end{bmatrix};$$

$$\tilde{\mathbb{T}}^{(3)}_{\vec{k}} = \begin{bmatrix} -\dfrac{(\mu - E^-_{\vec{k}})(2\mathrm{Re}\{\psi_{\vec{k}}\tau^*_{\vec{k}}\} + \sqrt{\Xi_{\vec{k}}})}{\tilde{\eta}^*_{\vec{k}}(\sqrt{\Xi_{\vec{k}}} + \Lambda_{\vec{k}})} \\ -\dfrac{2(\mu - E^-_{\vec{k}})(\mathrm{Re}\{\tau_{\vec{k}}\} - i\mathrm{Im}\{\psi_{\vec{k}}\})}{\sqrt{\Xi_{\vec{k}}} + \Lambda_{\vec{k}}} \\ -\dfrac{\Sigma^+_{\vec{k}}}{\tilde{\eta}^*_{\vec{k}}(\sqrt{\Xi_{\vec{k}}} + \Lambda_{\vec{k}})} \\ 1 \end{bmatrix}; \quad \tilde{\mathbb{T}}^{(4)}_{\vec{k}} = \begin{bmatrix} \dfrac{(\mu - E^+_{\vec{k}})(2\mathrm{Re}\{\psi_{\vec{k}}\tau^*_{\vec{k}}\} - \sqrt{\Xi_{\vec{k}}})}{\tilde{\eta}^*_{\vec{k}}(\sqrt{\Xi_{\vec{k}}} - \Lambda_{\vec{k}})} \\ \dfrac{2(\mu - E^+_{\vec{k}})(\mathrm{Re}\{\tau_{\vec{k}}\} - i\mathrm{Im}\{\psi_{\vec{k}}\})}{\sqrt{\Xi_{\vec{k}}} - \Lambda_{\vec{k}}} \\ \dfrac{\Sigma^-_{\vec{k}}}{\tilde{\eta}^*_{\vec{k}}(\sqrt{\Xi_{\vec{k}}} - \Lambda_{\vec{k}})} \\ 1 \end{bmatrix}.$$

We cannot construct $\mathbb{T}_{\vec{k}}$ from the unnormalized eigenvectors $\tilde{\mathbb{T}}^{(i)}_{\vec{k}}$, as the resulting matrix would not be pseudo-unitary. As discussed in Section 2.2.2, pseudo-unitarity is required in order for the diagonalization to preserve the bosonic commutation relations. The normalized eigenvectors $\mathbb{T}^{(i)}_{\vec{k}}$ are obtained by applying the normalization condition (2.42) to the eigenvectors $\tilde{\mathbb{T}}^{(i)}_{\vec{k}}$. The canonical transformation matrix can then be written compactly as

$$\mathbb{T}_{\vec{k}} = \begin{bmatrix} \mathbb{T}^{(1)}_{\vec{k}} & \mathbb{T}^{(2)}_{\vec{k}} & \mathbb{T}^{(3)}_{\vec{k}} & \mathbb{T}^{(4)}_{\vec{k}} \end{bmatrix}. \tag{7.97}$$

The eigenvectors $\mathbb{T}^{(i)}_{\vec{k}}$ generally form an orthonormal set, in the pseudo-unitary sense. However, something interesting happens if $\Xi_{\vec{k}} = 0$ for all $\vec{k}$. We then have $E^+_{\vec{k}} = E^-_{\vec{k}}$ and $\Sigma^+_{\vec{k}} = \Sigma^-_{\vec{k}}$, and by examining the unnormalized eigenvectors presented above, we see that $\mathbb{T}^{(1)}_{\vec{k}} = \mathbb{T}^{(2)}_{\vec{k}}$ and $\mathbb{T}^{(3)}_{\vec{k}} = \mathbb{T}^{(4)}_{\vec{k}}$. In the absence of a magnetic field, we therefore have a single, two-fold degenerate band with corresponding eigenvectors $\mathbb{T}^{(1)}_{\vec{k}}$ and $\mathbb{T}^{(3)}_{\vec{k}}$. A finite magnetic field will lift this degeneracy, giving rise to a two-band excitation spectrum. This does not alter the eigenvectors, since the magnetic field appears as a scalar matrix in $\mathbb{K}$. We note that $\Xi_{\vec{k}}$ is always zero for certain points $\vec{k}$ in the Brillouin zone, since it is defined in terms of (7.50), (7.57) and (7.61).

8 Numerical Results and Discussion

In order to study the physics of the ground state and the low-lying excitations of the Kane-Mele-Hubbard model, we must first solve the self-consistency equations for the mean field parameters. We will consider our system to be at zero temperature, where the free energy is the ground state energy of the system. The self-consistency equations are then given by

$$\begin{aligned}
\sum_{\vec{k}} \left(\frac{\partial E^+_{\vec{k}}}{\partial \mu} + \frac{\partial E^-_{\vec{k}}}{\partial \mu} \right) &= (2 + 2\kappa)N, \\
\sum_{\vec{k}} \left(\frac{\partial E^+_{\vec{k}}}{\partial \Delta_1} + \frac{\partial E^-_{\vec{k}}}{\partial \Delta_1} \right) &= -3NJ_1\Delta_1, \\
\sum_{\vec{k}} \left(\frac{\partial E^+_{\vec{k}}}{\partial \Delta_2} + \frac{\partial E^-_{\vec{k}}}{\partial \Delta_2} \right) &= -6NJ_2\Delta_2, \\
\sum_{\vec{k}} \left(\frac{\partial E^+_{\vec{k}}}{\partial \Delta_3} + \frac{\partial E^-_{\vec{k}}}{\partial \Delta_3} \right) &= -6N\Gamma\Delta_3.
\end{aligned} \tag{8.1}$$

We solve these numerically using an iterative Levenberg-Marquardt Monte Carlo approach. We first generate sets of four uniformly distributed random initial guesses for the mean field parameters. Each of these sets are passed to the Levenberg-Marquardt algorithm, and the solution obtained is used as an initial guess for the algorithm in the next step. This iterative procedure is repeated until the solutions obtained in two consecutive steps converge. In each step, we require that the solution satisfies the constraints (7.85) and (7.88). If it does not, the solution is discarded and we retry with a new set of random initial guesses. When the solution has converged for all sets of initial guesses, we calculate the ground state energy (7.91) for each of them. The mean field parameters are chosen to be the solution that minimizes it. With these obtained, we can investigate the dynamical properties, phase diagrams and transport signatures of the Kane-Mele-Hubbard model.

8.1 Spinon Dispersions and Physical Phases

The Schwinger boson mean field theory predicts that the ground state of the Kane-Mele-Hubbard model exhibits three different physical phases: a commensurate magnetic Néel order phase, an incommensurate magnetic Néel order phase and a quantum spin liquid phase. For a given boson density κ, the phases are determined by the properties of the lowest dispersion branch $E^{(1)}_{\vec{k}}$, which depends on the exchange parameters J_1, J_2, Γ, $D_\perp$ and B and the mean field parameters μ, Δ_1, Δ_2 and Δ_3. The two magnetic phases are characterized by gapless Schwinger boson excitations, whereas the quantum spin liquid phase is gapped. In the commensurate phase, the spinons condense at the Γ point in the first Brillouin zone. In the incommensurate phase, the condensation occurs at the K and K$'$ points. The bosons are in that case excited with a finite momentum, giving rise to a long-range magnetic order that does not have the same periodicity as the lattice [7]. Figures 7a, 7b and 7c show the two lowest spinon dispersion branches for the commensurate phase, the incommensurate phase

and the quantum spin liquid phase, respectively, along the high symmetry line $\Gamma-\mathrm{K}-\mathrm{M}-\Gamma$ in the first Brillouin zone. The left panels of these Figures show the dispersions in the absence of DMI, and the right panels show the dispersions for $D_\perp = 0.5J_1$. The magnetic field is set to $B = 0$, and a boson density of $\kappa = 0.35$ is used in all cases - we will comment on this choice in the next Section. In the absence of the magnetic field, the dispersion branches shown form the complete excitation spectrum of the Kane-Mele-Hubbard model, but they are representative for the low-lying spectrum if $B \neq 0$ as well.

As we can see, the Schwinger bosons are not exactly gapless in the two magnetic phases. We attribute this to the finite dimensionality of the lattice used in the numerical calculations. As the number of (reciprocal) lattice sites is increased, so too is the accuracy of the sums in (8.1), and the dispersion minima are lowered. However, because the computational cost of solving the self-consistency equations increases with the system size, it is unfeasible to consider the case where N is very large. Nevertheless, to ensure that our approach is as accurate as possible even for moderately large systems, it is important that the dispersion minima lie on the grid of reciprocal lattice points comprising the numerical Brillouin zone. To see why, consider the left hand side of the first self-consistency equation in (8.1):

$$\sum_{\vec{k}} \left(\frac{\partial E^+_{\vec{k}}}{\partial \mu} + \frac{\partial E^-_{\vec{k}}}{\partial \mu} \right) = \sum_{\vec{k}} \left(\frac{\mu}{E^+_{\vec{k}}} + \frac{\mu}{E^-_{\vec{k}}} \right). \tag{8.2}$$

The largest contribution will be from the point $\vec{k}$ where $E^-_{\vec{k}}$ has a minimum, and therefore we must ensure that it is included in the evaluation of the sum. From (7.11), we see that the inclusion of the Γ point is trivial. The inclusion of

$$\mathrm{K} = \left(\frac{2\pi}{3\sqrt{3}a}, \frac{2\pi}{3a} \right) \quad \text{and} \quad \mathrm{K}' = \left(-\frac{2\pi}{3\sqrt{3}a}, \frac{2\pi}{3a} \right), \tag{8.3}$$

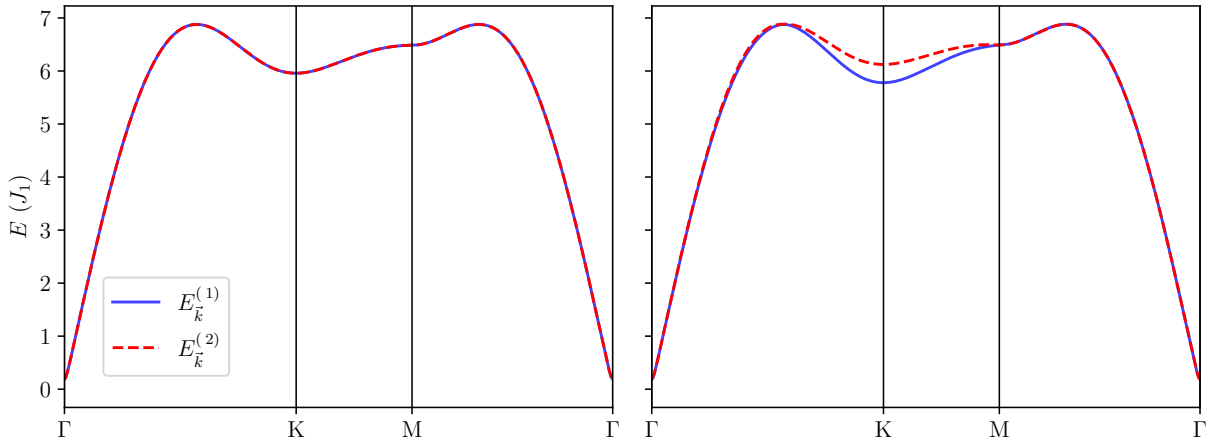
however, places a constraint on the number of lattice sites $\mathcal{N}$ in the direction of the primitive lattice vectors. Setting $\vec{k}$, as defined in (7.11), equal to K and K$'$, we find

$$(n_1, n_2) = \begin{cases} \left(\dfrac{2\mathcal{N}}{3}, \dfrac{\mathcal{N}}{3} \right) & \text{if } \vec{k} = \mathrm{K}, \\ \left(\dfrac{\mathcal{N}}{3}, \dfrac{2\mathcal{N}}{3} \right) & \text{if } \vec{k} = \mathrm{K}', \end{cases} \tag{8.4}$$

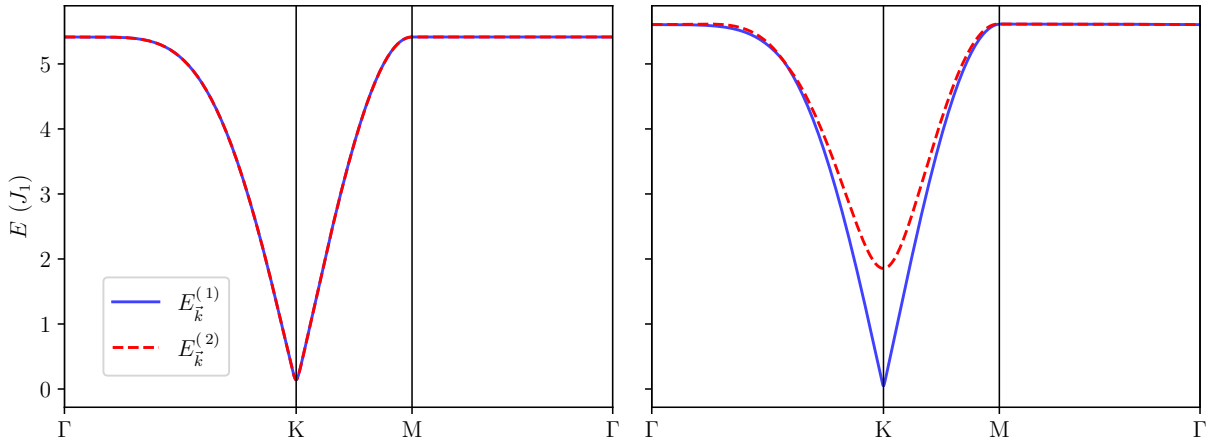
which means that $\mathcal{N}$ must be divisible by 3 in order for the numerical Brillouin zone to contain the dispersion minima. For the dispersions shown in Figures 7a, 7b and 7c, we set $\mathcal{N} = 30$, such that each of the sublattices A and B is comprised of $N = 900$ lattice sites. We consider this system size in all numerical calculations that follow.

The existence of the small, finite gap in the magnetically ordered phases calls for the definition of a transition parameter we can use to distinguish them from the quantum spin liquid phase. To this end, we define the relative gap

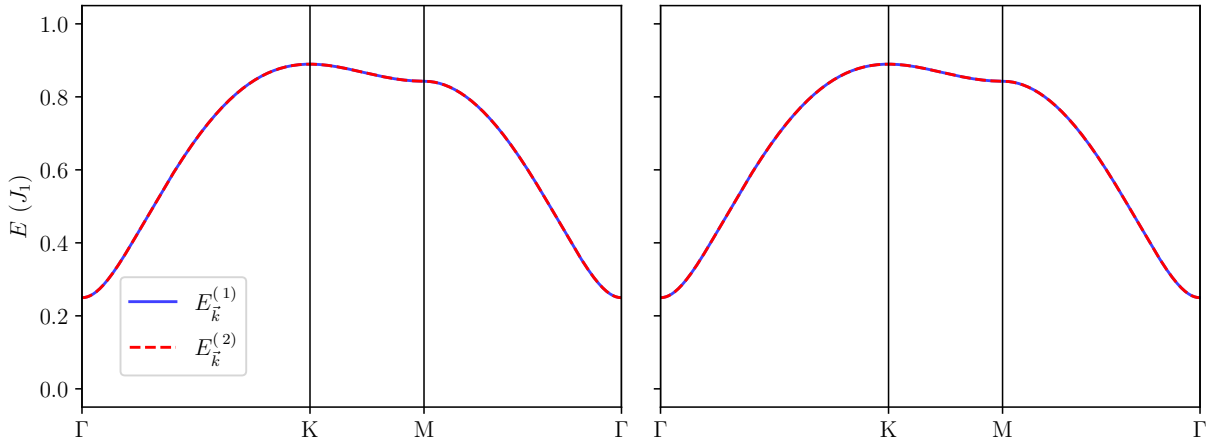
$$\delta = \frac{E^{(1)}_{\vec{k}_0}}{\mu}, \tag{8.5}$$



(a)



(b)



(c)

Figure 7: Spinon dispersions for the a) commensurate magnetic order phase, b) incommensurate magnetic order phase and c) quantum spin liquid phase. The left and right panels show the dispersions for $D_\perp = 0$ and $D_\perp = 0.5J_1$, respectively. The dispersions are obtained for a) $J_2 = 2J_1$, $\Gamma = 5J_1$, b) $J_2 = 5J_1$, $\Gamma = 2J_1$ and c) $J_2 = \Gamma = 0.5J_1$.

where $\vec{k}_0$ denotes the point where $E^{(1)}_{\vec{k}}$ attains its minimum value. If $E^{(1)}_{\vec{k}}$ has a small gap, the chemical potential μ of the Schwinger bosons is comparable to the $\vec{k}$-dependent terms in the dispersion at $\vec{k}_0$, as we can see from (7.85) and (7.87). The relative gap will therefore be correspondingly small in the magnetically ordered phases. However, if $E^{(1)}_{\vec{k}}$ has a significant gap, the chemical potential μ dominates the dispersion at $\vec{k}_0$, leading to a large relative gap.

We see from the left panels in Figures 7a, 7b and 7c that the two lowest dispersion branches are identical in the absence of DMI. From our analysis in Section 7.8, we know that $E^{(1)}_{\vec{k}}$ and $E^{(2)}_{\vec{k}}$ then correspond to the same eigenvector $\mathbb{T}^{(1)}_{\vec{k}}$, which means they represent the same band. This is also true for the two upper dispersion branches $E^{(3)}_{\vec{k}}$ and $E^{(4)}_{\vec{k}}$, which represent a single band corresponding to the eigenvector $\mathbb{T}^{(3)}_{\vec{k}}$. The magnetic field opens a gap between the two bands $E^{(1)}_{\vec{k}}$ and $E^{(3)}_{\vec{k}}$, which allows for a well-defined Berry curvature in all three phases. From the right panels, we see that the DMI gaps the dispersion branches near the K point in the two magnetic phases. As we move away from the K point, the two branches converge until they eventually overlap to form a single band, which makes the Berry curvature ill-defined. This is not the case for the quantum spin liquid phase. We can explain this by examining the numerically obtained values for the mean field parameters. In both magnetic phases, Δ_1 is zero. In the commensurate phase, the mean field parameter Δ_2 is zero, while Δ_3 takes on a finite value; in the incommensurate phase, the opposite is true. In the quantum spin liquid phase, Δ_1 is finite, while both Δ_2 and Δ_3 are zero. The DMI couples exclusively to Δ_2 and Δ_3 through (7.71), which means it is suppressed in the quantum spin liquid phase. The dynamical properties of this phase are therefore the same regardless of the presence of DMI, which allows for a well-defined Berry curvature in both cases. Strictly speaking, none of these mean field parameters are identically zero in any of the phases, but since their orders of magnitude are -10 or smaller and we are working in the mean field, we will consider them as such when we study the dynamics of the system. For comparison, the order of magnitude of the finite parameters is 1. We remark, however that a finite Δ_1 is required to couple the two sublattices A and B in an antiferromagnetic fashion, as J_1 is proportional to Δ_1 in the mean field Hamiltonian (7.72).

The values of the mean field parameters provide insight into the nature of the predicted phases. Δ_1 and Δ_2 correspond to nearest and next nearest neighbor spin singlets, respectively, and are formed by the operator $\hat{A}^\dagger_{ij}$ defined in (7.24). Δ_3 corresponds to a next nearest neighbor spin singlet formed by $\hat{E}^\dagger_{ij}$ defined in (7.28). In the magnetic phases, condensed Schwinger bosons form next nearest neighbor spin singlets, which gives rise to the long-range order parameter. Next nearest neighbor singlets do not form in the quantum spin liquid whatsoever, which fits well with the expected lack of magnetic order. Δ_1, on the other hand, is found to be finite in the quantum spin liquid, which means that gapped spinons can be excited as nearest neighbor spin singlets. The quantum spin liquid therefore bears the hallmarks of a short-range resonating valence bond state, as expected from the zero-flux mean field Ansatz [4, 34, 35].

8.2 Phase Diagrams

Having established the relative gap (8.5) as a way to distinguish the physical phases of the Kane-Mele-Hubbard model, we now study the nature of the ground state using phase diagrams. Based on the findings presented in Figure 7, a natural choice for the parameter space is (J_2, Γ). Through numerical experimentation and comparison with the schematic phase diagram found by Vaezi et al. [7], we define a phase as gapless if $\delta < 0.05$. This is quite a relaxed threshold; the magnetically ordered phases shown in Figures 7a and 7b have relative gaps $\delta \sim 0.03$, whereas the quantum spin liquid phase in Figure 7c has $\delta \sim 0.28$. Different choices for the threshold value will inevitably lead to different phase diagrams. The phase diagrams presented here are therefore by no means exact, but rather aim to provide qualitative information about the phases that can exist in the model for different interaction parameters.

The phase diagrams for six different values of $D_\perp$ are shown in Figure 8. Before we discuss the physics, some comments are in order. Due to computational cost, the resolution of the diagrams is only 10×10 points. Each point corresponds to the lower left vertex of a colored square in the phase diagram. This means that the phase reported in a given square may not be representative for the square as a whole, and caution should be exercised when trying to identify phase transition lines. Furthermore, there is an element of stochasticity in the determination of the physical phases. For each set of mean field parameters, we generated 100 initial guesses. Ideally, we would have considered several thousands of initial guesses to ensure a small numerical error, but this would increase the already high computational cost.

The phase diagrams depend critically on the boson density κ. In the extreme quantum limit $\kappa \to 0$, the only phase present is the quantum spin liquid. In the classical limit $\kappa \to \infty$, the phase diagram consists purely of magnetically ordered states. From the Schwinger boson constraint (7.17), we expect $\kappa = 2S = 1$ for a spin-1/2 electron system. In the mean field, however, this is not necessarily true. In the Schwinger boson mean field study conducted by Bauer and Fjærestad [36], it is found that

$$\langle \hat{\vec{S}}^2 \rangle = \frac{3}{8}\hbar^2\kappa(\kappa+2) = \frac{3}{2}\hbar^2 S(S+1), \tag{8.6}$$

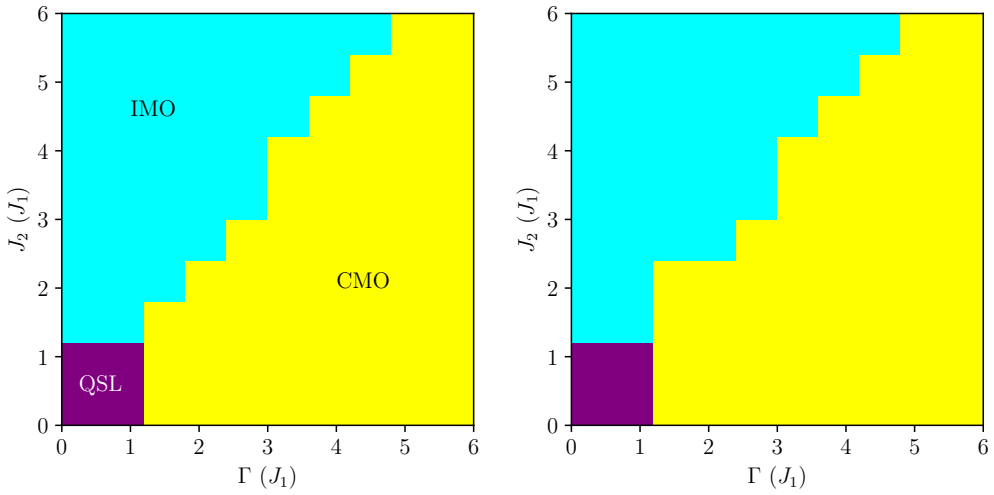
which overestimates the correct result by a factor 3/2. To obtain the correct result, the boson density must be adjusted to $\kappa = 0.73$. It is suggested by Bauer and Fjærestad [36] that the mean field theory underestimates quantum fluctuations, and that a lower value for κ can be used to compensate for this. We therefore treat κ as a continuous parameter in this thesis, choosing values that allow us to solve the self-consistency equations (8.1) with relative computational ease. The intermediate boson density $\kappa = 0.35$ is a good choice in this regard, as all three phases are present in the phase diagram. We have, however, conducted a parameter scan of κ in the absence of DMI, and found a critical value $\kappa_c \approx 0.59$ at which the quantum spin liquid phase vanishes.

Figure 8a shows the phase diagram in the absence of DMI. It agrees relatively well with the schematic phase diagram found by Vaezi et al. [7]. The main difference is that Vaezi et al. [7] predicted the quantum spin liquid to exist for $J_2 \sim J_1 - 5J_1$, whereas we find it closer to the origin. This may be due to differences in the numerical approach or the method used to distinguish the phases; no comments were made on this by Vaezi et al. [7]. Figure 8a

agrees even better with the phase diagram found for $D_\perp = 0$ by Ushakov [6], who used a transition parameter similar to the relative gap (8.5). Note that Ushakov [6] found a much lower critical boson density $\kappa_c \approx 0.3139$ than us. It is harder to compare the phase diagrams shown in Figure 8 for a finite DMI directly to those obtained by Ushakov [6], as $D_\perp$ and κ were varied simultaneously. However, we do see similar trends in our phase diagrams: As the strength of the DMI is increased, the commensurate magnetic phase is realized for lower values of Γ and higher values of J_2. At the same time, the quantum spin liquid materializes for higher values of J_2 and occupies an increasingly significant part of the phase diagram. As a result, the incommensurate magnetic phase becomes less prominent, being achievable only for high values of J_2. While we should be careful in drawing conclusions based on the phase diagrams, this does suggest that the DMI can stabilize the quantum spin liquid phase in honeycomb materials with strong next nearest neighbor spin interactions.

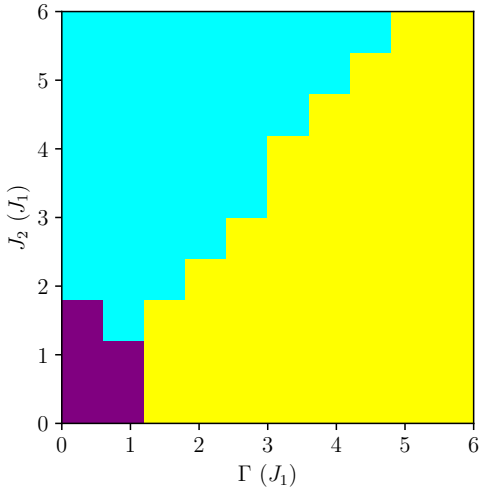
It would also be interesting to investigate the effect of the magnetic field B on the phase diagrams, as well as the combined effect of B and $D_\perp$. However, as the magnitude of the magnetic field is increased, obtaining solutions for the mean field parameters becomes very difficult in certain regions of the parameter space. We therefore chose not to pursue it in this thesis. Nevertheless, it could be an interesting avenue for further study of the phase diagrams of the Kane-Mele-Hubbard model.

We did obtain a phase diagram for $B = 0.1J_1$, which found the quantum spin liquid for $J_2 \sim J_1 - 2J_1$, $\Gamma \sim 0 - J_1$. This suggests that a weak magnetic field may contribute to increase the magnetic frustration. A strong magnetic field is nonetheless expected to induce long-range magnetic order in the system.

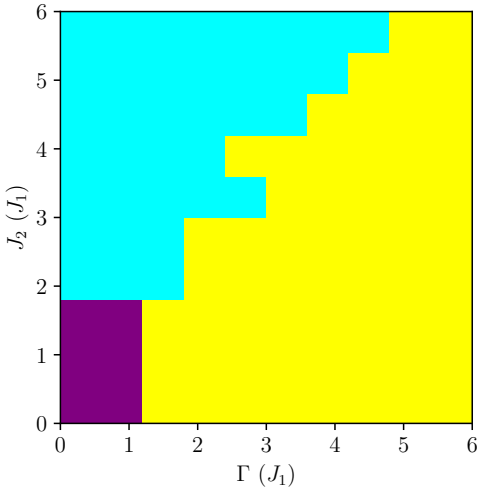


(a) $D_\perp = 0$.

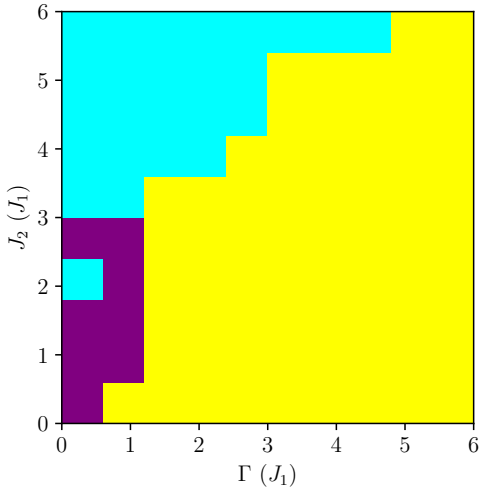
(b) $D_\perp = 0.01J_1$.



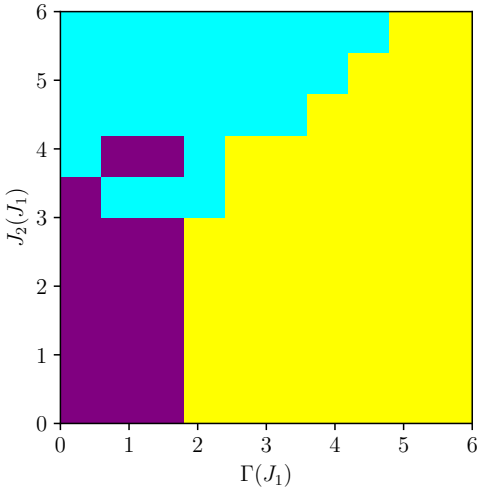
(c) $D_\perp = 0.05J_1$.



(d) $D_\perp = 0.1J_1$.



(e) $D_\perp = 0.5J_1$.



(f) $D_\perp = J_1$.

Figure 8: Phase diagrams for the Kane-Mele-Hubbard for six different values of $D_\perp$. In all cases, the boson density is $\kappa = 0.35$ and the magnetic field is $B = 0$.

8.3 Transport Signatures

While the relative gap serves as a useful theoretical tool to separate the quantum spin liquid phase from the magnetic phases, it is not suited for experimental probing. In order to detect the quantum spin liquid in an experiment, we have to identify measurable quantities that capture its distinct characteristics. Since a quantum spin liquid is expected to exhibit non-trivial topology, the topology-dependent conductivity formulas introduced in Section 6.2 are a good starting point. We now study the spin Nernst effect and thermal Hall effect in the predicted phases of the Kane-Mele-Hubbard model to see if the quantum spin liquid phase can be distinguished from the others based on its transport signatures.

We start off by examining the Berry curvatures of the three predicted phases. The Berry curvatures of the bands are computed numerically using (6.59), with the eigenvectors given by (7.97). Figure 9 shows the numerically computed Berry curvatures of the lowest

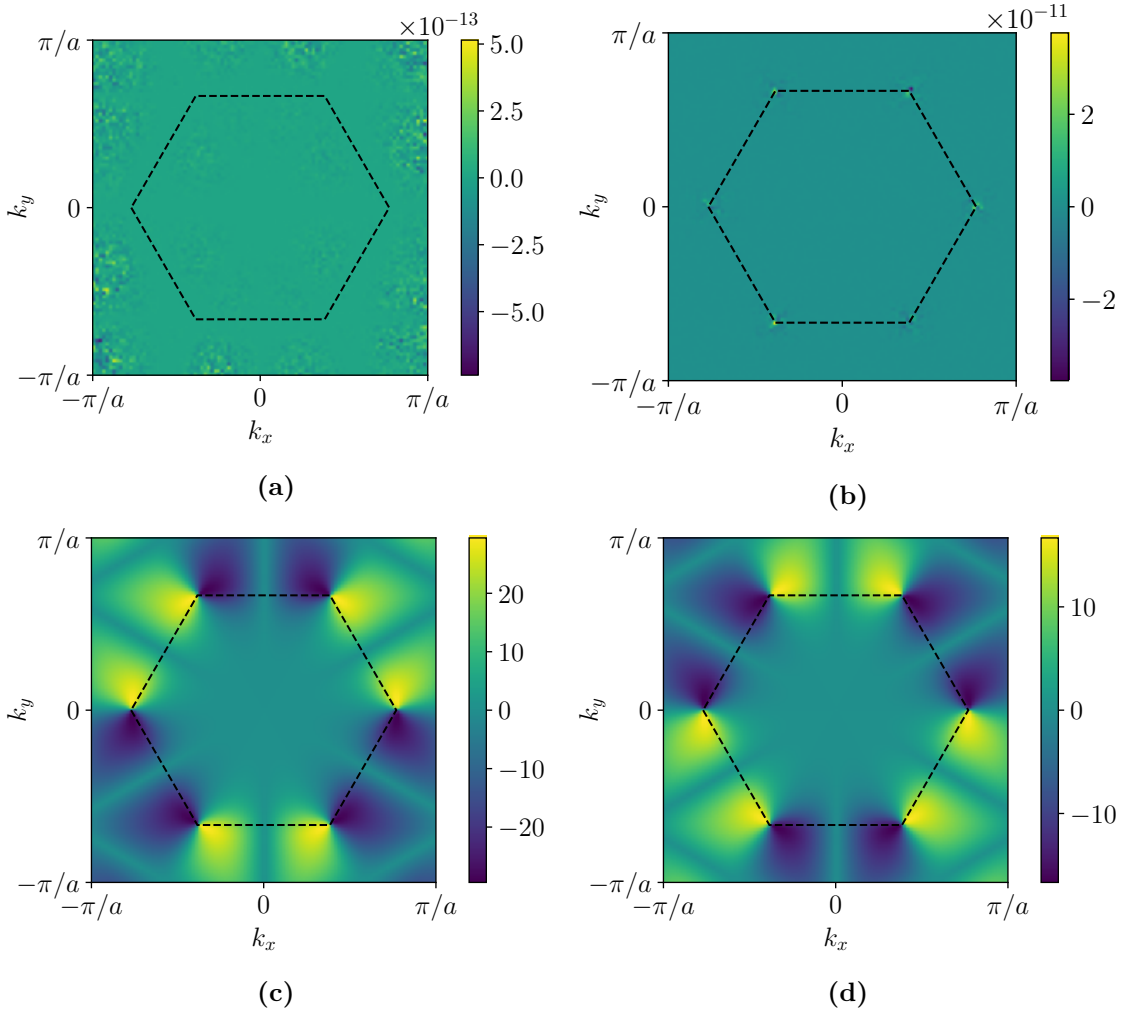


Figure 9: Berry curvatures of the lowest bands in the a) commensurate magnetic order phase, b) incommensurate magnetic order phase, c) quantum spin liquid phase, and d) quantum spin liquid phase for $D_\perp = 0.5J_1$. The applied magnetic field is $B = 0.1J_1$. The dashed hexagon outline represents the edges of the first Brillouin zone.

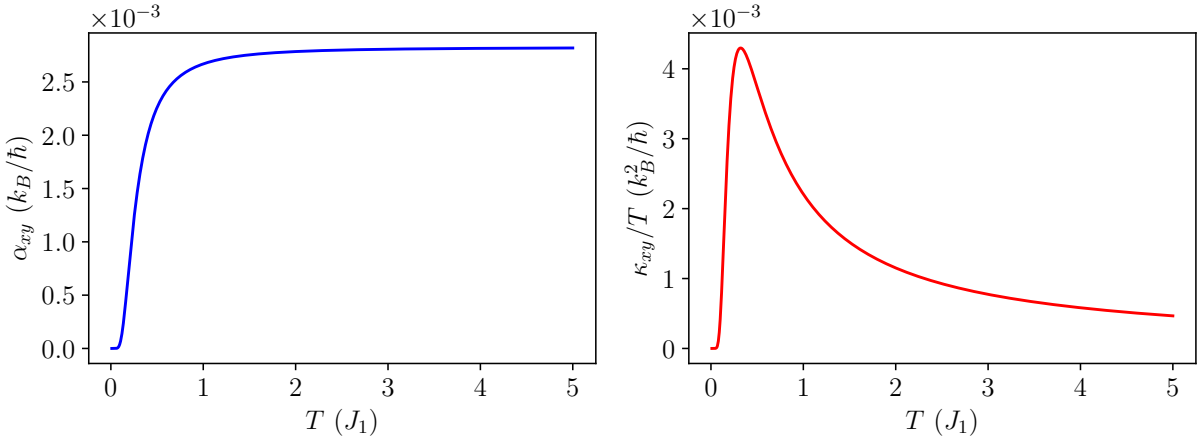
band in the cases where they are well-defined, according to Section 8.1, for a magnetic field $B = 0.1J_1$. Figures 9a, 9b and 9c show the Berry curvatures of the commensurate phase, the incommensurate phase and the quantum spin liquid phase, respectively, in the absence of DMI. Figure 9d shows the Berry curvature for the quantum spin liquid phase for a DMI strength $D_\perp = 0.5J_1$. While the two magnetic phases are seen to be topologically trivial, the Berry curvature takes on significant values near the K and K$'$ points in the quantum spin liquid phase. Even though the Chern numbers of the bands are zero, it is possible to for the quantum spin liquid to display the spin Nernst and thermal Hall effects, as the Berry curvature is weighted by $c_1(x)$ and $c_2(x)$ in the corresponding conductivity formulas (6.56) and (6.64).

We briefly comment on the apparent differences between the Berry curvatures in the presence and absence of DMI. In Section 8.1, we argued that very small mean field parameters can be considered trivial when we study the dynamics of the system. Δ_1 is found to be identical in the two cases shown in Figures 9c and 9d, and the values obtained for μ differ only at the 16th decimal. Δ_2 and Δ_3 are found to be of orders of magnitude -10 or less, but their values are different in the two cases. If we assume that such small mean field parameters can be considered zero, the observed differences between the Berry curvatures must be a purely numerical artifact. Under this assumption, the Berry curvature shown in Figure 9c is also representative for the finite DMI case, as the dynamical properties of the system are the same regardless of the presence of DMI.

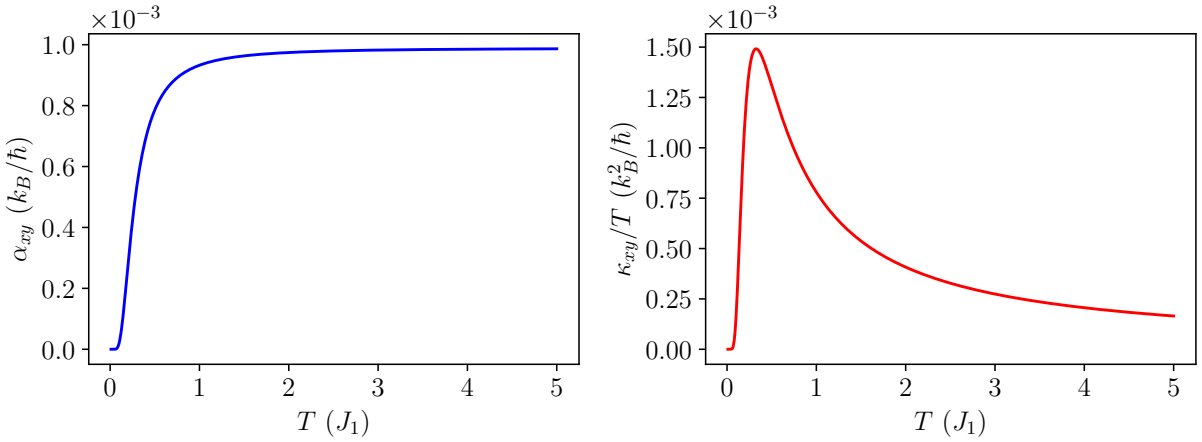
We therefore proceed to compute the conductivities (6.56) and (6.64) in the quantum spin liquid phase for a DMI strength $D_\perp = 0$. For completeness, the numerical artifacts introduced when computing the conductivities at finite DMI strengths are shown in Appendix´ C. Figures 10a, 10b and 10c show the transport signatures as a function of temperature for magnetic fields $B = 0.1J_1$, $B = 0.3J_1$ and $B = 0.5J_1$, respectively. The left panels of these Figures show the spin Nernst conductivity, while the right panels show the thermal Hall conductivity. Both effects are present, providing two different ways of probing the quantum spin liquid phase.

From Figure 10, we see that the magnitude of the response is diminished when the magnetic field strength is increased. This is in line with the interpretation of the Berry curvature as an interaction between the eigenstates of the system, which becomes weaker when the gap between the bands is increased. At high temperatures, the transport coefficients are seen to flatten out and attain temperature independent finite values. This is a result of solving the self-consistency equations (8.1) at zero temperature, and is not actual physical behavior. Our results for the conductivities are therefore only valid at low temperatures.

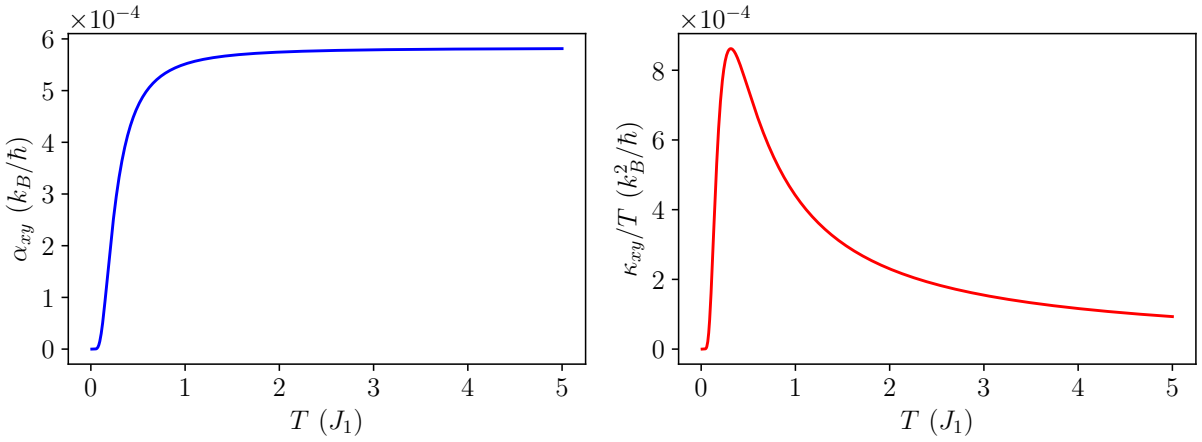
It is useful to compare the results obtained here to similar results found in the literature. Kovalev and Zyuzin [28] studied the magnon spin Nernst effect in a honeycomb antiferromagnet with nearest neighbor Heisenberg interactions and out-of-plane DMI using the Holstein-Primakoff transformation. The temperature dependence of the spin Nernst conductivity shown in Figure 10 is in good qualitative agreement with the one they report. Zhang et al. [37] studied the same model in the presence of a staggered magnetic field using Schwinger boson mean field theory. In the temperature range $T \sim 0 - 0.5J_1$, our results for the spin Nernst effect are in good quantitative agreement with theirs in the disordered state. Because they have obtained temperature dependent mean field parameters, they report a quenching of the spin Nernst conductivity at $T \approx 0.8J_1$, which our approach fails to describe.



(a)



(b)



(c)

Figure 10: Spin Nernst conductivities (left) and thermal Hall conductivities (right) in applied magnetic fields a) $B = 0.1J_1$, b) $B = 0.3J_1$ and c) $B = 0.5J_1$. The boson density is $\kappa = 0.3$.

Zhang et al. [37] also found that the thermal Hall effect vanishes due to the combined time reversal and inversion symmetry of their Hamiltonian. Since the applied magnetic field is not staggered in our model, this combined symmetry is not a symmetry of the Hamiltonian, which allows for a finite thermal Hall response.

Laurell and Fiete [38] studied the thermal Hall effect in a kagome antiferromagnet with a spin Hamiltonian identical to ours, with the addition of an in-plane DMI, using the Holstein-Primakoff transformation. Again, there is good qualitative agreement between their conductivity results and ours. The dispersions of the topological magnons carrying the thermal Hall current in their case bear very close resemblance to the spinon dispersion shown in Figure 7c. Our results also agree with those of Samajdar et al. [39], in which the Schwinger boson approach was used to study a square lattice antiferromagnet with nearest and next nearest neighbor Heisenberg interactions, with and without nearest neighbor in-plane DMI, in the presence of a magnetic field. Most importantly, our results for the thermal Hall conductivity are in qualitative agreement with experiments [40, 41]. We thus conclude that the spin Nernst and thermal Hall conductivities obtained provide ways to probe the predicted quantum spin liquid phase in the Kane-Mele-Hubbard model.

Before we conclude the thesis, some more remarks on the effect of DMI are in order. In all of the theoretical studies presented above, the DMI plays some role in the transport. The magnon spin Nernst effect reported by Kovalev and Zyuzin [28] is found to vanish in the absence of DMI, and the spin Nernst conductivity found for Schwinger bosons by Zhang et al. [37] increases with $D_\perp$ for a fixed magnetic field B. Laurell and Fiete [38] found that the thermal Hall conductivity attains its maximum value for an in-plane DMI $D_\parallel \sim 0.6J_1$, while it was found to be decreased for a finite in-plane DMI on the square lattice studied by Samajdar et al. [39]. Furthermore, Lee et al. [42] studied the kagome antiferromagnet with nearest neighbor Heisenberg interactions and found that it is possible to obtain a finite thermal Hall conductivity due to an out-of-plane DMI even in the absence of the Zeeman field.

The lacking influence of the DMI in our system can be attributed to the choice of mean field Ansatz. In the Ansatz employed in this thesis, we chose the mean field parameters $\langle \hat{A}_{\langle\langle i,j\rangle\rangle}\rangle$ and $\langle \hat{E}_{\langle\langle i,j\rangle\rangle}\rangle$ to be finite while setting $\langle \hat{B}_{\langle\langle i,j\rangle\rangle}\rangle = \langle \hat{D}_{\langle\langle i,j\rangle\rangle}\rangle = 0$. With this choice, the mean field DMI term (7.36) produces only off-diagonal contributions to the mean field Hamiltonian matrix. As we have seen, these are only able to gap the dispersion near the K and K$'$ points in the magnetic phases, and are totally suppressed in the quantum spin liquid phase. If either $\langle \hat{B}_{\langle\langle i,j\rangle\rangle}\rangle$ or $\langle \hat{D}_{\langle\langle i,j\rangle\rangle}\rangle$, or both, are considered finite, the DMI term also has diagonal contributions to the mean field Hamiltonian. This might give rise to gaps in the energy spectrum that are not present in the current Ansatz, which could influence the thermal transport for finite $D_\perp$.

9 Conclusion and Outlook

In this thesis, we have investigated the possibility of identifying quantum spin liquids based on their transport signatures. The starting point for such an investigation is a quantum spin Hamiltonian, where interactions between spins at different lattice sites are described by spin operators. The spin operator algebra makes the analytical treatment of such Hamiltonians difficult, which motivated us to represent the interactions in terms of other operators. To this end, we introduced the concepts of fermionization and bosonization, which let us represent the spin Hamiltonian in terms of fermionic and bosonic operators. These operators obey much simpler algebras, and allow us to interpret the spin interactions in terms of excitations of quasiparticles. The excitation spectra, which provide information about the dynamics of the spin system, can be obtained by diagonalizing the resulting fermionic or bosonic Hamiltonians. We presented general diagonalization procedures for such non-interacting, quadratic Hamiltonians. To cover the cases where the fermionic or bosonic Hamiltonians are interacting, we introduced a mean field approximation scheme that can be used to make quartic Hamiltonians quadratic. We demonstrated these procedures by studying the XY model using four different types of spin operator representations, each with its own advantages and disadvantages. Of particular interest was the Schwinger boson transformation, whose isotropic nature makes it well suited for studies of disordered spin systems such as quantum spin liquids.

We then set out to find a way to probe this exotic phase's response to an applied perturbation. To this end, we derived the Kubo formula for the linear response of an observable, which we used to compute a general expression for a system's current response to a perturbation described by a vector potential. We used this expression to compute the Hall conductivity for electronic insulators in great detail. The calculation showed that the topology of the system's energy spectrum, which is related to the geometric phase of its wave functions, can manifest itself in transport phenomena. Since non-trivial topology is one of the hallmarks of quantum spin liquids, we proposed the spin Nernst and thermal Hall effects as candidates for their transport signatures.

The subject of our study was the spin Hamiltonian of the strongly correlated Kane-Mele-Hubbard model on the honeycomb lattice. We used Schwinger boson mean field theory to study its dynamical properties and phase diagrams. In particular, we demonstrated that the ground state of the model can be a quantum spin liquid. We argued that the dynamical properties of the quantum spin liquid are unchanged by the inclusion of out-of-plane Dzyaloshinskii-Moriya interactions in the Ansatz employed. However, they do enlarge the quantum spin liquid region of the phase diagrams, which suggests they can stabilize the quantum spin liquid for stronger next nearest neighbor interactions.

The quantum spin liquid phase found has zero Chern numbers, but displays both the spin Nernst effect and the thermal Hall effect in the presence of an applied magnetic field. This is due to its significant Abelian Berry curvature contributions at the K and K$'$ points. The spin Nernst and thermal Hall conductivities obtained are found to be in good agreement with those found in theoretical studies and in experiments. We have therefore found two distinct ways of probing the quantum spin liquid phase in the Kane-Mele-Hubbard model.

We end the thesis by proposing topics for further work. Firstly, since Dzyaloshinskii-Moriya interactions are found to impact the transport properties of disordered spin systems

in References [28, 37–39, 42], their effects on the transport in the Kane-Mele-Hubbard model should be investigated further. We therefore propose the exploration of different mean field Ansätze, as they are critically important for the obtained physics. In particular, we suggest that the mean field Ansatz be extended or altered to include Dzyaloshinskii-Moriya interactions diagonally in the Hamiltonian.

Secondly, we propose the introduction of a scalar spin chirality term [26, 39, 43, 44]

$$\hat{\vec{S}}_i \cdot \left(\hat{\vec{S}}_j \times \hat{\vec{S}}_k\right),$$

whose presence can be caused by an applied magnetic field or the native spin-orbit coupling in the Kane-Mele-Hubbard model. Such a term is particularly interesting in the context of the thermal Hall effect, as the thermal current is directly related to the spin chirality [42]. The scalar spin chirality term may also be a useful tool for the study of potential chiral spin liquid phases in the Kane-Mele-Hubbard model.

This would also encourage further study of the Abrikosov fermion mean field theory of the model. The Abrikosov fermion transformation is the fermionic counterpart of the Schwinger boson transformation (7.16). Ushakov [6] and Vaezi et al. [7] used the Abrikosov fermion approach to study the Schwinger boson-predicted quantum spin liquid phase in more detail, and found that it has a chiral nature in a small region of the (J_2, Γ) parameter space. It would also be interesting to investigate the transport properties of the Abrikosov fermions.

Finally, we propose the inclusion of spin-phonon interactions in the Hamiltonian of the system. Spin-phonon hybridized energy bands have been shown to exhibit large Berry curvatures and finite Chern numbers, leading to a renormalization of the phonon thermal Hall conductivity [45]. Including such interactions in the Hamiltonian could bring us closer to accurate predictions of the thermal Hall conductivity in situations where both spin and phonon excitations are important. It would also be interesting to investigate how the phonons impact the phase diagrams.

Appendix A The Pauli Matrices

The Pauli matrices are a set of three 2×2 Hermitian matrices given by

$$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}; \quad \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}; \quad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{A.1}$$

They satisfy

$$\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = I, \tag{A.2}$$

$$[\sigma_i, \sigma_j] = 2i \sum_{k=1}^{3} \epsilon_{ijk} \sigma_k, \tag{A.3}$$

$$\{\sigma_i, \sigma_j\} = 2\delta_{ij} I, \tag{A.4}$$

and the Pauli matrix vector

$$\vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3) = \sigma_1 \hat{e}_1 + \sigma_2 \hat{e}_2 + \sigma_3 \hat{e}_3 \tag{A.5}$$

satisfies the completeness relation

$$\vec{\sigma}_{\alpha\beta} \cdot \vec{\sigma}_{\gamma\delta} = \sum_{i=1}^{3} (\sigma_i)_{\alpha\beta} (\sigma_i)_{\alpha\beta} = 2\delta_{\alpha\delta}\delta_{\beta\gamma} - \delta_{\alpha\beta}\delta_{\gamma\delta}. \tag{A.6}$$

Appendix B The Rashba Coupling

The treatment of the Kane-Mele-Hubbard model (7.12) in Section 7 did not include the Rashba coupling

$$\hat{H}_R = i\lambda_R \sum_{\langle i,j\rangle} \sum_{\alpha} (\vec{\sigma} \times \hat{d}_{ij})_3 \hat{c}^{\dagger}_{i\alpha} \hat{c}_{j\alpha}, \tag{B.1}$$

since inversion symmetry was assumed to be preserved. If this is not the case, as it may very well be in an experimental setup, its effect on the ground state should not be ignored. In this Appendix, we consider the Rashba term in its spin Hamiltonian form, which we express in terms of Schwinger bosons. The resulting quartic boson Hamiltonian is expressed in terms of decoupling operators, which can serve as a starting point for a mean field study of the inversion symmetry-broken Kane-Mele-Hubbard model.

Expanding to lowest order in t_1/U, the Rashba coupling may be written in terms of spin operators as [46]

$$\hat{H}_R = \sum_{\langle i,j\rangle} \vec{D}_{\parallel,ij} \cdot \left(\hat{\vec{S}}_i \times \hat{\vec{S}}_j \right) + \sum_{\langle i,j\rangle} \hat{\vec{S}}_i \mathbf{\Gamma}_{ij} \hat{\vec{S}}_j. \tag{B.2}$$

The first term in (B.2), which we will denote $\hat{H}_{D_\parallel}$, is an in-plane nearest neighbor DMI term, with the DMI vectors given by

$$\vec{D}_{\parallel,ij} = D_\parallel (\hat{d}_{ij} \times \hat{e}_3), \tag{B.3}$$

where the magnitude of the interaction is given by

$$D_\parallel = \frac{4t_1\lambda_R}{U}. \tag{B.4}$$

The DMI vectors on sublattice A are found using the definition of the nearest neighbor vectors (7.1):

$$\vec{D}^{(A)}_{\parallel,1} = \frac{D_\parallel}{2}(\hat{e}_1 - \sqrt{3}\hat{e}_2); \quad \vec{D}^{(A)}_{\parallel,2} = \frac{D_\parallel}{2}(\hat{e}_1 + \sqrt{3}\hat{e}_2); \quad \vec{D}^{(A)}_{\parallel,3} = -D_\parallel \hat{e}_1, \tag{B.5}$$

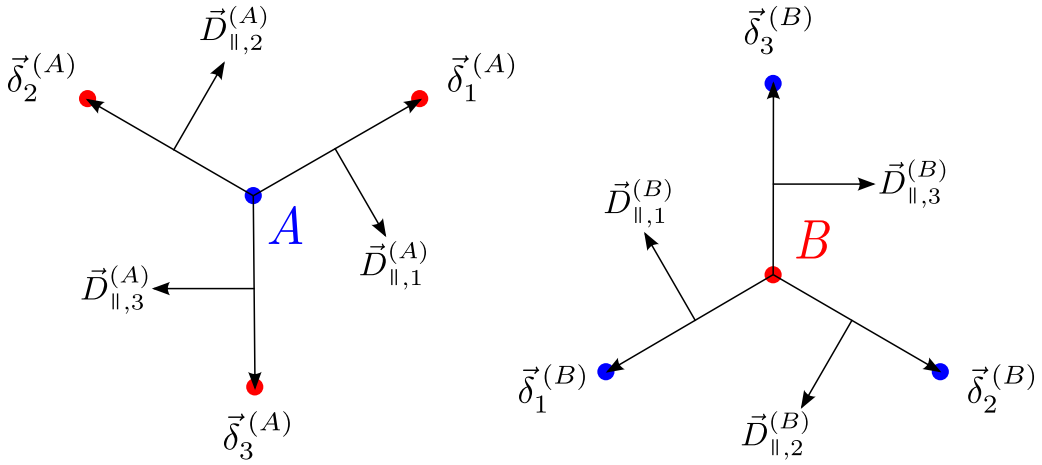


Figure B.1: The Rashba coupling gives rise to in-plane DMI. The DMI vectors are perpendicular to the nearest neighbor vectors and are opposite on the two sublattices A and B.

and it follows from $\vec{\delta}_l^{(B)} = -\vec{\delta}_l^{(A)}$ that $\vec{D}_{\parallel,l}^{(B)} = -\vec{D}_{\parallel,l}^{(A)}$ for $l \in \{1,2,3\}$. The six DMI vectors are shown in Figure B.1. The system is clearly not inversion-symmetric when these are present. The second term in (B.2), which we will denote $\hat{H}_\Gamma$, is an anisotropic nearest neighbor Heisenberg term, analogous to the anisotropic Heisenberg term present in (7.13), with the matrices $\mathbf{\Gamma}_{ij}$ describing the interactions. Defining the vector $\vec{\Delta}_{ij} = -i\lambda_R(\hat{d}_{ij} \times \hat{e}_3)$, the elements of these matrices are given compactly by

$$\Gamma_{ij}^{\alpha\beta} = \frac{2}{U}(\delta_{\alpha\beta}\vec{\Delta}_{ij}^2 - 2\Delta_{ij}^\alpha \Delta_{ij}^\beta). \tag{B.6}$$

We first consider the in-plane DMI term $\hat{H}_{D_\parallel}$. Having established what the DMI vectors are, we now write the first two components of the cross product $\hat{\vec{S}}_i \times \hat{\vec{S}}_j$ in terms of Schwinger bosons:

$$\begin{aligned}
(\hat{\vec{S}}_i \times \hat{\vec{S}}_j)_1 &= \hat{S}_i^y \hat{S}_j^z - \hat{S}_i^z \hat{S}_j^y \\
&= \frac{i\hbar^2}{4}(\hat{b}_{i\downarrow}^\dagger \hat{b}_{i\uparrow} - \hat{b}_{i\uparrow}^\dagger \hat{b}_{i\downarrow})(\hat{b}_{j\uparrow}^\dagger \hat{b}_{j\uparrow} - \hat{b}_{j\downarrow}^\dagger \hat{b}_{j\downarrow}) - \frac{i\hbar^2}{4}(\hat{b}_{i\uparrow}^\dagger \hat{b}_{i\uparrow} - \hat{b}_{i\downarrow}^\dagger \hat{b}_{i\downarrow})(\hat{b}_{j\downarrow}^\dagger \hat{b}_{j\uparrow} - \hat{b}_{j\uparrow}^\dagger \hat{b}_{j\downarrow}) \\
&= \frac{i\hbar^2}{4}\Big[\hat{b}_{i\downarrow}^\dagger \hat{b}_{i\uparrow}\hat{b}_{j\uparrow}^\dagger \hat{b}_{j\uparrow} - \hat{b}_{i\downarrow}^\dagger \hat{b}_{i\uparrow}\hat{b}_{j\downarrow}^\dagger \hat{b}_{j\downarrow} - \hat{b}_{i\uparrow}^\dagger \hat{b}_{i\downarrow}\hat{b}_{j\uparrow}^\dagger \hat{b}_{j\uparrow} + \hat{b}_{i\uparrow}^\dagger \hat{b}_{i\downarrow}\hat{b}_{j\downarrow}^\dagger \hat{b}_{j\downarrow} \\
&\qquad - \hat{b}_{i\uparrow}^\dagger \hat{b}_{i\uparrow}\hat{b}_{j\downarrow}^\dagger \hat{b}_{j\uparrow} + \hat{b}_{i\uparrow}^\dagger \hat{b}_{i\uparrow}\hat{b}_{j\uparrow}^\dagger \hat{b}_{j\downarrow} + \hat{b}_{i\downarrow}^\dagger \hat{b}_{i\downarrow}\hat{b}_{j\downarrow}^\dagger \hat{b}_{j\uparrow} - \hat{b}_{i\downarrow}^\dagger \hat{b}_{i\downarrow}\hat{b}_{j\uparrow}^\dagger \hat{b}_{j\downarrow}\Big],
\end{aligned} \tag{B.7}$$

$$\begin{aligned}
(\hat{\vec{S}}_i \times \hat{\vec{S}}_j)_2 &= \hat{S}_i^z \hat{S}_j^x - \hat{S}_i^x \hat{S}_j^z \\
&= \frac{\hbar^2}{4}(\hat{b}_{i\uparrow}^\dagger \hat{b}_{i\uparrow} - \hat{b}_{i\downarrow}^\dagger \hat{b}_{i\downarrow})(\hat{b}_{j\uparrow}^\dagger \hat{b}_{j\downarrow} + \hat{b}_{j\downarrow}^\dagger \hat{b}_{j\uparrow}) - \frac{\hbar^2}{4}(\hat{b}_{i\uparrow}^\dagger \hat{b}_{i\downarrow} + \hat{b}_{i\downarrow}^\dagger \hat{b}_{i\uparrow})(\hat{b}_{j\uparrow}^\dagger \hat{b}_{j\uparrow} - \hat{b}_{j\downarrow}^\dagger \hat{b}_{j\downarrow}) \\
&= \frac{\hbar^2}{4}\Big[\hat{b}_{i\uparrow}^\dagger \hat{b}_{i\uparrow}\hat{b}_{j\uparrow}^\dagger \hat{b}_{j\downarrow} + \hat{b}_{i\uparrow}^\dagger \hat{b}_{i\uparrow}\hat{b}_{j\downarrow}^\dagger \hat{b}_{j\uparrow} - \hat{b}_{i\downarrow}^\dagger \hat{b}_{i\downarrow}\hat{b}_{j\uparrow}^\dagger \hat{b}_{j\downarrow} - \hat{b}_{i\downarrow}^\dagger \hat{b}_{i\downarrow}\hat{b}_{j\downarrow}^\dagger \hat{b}_{j\uparrow} \\
&\qquad - \hat{b}_{i\uparrow}^\dagger \hat{b}_{i\downarrow}\hat{b}_{j\uparrow}^\dagger \hat{b}_{j\uparrow} + \hat{b}_{i\uparrow}^\dagger \hat{b}_{i\downarrow}\hat{b}_{j\downarrow}^\dagger \hat{b}_{j\downarrow} - \hat{b}_{i\downarrow}^\dagger \hat{b}_{i\uparrow}\hat{b}_{j\uparrow}^\dagger \hat{b}_{j\uparrow} + \hat{b}_{i\downarrow}^\dagger \hat{b}_{i\uparrow}\hat{b}_{j\downarrow}^\dagger \hat{b}_{j\downarrow}\Big].
\end{aligned} \tag{B.8}$$

The third component is redundant since all DMI vectors lie in the plane. We now decouple these interaction terms in the same way as before. Defining

$$\hat{F}_{ij} = \hat{b}_{i\uparrow}\hat{b}_{j\uparrow} + \hat{b}_{i\downarrow}\hat{b}_{j\downarrow}; \quad \hat{G}_{ij} = \hat{b}_{i\uparrow}^\dagger \hat{b}_{j\downarrow} + \hat{b}_{i\downarrow}^\dagger \hat{b}_{j\uparrow}; \quad \hat{H}_{ij} = \hat{b}_{i\uparrow}^\dagger \hat{b}_{j\downarrow} - \hat{b}_{i\downarrow}^\dagger \hat{b}_{j\uparrow}, \tag{B.9}$$

the first component may be written

$$\begin{aligned}
(\hat{\vec{S}}_i \times \hat{\vec{S}}_j)_1 &= \frac{i\hbar^2}{4}\Big[\hat{F}_{ij}^\dagger \hat{E}_{ij} - \hat{E}_{ij}^\dagger \hat{F}_{ij}\Big] = \frac{i\hbar^2}{4}\Big[\hat{B}_{ij}^\dagger \hat{G}_{ij} - \hat{G}_{ij}^\dagger \hat{B}_{ij}\Big] \\
&= \frac{i\hbar^2}{8}\Big[\hat{F}_{ij}^\dagger \hat{E}_{ij} - \hat{E}_{ij}^\dagger \hat{F}_{ij} + \hat{B}_{ij}^\dagger \hat{G}_{ij} - \hat{G}_{ij}^\dagger \hat{B}_{ij}\Big]
\end{aligned} \tag{B.10}$$

and the second component

$$\begin{aligned}
(\hat{\vec{S}}_i \times \hat{\vec{S}}_j)_2 &= \frac{\hbar^2}{4}\Big[\hat{F}_{ij}^\dagger \hat{A}_{ij} + \hat{A}_{ij}^\dagger \hat{F}_{ij}\Big] = \frac{\hbar^2}{4}\Big[\hat{B}_{ij}^\dagger \hat{H}_{ij} + \hat{H}_{ij}^\dagger \hat{B}_{ij}\Big] \\
&= \frac{\hbar^2}{8}\Big[\hat{F}_{ij}^\dagger \hat{A}_{ij} + \hat{A}_{ij}^\dagger \hat{F}_{ij} + \hat{B}_{ij}^\dagger \hat{H}_{ij} + \hat{H}_{ij}^\dagger \hat{B}_{ij}\Big],
\end{aligned} \tag{B.11}$$

with $\hat{A}_{ij}$ and $\hat{B}_{ij}$ defined in (7.24) and $\hat{E}_{ij}$ defined in (7.28). Before inserting (B.10) and (B.11) into the DMI term $\hat{H}_{D_\parallel}$, it is useful to write it out more explicitly to reveal its structure:

$$\begin{aligned}\hat{H}_{D_\parallel} &= \sum_{\langle i,j\rangle} \vec{D}_{\parallel,ij} \cdot \left(\hat{\vec{S}}_i \times \hat{\vec{S}}_j\right) = \sum_{\vec{r}} \sum_{l=1}^{3} \sum_{X} \vec{D}^{(X)}_{\parallel,l} \cdot \left(\hat{\vec{S}}_{\vec{r}} \times \hat{\vec{S}}_{\vec{r}+\vec{\delta}_l^{(X)}}\right) \\ &= \sum_{\vec{r}} \sum_{l=1}^{3} \left[\vec{D}^{(A)}_{\parallel,l} \cdot \left(\hat{\vec{S}}_{\vec{r}} \times \hat{\vec{S}}_{\vec{r}+\vec{\delta}_l^{(A)}}\right) + \vec{D}^{(B)}_{\parallel,l} \cdot \left(\hat{\vec{S}}_{\vec{r}} \times \hat{\vec{S}}_{\vec{r}+\vec{\delta}_l^{(B)}}\right)\right].\end{aligned} \tag{B.12}$$

We split this expression up into its A and B sublattice parts, $\hat{H}^{(A)}_{D_\parallel}$ and $\hat{H}^{(B)}_{D_\parallel}$. Since the latter is completely analogous to the former, with the appropriate changes of signs and indices, we will include this term only at the end of the analysis, and consider the former in detail in the following. Using (B.10) and (B.11), $\hat{H}^{(A)}_{D_\parallel}$ is expressed in terms of Schwinger bosons as:

$$\begin{aligned}\hat{H}^{(A)}_{D_\parallel} &= \sum_{\vec{r}} \sum_{l=1}^{3} \vec{D}^{(A)}_{\parallel,l} \cdot \left(\hat{\vec{S}}_{\vec{r}} \times \hat{\vec{S}}_{\vec{r}+\vec{\delta}_l^{(A)}}\right) \\ &= \sum_{\vec{r}} \left[\vec{D}^{(A)}_{\parallel,1} \cdot \left(\hat{\vec{S}}_{\vec{r}} \times \hat{\vec{S}}_{\vec{r}+\vec{\delta}_1^{(A)}}\right) + \vec{D}^{(A)}_{\parallel,2} \cdot \left(\hat{\vec{S}}_{\vec{r}} \times \hat{\vec{S}}_{\vec{r}+\vec{\delta}_2^{(A)}}\right) + \vec{D}^{(A)}_{\parallel,3} \cdot \left(\hat{\vec{S}}_{\vec{r}} \times \hat{\vec{S}}_{\vec{r}+\vec{\delta}_3^{(A)}}\right)\right] \\ &= D_\parallel \sum_{\vec{r}} \Bigg[\frac{1}{2}\frac{i\hbar^2}{8}\Bigg(\hat{F}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_1^{(A)}} \hat{E}_{\vec{r},\vec{r}+\vec{\delta}_1^{(A)}} - \hat{E}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_1^{(A)}} \hat{F}_{\vec{r},\vec{r}+\vec{\delta}_1^{(A)}} \\ &\qquad + \hat{B}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_1^{(A)}} \hat{G}_{\vec{r},\vec{r}+\vec{\delta}_1^{(A)}} - \hat{G}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_1^{(A)}} \hat{B}_{\vec{r},\vec{r}+\vec{\delta}_1^{(A)}}\Bigg) \\ &\quad - \frac{\sqrt{3}}{2}\frac{\hbar^2}{8}\Bigg(\hat{F}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_1^{(A)}} \hat{A}_{\vec{r},\vec{r}+\vec{\delta}_1^{(A)}} + \hat{A}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_1^{(A)}} \hat{F}_{\vec{r},\vec{r}+\vec{\delta}_1^{(A)}} \\ &\qquad + \hat{B}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_1^{(A)}} \hat{H}_{\vec{r},\vec{r}+\vec{\delta}_1^{(A)}} + \hat{H}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_1^{(A)}} \hat{B}_{\vec{r},\vec{r}+\vec{\delta}_1^{(A)}}\Bigg) \\ &\quad + \frac{1}{2}\frac{i\hbar^2}{8}\Bigg(\hat{F}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_2^{(A)}} \hat{E}_{\vec{r},\vec{r}+\vec{\delta}_2^{(A)}} - \hat{E}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_2^{(A)}} \hat{F}_{\vec{r},\vec{r}+\vec{\delta}_2^{(A)}} \\ &\qquad + \hat{B}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_2^{(A)}} \hat{G}_{\vec{r},\vec{r}+\vec{\delta}_2^{(A)}} - \hat{G}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_2^{(A)}} \hat{B}_{\vec{r},\vec{r}+\vec{\delta}_2^{(A)}}\Bigg) \\ &\quad + \frac{\sqrt{3}}{2}\frac{\hbar^2}{8}\Bigg(\hat{F}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_2^{(A)}} \hat{A}_{\vec{r},\vec{r}+\vec{\delta}_2^{(A)}} + \hat{A}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_2^{(A)}} \hat{F}_{\vec{r},\vec{r}+\vec{\delta}_2^{(A)}} \\ &\qquad + \hat{B}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_2^{(A)}} \hat{H}_{\vec{r},\vec{r}+\vec{\delta}_2^{(A)}} + \hat{H}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_2^{(A)}} \hat{B}_{\vec{r},\vec{r}+\vec{\delta}_2^{(A)}}\Bigg) \\ &\quad - \frac{i\hbar^2}{8}\Bigg(\hat{F}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_3^{(A)}} \hat{E}_{\vec{r},\vec{r}+\vec{\delta}_3^{(A)}} - \hat{E}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_3^{(A)}} \hat{F}_{\vec{r},\vec{r}+\vec{\delta}_3^{(A)}} \\ &\qquad + \hat{B}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_3^{(A)}} \hat{G}_{\vec{r},\vec{r}+\vec{\delta}_3^{(A)}} - \hat{G}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_3^{(A)}} \hat{B}_{\vec{r},\vec{r}+\vec{\delta}_3^{(A)}}\Bigg)\Bigg].\end{aligned} \tag{B.13}$$

Next, we consider the anisotropic term $\hat{H}_\Gamma$. At first glance, it may seem that there are six different matrices $\mathbf{\Gamma}_{ij}$ describing these interactions. However, since $\vec{\delta}_l^{(B)} = -\vec{\delta}_l^{(A)}$, it follows

that $\vec{\Delta}_l^{(B)} = -\vec{\Delta}_l^{(A)}$ for $l \in \{1,2,3\}$. The matrix elements (B.6) are given by products of components of $\vec{\Delta}_{ij}$, which means the interaction matrices $\mathbf{\Gamma}_l$ are the same for the two sublattices. The anisotropic term may then be written

$$\hat{H}_\Gamma = \sum_{\langle i,j \rangle} \hat{\vec{S}}_i \mathbf{\Gamma}_{ij} \hat{\vec{S}}_j = \sum_{\vec{r}} \sum_{l=1}^{3} \sum_{X} \hat{\vec{S}}_{\vec{r}} \mathbf{\Gamma}_l \hat{\vec{S}}_{\vec{r}+\vec{\delta}_l^{(X)}}. \tag{B.14}$$

By direct computation of the matrix elements, we find that the matrices are given by:

$$\begin{aligned}
\mathbf{\Gamma}_1 &= \frac{\lambda_R^2}{U} \begin{bmatrix} -1 & -\sqrt{3} & 0 \\ -\sqrt{3} & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix}, \\
\mathbf{\Gamma}_2 &= \frac{\lambda_R^2}{U} \begin{bmatrix} -1 & \sqrt{3} & 0 \\ \sqrt{3} & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix}, \\
\mathbf{\Gamma}_3 &= \frac{2\lambda_R^2}{U} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.
\end{aligned} \tag{B.15}$$

As for $\hat{H}_{D_\parallel}$, we split $\hat{H}_\Gamma$ into $\hat{H}_\Gamma^{(A)}$ and $\hat{H}_\Gamma^{(B)}$ and consider only the former in detail in what follows:

$$\hat{H}_\Gamma^{(A)} = \sum_{\vec{r}} \sum_{l=1}^{3} \hat{\vec{S}}_{\vec{r}} \mathbf{\Gamma}_l \hat{\vec{S}}_{\vec{r}+\vec{\delta}_l^{(A)}} = \sum_{\vec{r}} \left[\hat{\vec{S}}_{\vec{r}} \mathbf{\Gamma}_1 \hat{\vec{S}}_{\vec{r}+\vec{\delta}_1^{(A)}} + \hat{\vec{S}}_{\vec{r}} \mathbf{\Gamma}_2 \hat{\vec{S}}_{\vec{r}+\vec{\delta}_2^{(A)}} + \hat{\vec{S}}_{\vec{r}} \mathbf{\Gamma}_3 \hat{\vec{S}}_{\vec{r}+\vec{\delta}_3^{(A)}} \right]. \tag{B.16}$$

In order to express (B.16) in terms of Schwinger bosons, it is useful to compute the G tensor (7.20) for these matrices. For the matrix

$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} & 0 \\ A_{12} & -A_{11} & 0 \\ 0 & 0 & A_{33} \end{bmatrix}, \tag{B.17}$$

which has the same structure as $\mathbf{\Gamma}_1$ and $\mathbf{\Gamma}_2$, we find

$$G(\mathbf{A}) = \begin{bmatrix} A_{33} & 0 & 0 & -A_{33} \\ 0 & 2A_{11} - 2iA_{12} & 0 & 0 \\ 0 & 0 & 2A_{11} + 2iA_{12} & 0 \\ -A_{33} & 0 & 0 & A_{33} \end{bmatrix}, \tag{B.18}$$

and for the diagonal matrix $\mathbf{D} = \mathrm{diag}(D_1, -D_1, -D_1)$ with the same structure as $\mathbf{\Gamma}_3$ we have

$$G(\mathbf{D}) = \begin{bmatrix} -D_1 & 0 & 0 & D_1 \\ 0 & 2D_1 & 0 & 0 \\ 0 & 0 & 2D_1 & 0 \\ D_1 & 0 & 0 & -D_1 \end{bmatrix}. \tag{B.19}$$

We consider the Schwinger boson transformation of each term in (B.16) separately. For brevity of notation, we will write i for $\vec{r}$ and j for $\vec{r}+\vec{\delta}_l^{(A)}$ and omit the subscripts A and B on the bosonic operators in the following calculation. The first term is

$$\begin{aligned}\hat{\vec{S}}_i\mathbf{\Gamma}_1\hat{\vec{S}}_j &= \frac{\hbar^2}{4}\sum_{\alpha,\beta,\gamma,\delta}\hat{b}^\dagger_{i\alpha}\hat{b}_{i\beta}\hat{b}^\dagger_{j\gamma}\hat{b}_{j\delta}(\vec{\sigma}_{\alpha\beta}\cdot\mathbf{\Gamma}_1\vec{\sigma}_{\gamma\delta}) \\ &= \frac{\hbar^2\lambda_R^2}{2U}\big[-\hat{b}^\dagger_{i\uparrow}\hat{b}_{i\uparrow}\hat{b}^\dagger_{j\uparrow}\hat{b}_{j\uparrow}+\hat{b}^\dagger_{i\uparrow}\hat{b}_{i\uparrow}\hat{b}^\dagger_{j\downarrow}\hat{b}_{j\downarrow}+\hat{b}^\dagger_{i\downarrow}\hat{b}_{i\downarrow}\hat{b}^\dagger_{j\uparrow}\hat{b}_{j\uparrow}-\hat{b}^\dagger_{i\downarrow}\hat{b}_{i\downarrow}\hat{b}^\dagger_{j\downarrow}\hat{b}_{j\downarrow} \\ &\qquad -(\hat{b}^\dagger_{i\uparrow}\hat{b}_{i\downarrow}\hat{b}^\dagger_{j\uparrow}\hat{b}_{j\downarrow}+\hat{b}^\dagger_{i\downarrow}\hat{b}_{i\uparrow}\hat{b}^\dagger_{j\downarrow}\hat{b}_{j\uparrow})+i\sqrt{3}(\hat{b}^\dagger_{i\uparrow}\hat{b}_{i\downarrow}\hat{b}^\dagger_{j\uparrow}\hat{b}_{j\downarrow}-\hat{b}^\dagger_{i\downarrow}\hat{b}_{i\uparrow}\hat{b}^\dagger_{j\downarrow}\hat{b}_{j\uparrow})\big].\end{aligned} \tag{B.20}$$

The first four terms in (B.20) may be written as

$$\begin{aligned}\sum_{\alpha,\beta}(1-2\delta_{\alpha\beta})\hat{b}^\dagger_{i\alpha}\hat{b}_{i\alpha}\hat{b}^\dagger_{j\beta}\hat{b}_{j\beta} &= \sum_{\alpha,\beta}\hat{b}^\dagger_{i\alpha}\hat{b}_{i\alpha}\hat{b}^\dagger_{j\beta}\hat{b}_{j\beta}-2\sum_{\alpha}\hat{b}^\dagger_{i\alpha}\hat{b}_{i\alpha}\hat{b}^\dagger_{j\alpha}\hat{b}_{j\alpha} \\ &= 1-2(\hat{b}^\dagger_{i\uparrow}\hat{b}_{i\uparrow}\hat{b}^\dagger_{j\uparrow}\hat{b}_{j\uparrow}+\hat{b}^\dagger_{i\downarrow}\hat{b}_{i\downarrow}\hat{b}^\dagger_{j\downarrow}\hat{b}_{j\downarrow}),\end{aligned} \tag{B.21}$$

where we used (7.17) to perform the first sum in the last equality. Defining

$$\hat{K}_{ij} = \hat{b}_{i\uparrow}\hat{b}_{j\uparrow}-\hat{b}_{i\downarrow}\hat{b}_{j\downarrow}, \tag{B.22}$$

we find that (B.21) can be decoupled as

$$1-2(\hat{b}^\dagger_{i\uparrow}\hat{b}_{i\uparrow}\hat{b}^\dagger_{j\uparrow}\hat{b}_{j\uparrow}+\hat{b}^\dagger_{i\downarrow}\hat{b}_{i\downarrow}\hat{b}^\dagger_{j\downarrow}\hat{b}_{j\downarrow}) = 3-\hat{B}^\dagger_{ij}\hat{B}_{ij}-\hat{D}^\dagger_{ij}\hat{D}_{ij} = 1-\hat{F}^\dagger_{ij}\hat{F}_{ij}-\hat{K}^\dagger_{ij}\hat{K}_{ij}. \tag{B.23}$$

The two terms in the parentheses in (B.20) are decoupled as

$$\hat{b}^\dagger_{i\uparrow}\hat{b}_{i\downarrow}\hat{b}^\dagger_{j\uparrow}\hat{b}_{j\downarrow}+\hat{b}^\dagger_{i\downarrow}\hat{b}_{i\uparrow}\hat{b}^\dagger_{j\downarrow}\hat{b}_{j\uparrow} = \frac{1}{2}\bigg[\hat{F}^\dagger_{ij}\hat{F}_{ij}-\hat{K}^\dagger_{ij}\hat{K}_{ij}\bigg], \tag{B.24}$$

$$\hat{b}^\dagger_{i\uparrow}\hat{b}_{i\downarrow}\hat{b}^\dagger_{j\uparrow}\hat{b}_{j\downarrow}-\hat{b}^\dagger_{i\downarrow}\hat{b}_{i\uparrow}\hat{b}^\dagger_{j\downarrow}\hat{b}_{j\uparrow} = \frac{1}{2}\bigg[\hat{K}^\dagger_{ij}\hat{F}_{ij}-\hat{F}^\dagger_{ij}\hat{K}_{ij}\bigg], \tag{B.25}$$

which, after restoring the proper indices and subscripts, means we may write

$$\begin{aligned}\hat{\vec{S}}_{\vec{r}}\mathbf{\Gamma}_1\hat{\vec{S}}_{\vec{r}+\vec{\delta}_1^{(A)}} = \frac{\hbar^2\lambda_R^2}{4U}\bigg[&4-\hat{B}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_1^{(A)}}\hat{B}_{\vec{r},\vec{r}+\vec{\delta}_1^{(A)}}-\hat{D}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_1^{(A)}}\hat{D}_{\vec{r},\vec{r}+\vec{\delta}_1^{(A)}}-2\hat{F}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_1^{(A)}}\hat{F}_{\vec{r},\vec{r}+\vec{\delta}_1^{(A)}} \\ &+i\sqrt{3}\bigg(\hat{K}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_1^{(A)}}\hat{F}_{\vec{r},\vec{r}+\vec{\delta}_1^{(A)}}-\hat{F}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_1^{(A)}}\hat{K}_{\vec{r},\vec{r}+\vec{\delta}_1^{(A)}}\bigg)\bigg].\end{aligned} \tag{B.26}$$

Now, since the only difference between $\mathbf{\Gamma}_1$ and $\mathbf{\Gamma}_2$ is the sign of the matrix elements proportional to $\sqrt{3}$, the second term in (B.16) is readily obtained as

$$\begin{aligned}\hat{\vec{S}}_{\vec{r}}\mathbf{\Gamma}_2\hat{\vec{S}}_{\vec{r}+\vec{\delta}_2^{(A)}} = \frac{\hbar^2\lambda_R^2}{4U}\bigg[&4-\hat{B}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_2^{(A)}}\hat{B}_{\vec{r},\vec{r}+\vec{\delta}_2^{(A)}}-\hat{D}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_2^{(A)}}\hat{D}_{\vec{r},\vec{r}+\vec{\delta}_2^{(A)}}-2\hat{F}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_2^{(A)}}\hat{F}_{\vec{r},\vec{r}+\vec{\delta}_2^{(A)}} \\ &-i\sqrt{3}\bigg(\hat{K}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_2^{(A)}}\hat{F}_{\vec{r},\vec{r}+\vec{\delta}_2^{(A)}}-\hat{F}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_2^{(A)}}\hat{K}_{\vec{r},\vec{r}+\vec{\delta}_2^{(A)}}\bigg)\bigg].\end{aligned} \tag{B.27}$$

Finally, we have

$$\begin{aligned}\hat{\vec{S}}_i \mathbf{\Gamma}_3 \hat{\vec{S}}_j &= \frac{\hbar^2}{4}\sum_{\alpha\beta\gamma\delta} \hat{b}^\dagger_{i\alpha}\hat{b}_{i\beta}\hat{b}^\dagger_{j\gamma}\hat{b}_{j\delta}(\vec{\sigma}_{\alpha\beta}\cdot\mathbf{\Gamma}_3\vec{\sigma}_{\gamma\delta}) \\ &= \frac{\hbar^2\lambda_R^2}{2U}\Big[-\hat{b}^\dagger_{i\uparrow}\hat{b}_{i\uparrow}\hat{b}^\dagger_{j\uparrow}\hat{b}_{j\uparrow} + \hat{b}^\dagger_{i\uparrow}\hat{b}_{i\uparrow}\hat{b}^\dagger_{j\downarrow}\hat{b}_{j\downarrow} + \hat{b}^\dagger_{i\downarrow}\hat{b}_{i\downarrow}\hat{b}^\dagger_{j\uparrow}\hat{b}_{j\uparrow} - \hat{b}^\dagger_{i\downarrow}\hat{b}_{i\downarrow}\hat{b}^\dagger_{j\downarrow}\hat{b}_{j\downarrow} \\ &\qquad + 2(\hat{b}^\dagger_{i\uparrow}\hat{b}_{i\downarrow}\hat{b}^\dagger_{j\uparrow}\hat{b}_{j\downarrow} + \hat{b}^\dagger_{i\downarrow}\hat{b}_{i\uparrow}\hat{b}^\dagger_{j\downarrow}\hat{b}_{j\uparrow})\Big].\end{aligned} \tag{B.28}$$

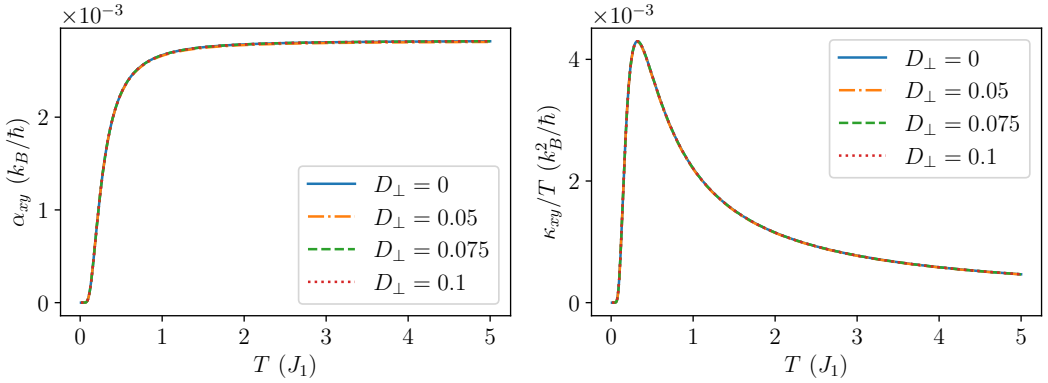
The first four terms are decoupled using (B.23), while the terms in the parenthesis are decoupled using (B.24). Restoring the indices and subscripts, we thus obtain

$$\begin{aligned}\hat{\vec{S}}_{\vec{r}}\mathbf{\Gamma}_3\hat{\vec{S}}_{\vec{r}+\vec{\delta}_3^{(A)}} = \frac{\hbar^2\lambda_R^2}{4U}\Big[4 &- \hat{B}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_3^{(A)}}\hat{B}_{\vec{r},\vec{r}+\vec{\delta}_3^{(A)}} - \hat{D}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_3^{(A)}}\hat{D}_{\vec{r},\vec{r}+\vec{\delta}_3^{(A)}} \\ &+ \hat{F}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_3^{(A)}}\hat{F}_{\vec{r},\vec{r}+\vec{\delta}_3^{(A)}} - 3\hat{K}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_3^{(A)}}\hat{K}_{\vec{r},\vec{r}+\vec{\delta}_3^{(A)}}\Big].\end{aligned} \tag{B.29}$$

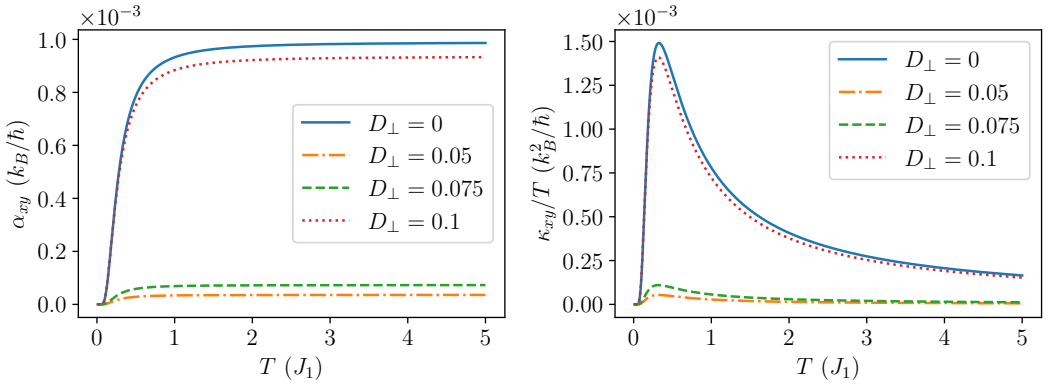
Inserting (B.26), (B.27) and (B.29) into (B.16), we obtain:

$$\begin{aligned}\hat{H}_\Gamma^{(A)} = \frac{\hbar^2\lambda_R^2}{4U}\sum_{\vec{r}}\Big[12 &- 3\hat{K}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_3^{(A)}}\hat{K}_{\vec{r},\vec{r}+\vec{\delta}_3^{(A)}} \\ &- \left(\hat{B}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_1^{(A)}}\hat{B}_{\vec{r},\vec{r}+\vec{\delta}_1^{(A)}} + \hat{B}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_2^{(A)}}\hat{B}_{\vec{r},\vec{r}+\vec{\delta}_2^{(A)}} + \hat{B}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_3^{(A)}}\hat{B}_{\vec{r},\vec{r}+\vec{\delta}_3^{(A)}}\right) \\ &- \left(\hat{D}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_1^{(A)}}\hat{D}_{\vec{r},\vec{r}+\vec{\delta}_1^{(A)}} + \hat{D}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_2^{(A)}}\hat{D}_{\vec{r},\vec{r}+\vec{\delta}_2^{(A)}} + \hat{D}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_3^{(A)}}\hat{D}_{\vec{r},\vec{r}+\vec{\delta}_3^{(A)}}\right) \\ &- 2\left(\hat{F}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_1^{(A)}}\hat{F}_{\vec{r},\vec{r}+\vec{\delta}_1^{(A)}} + \hat{F}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_2^{(A)}}\hat{F}_{\vec{r},\vec{r}+\vec{\delta}_2^{(A)}} - \frac{1}{2}\hat{F}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_3^{(A)}}\hat{F}_{\vec{r},\vec{r}+\vec{\delta}_3^{(A)}}\right) \\ &+ i\sqrt{3}\Big(\hat{K}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_1^{(A)}}\hat{F}_{\vec{r},\vec{r}+\vec{\delta}_1^{(A)}} - \hat{F}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_1^{(A)}}\hat{K}_{\vec{r},\vec{r}+\vec{\delta}_1^{(A)}} \\ &\qquad + \hat{F}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_2^{(A)}}\hat{K}_{\vec{r},\vec{r}+\vec{\delta}_2^{(A)}} - \hat{K}^\dagger_{\vec{r},\vec{r}+\vec{\delta}_2^{(A)}}\hat{F}_{\vec{r},\vec{r}+\vec{\delta}_2^{(A)}}\Big)\Big].\end{aligned} \tag{B.30}$$

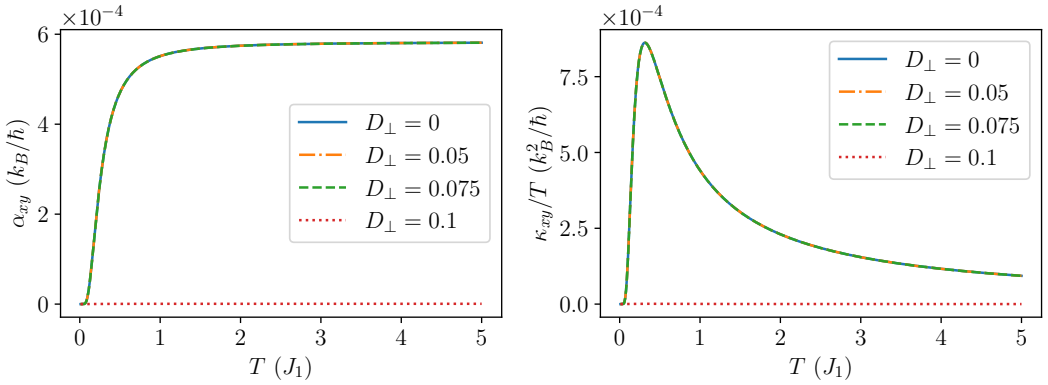
Appendix C Numerical Artifacts of Finite DMI



(a)



(b)



(c)

Figure C.1: Spin Nernst (left) and thermal Hall conductivities (right) in applied magnetic fields a) $B = 0.1J_1$, b) $B = 0.3J_1$ and c) $B = 0.5J_1$ for four different values of $D_\perp$ in units of J_1. The boson density is $\kappa = 0.3$. There is no clear trend to the variation in the transport as $D_\perp$ is increased, and we conclude that it is a numerical artifact.

References

[1] G. Binasch, P. Grünberg, F Saurenbach, and W Zinn, "Enhanced magnetoresistance in layered magnetic structures with antiferromagnetic interlayer exchange," Physical Review B **39**, 4828 (1989) (cited on page 1).

[2] M. N. Baibich, J. M. Broto, A. Fert, F. N. Van Dau, F. Petroff, P Etienne, G Creuzet, A Friederich, and J Chazelas, "Giant magnetoresistance of (001) Fe/(001) Cr magnetic superlattices," Physical Review Letters **61**, 2472 (1988) (cited on page 1).

[3] R. Weiss, R. Mattheis, and G. Reiss, "Advanced giant magnetoresistance technology for measurement applications," Measurement Science and Technology **24**, 082001 (2013) (cited on page 1).

[4] P. W. Anderson, "Resonating valence bonds: A new kind of insulator?" Materials Research Bulletin **8**, 153–160 (1973) (cited on pages 1, 84).

[5] Y. Zhou, K. Kanoda, and T.-K. Ng, "Quantum spin liquid states," Reviews of Modern Physics **89**, 025003 (2017) (cited on page 1).

[6] I. Ushakov, "Phase diagram of an extended Kane-Mele-Hubbard model in strongly correlated regime," Master's thesis (NTNU, 2020) (cited on pages 1, 5, 8, 61, 68–69, 86, 94).

[7] A. Vaezi, M. Mashkoori, and M. Hosseini, "Phase diagram of the strongly correlated Kane-Mele-Hubbard model," Physical Review B **85**, 195126 (2012) (cited on pages 1, 68–69, 81, 85, 94).

[8] J. Wen, S.-L. Yu, S. Li, W. Yu, and J.-X. Li, "Experimental identification of quantum spin liquids," npj Quantum Materials **4**, 1–9 (2019) (cited on page 1).

[9] S. K. L. Fosstveit, "Spin Dynamics in Quantum Spin Liquids," unpublished, n.d. (Cited on page 2).

[10] J. Van Hemmen, "A note on the diagonalization of quadratic boson and fermion hamiltonians," Zeitschrift für Physik B Condensed Matter **38**, 271–277 (1980) (cited on page 7).

[11] J. Colpa, "Diagonalization of the quadratic boson hamiltonian," Physica A: Statistical Mechanics and its Applications **93**, 327–353 (1978) (cited on page 9).

[12] H. Bruus and K. Flensberg, *Many-Body Quantum Theory in Condensed Matter Physics: An Introduction* (OUP Oxford, 2004) (cited on pages 10, 37, 49).

[13] A. Dutta, G. Aeppli, B. K. Chakrabarti, U. Divakaran, T. F. Rosenbaum, and D. Sen, *Quantum phase transitions in transverse field spin models: from statistical physics to quantum information* (Cambridge University Press, 2015) (cited on page 13).

[14] P. Coleman, *Introduction to Many-Body Physics* (Cambridge University Press, 2015) (cited on page 13).

[15] J. T. Haraldsen and R. S. Fishman, "Spin rotation technique for non-collinear magnetic systems: application to the generalized Villain model," Journal of Physics: Condensed Matter **21**, 216001 (2009) (cited on page 17).

[16] T. Oguchi, "The Equivalence of Hamiltonians of Holstein-Primakoff and Dyson in Spin-Wave Theory in Ferromagnetism," Progress of Theoretical Physics **25**, 721–722 (1961) (cited on page 20).

[17] M. Cieplak and L. Turski, "Solitons in quantum Heisenberg chain," Journal of Physics C: Solid State Physics **13**, 5741 (1980) (cited on page 26).

[18] M. V. Berry, "Quantal phase factors accompanying adiabatic changes," Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences **392**, 45–57 (1984) (cited on page 29).

[19] B. A. Bernevig, *Topological Insulators and Topological Superconductors* (Princeton University Press, 2013) (cited on page 29).

[20] F. D. M. Haldane, "Model for a quantum Hall effect without Landau levels: Condensed-matter realization of the "parity anomaly"," Physical Review Letters **61**, 2015 (1988) (cited on page 33).

[21] H. L. Ottesen, "Optical Conductivity of Dirac Fermions in Antiferromagnetic Semimetals," Master's thesis (NTNU, 2021) (cited on page 39).

[22] G. D. Mahan, *Many-Particle Systems* (Princeton University Press, 2008) (cited on page 39).

[23] J. Luttinger, "Theory of thermal transport coefficients," Physical Review **135**, A1505 (1964) (cited on page 56).

[24] V. A. Zyuzin and A. A. Kovalev, "Magnon Spin Nernst Effect in Antiferromagnets," Physical Review Letters **117**, 217203 (2016) (cited on page 57).

[25] R. Matsumoto, R. Shindou, and S. Murakami, "Thermal Hall effect of magnons in magnets with dipolar interaction," Physical Review B **89**, 054420 (2014) (cited on pages 57–58).

[26] J. H. Han and H. Lee, "Spin chirality and Hall-like transport phenomena of spin excitations," Journal of the Physical Society of Japan **86**, 011007 (2017) (cited on pages 57–58, 94).

[27] L Smrcka and P Streda, "Transport coefficients in strong magnetic fields," Journal of Physics C: Solid State Physics **10**, 2153 (1977) (cited on page 58).

[28] A. A. Kovalev and V. Zyuzin, "Spin torque and Nernst effects in Dzyaloshinskii-Moriya ferromagnets," Physical Review B **93**, 161106 (2016) (cited on pages 58, 89, 91, 94).

[29] B. Sheikhi, M. Kargarian, and A. Langari, "Hybrid topological magnon-phonon modes in ferromagnetic honeycomb and kagome lattices," Physical Review B **104**, 045139 (2021) (cited on page 58).

[30] R. Matsumoto and S. Murakami, "Rotational motion of magnons and the thermal Hall effect," Physical Review B **84**, 184406 (2011) (cited on page 59).

[31] G. Tatara, "Thermal vector potential theory of transport induced by a temperature gradient," Physical Review Letters **114**, 196601 (2015) (cited on page 60).

[32] C. Kittel and P. McEuen, *Introduction to solid state physics*, Vol. 8 (Wiley New York, 1996) (cited on page 62).

[33] J. Merino and A. Ralko, "Role of quantum fluctuations on spin liquids and ordered phases in the Heisenberg model on the honeycomb lattice," Physical Review B **97**, 205112 (2018) (cited on page 66).

[34] F. Wang, "Schwinger boson mean field theories of spin liquid states on a honeycomb lattice: Projective symmetry group analysis and critical field theory," Physical Review B **82**, 024419 (2010) (cited on pages 68–69, 84).

[35] T. Li, F. Becca, W. Hu, and S. Sorella, "Gapped spin-liquid phase in the J_1-J_2 Heisenberg model by a bosonic resonating valence-bond ansatz," Physical Review B **86**, 075111 (2012) (cited on page 84).

[36] D.-V. Bauer and J. Fjærestad, "Schwinger-boson mean-field study of the J_1-J_2 Heisenberg quantum antiferromagnet on the triangular lattice," Physical Review B **96**, 165141 (2017) (cited on page 85).

[37] Y. Zhang, S. Okamoto, and D. Xiao, "Spin-Nernst effect in the paramagnetic regime of an antiferromagnetic insulator," Physical Review B **98**, 035424 (2018) (cited on pages 89, 91, 94).

[38] P. Laurell and G. A. Fiete, "Magnon thermal Hall effect in kagome antiferromagnets with Dzyaloshinskii-Moriya interactions," Physical Review B **98**, 094419 (2018) (cited on pages 91, 94).

[39] R. Samajdar, S. Chatterjee, S. Sachdev, and M. S. Scheurer, "Thermal Hall effect in square-lattice spin liquids: A Schwinger boson mean-field study," Physical Review B **99**, 165126 (2019) (cited on pages 91, 94).

[40] M.-E. Boulanger, G. Grissonnanche, S. Badoux, A. Allaire, É. Lefrançois, A. Legros, A. Gourgout, M. Dion, C. Wang, X. Chen, et al., "Thermal Hall conductivity in the cuprate Mott insulators Nd_2CuO_4 and $Sr_2CuO_2Cl_2$," Nature Communications **11**, 1–9 (2020) (cited on page 91).

[41] M. Akazawa, M. Shimozawa, S. Kittaka, T. Sakakibara, R. Okuma, Z. Hiroi, H.-Y. Lee, N. Kawashima, J. H. Han, and M. Yamashita, "Thermal Hall Effects of Spins and Phonons in Kagome Antiferromagnet Cd-Kapellasite," Physical Review X **10**, 041059 (2020) (cited on page 91).

[42] H. Lee, J. H. Han, and P. A. Lee, "Thermal Hall effect of spins in a paramagnet," Physical Review B **91**, 125413 (2015) (cited on pages 91, 94).

[43] C. Hickey, L. Cincio, Z. Papić, and A. Paramekanti, "Haldane-Hubbard Mott insulator: From tetrahedral spin crystal to chiral spin liquid," Physical Review Letters **116**, 137202 (2016) (cited on page 94).

[44] Y. H. Gao, C. Hickey, T. Xiang, S. Trebst, and G. Chen, "Thermal Hall signatures of non-Kitaev spin liquids in honeycomb Kitaev materials," Physical Review Research **1**, 013014 (2019) (cited on page 94).

[45] S. Park and B.-J. Yang, "Topological magnetoelastic excitations in noncollinear antiferromagnets," Physical Review B **99**, 174435 (2019) (cited on page 94).

[46] J. M. Losada, "Ultrafast manipulation of Heisenberg exchange and Dzyaloshinskii–Moriya interactions in antiferromagnetic insulators," Master's thesis (NTNU, 2019) (cited on page 97).

Stein Kåre Lundeberg Fosstveit

Spin and Thermal Hall Effects in the Strongly Correlated Kane-Mele-Hubbard Model

NTNU
Norwegian University of
Science and Technology