Title

Electron-Phonon Coupling from GW Perturbation Theory and Electronic and Magnetic Properties of Novel Two-Dimensional Materials

Author

Li, Zhenglu

Publication Date

2019

Electron-Phonon Coupling from *GW* Perturbation Theory and Electronic and Magnetic Properties of Novel Two-Dimensional Materials

by

Zhenglu Li

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Physics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Steven G. Louie, Chair
Professor Alex Zettl
Professor Lin Lin

Summer 2019

Electron-Phonon Coupling from *GW* Perturbation Theory and Electronic and Magnetic Properties of Novel Two-Dimensional Materials

Abstract

Electron-Phonon Coupling from *GW* Perturbation Theory and Electronic and Magnetic Properties of Novel Two-Dimensional Materials

by

Zhenglu Li

Doctor of Philosophy in Physics

University of California, Berkeley

Professor Steven G. Louie, Chair

Condensed matter physics is a very broad and fast-developing field, which studies emerging phenomena, interactions, phases, and symmetries in materials, such as solids. *Predictive* first-principles, or *ab initio*, methodologies play a significant role in understanding various phenomena and new physics. This dissertation is aimed at developing new *ab initio* methodologies for the investigation of important novel phenomena and applying various *ab initio* methods combined with analytical approaches to a broad range of condensed matter systems, including the high-transition-temperature superconductor $Ba_{1-x}K_xBiO_3$, the two-dimensional (2D) ferromagnet $Cr_2Ge_2Te_6$, Dirac fermions generated in few-layer black phosphorus, defects in hexagonal boron nitride, and non-trivial topological surface states of antimony.

This dissertation is divided into two parts. Part I is focused on methods development, and Part II is a collection of theoretical and computational studies of novel materials. The dissertation is organized as follows:

Part I: Electronic structure methodologies for condensed matter

- In Chapter 1, we review some important *ab initio* methods to lay the foundation for the development of a new *ab initio* method − named *GW* perturbation theory (*GW*PT) − in Chapter 2, and for various applications to the materials studied in Part II. In Chapter 1, we review the basics of density functional theory (DFT), the *GW* method, the general phonon formalism and electron-phonon (*e*-ph) coupling formalism, density-functional perturbation theory (DFPT), and the Wannier representation of *e*-ph coupling.

- In Chapter 2, we present a new *ab initio* method, which we named the *GW* perturbation theory (*GW*PT). This method is a linear-response theory of the *GW* method, and it gives efficient and accurate access to all *e*-ph matrix elements at the many-electron

level in the full Brillouin zone and between any pairs of electronic states. We discuss its general formalism, implementation and verification in this Chapter.

- In Chapter 3, we develop a general renormalized spin-wave theory (RSWT) by including full sublattice dependence. This RSWT method includes magnon-magnon interactions, and therefore can give quantitative predictions of magnetic transition temperatures, especially in 2D. This method is solved numerically and self-consistently. We discuss its formalism, implementation, and behavior in this Chapter.

Part II: Studies of superconductivity, and electronic and magnetic interactions in novel materials

- In Chapter 4, we apply our newly developed GWPT method to study superconductivity in $\mathrm{Ba}_{1-x}\mathrm{K}_x\mathrm{BiO}_3$, which shows an experimental superconducting transition temperature (T_c) of $30-32$ K at optimal doping. Our GWPT calculations show that many-electron correlations significantly enhance the e-ph interactions compared to DFPT values for states near the Fermi surface and renormalize the e-ph coupling constant λ by a factor of 2.4, nicely explaining the high T_c as well as the doping dependence observed in this family of material.

- In Chapter 5, we present a collaborative work with experimental groups on the discovery of the 2D van der Waals ferromagnet $\mathrm{Cr}_2\mathrm{Ge}_2\mathrm{Te}_6$, probed using the scanning magneto-optic Kerr effect (MOKE) technique. We apply our RSWT method to this system, and our calculation nicely reproduces and explains the experimentally observed strong dimensionality effect in this 2D ferromagnet. Furthermore, our theory reveals an intriguing interplay between anisotropy and dimensionality, which leads to an unprecedented magnetic-field control of ferromagnetism in this system.

- In Chapter 6, we propose a strategy for the generation of novel anisotropic Dirac fermions in few-layer black phosphorus by applying inversely designed superlattice potentials. We show that these novel quasiparticles exhibit asymmetric Klein tunneling, in which the perfect transmission direction significantly differs from the normal incidence direction. These unusual states are highly tunable and accessible with experimentally achievable conditions. The findings revealed in this Chapter provide new platforms for device design.

- In Chapter 7, we present a collaborative work with an experimental group to study the electron-irradiation-induced triangular and hexagonal defects in hexagonal boron nitride, observed in transmission electron microscopy (TEM) measurements. We use DFT to calculate the formation enthalpy of different structures (as well as the edges and corners), to provide an overall diagram of preferred structures under different conditions at equilibrium. Our theory provides important insights into the formation of these defects.

- In Chapter 8, we present a collaborative work with experimental groups to study the unusual behavior of photoelectrons from the topological surface states of Sb(111), measured with spin- and angle-resolved photoemission spectroscopy (spin-ARPES). Our theory, using the *ab initio* tight-binding method, reproduces well the observed spin textures. Our theoretical analysis shows that the unexpected spin-polarization behavior comes from the interplay between strong spin-orbit coupling (SOC) and the symmetry requirement of the electron wavefunction in high symmetry regions of the Brillouin zone.

To Jing and Lillian,
for the lovely time we have spent together,
and will spend in the rest of my life.

Contents

I Electronic structure methodologies for condensed matter 1

List of Figures

List of Tables

Acknowledgments

First of all, I would like to deeply thank my Ph.D. advisor, Prof. Steven G. Louie. I always feel so grateful, lucky, and honored to complete my doctoral study under your guidance. Your sharp insights and the erudite knowledge of condensed matter physics make my Ph.D. study such a wonderful and joyful journey. Those moments of sitting together around the corner of the table in your office, working out ideas on papers, are always encouraging, creative, and full of determination to find solutions. I benefited endlessly learning from you to be very rigorous, be critical, and be collaborative in research. I have been given the most freedom I can ever imagine in learning and doing research. Your patience, support, and care for me and my family, are invaluable fortune to me. From you, I not only learn science, I also learn how to become a real scientist.

I would like to thank Mrs. Jane Louie for your kindness in preparing the annual Thanksgiving feast, and for the warm care for me and my wife. We feel so blessed.

I would also like to thank the people from the Louie group: Felipe and Diana, as senior students, you two have guided me a lot, and set up good examples for me to learn from. Thank you, and I enjoyed very much the discussions and collaborations with you. Ting and Meng, we probably spend the most time together, not only working out science ideas, but also experiencing abroad life in Berkeley. I feel a lot of fun exploring unknowns with you. Fangzhou and Jingwei, thanks for being thoughtful, and especially aggressive on basketball court. These splendid days will continue. Thank you, Gabriel, discussions with you are always intense and full of joyness. Yang-Hao, thanks for being so kind as my office mate. We often run into discussions on random research topics, and I always find them inspiring. Thank you, Chin-Shen, for always being supportive. It is always joyful to discuss with you. I would like to thank Sinisa, Johannes, Marco, Andrea, Yea-Lee, Liang, Jamal, Derek, and Brad. The times I spend with you and learn from you are full of sweetness and insights. Also, I would like to thank Katherine. The chats with you have taught me a lot of non-scientific knowledge. Arica, thank you for being so supportive to me. Your professionalism impresses me so much, and I try to learn as much as I can.

I would like to thank my experimental collaborators: Prof. Xiang Zhang, Prof. Alessandra Lanzara, Prof. Alex Zettl, Prof. Zi Qiang Qiu, Prof. James Analytis, Dr. Cheng Gong, Dr. Kenneth Gotlieb, and Dr. Thang Pham. Your expertise has expanded the horizon of my knowledge in condensed matter physics.

I would like to thank my undergraduate advisors Prof. Xingao Gong and Prof. Hongjun Xiang at Fudan University. Thank you for leading me into the field of computational condensed matter physics, and providing me with trust the excellent platform for performing research.

I would like to thank my parents for your unconditional love and support. I am proud to be your son.

Finally, I would like to thank my wife, Jing. Thank you for being so supportive and for sharing responsibilities in life. I feel so fortunate marrying you. Now we are full of happiness with the birth of our sweet little girl Lillian. It will be an amazing life in front of us.

Part I

Electronic structure methodologies for condensed matter

Chapter 1

Review of first-principles methodologies for electrons and phonons

1.1 Density-functional theory (DFT)

1.1.1 Kohn-Sham equation

To formulate and solve for the properties of materials, in particular, those of interacting electrons moving among nuclei, one builds a many-body Schrödinger equation involving all electrons and nuclei in the system. Density-functional theory (DFT) [1], as one of the most widely used *ab initio* method, reduces the many-body problem into a *mean-field* tractable problem for electrons, as the Hohenberg-Kohn theorem [2] states that any ground-state property of a system of many interacting electrons can be viewed as a functional of the ground state density $n(\mathbf{r})$. The core of static DFT is to solve self-consistently the Kohn-Sham [3] equation,

$$H^{\mathrm{DFT}}\psi_n(\mathbf{r}) = \left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial \mathbf{r}^2} + V^{\mathrm{KS}}(\mathbf{r})\right)\psi_n(\mathbf{r}) = \varepsilon_n \psi_n(\mathbf{r}), \tag{1.1}$$

where $\psi_n(\mathbf{r})$ is the non-interacting single-electron orbit, commonly used as quasiparticle wavefunctions, and ε_n is the eigenvalue, commonly interpreted as quasiparticle state energy. (These interpretations however can only be viewed as low-order estimates of the physical quantities.) The ground-state electron charge density is given by the single-electron orbit ($\psi_n(\mathbf{r})$) module square of all occupied states,

$$n(\mathbf{r}) = \sum_n^{\mathrm{occ}} |\psi_n(\mathbf{r})|^2. \tag{1.2}$$

In the Kohn-Sham equation, Eq. (1.1), the Kohn-Sham potential is written as,

$$\begin{aligned} V^{\mathrm{KS}}(\mathbf{r}) &= V_{\mathrm{ion}}(\mathbf{r}) + V_{\mathrm{H}}(\mathbf{r}) + V_{\mathrm{xc}}(\mathbf{r}) \\ &= V_{\mathrm{ion}}(\mathbf{r}) + V_{\mathrm{H}}[n(\mathbf{r})] + V_{\mathrm{xc}}[n(\mathbf{r})], \end{aligned} \tag{1.3}$$

which is a functional of the charge density. In the above equation, V_{ion} is the potential from ions (as well as any applied external potentials). V_H is the electrostatic Hartree potential that takes the form,

$$V_{\mathrm{H}}(\mathbf{r}) = e^2 \int d\mathbf{r}' \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}, \tag{1.4}$$

and V_{xc} is the exchange-correlation potential defined as,

$$V_{\mathrm{xc}}(\mathbf{r}) = \frac{\delta E_{\mathrm{xc}}[n(\mathbf{r})]}{\delta n(\mathbf{r})}, \tag{1.5}$$

where E_{xc} is the exchange-correlation energy, which is the most intriguing part and its exact form is still unknown. One of the most common (and also the most basic) approximation to the exchange-correlation energy is known as the local-density approximation (LDA) [4, 5] such that E_{xc} is approximated to depend only on the local density, i.e. only $n(\mathbf{r})$,

$$E_{\mathrm{xc}}^{\mathrm{LDA}}[n] = \int d\mathbf{r} n(\mathbf{r}) \epsilon_{\mathrm{xc}}^{\mathrm{hom}}\left(n(\mathbf{r})\right), \tag{1.6}$$

where $\epsilon_{\mathrm{xc}}^{\mathrm{hom}}(n(\mathbf{r}))$ is approximated to be the exchange-correlation energy per electron in a homogeneous electron gas at density n. The above equations can be solved self-consistently by minimizing the total energy.

1.1.2 Bloch theorem

Most systems that this dissertation focuses on are crystals hosting discrete translational symmetry. As a result, the quantum states are propagating states with n can be re-labeled as $n\mathbf{k}$ where $\mathbf{k}$ is a wavevector, and n denotes the other quantum numbers of the state. The energies of these states in crystals form band structures which are represented by $\varepsilon_{n\mathbf{k}}$. According to the Bloch theorem [6], the wavefunction in crystals is composed of a phase factor multiplied by a periodic function $u_{n\mathbf{k}}(\mathbf{r})$,

$$\psi_{n\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{N_l}} e^{i\mathbf{k}\cdot\mathbf{r}} u_{n\mathbf{k}}(\mathbf{r}), \tag{1.7}$$

where N_l is the number of unit cells under the periodic boundary condition, the phase factor carries a *crystal momentum* $\mathbf{k}$, and the periodic part satisfies,

$$u_{n\mathbf{k}}(\mathbf{r}) = u_{n\mathbf{k}}(\mathbf{r} + \mathbf{R}), \tag{1.8}$$

in which $\mathbf{R}$ is a lattice vector. Given the periodic property of $u_{n\mathbf{k}}(\mathbf{r})$, it is convenient to Fourier transform to reciprocal $\mathbf{G}$ space (this is the spirit of the many plane-wave basis codes),

$$u_{n\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}} c_{n\mathbf{k}}(\mathbf{G}) e^{i\mathbf{G}\cdot\mathbf{r}}, \tag{1.9}$$

where $\mathbf{G}$'s are reciprocal lattice vectors, and the summation is (usually) up to an energy cutoff in practical computations.

1.2 The *GW* method

DFT provides strong *predictive* power in the description of material ground-state properties, such as total energy, equilibrium structures, phonon frequencies and so on. However, static DFT is constructed for the ground state and thus fails in many cases, for example, on the properties of excited states phenomena, or for ground-state properties when the electron correlation becomes too strong [1, 7, 8] because of the approximate nature of current exchange-correlation functionals. A formal route to improve the *ab initio* prediction power is to pursue a many-body field-theoretic approach with the Green's function method.

The GW method has been shown to be very accurate in capturing the many-body correlation effects by evaluating the self-energy operator within the GW level [7–11]. This method was initially developed by Hedin in 1960s [9] for electron gas (nicely reviewed by Hedin and Lundqvist [10]), and was generalized to practical first-principles calculations in 1980s by Hybertsen and Louie [7, 8] showing that it correctly predicts the values of semiconductor and insulator quasiparticle band gaps (where DFT fails significantly). In this section, we review the core of the GW method.

1.2.1 Hedin equations

First, we introduce the one-electron Green's function [9, 10]. We consider a system of N electrons in its ground state and define the one-electron Green's function in coordinate space:

$$G(\mathbf{r}t, \mathbf{r}'t') = -i \langle N| T_t[\psi(\mathbf{r}, t)\psi^+(\mathbf{r}', t')] |N\rangle . \tag{1.10}$$

Here $|N\rangle$ denotes the ground state, and $\psi^+(\mathbf{r}, t)$ and $\psi(\mathbf{r}, t)$ are the Heisenberg field operators for creation and annihilation of an electron, respectively. T_t is the time-ordering operator. For $t > t'$, the Green's function describes the propagation of an additional electron injected at time t', whereas for $t < t'$ it describes the propagation of a hole (extraction of an electron). In the next, we introduce the short-hand index for the combined real-space position and time, such that $1 \equiv (\mathbf{r}_1, t_1)$, $1^+ = (\mathbf{r}_1, t_1 + \delta)$ with positive infinitesimal $\delta = 0^+$ here.

In the many-body problem, the fully interacting Green's function G can be casted into the Dyson's equation [12], with which by starting from a non-interacting Green's function G_0, all unknown interactions can be mapped into the *self-energy* term Σ, which is in general non-local and frequency-dependent. The Dyson's equation reads,

$$G(1,2) = G_0(1,2) + \int d(34) G_0(1,3)\Sigma(3,4)G(4,2). \tag{1.11}$$

The major goal is to solve for the electron self-energy operator Σ. Hedin [9, 10] expanded the self-energy operator in terms of the interacting Green's function and the screened Coulomb interaction W (which is much weaker than the bare Coulomb interaction v), and derived the following complete set of coupled integral equations (together with the Dyson's equation Eq. (1.11)), known as the Hedin equations [9, 10],

$$W(1,2) = v(1,2) + \int d(34) v(1,3)P(3,4)W(4,2), \tag{1.12}$$

$$P(1,2) = -i \int d(34) G(1,3)G(4,1^+)\Gamma(3,4;2), \tag{1.13}$$

$$\Sigma(1,2) = i \int d(34) G(1,3)\Gamma(3,2;4)W(1^+,4), \tag{1.14}$$

$$\Gamma(1,2;3) = \delta(1,2)\delta(1,3) + \int d(4567) \frac{\delta\Sigma(1,2)}{\delta G(4,5)} G(4,6)G(7,5)\Gamma(6,7;3). \tag{1.15}$$

In the above Hedin equations, P denotes the irreducible polarizability, and Γ is the vertex function.

1.2.2 The GW approximation

The above Hedin equations provide a systematic way for evaluating the many-body problem for one-particle excitations, but it is still extremely complicated especially with the involvement of the vertex function Γ. Therefore, Hedin [9, 10] takes an important approximation that only keeps the vertex function in its lowest order (essentially eliminating the vertex function):

$$\Gamma(1,2;3) = \delta(1,2)\delta(1,3). \tag{1.16}$$

This approximation directly generates two consequences. First, the polarizability is now rewritten as

$$P(1,2) = -iG(1,2^+)G(2,1), \tag{1.17}$$

and this form corresponds to the random phase approximation (RPA) for the dielectric matrix. Moreover, now the self-energy Σ is given to first order in G and W as,

$$\Sigma(1,2) = iG(1,2)W(1^+,2). \tag{1.18}$$

This is the well-known GW approximation [9, 10].

It is convenient to work in the frequency (energy) domain by taking advantage of the time transnational symmetry with reducing the two time arguments t and t' to $t = t - t'$. With the Fourier transform,

$$\Sigma(\mathbf{r}, \mathbf{r}'; \varepsilon) = \int dt \; e^{i\varepsilon t} \; \Sigma(\mathbf{r}, \mathbf{r}'; t), \tag{1.19}$$

and pluging Eq. (1.18) into the above equation and using the similar Fourier transforms for G and W, one can work out the following expression,

$$\Sigma(\mathbf{r}, \mathbf{r}'; \varepsilon) = i \int \frac{d\varepsilon'}{2\pi} e^{-i\delta\varepsilon'} G(\mathbf{r}, \mathbf{r}'; \varepsilon - \varepsilon') W(\mathbf{r}, \mathbf{r}'; \varepsilon'). \tag{1.20}$$

Within the GW approximation, we can solve the following quasiparticle equation [7, 9, 10],

$$\left(-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial \mathbf{r}^2} + V_{\text{ion}}(\mathbf{r}) + V_{\text{H}}(\mathbf{r}) \right) \psi_{n\mathbf{k}}(\mathbf{r}) + \int d\mathbf{r}' \Sigma(\mathbf{r}, \mathbf{r}'; \varepsilon_{n\mathbf{k}}^{GW}) \psi_{n\mathbf{k}}(\mathbf{r}') = \varepsilon_{n\mathbf{k}}^{GW} \psi_{n\mathbf{k}}(\mathbf{r}), \tag{1.21}$$

where $\varepsilon_{n\mathbf{k}}^{GW}$ is the solution, which corresponds to the quasiparticle energy within the GW method. In practice, many calculations of $\varepsilon_{n\mathbf{k}}^{GW}$ are done by [7, 8],

$$\varepsilon_{n\mathbf{k}}^{GW} = \varepsilon_{n\mathbf{k}}^{\text{DFT}} + \langle \psi_{n\mathbf{k}} | \Sigma(\varepsilon_{n\mathbf{k}}^{GW}) - V_{\text{xc}} | \psi_{n\mathbf{k}} \rangle \,. \tag{1.22}$$

1.2.3 Practical evaluation of G, W, and Σ

The fully interacting Green's function is unknown, therefore it is common to do a *one-shot* calculation, i.e. using DFT wavefunctions and eigenvalues to construct G, W, and Σ [7, 8, 11]. This is called the G_0W_0 approximation. Note that this G_0 indeed represents the non-interacting G_0 defined above if one were to call DFT as the mean-field starting point, and call $(\Sigma - V_{\text{xc}})$ the interaction part. This practice is justified [7, 8] with two important arguments: 1) it is found that DFT wavefunctions in most cases nicely represent the true quasiparticle wavefunctions, and 2) the negligence of vertex correction and the underestimation of band gaps from DFT (i.e. using DFT eigenvalues to construct W_0) together form an error cancellation. Therefore, practical G_0W_0 results from DFT starting mean field usually agree well with experiments.

The Green's function defined in Eq. (1.10) can be built with the Bloch states if we use the mean-field DFT results [7, 8, 10],

$$G(\mathbf{r}, \mathbf{r}'; \varepsilon) = \sum_{n\mathbf{k}} \frac{\psi_{n\mathbf{k}}(\mathbf{r}) \psi_{n\mathbf{k}}^*(\mathbf{r}')}{\varepsilon - \varepsilon_{n\mathbf{k}} - i\delta_{n\mathbf{k}}}, \tag{1.23}$$

where $\delta_{n\mathbf{k}} = 0^+$ for $\varepsilon_{n\mathbf{k}} < \varepsilon_F$ and $\delta_{n\mathbf{k}} = 0^-$ for $\varepsilon_{n\mathbf{k}} > \varepsilon_F$, and ε_F is the Fermi energy.

The screened Coulomb interaction W within RPA [7, 8, 10, 13] is constructed from the RPA dielectric matrix ϵ from Eq. (1.12) as,

$$W(\mathbf{r},\mathbf{r}';\omega) = \int d\mathbf{r}''\epsilon^{-1}(\mathbf{r},\mathbf{r}'';\omega)v(\mathbf{r}'',\mathbf{r}'), \tag{1.24}$$

where the bare Coulomb interaction is,

$$v(\mathbf{r},\mathbf{r}') = \frac{e^2}{|\mathbf{r}-\mathbf{r}'|}. \tag{1.25}$$

The RPA dielectric matrix therefore is related to the RPA polarizability as,

$$\epsilon(\mathbf{r},\mathbf{r}';\omega) = \delta(\mathbf{r}-\mathbf{r}') - \int d\mathbf{r}''v(\mathbf{r},\mathbf{r}'')P(\mathbf{r}'',\mathbf{r}';\omega). \tag{1.26}$$

The polarizability can be constructed with Eq. (1.13) and Eq. (1.23).

With a plane-wave basis set [8, 13], we can further write the polarizability for an insulator in reciprocal space as,

$$\begin{aligned} P_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega) = \frac{2}{V}\sum_{vc\mathbf{k}} & \langle\psi_{c\mathbf{k}}|e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}}|\psi_{v\mathbf{k}+\mathbf{q}}\rangle \langle\psi_{v\mathbf{k}+\mathbf{q}}|e^{i(\mathbf{q}+\mathbf{G}')\cdot\mathbf{r}'}|\psi_{c\mathbf{k}}\rangle \\ & \times\left(\frac{1}{\omega-(\varepsilon_{c\mathbf{k}}-\varepsilon_{v\mathbf{k}+\mathbf{q}})+i\delta} - \frac{1}{\omega+(\varepsilon_{c\mathbf{k}}-\varepsilon_{v\mathbf{k}+\mathbf{q}})-i\delta}\right), \end{aligned} \tag{1.27}$$

where V is the crystal volume, $\mathbf{q}$ is a wavevector, v and c represent summation over valence and conduction bands, respectively. The dielectric matrix in the reciprocal space then becomes,

$$\epsilon_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega) = \delta_{\mathbf{G}\mathbf{G}'} - v(\mathbf{q}+\mathbf{G})P_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega), \tag{1.28}$$

where

$$v(\mathbf{q}+\mathbf{G}) = \frac{4\pi e^2}{|\mathbf{q}+\mathbf{G}|^2}. \tag{1.29}$$

The screened Coulomb interaction in reciprocal space can be constructed after inverting the dielectric matrix to get $\epsilon^{-1}_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega)$, that is,

$$W_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega) = \epsilon^{-1}_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega)v(\mathbf{q}+\mathbf{G}'). \tag{1.30}$$

The frequency dependence in the dielectric matrix can be either sampled directly (the full-frequency calculation) which may be expensive, or treated using models based on the static dielectric matrix, i.e. with $\omega = 0$, that is directly calculated. In this dissertation, we often adopt the Hybertsen-Louie generalized plasmon-pole model [8] for the frequency convolution in Eq. (1.20).

Finally, we are interested in calculating the quasiparticle energies using Eq. (1.22), in which an important ingredient, the matrix element of Σ, is needed. Using the Fourier transform relation of W, Eq. (1.20), and Eq. (1.23), we arrive at [8],

$$\begin{aligned}&\langle\psi_{n\mathbf{k}}|\Sigma(\mathbf{r},\mathbf{r}';\varepsilon)|\psi_{n\mathbf{k}}\rangle\\ =\ &\frac{i}{2\pi}\sum_{n'}\sum_{\mathbf{qGG'}}\langle\psi_{n\mathbf{k}}|e^{i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}}|\psi_{n'\mathbf{k}-\mathbf{q}}\rangle\,\langle\psi_{n'\mathbf{k}-\mathbf{q}}|e^{-i(\mathbf{q}+\mathbf{G}')\cdot\mathbf{r}'}|\psi_{n\mathbf{k}}\rangle\\ &\times\int d\varepsilon'\frac{W_{\mathbf{GG}'}(\mathbf{q},\varepsilon')e^{-i\delta\varepsilon'}}{\varepsilon-\varepsilon_{n'\mathbf{k}-\mathbf{q}}-i\delta_{n'\mathbf{k}-\mathbf{q}}-\varepsilon'}.\end{aligned}\tag{1.31}$$

The quasiparticle band energies within the GW method can then be calculated using Eq. (1.22) with the matrix elements obtained in the above equation.

1.3 Phonons and electron-phonon coupling

1.3.1 Phonons: lattice vibrations in crystals

In crystals, quantum lattice vibrations are commonly treated in the language of *phonons* [14–16]. As mentioned before, a fundamental treatment of crystals is done with periodic boundary conditions, also known as the Born-von Kármán (BvK) conditions. Under these conditions, a BvK supercell contains N_l unit cells which corresponds to the number of degrees of freedom in wavevectors ($\mathbf{k}$, $\mathbf{q}$ points). These unit cells are positioned at $\mathbf{R}_l$, which are the lattice vectors. The summation over all BvK supercells are *implicitly included* in this dissertation, or equivalently, by conceptually *imposing* BvK boundary conditions for all quantities.

Now we define the notations [16] for ions, which are the core objects for phonons. Within each unit cell, the position of the nucleus κ with mass M_κ is defined as $\boldsymbol{\tau}_\kappa$ with three Cartesian components $\alpha = x, y, z$. We also introduce the position vectors for all ions in the BvK supercell as $\boldsymbol{\tau}_{\kappa l} = \mathbf{R}_l + \boldsymbol{\tau}_\kappa$. We denote the set of equilibrium positions as $\{\boldsymbol{\tau}^0_{\kappa l}\}$.

Within the Born-Oppenheimer and harmonic approximations [16], the total potential energy of electrons and nuclei U can be expanded over small atomic displacements away from equilibrium positions,

$$U = U_0 + \frac{1}{2}\sum_{\substack{\kappa\alpha l\\ \kappa'\alpha' l'}} C_{\kappa\alpha l,\kappa'\alpha' l'}\Delta\tau_{\kappa\alpha l}\Delta\tau_{\kappa'\alpha' l'},\tag{1.32}$$

where U_0 is the total potential energy at equilibrium, and the interatomic force constants are defined as,

$$C_{\kappa\alpha l,\kappa'\alpha' l'} = \frac{\partial^2 U}{\partial\tau_{\kappa\alpha l}\partial\tau_{\kappa'\alpha' l'}}.\tag{1.33}$$

The Fourier transform of the interatomic force constants is the dynamical matrix, a central quantity in phonon calculations that takes into account the crystal translational symmetry. The dynamical matrix at a given phonon wavevector $\mathbf{q}$ is,

$$D_{\kappa\alpha,\kappa'\alpha'}(\mathbf{q}) = (M_\kappa M_{\kappa'})^{-1/2} \sum_{l'} C_{\kappa\alpha 0,\kappa'\alpha' l'}\ e^{i\mathbf{q}\cdot\mathbf{R}_{l'}}. \tag{1.34}$$

Diagonalization [16] of the dynamical matrix yields the phonon frequencies and eigenvectors, such that,

$$\sum_{\kappa'\alpha'} D_{\kappa\alpha,\kappa'\alpha'}(\mathbf{q}) e_{\kappa'\alpha',\nu}(\mathbf{q}) = \omega^2_{\mathbf{q}\nu} e_{\kappa\alpha,\nu}(\mathbf{q}), \tag{1.35}$$

where ν labels the phonon branches, and the square root of the eigenvalue $\omega_{\mathbf{q}\nu}$ is the phonon frequency. When a structure is stable, all $\omega^2_{\mathbf{q}\nu}$ would be equal or greater than zero. When negative $\omega^2_{\mathbf{q}\nu}$ exists, this is the case with imaginary phonon frequency, indicating that the structure is unstable and the system tends to form a structural transition.

By introducing creation and annihilation operators for phonons $a^\dagger_{\mathbf{q}\nu}$ and $a_{\mathbf{q}\nu}$ [16], the lattice vibrations become quantized, known as phonons. These operators obey Bose-Einstein statistics,

$$\begin{aligned} &[a_{\mathbf{q}\nu}, a^\dagger_{\mathbf{q}'\nu'}] = \delta_{\nu\nu'}\delta_{\mathbf{q}\mathbf{q}'}, \\ &[a_{\mathbf{q}\nu}, a_{\mathbf{q}'\nu'}] = [a^\dagger_{\mathbf{q}\nu}, a^\dagger_{\mathbf{q}'\nu'}] = 0. \end{aligned} \tag{1.36}$$

Importantly, the atom displacement can be written in terms of these operators,

$$\Delta\tau_{\kappa\alpha l} = \frac{1}{\sqrt{N_l M_\kappa}} \sum_{\mathbf{q}\nu} \sqrt{\frac{\hbar}{2\omega_{\mathbf{q}\nu}}} e_{\kappa\alpha,\nu}(\mathbf{q}) e^{i\mathbf{q}\cdot\mathbf{R}_l} (a_{\mathbf{q}\nu} + a^\dagger_{-\mathbf{q}\nu}). \tag{1.37}$$

The phonon Hamiltonian can be written as [16],

$$H_{\text{ph}} = \sum_{\mathbf{q}\nu} \hbar\omega_{\mathbf{q}\nu}(a^\dagger_{\mathbf{q}\nu} a_{\mathbf{q}\nu} + \frac{1}{2}). \tag{1.38}$$

1.3.2 Electron-phonon coupling Hamiltonian based on DFT

Based on DFT and assuming that the Kohn-Sham eigenvalues and eigenfunctions are the corresponding quasiparticle quantities, we derive the electron-phonon (e-ph) Hamiltonian [16]. In this context, electron-phonon coupling means when an atom displacement or phonon perturbation is introduced into the system, how the electron states would respond to this perturbation. Since rigorously the Kohn-Sham states are fictitious independent particles that are used to construct only the density, the DFT electron-phonon coupling derived below is just an approximation to the real electron-phonon coupling. Recall that within DFT, the governing equation for the fictitious independent electrons is the Kohn-Sham equation, Eq.

(1.1). Now we would explicitly include the atom position dependence in the Kohn-Sham potential $V^{\mathrm{KS}}(\mathbf{r};\{\tau_{\kappa\alpha l}\})$. In terms of second quantization, we write the electron Hamiltonian as [16],

$$H_e^{\mathrm{DFT}} = \sum_{n\mathbf{k},n'\mathbf{k}'} \langle\psi_{n\mathbf{k}}|\, H^{\mathrm{DFT}}\, |\psi_{n'\mathbf{k}'}\rangle\, c_{n\mathbf{k}}^{\dagger} c_{n'\mathbf{k}'} = \sum_{n\mathbf{k}} \varepsilon_{n\mathbf{k}}^{\mathrm{DFT}} c_{n\mathbf{k}}^{\dagger} c_{n\mathbf{k}}, \tag{1.39}$$

where $c_{n\mathbf{k}}^{\dagger}$ and $c_{n'\mathbf{k}'}$ are the creation and annihilation operators for Kohn-Sham electrons. The implicit approximation made here is that the electronic system can be described by well-defined quasiparticles and they are given the Kohn-Sham eigenvalues and eigenfunctions.

Expand the Kohn-Sham potential near the equilibrium atom positions up to the first order, i.e. within linear response, we have,

$$V^{\mathrm{KS}}(\{\tau_{\kappa\alpha l}\}) = V^{KS}(\{\tau_{\kappa\alpha l}^{0}\}) + \sum_{\kappa\alpha l} \frac{\partial V^{\mathrm{KS}}}{\partial\tau_{\kappa\alpha l}} \Delta\tau_{\kappa\alpha l}. \tag{1.40}$$

We denote the last term in the above equation as $\Delta V^{\mathrm{KS}}(\{\tau_{\kappa\alpha l}\})$, and it can be rewritten as phonon mode-decomposed perturbation by plugging Eq. (1.37) into the above expression [16],

$$\begin{aligned}
\Delta V^{\mathrm{KS}}(\{\tau_{\kappa\alpha l}\}) &= \sum_{\kappa\alpha l} \frac{\partial V^{\mathrm{KS}}(\{\tau_{\kappa\alpha l}\})}{\partial\tau_{\kappa\alpha l}} \Delta\tau_{\kappa\alpha l} \\
&= \sum_{\kappa\alpha l} \frac{1}{N_l\sqrt{M_\kappa}} \sum_{\mathbf{q}\nu} \sqrt{\frac{\hbar}{2\omega_{\mathbf{q}\nu}}} e_{\kappa\alpha,\nu}(\mathbf{q}) e^{i\mathbf{q}\cdot\mathbf{R}_l} (a_{\mathbf{q}\nu} + a_{-\mathbf{q}\nu}^{\dagger}) \frac{\partial V^{\mathrm{KS}}(\{\tau_{\kappa\alpha l}\})}{\partial\tau_{\kappa\alpha l}} \\
&= \frac{1}{\sqrt{N_l}} \sum_{\mathbf{q}\nu} (a_{\mathbf{q}\nu} + a_{-\mathbf{q}\nu}^{\dagger}) \Delta_{\mathbf{q}\nu} V^{\mathrm{KS}}.
\end{aligned} \tag{1.41}$$

In this above expression, we have introduced the phonon-mode differential perturbation operator $\Delta_{\mathbf{q}\nu}$, which is defined as [17],

$$\Delta_{\mathbf{q}\nu} = \sqrt{\frac{\hbar}{2\omega_{\mathbf{q}\nu}}} \sum_{\kappa\alpha} \frac{1}{\sqrt{M_\kappa}} e_{\kappa\alpha,\nu}(\mathbf{q}) \sum_{l}^{N_l} e^{i\mathbf{q}\cdot\mathbf{R}_l} \frac{\partial}{\partial\tau_{\kappa\alpha l}}. \tag{1.42}$$

It is very important to note that the operator $\Delta_{\mathbf{q}\nu}$ defined here carries the crystal momentum of $\mathbf{q}$, which means that $\Delta_{\mathbf{q}\nu}V^{\mathrm{KS}}$ is just in a form as a phase factor $e^{i\mathbf{q}\cdot\mathbf{r}}$ multiplied by a unit-cell lattice-periodic function (see next Section 1.4 for detailed analysis). Importantly, we note that the perturbations from phonons to the Kohn-Sham potential can be decomposed into independent components with respect to each phonon mode as $\Delta_{\mathbf{q}\nu}V^{\mathrm{KS}}$.

The e-ph Hamiltonian is defined as [16]

$$H_{e\text{-ph}}^{\mathrm{DFT}} = \sum_{n\mathbf{k},n'\mathbf{k}'} \langle\psi_{n\mathbf{k}}|\, \Delta V^{\mathrm{KS}}(\{\tau_{\kappa\alpha l}\})\, |\psi_{n'\mathbf{k}'}\rangle\, c_{n\mathbf{k}}^{\dagger} c_{n'\mathbf{k}'}. \tag{1.43}$$

By plugging Eq. (1.41) into the above equation and applying crystal momentum conservation, we arrive at,

$$H_{e\text{-ph}}^{\mathrm{DFT}} = \frac{1}{\sqrt{N_l}} \sum_{mn\nu} \sum_{\mathbf{kq}} g_{mn\nu}^{\mathrm{DFT}}(\mathbf{k},\mathbf{q}) c_{m\mathbf{k}+\mathbf{q}}^{\dagger} c_{n\mathbf{k}} (a_{\mathbf{q}\nu} + a_{-\mathbf{q}\nu}^{\dagger}), \tag{1.44}$$

in which the e-ph matrix element at the DFT level, is defined as,

$$g_{mn\nu}^{\mathrm{DFT}}(\mathbf{k},\mathbf{q}) = \langle \psi_{m\mathbf{k}+\mathbf{q}} | \Delta_{\mathbf{q}\nu} V^{\mathrm{KS}} | \psi_{n\mathbf{k}} \rangle . \tag{1.45}$$

The e-ph matrix element $g_{mn\nu}(\mathbf{k},\mathbf{q})$ is the core quantity in e-ph physics, because it serves as the building blocks of all microscopic formulations of various important phenomena such as phonon-induced superconductivity [18], electrical and thermal transport [19], temperature-dependent direct and indirect optical absorption [20], and so on. Each matrix element $g_{mn\nu}(\mathbf{k},\mathbf{q})$ represents the scattering magnitude from the initial state $|\psi_{n\mathbf{k}}\rangle$ to the final state $|\psi_{m\mathbf{k}+\mathbf{q}}\rangle$ by a phonon mode labeled by $\mathbf{q}\nu$. Note that in the integral in Eq. (1.45), the two wavefunctions and the first-order change in Kohn-Sham potential all have phases that can be factored out and they together conserve the crystal momentum, therefore the integral is indeed done with only unit-cell lattice-periodic functions. Consequently, the e-ph matrix elements $g_{mn\nu}^{\mathrm{DFT}}(\mathbf{k},\mathbf{q})$ can be calculated within *primitive unit cell* calculations from first principles [15, 16]. We will discuss this feature in more details in the next Section 1.4.

1.4 Density-functional perturbation theory (DFPT)

In this section, we review density-functional perturbation theory (DFPT)[15, 21–23], which is a linear response theory of DFT. DFPT takes the advantage that it *self-consistently* calculates the change (response function) in charge density, potential, wavefunctions, etc. within a single *primitive unit cell*, enabled by linear response theory. Phonon [15] is among the most important applications of DFPT, and any phonon mode $\mathbf{q}\nu$ is independent from each other, and the calculation time is about the same, also similar to its corresponding single DFT self-consistent calculation.

1.4.1 Sternheimer equation

By linearizing the Kohn-Sham equation in Eq. (1.1), we can derive the Sternheimer equation [24], which is the main equation in DFPT [15, 21, 22, 24, 25],

$$P_c^{\mathbf{k}+\mathbf{q}} \left(H^{\mathrm{DFT}} - \varepsilon_{v\mathbf{k}} \right) P_c^{\mathbf{k}+\mathbf{q}} \left| \Delta_{\mathbf{q}\nu} \psi_{v\mathbf{k}} \right\rangle = -P_c^{\mathbf{k}+\mathbf{q}} \Delta_{\mathbf{q}\nu} V^{\mathrm{KS}} \left| \psi_{v\mathbf{k}} \right\rangle , \tag{1.46}$$

where v and c represent valence and conduction bands, respectively, and

$$P_c^{\mathbf{k}+\mathbf{q}} = \sum_c \left| \psi_{c\mathbf{k}+\mathbf{q}} \right\rangle \left\langle \psi_{c\mathbf{k}+\mathbf{q}} \right| . \tag{1.47}$$

The projection operator $P_c^{\mathbf{k}+\mathbf{q}}$ is to remove the singularity in $\left(H^{\mathrm{DFT}} - \varepsilon_{v\mathbf{k}}\right)$ [15]. Note that the separation using valence (v) and conduction (c) states are most commonly seen in the literature. This is because, as we will see in the next, the first-order change in the charge density $\Delta_{\mathbf{q}\nu} n(\mathbf{r})$ and Kohn-Sham potential $\Delta_{\mathbf{q}\nu} V^{\mathrm{KS}}(\mathbf{r})$ requires only first-order change in valence wavefunctions $\Delta_{\mathbf{q}\nu}\psi_{v\mathbf{k}}(\mathbf{r})$. For metals, a similar formalism applies by differentiating $\mathbf{k}$-point dependent occupied and unoccupied states, and practically a smearing scheme is often adopted [15]. This separation into two orthogonal subspaces can be generalized in other way [25], and we will discuss this generalization in the next chapter in Section 2.3.2 for the development of GW perturbation theory (GWPT) [17]. The first-order change in charge density is [15],

$$\begin{aligned}\Delta_{\mathbf{q}\nu} n(\mathbf{r}) &= 2\sum_{v\mathbf{k}}\Big(\psi_{v\mathbf{k}}^*(\mathbf{r})\Delta_{\mathbf{q}\nu}\psi_{v\mathbf{k}}(\mathbf{r}) + [\Delta_{-\mathbf{q}\nu}\psi_{v\mathbf{k}}(\mathbf{r})]^*\psi_{v\mathbf{k}}(\mathbf{r})\Big)\\ &= 4\sum_{v\mathbf{k}}\psi_{v\mathbf{k}}^*(\mathbf{r})\Delta_{\mathbf{q}\nu}\psi_{v\mathbf{k}}(\mathbf{r}),\end{aligned} \tag{1.48}$$

where we have used the fact that the summation over $\mathbf{k}$ is in the full Brillouin zone, and have applied the following relations,

$$\Delta_{\mathbf{q}\nu}[\psi_{n\mathbf{k}}^*(\mathbf{r})] = [\Delta_{-\mathbf{q}\nu}\psi_{n\mathbf{k}}(\mathbf{r})]^*, \tag{1.49}$$

and

$$[\Delta_{-\mathbf{q}\nu}\psi_{n-\mathbf{k}}(\mathbf{r})]^* = \Delta_{\mathbf{q}\nu}\psi_{n\mathbf{k}}(\mathbf{r}), \tag{1.50}$$

due to *time-reversal symmetry*, which is one property of the systems we restricted to (i.e. non-magnetic systems) when using DFPT and also later GWPT. In case of magnetic materials, time-reversal symmetry is broken, and the above formalism will need to be generalized because $\mathbf{q}$ and $-\mathbf{q}$ become no longer time-reversal related. The first-order change in the Kohn-Sham potential, by linearizing Eq. (1.3), can be written as [15],

$$\Delta_{\mathbf{q}\nu} V^{\mathrm{KS}}(\mathbf{r}) = \Delta_{\mathbf{q}\nu} V_{\mathrm{ion}}(\mathbf{r}) + \Delta_{\mathbf{q}\nu} V^{\mathrm{H}}[\Delta_{\mathbf{q}\nu} n(\mathbf{r})] + V^{\mathrm{xc}}[\Delta_{\mathbf{q}\nu} n(\mathbf{r})]. \tag{1.51}$$

The three equations Eq. (1.46), Eq. (1.48), and Eq. (1.51) form a set of closed equations, from which the three unknown quantities $\Delta_{\mathbf{q}\nu} n(\mathbf{r})$, $\Delta_{\mathbf{q}\nu} V^{\mathrm{KS}}(\mathbf{r})$, and $\Delta_{\mathbf{q}\nu}\psi_{v\mathbf{k}}(\mathbf{r})$ can be solved self-consistently [21].

1.4.2 Factorization of crystal momentum phase factors and calculation within primitive unit cell

We would discuss one important essence here, as we mentioned before, that the whole set of DFPT equations based on the Sternheimer equation can be solved in a *primitive unit cell* for phonon perturbations [15, 16, 21–23, 25]. The key is to factor out the phase factor. We already know this fact for the wavefunctions in the Bloch form Eq. (1.7) , and this is true for the change in wavefunctions as well from the definition of the $\Delta_{\mathbf{q}\nu}$ operator Eq. (1.37), so is the change in the charge density $\Delta_{\mathbf{q}\nu}n$. Now we show that for $\Delta_{\mathbf{q}\nu}V^{\mathrm{KS}}(\mathbf{r})$, we also can write it as [15, 16],

$$\Delta_{\mathbf{q}\nu}V^{\mathrm{KS}}(\mathbf{r}) = e^{i\mathbf{q}\cdot\mathbf{r}}\Delta v^{\mathrm{KS}}_{\mathbf{q}\nu}, \tag{1.52}$$

where,

$$\Delta v^{\mathrm{KS}}_{\mathbf{q}\nu}(\mathbf{r}) = \sqrt{\frac{\hbar}{2\omega_{\mathbf{q}\nu}}}\sum_{\kappa\alpha}\frac{1}{\sqrt{M_\kappa}}e_{\kappa\alpha,\nu}(\mathbf{q})\sum_{l}e^{-i\mathbf{q}\cdot(\mathbf{r}-\mathbf{R}_l)}\frac{\partial V^{\mathrm{KS}}(\mathbf{r}-\mathbf{R}_l)}{\partial\tau_{\kappa\alpha 0}}, \tag{1.53}$$

which is a unit-cell lattice periodic function satisfying,

$$\Delta v^{\mathrm{KS}}_{\mathbf{q}\nu}(\mathbf{r}) = \Delta v^{\mathrm{KS}}_{\mathbf{q}\nu}(\mathbf{r}+\mathbf{R}). \tag{1.54}$$

In Eq. (1.53), we have utilized the lattice periodicity of $V^{\mathrm{KS}}(\mathbf{r})$ that gives,

$$\frac{\partial V^{\mathrm{KS}}(\mathbf{r})}{\partial\tau_{\kappa\alpha l}} = \frac{\partial V^{\mathrm{KS}}(\mathbf{r}-\mathbf{R}_l)}{\partial\tau_{\kappa\alpha 0}}. \tag{1.55}$$

Therefore, all quantities within the linear-response DFPT can be formulated and calculated in a single primitive unit cell with arbitrary phonon perturbation with wavevector $\mathbf{q}$.

It is noted that DFPT can serve as a straightforward and very convenient method to calculate the e-ph matrix elements $g^{\mathrm{DFT}}_{mn\nu}(\mathbf{k},\mathbf{q})$ in Eq. (1.45), once the first-order change in Kohn-Sham potential $\Delta_{\mathbf{q}\nu}V^{\mathrm{KS}}(\mathbf{r})$ is solved in DFPT [16]. Equivalently, we can write the e-ph matrix element as,

$$g^{\mathrm{DFT}}_{mn\nu}(\mathbf{k},\mathbf{q}) = \langle u_{m\mathbf{k}+\mathbf{q}}|\Delta v^{\mathrm{KS}}_{\mathbf{q}\nu}|u_{n\mathbf{k}}\rangle_{\text{unit cell}}, \tag{1.56}$$

where note that the spatial integral is done with all lattice-periodic functions, and therefore the integral is performed within one primitive unit cell [16]. Till now, we can study the e-ph physics from first principles at the DFT level.

1.5 Electron-phonon coupling in Wannier representation

Now an important ingredient for e-ph physics, the e-ph matrix elements $g^{\mathrm{DFT}}_{mn\nu}(\mathbf{k},\mathbf{q})$, can be calculated from *ab initio* DFPT method. However, it is found that often the numerical convergence of various e-ph properties requires very fine $\mathbf{k}$ and $\mathbf{q}$ mesh and straight forward application of DFPT to all the fine $\mathbf{k}$ and $\mathbf{q}$ points is very computationally intensive..

To overcome this difficulty, Wannier representation [26–28] of e-ph coupling has been developed and used to very efficiently interpolate the e-ph matrix elements [16, 29, 30] from a coarse to a fine grid while keeping high *ab initio* accuracy. Wannier interpolation for e-ph coupling is a rather physical interpolation (than a numerical interpolation) because it captures the dominant characters in the Bloch states with proper Wannier functions.

Furthermore, by specifically using the maximally localized Wannier functions [26–28] which decay relatively fast in real space, the interpolation becomes very efficient with a reasonable coarse grid in the DFT and DFPT calculations.

We introduce the e-ph matrix elements in the Wannier representation [30], $g_{st\kappa\alpha}(\mathbf{R}_e, \mathbf{R}_p)$, where s and t label the Wannier functions, and $\mathbf{R}_e$ and $\mathbf{R}_p$ are lattice vectors for electron Wannier functions and phonon Wannier functions (i.e. atom positions), respectively. We introduce the basis transformation matrix between electron Bloch waves and electron Wannier functions as $U_{nt\mathbf{k}}$ which can be derived by projecting DFT wavefunctions onto Wannier basis [26–28]. Similarly, we introduce the basis transformation matrix between phonon eigenmodes and phonon Wannier functions, which are just the phonon eigenvectors. Then the e-ph matrix elements in the Bloch representation can be written in the Wannier representation as [16, 30],

$$g_{mn\nu}(\mathbf{k}, \mathbf{q}) = \sum_{\mathbf{R}_e \mathbf{R}_p} e^{i\mathbf{k}\cdot\mathbf{R}_e + i\mathbf{q}\cdot\mathbf{R}_p} \sum_{st\kappa\alpha} \sqrt{\frac{\hbar}{2M_\kappa \omega_{\mathbf{q}\nu}}} U_{ms\mathbf{k}+\mathbf{q}} g_{st\kappa\alpha}(\mathbf{R}_e, \mathbf{R}_p) U^\dagger_{tn\mathbf{k}} e_{\kappa\alpha,\nu}(\mathbf{q}), \tag{1.57}$$

and the inverse transformation is [16, 30],

$$g_{st\kappa\alpha}(\mathbf{R}_e, \mathbf{R}_p) = \frac{1}{N_e N_p} \sum_{\mathbf{k}\mathbf{q}} e^{-i\mathbf{k}\cdot\mathbf{R}_e - i\mathbf{q}\cdot\mathbf{R}_p} \sum_{mn\nu} \sqrt{\frac{2M_\kappa \omega_{\mathbf{q}\nu}}{\hbar}} e^*_{\kappa\alpha,\nu}(\mathbf{q}) U^\dagger_{sm\mathbf{k}+\mathbf{q}} g_{mn\nu}(\mathbf{k}, \mathbf{q}) U_{nt\mathbf{k}}, \tag{1.58}$$

where N_e and N_p have the same meaning as N_l, but are for electrons and phonons BvK supercells, respectively.

With these two relations, we can transform $g_{mn\nu}(\mathbf{k}_{\mathrm{co}}, \mathbf{q}_{\mathrm{co}})$ on uniform coarse grids (denoted by subscript ‘co’) into the Wannier basis $g_{st\kappa\alpha}(\mathbf{R}_e, \mathbf{R}_p)$ using Eq. (1.58), and then perform the interpolation by transforming back to get $g_{mn\nu}(\mathbf{k}_{\mathrm{fi}}, \mathbf{q}_{\mathrm{fi}})$ on arbitrary points or fine grids (denoted by subscript ‘fi’) by using Eq. (1.57). With this capability, e-ph physics quantities can be calculated efficiently with *ab initio* accuracy.

Chapter 2

Development of GW perturbation theory (GWPT) for electron-phonon interactions at many-electron level

In this Chapter, we develop a *new* first-principles method called GW perturbation theory (GWPT) [17, 31]. It is a linear-response theory of the GW self-energy, in similar spirit of density-functional perturbation theory, but at the GW level that thus gives the coupling of phonons to real quasiparticles as opposed to Kohn-Sham fictitious electrons. In the GWPT method, we deal with the non-local and frequency-dependent change in self-energy. With GWPT, one can access efficiently and accurately the e-ph matrix elements at the many-electron GW level. We present the formalism, implementations, verification, and discussions in this Chapter. Its application to a correlated superconductor with high transition temperature is presented in Chapter 4. The formalism of GWPT is first published in Ref. [17], and the manuscript with more details and discussions is in preparation [31].

2.1 Introduction

First-principles calculation of electron-phonon (e-ph) coupling [16, 17] is of tremendous interest as it serves as a non-empirical approach to predict and understand a number of phenomena in condensed matter physics and materials physics, such as phonon-mediated superconductivity [18], electrical and thermal transport [19], quasiparticle energy renormalization [32], charge-density wave (CDW) [33], and vibrational features in optical spectra [20]. By formulating a linear-response theory of density functional theory (DFT) [1] to phonon perturbations, density-functional perturbation theory (DFPT) [15, 21–23] has been the prevailing and most efficient *ab initio* method to study the e-ph interactions within DFT. The e-ph coupling treated in DFPT is at the same level as the DFT Kohn-Sham eigenvalues which are not the true electron (or quasiparticle) energies. This is reflected in that, in general, the Kohn-Sham eigenvalues do not yield accurate band gaps and band widths nor information on lifetimes [7, 8]. The exchange-correlation potentials V_{xc} in DFT (such as those in the local-density approximation (LDA) [4] or the generalized gradient approximation (GGA) [5]) can only be at best considered as an approximation to the nonlocal, frequency-dependent self-energy operator .

The *GW* approximation [7–11] has proven, for many materials, to be an accurate *ab initio* method in capturing the many-electron correlation effects in the evaluation of the quasiparticle energies. In the *GW* approximation, the self-energy operator Σ is expanded in terms of the single-particle Greens function G and the screened Coulomb interaction W to first order, i.e. $\Sigma = iGW$, hence named the *GW* method. By combining frozen-phonon technique with *GW* calculations, previous studies [34–40] have found that many-electron corrections to the DFT *e*-ph coupling strength are essential to accurately describe a number of phenomena, such as the phonon dispersion in graphene and graphite [34, 35], the temperature-dependent band gap in diamond [38], and superconductivity in $\mathrm{Ba_{0.6}K_{0.4}BiO_3}$ [37]. However, the frozen-phonon technique is limited to only investigate couplings to phonon wavevectors that are commensurate to a large supercell, which makes it prohibitive to achieve a fine sampling of the Brillouin zone. More importantly, frozen-phonon calculations can only provide some intra-band part of the *e*-ph matrix elements indirectly and an overall *e*-ph coupling strength by examining band energy shifts; the *e*-ph matrix elements among all bands and across the full Brillouin zone the essential ingredient of microscopic *e*-ph formulations of many physical phenomena are not available [16, 37, 39, 40]. The importance of self-energy effects in *e*-ph coupling and severe limitations of the frozen-phonon *GW* technique thus point to a strong necessity for a linear-response *GW* theory (similar in spirit as DFPT [15, 21–23]) to efficiently and accurately calculate the quasiparticle *e*-ph interactions at the *GW* level [16, 37, 39, 40].

In this Chapter, we present our development [17] of the first-principles linear-response *GW* method to external perturbations, which we call the *GW* perturbation theory (*GW*PT). In this scheme, the first-order change of the self-energy operator to a phonon perturbation $\Delta_{\mathbf{q}\nu}\Sigma$ is constructed from a linear-response calculation, which is performed within a single primitive unit cell for any phonon wavevector $\mathbf{q}$ and phonon branch ν. This method avoids the use of supercells, and the computational cost naturally scales linearly with the number of phonon modes needed. It provides the *e*-ph matrix elements at the *GW* level for any pairs of electronic states efficiently, making *GW*PT a desirable *ab initio* method to systematically study *e*-ph interactions including many-electron self-energy effects.

2.2 Formalism of *GW* perturbation theory (*GW*PT)

We derive the *GW*PT formalism in this section.

2.2.1 Derivation starting from quasiparticle equation

We start from the quasiparticle equation Eq. (1.21). By linearizing Eq. (1.21), and by generalizing the e-ph Hamiltonian at the DFT level to the GW level, we can have,

$$
\begin{aligned}
H^{GW}_{\text{e-ph}} &= \sum_{n\mathbf{k},n'\mathbf{k}'} \langle\psi_{n\mathbf{k}}|\, \Delta V_{\text{ion}}(\{\tau_{\kappa\alpha l}\}) + \Delta V_{\text{H}}(\{\tau_{\kappa\alpha l}\}) + \Delta\Sigma(\{\tau_{\kappa\alpha l}\})\, |\psi_{n'\mathbf{k}'}\rangle\, c^{\dagger}_{n\mathbf{k}} c_{n'\mathbf{k}'} \\
&= \sum_{n\mathbf{k},n'\mathbf{k}'} \langle\psi_{n\mathbf{k}}|\, \Delta V^{\text{KS}}(\{\tau_{\kappa\alpha l}\}) - \Delta V_{\text{xc}}(\{\tau_{\kappa\alpha l}\}) + \Delta\Sigma(\{\tau_{\kappa\alpha l}\})\, |\psi_{n'\mathbf{k}'}\rangle\, c^{\dagger}_{n\mathbf{k}} c_{n'\mathbf{k}'} \\
&= \frac{1}{\sqrt{N_l}} \sum_{mn\nu} \sum_{\mathbf{kq}} \left[g^{\text{DFT}}_{mn\nu}(\mathbf{k},\mathbf{q}) - \langle\psi_{m\mathbf{k}+\mathbf{q}}|\, \Delta_{\mathbf{q}\nu} V_{\text{xc}}\, |\psi_{n\mathbf{k}}\rangle \right] c^{\dagger}_{m\mathbf{k}+\mathbf{q}} c_{n\mathbf{k}} (a_{\mathbf{q}\nu} + a^{\dagger}_{-\mathbf{q}\nu}) \\
&\quad + \sum_{n\mathbf{k},n'\mathbf{k}'} \langle\psi_{n\mathbf{k}}|\, \Delta\Sigma\, |\psi_{n'\mathbf{k}'}\rangle\, c^{\dagger}_{n\mathbf{k}} c_{n'\mathbf{k}'},
\end{aligned}
\tag{2.1}
$$

The goal is to write the last term in the last line of Eq. (2.1) into phonon mode decomposition associated with $c^{\dagger}_{m\mathbf{k}+\mathbf{q}} c_{n\mathbf{k}} (a_{\mathbf{q}\nu} + a^{\dagger}_{-\mathbf{q}\nu})$ for coupling between electrons and phonons.

Note that unlike $V^{\text{KS}}(\mathbf{r})$ (in which $V_{\text{xc}}(\mathbf{r})$ is one part) which is a local potential, the self-energy operator $\Sigma(\mathbf{r},\mathbf{r}')$ is a *non-local* operator. Similar to Eq. (1.41) , the total change in self-energy $\Delta\Sigma$ due to atom displacements is defined as,

$$
\begin{aligned}
\Delta\Sigma(\mathbf{r},\mathbf{r}';\{\tau_{\kappa\alpha l}\}) &= \Sigma(\mathbf{r},\mathbf{r}';\{\tau_{\kappa\alpha l}\}) - \Sigma(\mathbf{r},\mathbf{r}';\{\tau^{0}_{\kappa\alpha l}\}) \\
&= \sum_{\kappa\alpha l} \frac{\partial\Sigma(\mathbf{r},\mathbf{r}';\{\tau_{\kappa\alpha l}\})}{\partial\tau_{\kappa\alpha l}} \Delta\tau_{\kappa\alpha l}.
\end{aligned}
\tag{2.2}
$$

2.2.2 Constant-screening approximation

From the time and position domain Eq. (1.18), the change in the self-energy operator in principle involves two terms,

$$\Delta\Sigma = i\Delta G W + iG\Delta W = i\Delta G W. \tag{2.3}$$

where we have used the constant-screening approximation [39] such that ΔW may be neglected compared to ΔG against small perturbations,

$$\Delta W \approx 0. \tag{2.4}$$

The validity of this approximation has been verified by using frozen-phonon calculations in a previous study [39] and by our own calculations. It is expected to be generally valid in semiconductors where the charges are bounded in bonds, and in metals with large Fermi surfaces. This constant-screening approximation is equivalent to the important approximation adopted in the widely used Bethe-Salpeter equation (BSE) [41–43] for solving excitonic properties. The approximation adopted in the BSE formalism is [41–43],

$$\frac{\delta W}{\delta G} \approx 0. \tag{2.5}$$

Note that the charge density is just,

$$n(\mathbf{r}_1) = -iG(1, 1^+). \tag{2.6}$$

Therefore, the change in W due to the atomic position perturbation indeed can be approximated as,

$$\frac{\partial W}{\partial \tau_{\kappa\alpha l}} = \frac{\delta W}{\delta n}\frac{\partial n}{\partial \tau_{\kappa\alpha l}} = i\frac{\delta W}{\delta G(1,1^+)}\frac{\partial n}{\partial \tau_{\kappa\alpha l}} \approx 0. \tag{2.7}$$

Therefore, the constant-screening approximation in Eq. (2.4) is also justified by the wide successful applications with BSE [11, 41–43].

2.2.3 Phonon-mode decomposed changes $\Delta_{\mathbf{q}\nu}G$ and $\Delta_{\mathbf{q}\nu}\Sigma$

Using constant-screening approximation, and working in the frequency domain based on Eq. (1.20), within linear response, we have,

$$\Delta\Sigma(\mathbf{r},\mathbf{r}';\varepsilon) = i\int\frac{d\varepsilon'}{2\pi}e^{-i\delta\varepsilon'}\Delta G(\mathbf{r},\mathbf{r}';\varepsilon-\varepsilon')W(\mathbf{r},\mathbf{r}';\varepsilon'). \tag{2.8}$$

Now, we need to work out the expression for $\Delta G(\mathbf{r},\mathbf{r}';\varepsilon-\varepsilon')$. Propagating the differential operator into the Green's function in Eq. (1.23), and consider first the numerator part, we need to work out $\Delta[\psi_{n\mathbf{k}}(\mathbf{r})\psi^*_{n\mathbf{k}}(\mathbf{r})]$. Let us look at the change in a single wavefunction,

$$\psi_{n\mathbf{k}}(\mathbf{r};\{\boldsymbol{\tau}_{\kappa p}\}) = \psi_{n\mathbf{k}}(\mathbf{r};\{\boldsymbol{\tau}^0_{\kappa p}\}) + \Delta\psi_{n\mathbf{k}}(\mathbf{r};\{\boldsymbol{\tau}_{\kappa p}\}). \tag{2.9}$$

Since $\psi^*_{n\mathbf{k}}(\mathbf{r})$ is *not* an independent quantity, and *it is* the complex conjugate of a wavefunction $\psi_{n\mathbf{k}}(\mathbf{r})$, therefore,

$$\psi^*_{n\mathbf{k}}(\mathbf{r};\{\boldsymbol{\tau}_{\kappa p}\}) = \psi^*_{n\mathbf{k}}(\mathbf{r};\{\boldsymbol{\tau}^0_{\kappa p}\}) + [\Delta\psi_{n\mathbf{k}}(\mathbf{r};\{\boldsymbol{\tau}_{\kappa p}\})]^*. \tag{2.10}$$

Note the order of applying Δ operator and applying complex conjugate in the above equation. Using these two equation, then within the linear-response regime by keeping linear order, we have,

$$\Delta[\psi_{n\mathbf{k}}(\mathbf{r})\psi^*_{n\mathbf{k}}(\mathbf{r})] = \Delta\psi_{n\mathbf{k}}(\mathbf{r})\psi^*_{n\mathbf{k}}(\mathbf{r}) + \psi_{n\mathbf{k}}(\mathbf{r})[\Delta\psi_{n\mathbf{k}}(\mathbf{r};\{\boldsymbol{\tau}_{\kappa p}\})]^*. \tag{2.11}$$

According to the first-order perturbation theory [15, 16],

$$\begin{aligned}\Delta\psi_{n\mathbf{k}}(\mathbf{r}) &= \sum_{n'\mathbf{k}'\neq n\mathbf{k}}\psi_{n'\mathbf{k}'}(\mathbf{r})\frac{\langle\psi_{n'\mathbf{k}'}|\,\Delta V^{\mathrm{KS}}\,|\psi_{n\mathbf{k}}\rangle}{\varepsilon_{n\mathbf{k}}-\varepsilon_{n'\mathbf{k}'}}\\ &= \frac{1}{\sqrt{N_l}}\sum_m\sum_{\mathbf{q}\nu}\psi_{m\mathbf{k}+\mathbf{q}}(\mathbf{r})\frac{(a_{\mathbf{q}\nu}+a^\dagger_{-\mathbf{q}\nu})}{\varepsilon_{n\mathbf{k}}-\varepsilon_{m\mathbf{k}+\mathbf{q}}}\langle\psi_{m\mathbf{k}+\mathbf{q}}|\,\Delta_{\mathbf{q}\nu}V^{\mathrm{KS}}\,|\psi_{n\mathbf{k}}\rangle\,,\end{aligned} \tag{2.12}$$

where we have used the phonon-mode decomposed expression for ΔV^{KS} in Eq. (1.41), and applied crystal momentum conservation. We further introduce the following phonon-mode decomposed first-order change in the wavefunction $\Delta_{\mathbf{q}\nu}\psi_{n\mathbf{k}}$, with noticing that the above integral just gives the e-ph matrix element, and then we have,

$$\Delta_{\mathbf{q}\nu}\psi_{n\mathbf{k}}(\mathbf{r}) = \sum_{m} \frac{g^{\mathrm{DFT}}_{mn\nu}(\mathbf{k},\mathbf{q})}{\varepsilon_{n\mathbf{k}} - \varepsilon_{m\mathbf{k}+\mathbf{q}}} \psi_{m\mathbf{k}+\mathbf{q}}(\mathbf{r}). \tag{2.13}$$

Then we arrive at,

$$\Delta\psi_{n\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{N_l}} \sum_{\mathbf{q}\nu} (a_{\mathbf{q}\nu} + a^{\dagger}_{-\mathbf{q}\nu}) \Delta_{\mathbf{q}\nu}\psi_{n\mathbf{k}}(\mathbf{r}). \tag{2.14}$$

This is result is important because it first shows that the phonon-mode decomposed first-order change in wavefunction can be derived by the standard sum-over-states method from Eq. (2.13), and second it shows that the total change in wavefunction indeed can be decomposed into independent perturbations from each phonon mode, as shown in Eq. (2.14). Note that both Eq. (2.13) and the DFPT method with Sternheimer equation Eq. (1.46) can give $\Delta\psi_{n\mathbf{k}}(\mathbf{r})$, but there is a critical *difference* between the two. We will discuss in more details later in Section 2.3.2 later in this Chapter. For *GW*PT, Eq. (2.13) should be used. We will show later in Section 2.3.2 how we transform the output from solving Sternheimer equation to Eq. (2.13). Now, we take complex conjugate of Eq. (2.14), and get,

$$\begin{aligned} [\Delta\psi_{n\mathbf{k}}(\mathbf{r})]^* &= \frac{1}{\sqrt{N_l}} \sum_{\mathbf{q}\nu} (a^{\dagger}_{\mathbf{q}\nu} + a_{-\mathbf{q}\nu}) [\Delta_{\mathbf{q}\nu}\psi_{n\mathbf{k}}(\mathbf{r})]^* \\ &= \frac{1}{\sqrt{N_l}} \sum_{\mathbf{q}\nu} (a^{\dagger}_{-\mathbf{q}\nu} + a_{\mathbf{q}\nu}) [\Delta_{-\mathbf{q}\nu}\psi_{n\mathbf{k}}(\mathbf{r})]^*, \end{aligned} \tag{2.15}$$

where the last step is done by noting the summation is over the full Brillouin zone of phonon $\mathbf{q}$, and it is to match the single $\mathbf{q}$ decomposition which comes into the e-ph Hamiltonian as $c^{\dagger}_{m\mathbf{k}+\mathbf{q}} c_{n\mathbf{k}} (a_{\mathbf{q}\nu} + a^{\dagger}_{-\mathbf{q}\nu})$ at the end. Taken together, we have,

$$\Delta[\psi_{n\mathbf{k}}(\mathbf{r})\psi^*_{n\mathbf{k}}(\mathbf{r})] = \frac{1}{\sqrt{N_l}} \sum_{\mathbf{q}\nu} (a_{\mathbf{q}\nu} + a^{\dagger}_{-\mathbf{q}\nu}) \Big(\Delta_{\mathbf{q}\nu}\psi^*_{n\mathbf{k}}(\mathbf{r})\psi_{n\mathbf{k}}(\mathbf{r}) + \psi_{n\mathbf{k}}(\mathbf{r})[\Delta_{-\mathbf{q}\nu}\psi_{n\mathbf{k}}(\mathbf{r})]^* \Big). \tag{2.16}$$

To get ΔG in terms of $\Delta_{\mathbf{q}\nu} G$, according to Eq. (1.23) and the chain rule of taking derivatives, the denominator needs to be taken care of. One will end up with dealing the quantity $\Delta_{\mathbf{q}\nu}\varepsilon_{n\mathbf{k}}$. In general, $\frac{\partial \varepsilon_{n\mathbf{k}}}{\partial \tau_{\kappa\alpha l}} \neq 0$. However, by nicely utilizing the periodic perturbation nature embedded in phonon wavevector $\mathbf{q}$, i.e. $\Delta_{\mathbf{q}\nu}$ operator carries a crystal momentum of $\mathbf{q}$, we show from the crystal momentum conservation that,

$$\Delta_{\mathbf{q}\nu}\varepsilon_{n\mathbf{k}} = \langle \psi_{n\mathbf{k}} | \Delta_{\mathbf{q}\nu} H | \psi_{n\mathbf{k}} \rangle = 0, \quad \forall \mathbf{q} \neq 0, \tag{2.17}$$

where we have also used the Hellmann-Feynman theorem [15]. Note that the special case arises with $\mathbf{q} = 0$ where the above equality cannot be guaranteed, and indeed it will depend closely on the detailed crystal symmetry. But since in the general formulations of various e-ph quantities, the $\mathbf{q}=0$ point (as other $\mathbf{q}$ points) represents a certain region in the Brillouin zone associated with a weight, and indeed the weight of the exact $\mathbf{q}=0$ point becomes zero in the region it represents. Therefore, for the relation Eq. (2.17), we treat all the $\mathbf{q}$ points on the same footing.

Then the next steps become quite straightforward that using Eq. (2.16), Eq. (2.17), and Eq. (1.23), we can write

$$\Delta G(\mathbf{r},\mathbf{r}';\varepsilon)=\frac{1}{\sqrt{N_l}}\sum_{\mathbf{q}\nu}(a_{\mathbf{q}\nu}+a^{\dagger}_{-\mathbf{q}\nu})\Delta_{\mathbf{q}\nu}G(\mathbf{r},\mathbf{r}';\varepsilon), \tag{2.18}$$

where the phonon-mode decomposed first-order change in the Green's function is,

$$\Delta_{\mathbf{q}\nu}G(\mathbf{r},\mathbf{r}';\varepsilon)=\sum_{n\mathbf{k}}\frac{\Delta_{\mathbf{q}\nu}\psi_{n\mathbf{k}}(\mathbf{r})\psi_{n\mathbf{k}}(\mathbf{r})+\psi_{n\mathbf{k}}(\mathbf{r})[\Delta_{-\mathbf{q}\nu}\psi_{n\mathbf{k}}(\mathbf{r})]^{*}}{\varepsilon-\varepsilon_{n\mathbf{k}}-i\delta_{n\mathbf{k}}}. \tag{2.19}$$

We can check that $\Delta_{\mathbf{q}\nu}G$ still carries the crystal momentum of $\mathbf{q}$, as $[\Delta_{-\mathbf{q}\nu}\psi_{n\mathbf{k}}(\mathbf{r})]^{*}$ also carries crystal momentum of $\mathbf{q}-\mathbf{k}$ by using Eq. (1.49). Then similarly, from Eq. (2.8), we have,

$$\Delta\Sigma(\mathbf{r},\mathbf{r}';\varepsilon)=\frac{1}{\sqrt{N_l}}\sum_{\mathbf{q}\nu}(a_{\mathbf{q}\nu}+a^{\dagger}_{-\mathbf{q}\nu})\Delta_{\mathbf{q}\nu}\Sigma(\mathbf{r},\mathbf{r}';\varepsilon), \tag{2.20}$$

where the phonon-mode decomposed first-order change in the self-energy operator is,

$$\Delta_{\mathbf{q}\nu}\Sigma(\mathbf{r},\mathbf{r}';\varepsilon)=i\int\frac{d\varepsilon'}{2\pi}e^{-i\delta\varepsilon'}\Delta G_{\mathbf{q}\nu}(\mathbf{r},\mathbf{r}';\varepsilon-\varepsilon')W(\mathbf{r},\mathbf{r}';\varepsilon'). \tag{2.21}$$

2.2.4 Electron-phonon matrix element at the *GW* level from *GW*PT

Now we put Eq. (2.20) and Eq. (2.21) back to the e-ph Hamiltonian at the GW level in Eq. (2.1), and by utilizing the crystal momentum conservation, we arrive at,

$$H^{GW}_{e\text{-ph}}=\frac{1}{\sqrt{N_l}}\sum_{mn\nu}\sum_{\mathbf{k}\mathbf{q}}g^{GW}_{mn\nu}(\mathbf{k},\mathbf{q})c^{\dagger}_{m\mathbf{k}+\mathbf{q}}c_{n\mathbf{k}}(a_{\mathbf{q}\nu}+a^{\dagger}_{-\mathbf{q}\nu}), \tag{2.22}$$

where the e-ph matrix element at the GW level is written as,

$$g^{GW}_{mn\nu}(\mathbf{k},\mathbf{q})=g^{\mathrm{DFT}}_{mn\nu}(\mathbf{k},\mathbf{q})-\langle\psi_{m\mathbf{k}+\mathbf{q}}|\,\Delta_{\mathbf{q}\nu}V_{\mathrm{xc}}\,|\psi_{n\mathbf{k}}\rangle+\langle\psi_{m\mathbf{k}+\mathbf{q}}|\,\Delta_{\mathbf{q}\nu}\Sigma\,|\psi_{n\mathbf{k}}\rangle\,. \tag{2.23}$$

Finally, we need to calculate the matrix elements of $\Delta_{\mathbf{q}\nu}\Sigma$. By utilizing crystal momentum conservation, we arrive at,

$$\begin{aligned}
&\langle\psi_{m\mathbf{k}+\mathbf{q}}|\Delta_{\mathbf{q}\nu}\Sigma(\mathbf{r},\mathbf{r}';\varepsilon)|\psi_{n\mathbf{k}}\rangle \\
=& \frac{i}{2\pi}\sum_{n'}\sum_{\mathbf{p}\mathbf{G}\mathbf{G}'}\Bigg(\langle\psi_{m\mathbf{k}+\mathbf{q}}|e^{i(\mathbf{p}+\mathbf{G})\cdot\mathbf{r}}|\Delta_{\mathbf{q}\nu}\psi_{n'\mathbf{k}-\mathbf{p}}\rangle\,\langle\psi_{n'\mathbf{k}-\mathbf{p}}|e^{-i(\mathbf{p}+\mathbf{G}')\cdot\mathbf{r}'}|\psi_{n\mathbf{k}}\rangle \\
&\times\int d\varepsilon'\frac{W_{\mathbf{G}\mathbf{G}'}(\mathbf{p},\varepsilon')e^{-i\delta\varepsilon'}}{\varepsilon-\varepsilon_{n'\mathbf{k}-\mathbf{p}}-i\delta_{n'\mathbf{k}-\mathbf{p}}-\varepsilon'} \\
&+\ \langle\psi_{m\mathbf{k}+\mathbf{q}}|e^{i(\mathbf{p}+\mathbf{G})\cdot\mathbf{r}}|\psi_{n'\mathbf{k}+\mathbf{q}-\mathbf{p}}\rangle\,\langle\Delta_{-\mathbf{q}\nu}\psi_{n'\mathbf{k}+\mathbf{q}-\mathbf{p}}|e^{-i(\mathbf{p}+\mathbf{G}')\cdot\mathbf{r}'}|\psi_{n\mathbf{k}}\rangle \\
&\times\int d\varepsilon'\frac{W_{\mathbf{G}\mathbf{G}'}(\mathbf{p},\varepsilon')e^{-i\delta\varepsilon'}}{\varepsilon-\varepsilon_{n'\mathbf{k}+\mathbf{q}-\mathbf{p}}-i\delta_{n'\mathbf{k}+\mathbf{q}-\mathbf{p}}-\varepsilon'}\Bigg).
\end{aligned} \tag{2.24}$$

The energy dependence of $\Delta_{\mathbf{q}\nu}\Sigma(\varepsilon)$ is treated with the strategy that every matrix element is evaluated at both $\varepsilon_{n\mathbf{k}}$ and $\varepsilon_{m\mathbf{k}+\mathbf{q}}$, and the average value is taken. We now have the matrix element of $\Delta_{\mathbf{q}\nu}\Sigma$. Eq. (2.24) completes Eq. (2.23) to get $g^{GW}_{mn\nu}(\mathbf{k},\mathbf{q})$.

Note that in practice, in the construction of $\Delta_{\mathbf{q}\nu}$ operator, the use of Kohn-Sham eigenvalues, DFPT e-ph matrix elements, and the unperturbed and first-order wavefunctions from DFT and DFPT, respectively, makes the current calculations at the level of one-shot G_0W_0PT, consistent with the conventional one-shot G_0W_0 calculations. In principle, further iterations can be applied.

2.2.5 Symmetries in e-ph matrix elements $g^{GW}_{mn\nu}(\mathbf{k},\mathbf{q})$

In this section, we discuss the symmetries in the e-ph matrix elements at the GW level. It will be very helpful to save computation time as only phonon wave vectors $\mathbf{q}$ in the irreducible wedge in the Brillouin zone are needed to be calculated. The full phonon Brillouin zone information can then be unfolded.

We introduce the symmetry operation [14, 30],

$$\{\mathcal{S}|\mathbf{v}\}\mathbf{r}=\mathcal{S}\mathbf{r}+\mathbf{v}, \tag{2.25}$$

where $\mathcal{S}$ is the rotational part and $\mathbf{v}$ the fractional translation. $\{\mathcal{S}|\mathbf{v}\}$ is a symmetry under which the system is invariant. It has been shown that the DFT e-ph matrix elements follow the relation [30],

$$g^{\mathrm{DFT}}_{mn\nu}(\mathbf{k},\mathcal{S}\mathbf{q})=g^{\mathrm{DFT}}_{mn\nu}(\mathcal{S}^{-1}\mathbf{k},\mathbf{q}), \tag{2.26}$$

by using the following symmetry relation,

$$\Delta_{\mathcal{S}\mathbf{q}}V(\mathbf{r})=\Delta_{\mathbf{q}}V^{\mathrm{KS}}(\{\mathcal{S}|\mathbf{v}\}^{-1}\mathbf{r}), \tag{2.27}$$

and,

$$\psi_{m\mathbf{k}+\mathcal{S}\mathbf{q}}(\{\mathcal{S}|\mathbf{v}\}\mathbf{r})=\psi_{m\mathcal{S}^{-1}\mathbf{k}+\mathbf{q}}(\mathbf{r}), \tag{2.28}$$

and,

$$\varepsilon_{n\mathbf{k}+\mathcal{S}\mathbf{q}} = \varepsilon_{n\mathcal{S}^{-1}\mathbf{k}+\mathbf{q}}. \tag{2.29}$$

The GW e-ph matrix element $g^{GW}_{mn\nu}(\mathbf{k},\mathbf{q})$ is being calculated from the non-local $\Delta_{\mathbf{q}\nu}\Sigma(\mathbf{r},\mathbf{r}')$ operator, and we derive its symmetry relation here. First, the change in the wavefunction follows,

$$\begin{aligned}\Delta_{\mathcal{S}\mathbf{q}\nu}\psi_{n\mathbf{k}}(\mathbf{r}) &= \sum_m \frac{g^{\mathrm{DFT}}_{mn\nu}(\mathbf{k},\mathcal{S}\mathbf{q})}{\varepsilon_{n\mathbf{k}} - \varepsilon_{m\mathbf{k}+\mathcal{S}\mathbf{q}}}\psi_{m\mathbf{k}+\mathcal{S}\mathbf{q}}(\mathbf{r}) \\ &= \sum_m \frac{g^{\mathrm{DFT}}_{mn\nu}(\mathcal{S}^{-1}\mathbf{k},\mathbf{q})}{\varepsilon_{n\mathcal{S}^{-1}\mathbf{k}} - \varepsilon_{m\mathcal{S}^{-1}\mathbf{k}+\mathbf{q}}}\psi_{m\mathcal{S}^{-1}\mathbf{k}+\mathbf{q}}(\{\mathcal{S}|\mathbf{v}\}^{-1}\mathbf{r}) \\ &= \Delta_{\mathbf{q}\nu}\psi_{n\mathcal{S}^{-1}\mathbf{k}}(\{\mathcal{S}|\mathbf{v}\}^{-1}\mathbf{r}).\end{aligned} \tag{2.30}$$

For the change in the Green's function, using Eq. (2.28), Eq. (2.30) and by noting the summation is in full $\mathbf{k}$ Brillouin zone, we can show that,

$$\Delta_{\mathcal{S}\mathbf{q}\nu}G(\mathbf{r},\mathbf{r}') = \Delta_{\mathbf{q}\nu}G(\{\mathcal{S}|\mathbf{v}\}^{-1}\mathbf{r},\{\mathcal{S}|\mathbf{v}\}^{-1}\mathbf{r}'). \tag{2.31}$$

The screened Coulomb interaction satisfies the following equation by inheriting the symmetry relation in dielectric matrix and the bare Coulomb interaction,

$$W\left(\{\mathcal{S}|\mathbf{v}\}\mathbf{r},\{\mathcal{S}|\mathbf{v}\}\mathbf{r}'\right) = W\left(\mathbf{r},\mathbf{r}'\right). \tag{2.32}$$

Now we label the GW correction to the matrix element in Eq. (2.24), i.e. the contribution from $\Delta_{\mathbf{q}\nu}\Sigma$ as g^{Σ},

$$g^{\Sigma}_{mn\nu}(\mathbf{k},\mathbf{q}) = g^{GW}_{mn\nu}(\mathbf{k},\mathbf{q}) - g^{\mathrm{DFT}}_{mn\nu}(\mathbf{k},\mathbf{q}). \tag{2.33}$$

Using Eq. (1.20), Eq. (2.31), and Eq. (2.32), we have the following relation,

$$\Delta_{\mathcal{S}\mathbf{q}\nu}\Sigma(\mathbf{r},\mathbf{r}') = \Delta_{\mathbf{q}\nu}\Sigma(\{\mathcal{S}|\mathbf{v}\}^{-1}\mathbf{r},\{\mathcal{S}|\mathbf{v}\}^{-1}\mathbf{r}') \tag{2.34}$$

Then we have,

$$\begin{aligned}g^{\Sigma}_{mn\nu}(\mathbf{k},\mathcal{S}\mathbf{q}) &= \langle\psi_{m\mathbf{k}+\mathcal{S}\mathbf{q}}(\mathbf{r})|\,\Delta_{\mathcal{S}\mathbf{q}\nu}\Sigma(\mathbf{r},\mathbf{r}')\,|\psi_{n\mathbf{k}}(\mathbf{r}')\rangle \\ &= \langle\psi_{m\mathcal{S}^{-1}\mathbf{k}+\mathbf{q}}(\{\mathcal{S}|\mathbf{v}\}^{-1}\mathbf{r})|\,\Delta_{\mathbf{q}\nu}\Sigma(\{\mathcal{S}|\mathbf{v}\}^{-1}\mathbf{r},\{\mathcal{S}|\mathbf{v}\}^{-1}\mathbf{r}')\,|\psi_{n\mathcal{S}^{-1}\mathbf{k}}(\{\mathcal{S}|\mathbf{v}\}^{-1}\mathbf{r}')\rangle \\ &= g^{\Sigma}_{mn\nu}(\mathcal{S}^{-1}\mathbf{k},\mathbf{q}).\end{aligned} \tag{2.35}$$

Then combined with Eq. (2.26), we have

$$g^{GW}_{mn\nu}(\mathbf{k},\mathcal{S}\mathbf{q}) = g^{GW}_{mn\nu}(\mathcal{S}^{-1}\mathbf{k},\mathbf{q}). \tag{2.36}$$

The e-ph matrix elements at the GW level follow the same symmetry relation as those at the DFT level. We can nicely utilize these relations to reduce computational efforts.

2.3 Further remarks on GWPT method

2.3.1 Time-reversal symmetry treatment and gauge

In Eq. (2.24), the wavefunction $\Delta_{-\mathbf{q}\nu}\psi_{n\mathbf{k}}$ is involved in the expression for the matrix element of $\Delta_{\mathbf{q}\nu}\Sigma$. Because the whole set of calculations (from DFPT to GWPT) is for a monochromatic $\mathbf{q}$, therefore we would like to utilize TRS to generate $\Delta_{-\mathbf{q}\nu}\psi_{n\mathbf{k}}$ using Eq. (1.50). However, note that the operator Σ itself is *gauge invariant*, and therefore $\Delta_{\mathbf{q}\nu}\Sigma$ should only carry a phase of $e^{i\mathbf{q}\cdot\mathbf{r}}$ without any gauges. The gauge associated with wavefunctions usually arises from diagonalization of a matrix, and the resulting gauge is usually *arbitrary*. But by definition in Eq. (2.12), this arbitrary phase is the same for $\psi_{n\mathbf{k}}$ and $\Delta_{\mathbf{q}\nu}\psi_{n\mathbf{k}}$ except that the latter carries an extra definitive $e^{i\mathbf{q}\cdot\mathbf{r}}$ phase. Therefore, in the construction of the Green's function, the associated wavefunction and first-order change in the wavefunction must be generated from TRS simultaneously, i.e. the following substitution is used in Eq. (2.24),

$$|\psi_{n\mathbf{k}}\rangle\,\langle\Delta_{-\mathbf{q}\nu}\psi_{n\mathbf{k}}| = |T\psi_{n-\mathbf{k}}\rangle\,\langle T(\Delta_{\mathbf{q}\nu}\psi_{n-\mathbf{k}})|\,. \tag{2.37}$$

In this way, the e-ph matrix elements (which are gauge-dependent) at both DFT and GW level, $g^{\mathrm{DFT}}_{mn\nu}(\mathbf{k},\mathbf{q})$ and $g^{GW}_{mn\nu}(\mathbf{k},\mathbf{q})$, are having the same gauge that is completely determined by the wavefunctions used to evaluate (not construct) the operators, i.e. $\langle\psi_{m\mathbf{k}+\mathbf{q}}|$ and $|\psi_{n\mathbf{k}}\rangle$. Note that this is a *requirement* to correctly applying the GWPT correction to the DFPT e-ph matrix elements.

2.3.2 Complete first-order change in wavefunction $\Delta_{\mathbf{q}\nu}\psi_{n\mathbf{k}}$

We have briefly touched before that the first-order change in wavefunction $\Delta_{\mathbf{q}\nu}\psi_{n\mathbf{k}}$ solved from Sternheimer equation Eq. (1.46) and constructed by sum-over-states Eq. (2.13) are *different*. We have also mentioned the in Eq. (1.46), the separation into valence v and conduction c subspaces are not unique. We will address these two related points here.

First, let us differentiate the two $\Delta_{\mathbf{q}\nu}\psi_{n\mathbf{k}}$. The one solved from Sternheimer equation Eq. (1.46) is in the parallel-transport gauge [25], therefore we label it as $\Delta_{\mathbf{q}\nu}\psi^{\mathrm{pt}}_{n\mathbf{k}}$. The one constructed by sum-over-states Eq. (2.13) is in the diagonal gauge [25], therefore we label it as $\Delta_{\mathbf{q}\nu}\psi^{\mathrm{d}}_{n\mathbf{k}}$. Note that $\Delta_{\mathbf{q}\nu}\psi^{\mathrm{d}}_{n\mathbf{k}}$ and its definition Eq. (2.13) is the standard definition (by summing over all states) and represents the true first-order change in wavefunctions.

The Sternheimer equation in Eq. (1.46) with the separation into v and c subspaces is the most commonly seen one. In principle, when solving for one particular band n, as long as its own component is projected out (by $1-P_n$), the Sternheimer equation has no singularity and can be solved. In this case, one gets $\Delta_{\mathbf{q}\nu}\psi^{\mathrm{d}}_{n\mathbf{k}}$ in diagonal gauge. In practical DFPT, there is a *minimum* requirement, that only and the changes in all the valence wavefunctions $\Delta_{\mathbf{q}\nu}\psi_{v\mathbf{k}}$ are needed to construct $\Delta_{\mathbf{q}\nu}n$ from Eq. (1.48) and $\Delta_{\mathbf{q}\nu}V^{\mathrm{KS}}$ from Eq. (1.51). So now all valence states are in the same situation, but this does *not* mean that they can be grouped together and projected out. The reason that $P^{\mathbf{k}+\mathbf{q}}_c = 1 - P^{\mathbf{k}+\mathbf{q}}_v$ can be applied is *fundamentally* related to the density functional. It is important to notice that the construction of $\Delta_{\mathbf{q}\nu}n$

only requires the conduction manifolds in $\Delta_{\mathbf{q}\nu}\psi_{v\mathbf{k}}$, and therefore, within DFPT using the Sternheimer equation Eq. (1.46), one can all together project out all the valence states because it does not change $\Delta_{\mathbf{q}\nu}n$ and $\Delta_{\mathbf{q}\nu}V^{\mathrm{KS}}$, however just bringing the first-order change in wavefunction into a the parallel-transport gauge. We elaborate this property of $\Delta_{\mathbf{q}\nu}n$ here based on Eq. (1.48) and Eq. (2.13),

$$\begin{aligned}
\Delta_{\mathbf{q}\nu}n(\mathbf{r}) &= 4\sum_{n\mathbf{k}}^{\mathrm{occ}} \psi^*_{n\mathbf{k}}(\mathbf{r})\Delta_{\mathbf{q}\nu}\psi^{\mathrm{d}}_{n\mathbf{k}}(\mathbf{r}) \\
&= 4\sum_{n\mathbf{k}}^{\mathrm{occ}}\sum_{m} \psi^*_{n\mathbf{k}}(\mathbf{r})\psi_{m\mathbf{k}+\mathbf{q}}(\mathbf{r})\frac{\langle\psi_{m\mathbf{k}+\mathbf{q}}|\Delta_{\mathbf{q}\nu}V^{\mathrm{KS}}|\psi_{n\mathbf{k}}\rangle}{\varepsilon_{n\mathbf{k}}-\varepsilon_{m\mathbf{k}+\mathbf{q}}} \\
&= 4\sum_{n\mathbf{k}}^{\mathrm{occ}}\sum_{m}^{\mathrm{unocc}} \psi^*_{n\mathbf{k}}(\mathbf{r})\psi_{m\mathbf{k}+\mathbf{q}}(\mathbf{r})\frac{\langle\psi_{m\mathbf{k}+\mathbf{q}}|\Delta_{\mathbf{q}\nu}V^{\mathrm{KS}}|\psi_{n\mathbf{k}}\rangle}{\varepsilon_{n\mathbf{k}}-\varepsilon_{m\mathbf{k}+\mathbf{q}}} \\
&= 4\sum_{n\mathbf{k}}^{\mathrm{occ}} \psi^*_{n\mathbf{k}}(\mathbf{r})\Delta_{\mathbf{q}\nu}\psi^{\mathrm{pt}}_{n\mathbf{k}}(\mathbf{r}),
\end{aligned} \tag{2.38}$$

where the summation range of m changes from all states to only unoccupied states, and this is done noting the $\mathbf{k}$ summation is over full Brillouin zone, and the factor $\frac{1}{\varepsilon_{n\mathbf{k}}-\varepsilon_{m\mathbf{k}+\mathbf{q}}}$ cancels all transitions from valence to valence states, and leaving only transitions from valence to conduction states. This is quite physical because for fermions, the unperturbed valence states are all fully occupied. Any perturbation can only perturb the density into the unoccupied states. In the last line of the above equation, we have used the relation that,

$$\Delta_{\mathbf{q}\nu}\psi^{\mathrm{pt}}_{n\mathbf{k}}(\mathbf{r}) = P_c^{\mathbf{k}+\mathbf{q}}\Delta_{\mathbf{q}\nu}\psi^{\mathrm{d}}_{n\mathbf{k}}(\mathbf{r}) = \sum_{m}^{\mathrm{unocc}}\frac{g^{\mathrm{DFT}}_{mn\nu}(\mathbf{k},\mathbf{q})}{\varepsilon_{n\mathbf{k}}-\varepsilon_{m\mathbf{k}+\mathbf{q}}}\psi_{m\mathbf{k}+\mathbf{q}}(\mathbf{r}), \tag{2.39}$$

which can be shown by combining first-order perturbation theory and the Sternheimer equation defined in Eq. (1.46). As a result, the solution of $\Delta_{\mathbf{q}\nu}\psi^{\mathrm{pt}}_{n\mathbf{k}}$ in DFPT does not alter the physical results. However, for *GW*PT calculation, one *cannot directly* use $\Delta_{\mathbf{q}\nu}\psi^{\mathrm{pt}}_{n\mathbf{k}}$, because what really goes into the formalism is $\Delta_{\mathbf{q}\nu}\psi^{\mathrm{d}}_{n\mathbf{k}}$.

Another point we would like to discuss here is the generalization of the separation by v and c. This is because in practical *GW*PT calculation, e.g. in Eq. (2.24), the summation over n' goes to “infinity” until a convergence cutoff is reached. Therefore we also need the change in the *conduction* wavefunctions to be calculated within DFPT using Sternheimer equation. Therefore, we separate the whole Hilbert space with a band index N_a, such that for $n \leq N_a$, we call it the active space, and for band index $n > N_a$, we call it the Sternheimer space. Note that we require $N_a \geq N_v$ so we still have all $\Delta_{\mathbf{q}\nu}\psi_{v\mathbf{k}}$ for $\Delta_{\mathbf{q}\nu}n$ and $\Delta_{\mathbf{q}\nu}V^{\mathrm{KS}}$. We then define the projection operator to project all states into the active space as P_a, and then the projection into the Sternheimer space P_s is,

$$P_s = 1 - P_a. \tag{2.40}$$

We then can generalize the Sternheimer equation in Eq. (1.46) as,

$$P_s^{\mathbf{k}+\mathbf{q}}\left(H^{\mathrm{DFT}}-\varepsilon_{n\mathbf{k}}\right)P_s^{\mathbf{k}+\mathbf{q}}\left|\Delta_{\mathbf{q}\nu}\psi_{n\mathbf{k}}^{\mathrm{pt}}\right\rangle=-P_s^{\mathbf{k}+\mathbf{q}}\Delta_{\mathbf{q}\nu}V^{\mathrm{KS}}\left|\psi_{n\mathbf{k}}\right\rangle. \tag{2.41}$$

Now we can solve for change in the conduction wavefunctions by setting N_s to included all conduction and valence states (labeled by n in the above equation) we would like to calculate. The calculation of $\Delta_{\mathbf{q}\nu}n$ and $\Delta_{\mathbf{q}\nu}V^{\mathrm{KS}}$ still use all $n=\{v\}$ as before. (Note, in practical calculation using Eq. (2.41), the first-order $\Delta_{\mathbf{q}\nu}\psi_{n\mathbf{k}}^{\mathrm{pt}}$ still needs to be completed as discussed in the next paragraph in order to get the correct $\Delta_{\mathbf{q}\nu}n$ using a similar equation to Eq. (2.42), but the summation over m usually goes from N_v+1 to N_a so that the full empty manifold is restored.)

Lastly, we construct the full $\Delta_{\mathbf{q}\nu}\psi_{n\mathbf{k}}^{\mathrm{d}}$ in diagonal gauge from the output of Sternheimer DFPT $\Delta_{\mathbf{q}\nu}\psi_{n\mathbf{k}}^{\mathrm{pt}}$ in the parallel-transport gauge using Eq. (2.41). This is done by the following relation,

$$\Delta_{\mathbf{q}\nu}\psi_{n\mathbf{k}}^{\mathrm{d}}(\mathbf{r})=\Delta_{\mathbf{q}\nu}\psi_{n\mathbf{k}}^{\mathrm{pt}}(\mathbf{r})+\sum_{m=1}^{N_a}\frac{g_{mn\nu}^{\mathrm{DFT}}(\mathbf{k},\mathbf{q})}{\varepsilon_{n\mathbf{k}}-\varepsilon_{m\mathbf{k}+\mathbf{q}}}\psi_{m\mathbf{k}+\mathbf{q}}(\mathbf{r}). \tag{2.42}$$

Note that in construction of $\Delta_{\mathbf{q}\nu}\psi_{n\mathbf{k}}^{\mathrm{d}}(\mathbf{r})$ using Eq. (2.42), according to the first-order perturbation theory, the sum-over-state part will exclude the degenerate subset such that $\varepsilon_{n\mathbf{k}}=\varepsilon_{m\mathbf{k}+\mathbf{q}}$. For numerical stability, this is done by adding a small non-zero imaginary part in the denominator as a standard treatment. Our tests show that this treatment gives negligible effects, and this imaginary part can be smaller than the smearing factor used in the density of states calculations, and should not be unphysically large. We adopt Eq. (2.42) in the *GW*PT calculations.

2.4 Code developments and practical workflow

2.4.1 Abinit code for DFT and DFPT calculations

We adopt Abinit code [44] for the DFT and DFPT calculation because it is implemented with the generalized separation of subspaces in Sternheimer equation Eq. (2.41). Therefore we can generate $\Delta_{\mathbf{q}\nu}\psi_{n\mathbf{k}}$ for as many empty states as needed. Note that in Abinit response-function calculations [44], the differential operator is taken as for each atom, moving atom along the three primitive cell vectors (not Cartesian coordinates). Therefore coordinate transformation is performed at the end to transform to Cartesian coordinates in definition, then transform into the phonon mode basis. Furthermore, we use Abinit to generate $\Delta_{\mathbf{q}\nu}V_{\mathrm{sc}}$, for the calculation of its matrix elements later.

We have developed the interface for DFPT-*GW*PT in the wrapper abi2bgw distributed with the BerkeleyGW package [45], to handle the format transformation of $\Delta_{\mathbf{q}\nu}\psi_{n\mathbf{k}}$ and $\Delta_{\mathbf{q}\nu}V_{\mathrm{sc}}$.

2.4.2 Implementation of *GW*PT in BerkeleyGW code

The main code development is in BerkeleyGW [45], in particular the Sigma part. With the converted $\Delta_{\mathbf{q}\nu}\psi_{n\mathbf{k}}$ and $\Delta_{\mathbf{q}\nu}V_{\mathrm{sc}}$ from abi2bgw, BerkeleyGW reads these two files in addition to the regular input data files. Then the matrix elements of $\Delta_{\mathbf{q}\nu}\Sigma$ as in Eq. (2.24) and $\Delta_{\mathbf{q}\nu}V_{\mathrm{xc}}$ are calculated.

The current implementation supports calculation of e-ph matrix elements between any specified pair of states in the scattering, for a given $\mathbf{q}\nu$. With the advantage of linear response, the scaling with phonon modes is perfectly linear. These calculations are still demanding because the combined degrees of freedom in $\mathbf{q}\nu$ is usually to the order of 1,000. Therefore the computation of a regular *GW*PT calculation is about three orders of magnitude of a conventional *GW* calculations.

2.4.3 Epw code for Wannier interpolation - development of interface to Abinit and BerkeleyGW

We have developed a general interface to the Epw code [46], which is originally interfaced and distributed with Quantum Espresso [47]. This new interface enables general input of e-ph matrix elements, band energies, phonon dynamical matrices, etc. to be used and interpolated by Epw. We use this interface to interpolate the e-ph matrix elements $g^{\mathrm{DFT}}_{mn\nu}(\mathbf{k},\mathbf{q})$ calculated using Abinit, and $g^{GW}_{mn\nu}(\mathbf{k},\mathbf{q})$ calculated using BerkeleyGW.

Note that the Wannierization step using Wannier90 code [48] requires a set of wavefunctions on full $\mathbf{k}$ Brillouin zone from DFT, and the rotation matrices $U_{nt\mathbf{k}}$ in Eq. (1.57) and Eq. (1.58) depend on the gauge of this given set of wavefunctions. Therefore, for Wannier interpolation purpose, we use full $\mathbf{k}$ and $\mathbf{q}$ grids through the DFPT and *GW*PT calculations, and do not use any symmetry (symmetry rotation introduces different gauges), to keep all the gauges consistent through this DFT → DFPT → *GW*PT → Wannier interpolation workflow. Note that if no Wannier interpolation is used, and direct calculations up to *GW*PT are adequate for convergence, symmetries in $g^{GW}_{mn\nu}(\mathbf{k},\mathbf{q})$ can be applied as discussed in the above Section 2.2.5.

2.5 Verification benchmarks against frozen-phonon results

Frozen-phonon technique calculates phonon response by extracting information from direct supercell calculations with displaced atoms. Such method can be extremely expensive when supercells become very large, therefore is only limited for some high-symmetry $\mathbf{q}$ points (often the zone-boundary points) that are accessible from existing computational resource. Furthermore, to extract e-ph matrix elements, straightforward frozen-phonon calculations can only extract $g_{mn\nu}(\mathbf{k},\mathbf{q})$ for scatterings between degenerate states by fitting band splitting with displacements [16, 37, 39, 40]. However, the linear-response DFPT and GWPT can easily access all scatterings on the same footing.

We can use frozen-phonon results to verify our GWPT method. In the frozen-phonon GW calculations, the supercell wavefunctions with displacements are used to construct the Σ operator, and the equilibrium wavefunctions are used for the evaluation of the quasiparticle energy. There is a subtlety here - that is, a straightforward/conventional frozen-phonon calculation, such as those in DFT, uses the same wavefunctions with displacements at all time. This is not an issue for the comparison between frozen-phonon DFT and DFPT, because both theories self-consistently update the wavefunctions. However, both the G_0W_0 method and G_0W_0PT method are not self-consistent theories (A full self-consistency within GW theory involves many theoretical and technical challenges and is beyond the scope of this work.), and this leads to some degrees of freedom to build the theory. We use the unperturbed initial and final states wavefunctions to evaluate the first-order change in self-energy operator $\Delta_{\mathbf{q}\nu}\Sigma$ in the GWPT theory (as presented in the main text), and this is consistent with the Hellmann-Feynman theorem and the formalism of DFPT. The frozen-phonon GW calculations are prepared and performed with the same idea in the supercells. In this way, we benchmark our implementation of the GWPT method against the finite-difference frozen-phonon calculations. Note that the assumption that the electrons respond instantaneously to the motion of the ions, i.e. the Born-Oppenheimer approximation [1, 15], is used in both DFPT and GWPT, therefore the comparison between the perturbation theory and the frozen-phonon method is meaningful.

We provide the verification of diamond by comparing frozen-phonon results and perturbation theory calculations. We choose a phonon wavevector $\mathbf{q} = L$, which corresponds to a $2 \times 1 \times 1$ supercell. In the frozen-phonon calculation, the energy of the degenerate states at the Brillouin zone boundary splits linearly with increasing displacement (when it is small enough). The slope in the change in energy with respect to displacement is given by a specific single e-ph matrix element that can be fitted from finite-difference frozen-phonon calculations, or directly calculated with the linear-response perturbation theory in a primitive unit cell. This type of e-ph matrix elements that connect degenerate states is the only one that frozen-phonon GW can relatively accurately calculate by making supercells [16, 37, 39, 40] from extracting band energies, but GWPT can access all inter-/intra-band e-ph matrix elements across the whole Brillouin zone with equal and high accuracy. In this example of diamond, the state of interest has quadruple degeneracy at the supercell Brillouin zone boundary as plotted in Fig. 2.1(a), and will split upon the atom displacement (moving one atom along the first lattice vector). The results are shown in Fig. 2.1(b). We find excellent agreement between frozen-phonon DFT and DFPT, and between frozen-phonon GW and GWPT, nicely verifying our GWPT method.

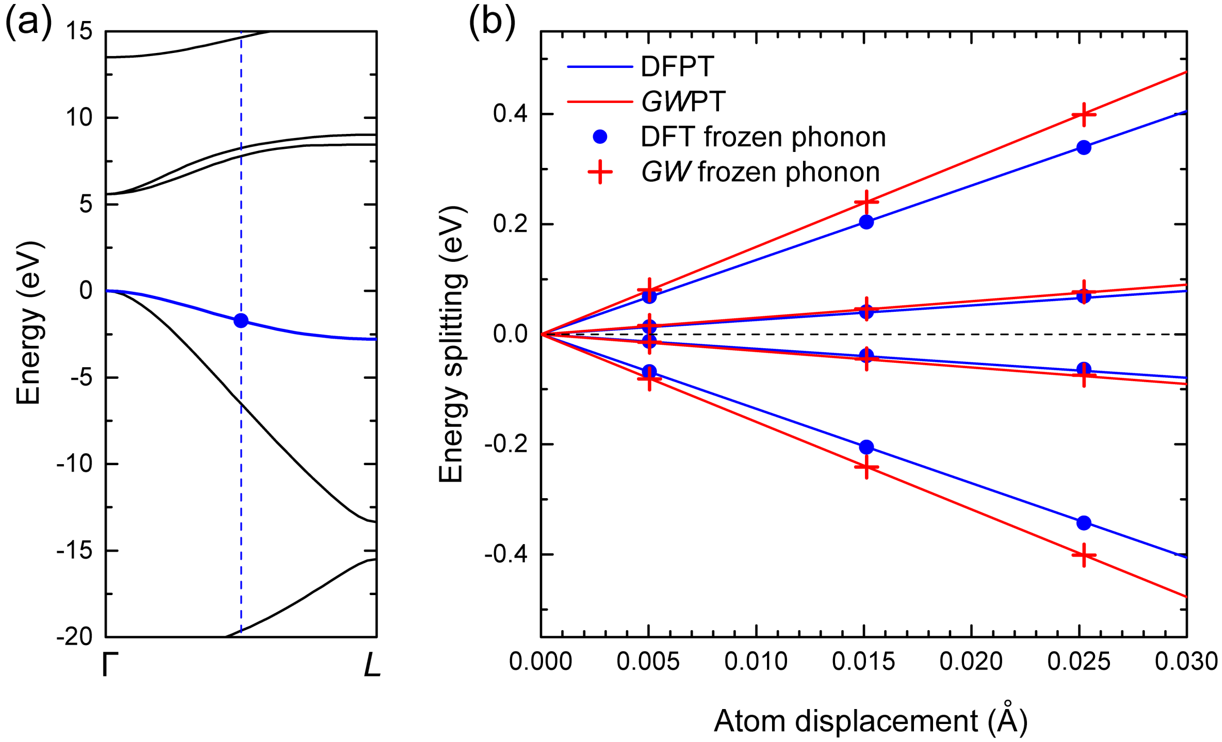

Figure 2.1: (a) DFT band structure of diamond, where the valence band maximum is set to zero energy. The high symmetry points are $\Gamma = (0.0, 0.0, 0.0)$ and $L = (0.5, 0.0, 0.0)$. We consider a phonon wavevector $\mathbf{q} = L$, which corresponds to a $2 \times 1 \times 1$ supercell, folding the Brillouin zone at the blue dashed line. The highlighted blue band is doubly degenerate, therefore the state of interest (at $\mathbf{k} = L/2$ in the primitive unit cell Brillouin zone, indicated by the blue dot) will have four degenerate states at the supercell Brillouin zone boundary. After applying an atom displacement (moving one atom along the first lattice vector), the four states will split. (b) Comparison of energy splitting-versus-displacement curves between perturbation theory and direct frozen-phonon (finite-difference) calculations.

2.6 Conclusion

In this Chapter, we have developed, for the first time, the *GW*PT method [17]. This method is a linear-response theory of the *GW* method, and is designed for calculating the *e*-ph coupling, in particular, any *e*-ph matrix elements of interests with the many-electron correlation effects included. We have presented detailed derivations with some sensible approximations. This method has been implemented into the BerkeleyGW code, and we have developed its interfaces to Abinit and Epw for the *e*-ph coupling calculations. We can calculate the general *e*-ph related quantities accurately and efficiently. Verification against frozen-phonon results gives excellent agreement.

*GW*PT in general is able to systematically and accurately investigate the rich *e*-ph physics at the *GW* level, beyond the accessibility of any other existing *ab initio* methods. The capability of *GW*PT demonstrates its great application potential to the study of the rich *e*-ph physics in a wide range of materials, going beyond DFT and DFPT. We expect a lot new physics including how *e*-ph coupling and many-electron correlations intertwine, and what is the role of *e*-ph coupling in correlated materials, to be revealed with the *GW*PT method.

Chapter 3

Renormalized spin-wave theory (RSWT) for magnetic materials

In this Chapter, we develop a computational method called renormalized spin-wave theory (RSWT). This method is developed in an collaborative work with experimental group on the first discovery of two-dimensional ferromagnetic van der Waals materials $Cr_2Ge_2Te_6$ in Ref. [49]. Discussions on experimental observations and materials properties are presented in Chapter 5. More theoretical details of RSWT are summarized in Ref. [50]. The RSWT method includes the higher order magnon-magnon interactions beyond linear spin-wave theory (LSWT), and explains well the experimentally observed behavior of the 2D ferromagnet – few-layer $Cr_2Ge_2Te_6$. We present the formalism, implementation, and behavior of RSWT in this Chapter.

3.1 Introduction

Recently, two-dimensional (2D) magnetism has been successfully achieved experimentally in atomically thin samples by exfoliating magnetic bulk van der Waals (vdW) crystals down to the few-layer limit, including $Cr_2Ge_2Te_6$ [49] and CrI_3 [51]. This new addition of atomically thin magnetic 2D materials expands the families of 2D vdW materials beyond the well-studied systems such as graphene [52] and transition metal dichalcogenides [53]. Like other novel properties that arise in the 2D limit, ferromagnetism may behave very differently in 2D than in three dimensions (3D). The Mermin-Wagner theorem [54] excludes ferromagnetic phases at any finite temperatures in a strictly 2D isotropic Heisenberg model. However, an easy-axis anisotropy could break the rotational symmetry and stabilize ferromagnetic order in 2D. Moreover, with the help of external magnetic field, it is discovered that the interplay between anisotropy and dimensionality can give rise to useful behavior for systems with small intrinsic anisotropy [49].

An accurate theoretical treatment of magnetic systems at finite temperature is generally difficult due to the complex behaviors of interacting spins. In the quasi 2D limit, it is even more challenging owing to the stronger role of thermal fluctuations and quantum effects (such as quantum statistics) in reduced dimensions. In this work, we present the details of a systematic approach that is suitable to study ferromagnetism in few-layer vdW crystals, such as $Cr_2Ge_2Te_6$ [49, 55, 56]. There have been several theoretical works on this family of materials, using combination of *ab initio* electronic structure calculations and classical Monte Carlo simulations or molecular-field theory to study their ferromagnetic behavior [57–59]. These previous studies treat spins as classical objects; however, in low dimensions and at low temperatures, the quantum nature of spins may manifest significantly.

In this project, we adopt a renormalized spin-wave theory (RSWT) [60] to study magnetism in few-layer systems. Within the RSWT approach, we solve the equation of motion of spins to obtain the spin-wave (magnon) excitation spectrum within a self-consistent Hartree-Fock approximation. The spin-wave excitation spectrum is numerically calculated as a function of temperature. We discuss the details of the numerical implementation and the behavior of the RSWT method. We apply the RSWT method to study $Cr_2Ge_2Te_6$ in its bulk and few-layer forms, which will be discussed in more details in Chapter 5.

3.2 Formalism of renormalized spin-wave theory

3.2.1 General Hamiltonian and Holstein-Primakoff transformation

In this work, we study collinear ferromagnetism within the framework of RSWT for systems with different number of atomic layers. This method can be extended to noncollinear systems by rotating the spin quantization axes [61]. We consider first an isotropic Heisenberg system with the exchange interactions between localized spins denoted as $J_{\nu\nu'}^{\mathbf{l}-\mathbf{l}'}$ where $\mathbf{l}$ are the lattice vectors, labeling the N unit cells under periodic boundary condition, and ν denotes the basis sites ($\nu = 1, ..., n$, where n is the number of magnetic basis atoms) in one unit cell. We then include a single-ion anisotropy, denoted by A, an effect from spin-orbit coupling, and ignore its site dependence because we will be dealing with only one type of magnetic ion in this work. Furthermore, we include the Zeeman term from an external magnetic field B, to explore the interplay between dimensionality and externally induced magnetic anisotropies on the magnetic behavior of the system.

The general Hamiltonian is

$$H = \frac{1}{2}\sum_{\mathbf{l}\mathbf{l}'}\sum_{\nu\nu'} J_{\nu\nu'}^{\mathbf{l}-\mathbf{l}'} \mathbf{S}_{\mathbf{l}\nu} \cdot \mathbf{S}_{\mathbf{l}'\nu'} + \sum_{\mathbf{l}}\sum_{\nu} A(S_{\mathbf{l}\nu}^z)^2 - g\mu_B \sum_{\mathbf{l}}\sum_{\nu} BS_{\mathbf{l}\nu}^z, \tag{3.1}$$

where $\mathbf{S}_{\mathbf{l}\nu} = (S_{\mathbf{l}\nu}^x, S_{\mathbf{l}\nu}^y, S_{\mathbf{l}\nu}^z)$ is the spin operator on site ν in the $\mathbf{l}$-th unit cell (Note that the position vector is $\mathbf{l}$). $J < 0$ ($J > 0$) represents ferromagnetic (antiferromagnetic) interaction, and g is the Landé g-factor. In this work, we consider an easy-axis single-ion anisotropy along the z direction (the direction normal to the plane of the atomic layers) so that $A < 0$. We also align the external magnetic field in the easy axis, i.e. the z direction.

Now we introduce the Holstein-Primakoff transformation [62], rewriting the spin operators with the so-called deviation creation and annihilation operators $a^\dagger$ and a, respectively:

$$\begin{cases} S^+_{\mathbf{l}\nu} = \sqrt{2S - a^\dagger_{\mathbf{l}\nu} a_{\mathbf{l}\nu}}\, a_{\mathbf{l}\nu} \\ S^-_{\mathbf{l}\nu} = a^\dagger_{\mathbf{l}\nu} \sqrt{2S - a^\dagger_{\mathbf{l}\nu} a_{\mathbf{l}\nu}} \\ S^z_{\mathbf{l}\nu} = S - a^\dagger_{\mathbf{l}\nu} a_{\mathbf{l}\nu} \end{cases} \tag{3.2}$$

where S is the total angular momentum of a spin, and $S^\pm = S^x \pm iS^y$ are spin ladder operators. The deviation operators strictly follow the bosonic commutation relations:

$$\begin{aligned} &[a_{\mathbf{l}\nu}, a^\dagger_{\mathbf{l}'\nu'}] = \delta_{\mathbf{l}\mathbf{l}'}\delta_{\nu\nu'}, \\ &[a_{\mathbf{l}\nu}, a_{\mathbf{l}'\nu'}] = [a^\dagger_{\mathbf{l}\nu}, a^\dagger_{\mathbf{l}'\nu'}] = 0. \end{aligned} \tag{3.3}$$

However, the deviation operators do not represent real bosons because $\langle n_{\mathbf{l}\nu} \rangle = \langle a^\dagger_{\mathbf{l}'\nu'} a_{\mathbf{l}\nu} \rangle$ cannot go to infinity and is upper-bounded by $2S$. In the case that $\langle a^\dagger a \rangle \ll 2S$, one can approximate these deviations as low-energy excitations – spin waves, or magnons – and they have bosonic nature, and are therefore treated as (and often called) bosons. To proceed, we expand Eq. (3.2) in different orders of $a^\dagger a/2S$ for different levels of approximations.

3.2.2 Linear spin-wave theory (LSWT) and Fourier transform

We start with linear spin-wave theory (LSWT) by expanding Eq. (3.2) to the lowest order,

$$\begin{cases} S^+_{\mathbf{l}\nu} \approx \sqrt{2S}\, a_{\mathbf{l}\nu} \\ S^-_{\mathbf{l}\nu} \approx \sqrt{2S}\, a^\dagger_{\mathbf{l}\nu} \\ S^z_{\mathbf{l}\nu} = S - a^\dagger_{\mathbf{l}\nu} a_{\mathbf{l}\nu} \end{cases} \tag{3.4}$$

We then rewrite the Hamiltonian (Eq. (3.1)) in terms of the deviation operators as

$$\begin{aligned} H_0 = \frac{S}{2} \sum_{\mathbf{l}\mathbf{l}'} \sum_{\nu\nu'} J^{\mathbf{l}-\mathbf{l}'}_{\nu\nu'} \left(a^\dagger_{\mathbf{l}\nu} a_{\mathbf{l}'\nu'} + a_{\mathbf{l}\nu} a^\dagger_{\mathbf{l}'\nu'} \right) - \sum_{\mathbf{l}\mathbf{l}'} \sum_{\nu\nu'} S J^{\mathbf{l}-\mathbf{l}'}_{\nu\nu'} a^\dagger_{\mathbf{l}\nu} a_{\mathbf{l}\nu} \\ - \sum_{\mathbf{l}} \sum_{\nu} \left(2AS - g\mu_B B \right) a^\dagger_{\mathbf{l}\nu} a_{\mathbf{l}\nu} + E_0, \end{aligned} \tag{3.5}$$

where

$$E_0 = nN \left(\frac{1}{2} \sum_{\mathbf{l}-\mathbf{l}'} \sum_{\nu'} J^{\mathbf{l}-\mathbf{l}'}_{\nu\nu'} S^2 + AS^2 - g\mu_B BS \right) \tag{3.6}$$

is the ground-state energy. Here, we use H_0 to denote the transformed Hamiltonian at the LSWT level. We introduce the Fourier transform of the deviation operators through:

$$\begin{aligned} a^{\dagger}_{\mathbf{l}\nu} &= \frac{1}{\sqrt{N}}\sum_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{l}} b^{\dagger}_{\mathbf{k}\nu} \\ a_{\mathbf{l}\nu} &= \frac{1}{\sqrt{N}}\sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{l}} b_{\mathbf{k}\nu} \end{aligned} \tag{3.7}$$

where $b^{\dagger}_{\mathbf{k}\nu}$ and $b_{\mathbf{k}\nu}$ are basis operators in Bloch wave form that will be used to solve for the spin-wave eigenmodes of the system. It is easy to show that they follow similar commutation relations as their lattice-space counterparts:

$$\begin{aligned} &[b_{\mathbf{k}\nu}, b^{\dagger}_{\mathbf{k}'\nu'}] = \delta_{\mathbf{k}\mathbf{k}'}\delta_{\nu\nu'}, \\ &[b_{\mathbf{k}\nu}, b_{\mathbf{k}'\nu'}] = [b^{\dagger}_{\mathbf{k}\nu}, b^{\dagger}_{\mathbf{k}'\nu'}] = 0. \end{aligned} \tag{3.8}$$

We also introduce the Fourier representation of the exchange interaction,

$$J^{\mathbf{k}}_{\nu\nu'} = \sum_{\mathbf{l}-\mathbf{l}'} e^{-i\mathbf{k}\cdot(\mathbf{l}-\mathbf{l}')} J^{\mathbf{l}-\mathbf{l}'}_{\nu\nu'}. \tag{3.9}$$

The isotropic pairwise interaction satisfies $J^{\mathbf{l}-\mathbf{l}'}_{\nu\nu'} = J^{\mathbf{l}'-\mathbf{l}}_{\nu'\nu}$, which leads to the symmetry $J^{\mathbf{k}}_{\nu\nu'} = J^{-\mathbf{k}}_{\nu'\nu}$. Using the relation $\frac{1}{N}\sum_{\mathbf{l}} e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{l}} = \delta_{\mathbf{k},\mathbf{k}'}$, the Hamiltonian of Eq. (3.5) can be rewritten (i.e. Eq. (3.1) at the LSWT level) in reciprocal space as

$$\begin{aligned} H_0 = &\frac{S}{2}\sum_{\nu\nu'}\sum_{\mathbf{k}}\left(J^{\mathbf{k}}_{\nu\nu'} b^{\dagger}_{\mathbf{k}\nu} b_{\mathbf{k}\nu'} + J^{-\mathbf{k}}_{\nu\nu'} b_{\mathbf{k}\nu} b^{\dagger}_{\mathbf{k}\nu'}\right) \\ &- \sum_{\nu}\left(S\tilde{J}^{\mathbf{k}=0}_{\nu} + 2AS - g\mu_B B\right)\sum_{\mathbf{k}} b^{\dagger}_{\mathbf{k}\nu} b_{\mathbf{k}\nu} + E_0, \end{aligned} \tag{3.10}$$

where $\tilde{J}^{\mathbf{k}=0}_{\nu} = \sum_{\mathbf{l}-\mathbf{l}'}\sum_{\nu'} J^{\mathbf{l}-\mathbf{l}'}_{\nu\nu'} = \sum_{\nu'} J^{\mathbf{k}=0}_{\nu\nu'}$ is a short-hand notation, denoting the sum of all J's on one specific spin site. At the LSWT level, the magnons (eigenstates of H_0) are free bosons without mutual interactions.

3.2.3 Derivation of renormalized spin-wave theory (RSWT) Hamiltonian

At low temperatures, the deviation of the magnetization from its saturated value versus temperature shows a power law behavior, with the exponent being 3/2. This behavior is well given by the LSWT, known as the Bloch-3/2 theorem [63]. However, as we will show later, at high temperature, LSWT fails to capture the physics. This is because as temperature increases, magnon-magnon interactions become indispensable, so that higher-order interaction effects neglected in LSWT must be taken into account. In this work, going beyond LSWT, we consider RSWT, where we include up to fourth order terms in the deviation operators in the Hamiltonian, and treat the magnon-magnon interaction at the self-consistent Hartree-Fock level.

We expand Eq. (3.2) to the second order [60], and get,

$$\begin{cases} S^{+}_{\mathbf{l}\nu} \approx \sqrt{2S}\left(a_{\mathbf{l}\nu} - \dfrac{a^{\dagger}_{\mathbf{l}\nu}a_{\mathbf{l}\nu}a_{\mathbf{l}\nu}}{4S}\right) \\ S^{-}_{\mathbf{l}\nu} \approx \sqrt{2S}\left(a^{\dagger}_{\mathbf{l}\nu} - \dfrac{a^{\dagger}_{\mathbf{l}\nu}a^{\dagger}_{\mathbf{l}\nu}a_{\mathbf{l}\nu}}{4S}\right) \\ S^{z}_{\mathbf{l}\nu} = S - a^{\dagger}_{\mathbf{l}\nu}a_{\mathbf{l}\nu} \end{cases} \tag{3.11}$$

Using Eq. (3.7) in Eq. (3.1), we next separate the total Hamiltonian H into two terms,

$$H = H_0 + H_1, \tag{3.12}$$

where H_0 represents the LSWT part (Eq. (3.5) or Eq. (3.10)), and the other term H_1 gives the interaction effects in what is called the RSWT. In lattice-space representation, H_1 reads

$$\begin{aligned} H_1 = \frac{1}{2}\sum_{\mathbf{l}\mathbf{l}'}\sum_{\nu\nu'} J^{\mathbf{l}-\mathbf{l}'}_{\nu\nu'}\Big(& a^{\dagger}_{\mathbf{l}\nu}a_{\mathbf{l}\nu}a^{\dagger}_{\mathbf{l}'\nu'}a_{\mathbf{l}'\nu'} - \frac{1}{4}a^{\dagger}_{\mathbf{l}\nu}a_{\mathbf{l}\nu}a_{\mathbf{l}\nu}a^{\dagger}_{\mathbf{l}'\nu'} - \frac{1}{4}a_{\mathbf{l}\nu}a^{\dagger}_{\mathbf{l}'\nu'}a^{\dagger}_{\mathbf{l}'\nu'}a_{\mathbf{l}'\nu'} \\ & - \frac{1}{4}a^{\dagger}_{\mathbf{l}\nu}a^{\dagger}_{\mathbf{l}\nu}a_{\mathbf{l}\nu}a_{\mathbf{l}'\nu'} - \frac{1}{4}a^{\dagger}_{\mathbf{l}\nu}a^{\dagger}_{\mathbf{l}'\nu'}a_{\mathbf{l}'\nu'}a_{\mathbf{l}'\nu'}\Big) + \sum_{\mathbf{l}}\sum_{\nu} A a^{\dagger}_{\mathbf{l}\nu}a_{\mathbf{l}\nu}a^{\dagger}_{\mathbf{l}\nu}a_{\mathbf{l}\nu}. \end{aligned} \tag{3.13}$$

Using Fourier transforms and the commutation relations of the deviation operators, we can rewrite all terms in Eq. (3.13) in the form of $b^{\dagger}_{\mathbf{k}'+\mathbf{q}}b^{\dagger}_{\mathbf{k}-\mathbf{q}}b^{\dagger}_{\mathbf{k}'}b_{\mathbf{k}}$ where $\mathbf{k}$, $\mathbf{k}'$, $\mathbf{q}$ are three independent wave vectors. Note that the terms in Eq. (3.13) have different sub-lattice dependence, and this leads to the complexity of the general formalism. We now convert the interacting part of the Hamiltonian given by Eq. (3.13) into the reciprocal space:

$$\begin{aligned} H_1 = \frac{1}{2N}\sum_{\nu\nu'}\sum_{\mathbf{k}\mathbf{k}'}\sum_{\mathbf{q}}\Big(& J^{\mathbf{q}}_{\nu\nu'}b^{\dagger}_{\mathbf{k}'+\mathbf{q}\nu}b^{\dagger}_{\mathbf{k}-\mathbf{q}\nu'}b_{\mathbf{k}'\nu}b_{\mathbf{k}\nu'} \\ & - \frac{1}{4}J^{-\mathbf{k}+\mathbf{q}}_{\nu\nu'}b^{\dagger}_{\mathbf{k}'+\mathbf{q}\nu}b^{\dagger}_{\mathbf{k}-\mathbf{q}\nu'}b_{\mathbf{k}'\nu}b_{\mathbf{k}\nu} - \frac{1}{4}J^{-\mathbf{k}'}_{\nu\nu'}b^{\dagger}_{\mathbf{k}'+\mathbf{q}\nu'}b^{\dagger}_{\mathbf{k}-\mathbf{q}\nu'}b_{\mathbf{k}'\nu}b_{\mathbf{k}\nu'} \\ & - \frac{1}{4}J^{\mathbf{k}}_{\nu\nu'}b^{\dagger}_{\mathbf{k}'+\mathbf{q}\nu}b^{\dagger}_{\mathbf{k}-\mathbf{q}\nu}b_{\mathbf{k}'\nu}b_{\mathbf{k}\nu'} - \frac{1}{4}J^{\mathbf{k}'+\mathbf{q}}_{\nu\nu'}b^{\dagger}_{\mathbf{k}'+\mathbf{q}\nu}b^{\dagger}_{\mathbf{k}-\mathbf{q}\nu'}b_{\mathbf{k}'\nu'}b_{\mathbf{k}\nu'}\Big) \\ & + \frac{1}{N}\sum_{\nu}\sum_{\mathbf{k}\mathbf{k}'}\sum_{\mathbf{q}} A b^{\dagger}_{\mathbf{k}'+\mathbf{q}\nu}b^{\dagger}_{\mathbf{k}-\mathbf{q}\nu}b_{\mathbf{k}'\nu}b_{\mathbf{k}\nu} + \sum_{\nu}\sum_{\mathbf{k}} A b^{\dagger}_{\mathbf{k}\nu}b_{\mathbf{k}\nu}. \end{aligned} \tag{3.14}$$

Interestingly, a linear order term in $b^{\dagger}_{\mathbf{k}\nu}b_{\mathbf{k}\nu}$ associated with the single-ion anisotropy (A) survives, which originates from the local commutation relation between the deviation operators.

3.2.4 Equation of motion and Hartree-Fock approximation

We now have a spin Hamiltonian with a non-interacting part H_0 and an interacting part H_1. To solve this Hamiltonian, i.e. calculate its spectrum (eigenstates and eigenvalues), we solve the equation of motion,

$$i\hbar\frac{\partial}{\partial t}b_{\mathbf{k}_0\nu_0} = [b_{\mathbf{k}_0\nu_0}, H], \tag{3.15}$$

where $b_{\mathbf{k}_0\nu_0}$ is for a specific label of $\mathbf{k}_0$ and ν_0. The commutator is in general an nn matrix (i.e., for all possible ν_0) at each given k_0, and the eigenvalues of this matrix correspond to the excitation energy $\hbar\omega_{\mathbf{k}_0}$ of the different branches of the magnons at k_0. We deal with non-interacting part first. The commutation dictating the evolution of deviation operator can be performed easily using the commutation relation of $b_{\mathbf{k}_0\nu_0}$ and $b^\dagger_{\mathbf{k}_0\nu_0}$,

$$[b_{\mathbf{k}_0\nu_0}, H_0] = S\sum_{\nu} J^{\mathbf{k}_0}_{\nu_0\nu} b_{\mathbf{k}_0\nu} - \left(S\tilde{J}^{\mathbf{k}=0}_{\nu_0} + 2AS - g\mu_B B\right) b_{\mathbf{k}_0\nu_0}. \tag{3.16}$$

By running $\nu_0 = 1, ..., n$, we can construct the full matrix of the equation of motion. The magnon spectrum obtained at this step is that of the LSWT, and has been widely discussed in the literature. The excitation spectrum within the LSWT is independent of the temperature, because the evolution of a given magnon does not correlate with excitations of other magnons at this level of approximation. As a result, the LSWT is usually acceptable only when the temperature of the system is much lower than the transition temperature, in which case the magnon-magnon interaction events are rare.

To include the effects of magnon-magnon interaction, we derive the commutator for the interacting part,

$$\begin{aligned}
[b_{\mathbf{k}_0\nu_0}, H_1] = \frac{1}{2N}\sum_{\nu}\sum_{\mathbf{k}}\sum_{\mathbf{q}}\Big(& J^{\mathbf{q}}_{\nu_0\nu} b^\dagger_{\mathbf{k}-\mathbf{q}\nu} b_{\mathbf{k}_0-\mathbf{q}\nu_0} b_{\mathbf{k}\nu} + J^{\mathbf{q}}_{\nu\nu_0} b^\dagger_{\mathbf{k}+\mathbf{q}\nu} b_{\mathbf{k}\nu} b_{\mathbf{k}_0+\mathbf{q}\nu_0} \\
& - \frac{1}{4} J^{-\mathbf{k}+\mathbf{q}}_{\nu_0\nu} b^\dagger_{\mathbf{k}-\mathbf{q}\nu} b_{\mathbf{k}_0-\mathbf{q}\nu_0} b_{\mathbf{k}\nu_0} - \frac{1}{4} J^{-\mathbf{k}_0}_{\nu\nu_0} b^\dagger_{\mathbf{k}+\mathbf{q}\nu} b_{\mathbf{k}\nu} b_{\mathbf{k}_0+\mathbf{q}\nu} \\
& - \frac{1}{4} J^{-\mathbf{k}_0+\mathbf{q}}_{\nu\nu_0} b^\dagger_{\mathbf{k}-\mathbf{q}\nu_0} b_{\mathbf{k}_0-\mathbf{q}\nu} b_{\mathbf{k}\nu_0} - \frac{1}{4} J^{-\mathbf{k}}_{\nu\nu_0} b^\dagger_{\mathbf{k}+\mathbf{q}\nu_0} b_{\mathbf{k}\nu} b_{\mathbf{k}_0+\mathbf{q}\nu_0} \\
& - \frac{1}{4} J^{\mathbf{k}}_{\nu_0\nu} b^\dagger_{\mathbf{k}-\mathbf{q}\nu_0} b_{\mathbf{k}_0-\mathbf{q}\nu_0} b_{\mathbf{k}\nu} - \frac{1}{4} J^{\mathbf{k}_0+\mathbf{q}}_{\nu_0\nu} b^\dagger_{\mathbf{k}+\mathbf{q}\nu_0} b_{\mathbf{k}\nu_0} b_{\mathbf{k}_0+\mathbf{q}\nu} \\
& - \frac{1}{4} J^{\mathbf{k}_0}_{\nu_0\nu} b^\dagger_{\mathbf{k}-\mathbf{q}\nu} b_{\mathbf{k}_0-\mathbf{q}\nu} b_{\mathbf{k}\nu} - \frac{1}{4} J^{\mathbf{k}+\mathbf{q}}_{\nu\nu_0} b^\dagger_{\mathbf{k}+\mathbf{q}\nu} b_{\mathbf{k}\nu_0} b_{\mathbf{k}_0+\mathbf{q}\nu_0}\Big) \\
& + \frac{1}{N}\sum_{\mathbf{k}}\sum_{\mathbf{q}}\Big(A b^\dagger_{\mathbf{k}-\mathbf{q}\nu_0} b_{\mathbf{k}_0-\mathbf{q}\nu_0} b_{\mathbf{k}\nu_0} + A b^\dagger_{\mathbf{k}+\mathbf{q}\nu_0} b_{\mathbf{k}\nu_0} b_{\mathbf{k}_0+\mathbf{q}\nu_0}\Big) \\
& + A b_{\mathbf{k}_0\nu_0}.
\end{aligned} \tag{3.17}$$

This many-body problem is very difficult to be accurately solved at finite temperature. Here we adopt the Hartree-Fock approximation in which we only keep the so-called diagonal terms that have the same $\mathbf{k}$ and ν indices [60]:

$$\begin{aligned} & b^{\dagger}_{\mathbf{k}_0\nu_0} b_{\mathbf{k}\nu} b_{\mathbf{k}'\nu'} \\ \approx\; & \delta_{\mathbf{k}_0\mathbf{k}} \delta_{\nu_0\nu} \langle b^{\dagger}_{\mathbf{k}_0\nu_0} b_{\mathbf{k}_0\nu_0} \rangle b_{\mathbf{k}'\nu'} + \delta_{\mathbf{k}_0\mathbf{k}'} \delta_{\nu_0\nu'} \langle b^{\dagger}_{\mathbf{k}_0\nu_0} b_{\mathbf{k}_0\nu_0} \rangle b_{\mathbf{k}\nu}. \end{aligned} \tag{3.18}$$

Applying this approximation to Eq. (3.17), we arrive at

$$\begin{aligned} & [b_{\mathbf{k}_0\nu_0}, H_1] \\ = & \frac{1}{N} \sum_{\mathbf{k}} \left(J^{\mathbf{k}-\mathbf{k}_0}_{\nu_0\nu_0} - J^{\mathbf{k}}_{\nu_0\nu_0} \right) \langle n_{\mathbf{k}\nu_0} \rangle b_{\mathbf{k}_0\nu_0} + \frac{1}{N} \sum_{\nu} \sum_{\mathbf{k}} J^{0}_{\nu_0\nu} \langle n_{\mathbf{k}\nu} \rangle b_{\mathbf{k}_0\nu_0} \\ - & \frac{1}{2N} \sum_{\nu} \sum_{\mathbf{k}} \left(J^{\mathbf{k}_0}_{\nu_0\nu} \langle n_{\mathbf{k}\nu} \rangle + J^{\mathbf{k}_0}_{\nu_0\nu} \langle n_{\mathbf{k}\nu_0} \rangle \right) b_{\mathbf{k}_0\nu_0} \\ + & \frac{1}{N} \sum_{\mathbf{k}} 4A \langle n_{\mathbf{k}\nu_0} \rangle b_{\mathbf{k}_0\nu_0} + A b_{\mathbf{k}_0\nu_0}, \end{aligned} \tag{3.19}$$

where we define,

$$\langle n_{\mathbf{k}\nu} \rangle = \langle b^{\dagger}_{\mathbf{k}\nu} b_{\mathbf{k}\nu} \rangle , \tag{3.20}$$

the occupation number of the Bloch wave of wave vector $\mathbf{k}$ on site ν. This occupation number can be obtained after the eigenmodes (a linear combination of the different Bloch waves) are solved with diagonalization. Till now, we have derived the essential formalism for solving the spin Hamiltonian with up to first- and second-order approximations. Combining Eq. (3.16) and Eq. (3.19), we can solve for the magnon spectrum with the inclusion of magnon-magnon interactions, and this level of theory is called RSWT. Note that $\langle n_{\mathbf{k}\nu} \rangle$ is temperature-dependent. This indicates temperature renormalization effects in the magnon spectrum within RSWT. The solution, i.e. the eigenvalues of $[b_{\mathbf{k}_0\nu_0}, H_0 + H_1]$, which are the magnons of the system, indeed requires self-consistent solutions, as will be discussed in the next section.

3.3 Code implementation

3.3.1 Self-consistent equations for magnetization

Basically, we are solving the combination of Eq. (3.16) and Eq. (3.19). Here, as an example, we discuss the case of $Cr_2Ge_2Te_6$, which is the main physical research object of Chapter 5. $Cr_2Ge_2Te_6$ is a layered vdW ferromagnetic semiconductor, and has recently been successfully exfoliated down to the 2D limit [49]. The detailed structural information and its theory study using a combined DFT plus RSWT approach is discussed in details in Chapter 5. Here, we focus on the algorithm behavior.

For the monolayer or the bulk $Cr_2Ge_2Te_6$, in one unit cell there are two magnetic sites ($n = 2$), so that $\nu = 1, 2$; and we introduce $\nu_1 = 1$ and $\nu_2 = 2$. Moreover, the two sites are symmetrically equivalent, therefore after diagonalization the eigenstates of the system are just in-phase acoustic mode (labeled by $-$) with the cell-periodic part of the eigenvector $\frac{1}{\sqrt{2}}(1,1)^T$ and anti-phase optical mode (labeled by $+$) with the cell-periodic part of the eigenvector $\frac{1}{\sqrt{2}}(1,-1)^T$. The total number of magnons excited at a given $\mathbf{k}$ point in the two magnon branches may be obtained from the Bose-Einstein distribution, i.e.:

$$\begin{aligned} \langle n_{\mathbf{k}} \rangle &= \langle n_{\mathbf{k}_+} \rangle + \langle n_{\mathbf{k}_-} \rangle \\ &= \sum_{\pm} \frac{1}{\exp\left(\frac{\hbar\omega_{\mathbf{k}_\pm}}{k_B T}\right) - 1}. \end{aligned} \tag{3.21}$$

In general, at a given k point, the relation between the total excitation number $\langle n_{\mathbf{k}} \rangle$ and the excitation number of each site $n_{\mathbf{k}\nu}$ is directly given by the eigenvalues and eigenvectors of the magnon modes at that $\mathbf{k}$ point. For this particular system of two equivalent magnetic sites, by symmetry we have

$$n_{\mathbf{k}\nu_1} = n_{\mathbf{k}\nu_2} = \frac{1}{2} n_{\mathbf{k}\nu}, \tag{3.22}$$

for all $\mathbf{k}$ points. (Note that the right hand side of Eq. (3.21) is in general an approximation for the excitation number, because the deviation operators are not real bosonic operators, as we discussed before.) Then, for the single-layer or the bulk $\mathrm{Cr_2Ge_2Te_6}$ system, the equation of motion of magnons is further simplified as

$$\begin{aligned} &[b_{\mathbf{k}_0\nu_0}, H] \\ &= S \sum_{\nu} J^{\mathbf{k}_0}_{\nu_0\nu} b_{\mathbf{k}_0\nu} - \left(S \tilde{J}^{\mathbf{k}=0}_{\nu_0} + 2AS - g\mu_B B\right) b_{\mathbf{k}_0\nu_0} \\ &+ \frac{1}{2N} \sum_{\mathbf{k}} \left(J^{\mathbf{k}-\mathbf{k}_0}_{\nu_0\nu_0} - J^{\mathbf{k}}_{\nu_0\nu_0} + J^{0}_{\nu_0\nu} - J^{\mathbf{k}_0}_{\nu_0\nu}\right) \langle n_{\mathbf{k}} \rangle\, b_{\mathbf{k}_0\nu_0} \\ &+ \frac{1}{N} \sum_{\mathbf{k}} 2A \langle n_{\mathbf{k}} \rangle\, b_{\mathbf{k}_0\nu_0} + A b_{\mathbf{k}_0\nu_0}. \end{aligned} \tag{3.23}$$

Eq. (3.21) and Eq. (3.23) form the self-consistent loop: at a given temperature, the occupation $\langle n_{\mathbf{k}} \rangle$ given by Eq. (3.21) depends on the spin-wave spectrum (i.e., the eigenvalues), and the spin-wave eigenvalues are derived from diagonalization of the matrix representation of the equation of motion (Eq. (3.23)), which explicitly depends on the occupation.

We solve this set of self-consistent equations in the following way: (a) at each given temperature, we start from LSWT and solve for the non-interacting spectrum; (b) with the spectrum solved, we calculate $\langle n_{\mathbf{k}} \rangle$ using Eq. (3.21); (c) we put the calculated $\langle n_{\mathbf{k}} \rangle$ back into Eq. (3.23) to solve for a new spectrum. These steps are repeated until the solutions are self-consistent. The convergence criterion can be defined either using the spectrum, or the total magnetization. In this work, we use the latter criterion,

$$\frac{M(T)}{M_0} = 1 - \frac{1}{nNS} \sum_{\mathbf{k}} \langle n_{\mathbf{k}} \rangle, \tag{3.24}$$

where $M(T)$ is the total magnetization as a function of temperature T, and $M_0 = M(0)$ denotes the fully polarized magnetization (at $T = 0$). The ferromagnetic Curie temperature T_C is simply determined here by the temperature at which the total magnetization becomes zero, i.e. $M(T_C) = 0$. Our calculations show that, at a given temperature, 20 self-consistent steps usually give very good convergence in $M(T)$.

3.3.2 Dimensionality reflected in numerical implementation

The summation over $\mathbf{k}$ points in our formalism is equivalent to an integral over the first Brillouin zone, in the limit that the number of $\mathbf{k}$ points used goes to infinity. However, numerically, we could only perform the summation over a grid of finite number of points. For this reason, care must be taken when dealing with some numerically diverging cases based on the physical behavior from analytical limit. For example, in the long wavelength regime, as $\mathbf{k} \to 0$, the acoustic spin-wave excitation in the ferromagnetic phase behaves as $\hbar\omega_{\mathbf{k}} \sim k^2$, whereas the optical mode has a much higher energy and its relative contribution is little. Therefore, the integration over the occupation numbers

$$\sum_{\mathbf{k}} \langle n_{\mathbf{k}} \rangle \to \frac{V}{(2\pi)^D} \int \frac{d^D\mathbf{k}}{\exp\left(\frac{\hbar\omega_{\mathbf{k}}}{k_B T}\right) - 1} \tag{3.25}$$

behaves quite differently in 2D ($D = 2$) and in 3D ($D = 3$). Considering the ideal case of a truly isotropic Hamiltonian with no spin-wave excitation gap, around the $\mathbf{k} = 0$ point, the integral diverges in 2D, but remains finite in 3D. This means from Eq. (3.24) that there cannot be any magnetization at any finite temperature in this magnetically isotropic 2D system, as shown by the Mermin-Wagner theorem. In 3D, since the integral over a small region around the point $\mathbf{k} = 0$ is finite, the contribution of the $\mathbf{k} = 0$ point vanish as the $\mathbf{k}$ grid becomes finer. In 2D, the contribution from integrating over this point is divergent, so that the deviation blows up and we have $T_C = 0$. In the case that anisotropy is present and a spin-wave excitation gap is established in a 2D system, the contribution of the $\mathbf{k} = 0$ is finite, and sometimes dominant (especially when the gap is small), we directly evaluate its contribution and include it in the summation. In our numerical implementation of the method (with and without a spin-wave energy gap), for the case of 2D calculation, we always keep the contribution of the $\mathbf{k} = 0$ point; for the case of 3D calculation, one can safely neglect the contribution of the $\mathbf{k} = 0$ point. In both cases, we need very fine $\mathbf{k}$ point to converge the calculation (e.g. 72×72 for 2D and 36×36×36 for 3D, for application in $\mathrm{Cr_2Ge_2Te_6}$ system discussed in Chapter 5). Spatial symmetries are used to reduce the computation time.

With an attentive treatment of $\mathbf{k} = 0$, we can investigate the layer-number dependence of ferromagnetism. The inclusion of interlayer couplings allows for more magnon modes at each $\mathbf{k}$ point in the 2D Brillouin zone, physically corresponding to additional confined modes in the normal direction. This additional modes along the normal direction are mimicked by introducing discrete k_z values, with the number of k_z points correspond to the number of layers n_z [64]. We replace the interlayer coupling J_z of the bulk by $\frac{n_z-1}{n_z} J_z$ to give the correct ground-state total energy, as the coupling at the two surfaces (to the upper or lower layer) of a multilayer sample are physically cut off. With this method, we are always dealing with a unit cell with two basis sites. In principle, in the multilayer cases, one would include the multiples of basis sites, then construct and diagonalize bigger matrices to obtain the different magnon modes. Our approximate treatment here (by introducing discrete k_z and modifying interlayer coupling) can be shown to yield the same results to the direct diagonalization method in the monolayer, bilayer (two layers are equivalent), and bulk $Cr_2Ge_2Te_6$ systems. For other multilayer cases, this simple treatment gives satisfactory description of the multilayer behavior of vdW quasi 2D systems with weak interlayer coupling.

In the next section, we present our benchmark results on quasi 2D magnetism within RSWT. Note that in the presence of an external magnetic field, rigorously speaking, there is no well-defined phase transition temperature T_C; however, experimentally, a transition between ferromagnetic-like and paramagnetic-like states can be distinguished from the tail effect [49], and theoretically under the spin-wave theory framework, we can locate where the magnetization vanishes and thus determine an effective T_C. Therefore, we use the notation T_C^* to represent the transition temperature in the presence of an external field.

3.3.3 Self-consistent behavior of RSWT and temperature dependent magnon spectrum

Before presenting and discussing more results, we first show the self-consistent behavior of the RSWT method. The example used here is the monolayer $Cr_2Ge_2Te_6$ system, setting $A = 0$, $B = 1.0$ T. As we can see in Fig. 3.1(a), the magnetization curve of the monolayer system converges rapidly with the number of iterations, giving $T_C^* = 35$ K with 20 iterations. The line with zero iteration represents LSWT, and it gives a $T_C^* = 110$ K, showing that LSWT overestimates T_C (and T_C^*) dramatically, as was also observed in studies of other systems [60]. However, it is noted that the converged magnetization-temperature curve from RSWT show a first-order-like phase transition (a nearly sudden drop of $M(T)$ to zero magnetization), whereas ferromagnetic transition is known to be second order. This feature is a known artifact of RSWT, as has been discussed in Refs. [60, 65]. Physically, RSWT can describe the general temperature behavior of ferromagnetic systems, but not necessarily the exact behavior near the vicinity of T_C (e.g., $T > 0.9T_C$), where the approximate boson approximation breaks down and the constraint of having a finite total spin plays an important role. However, the determination of T_C (and T_C^*) using vanishing magnetization from RSWT usually gives rather accurate estimation of T_C (and T_C^*) [60, 65].

The inclusion of spin-wave interactions is important, and RSWT captures the essence of magnon softening or the temperature renormalization of the magnon dispersion. It has been observed experimentally [66, 67], that the magnon excitation energy decreases as temperature increases, and the renormalization is especially strong when approaching T_C. Fig. 3.1(b) shows the temperature-dependent magnon dispersions of monolayer $Cr_2Ge_2Te_6$ under $B = 1.0$ T with a converged $T_C^* = 35$ K from RSWT. Shown are the calculated temperature-

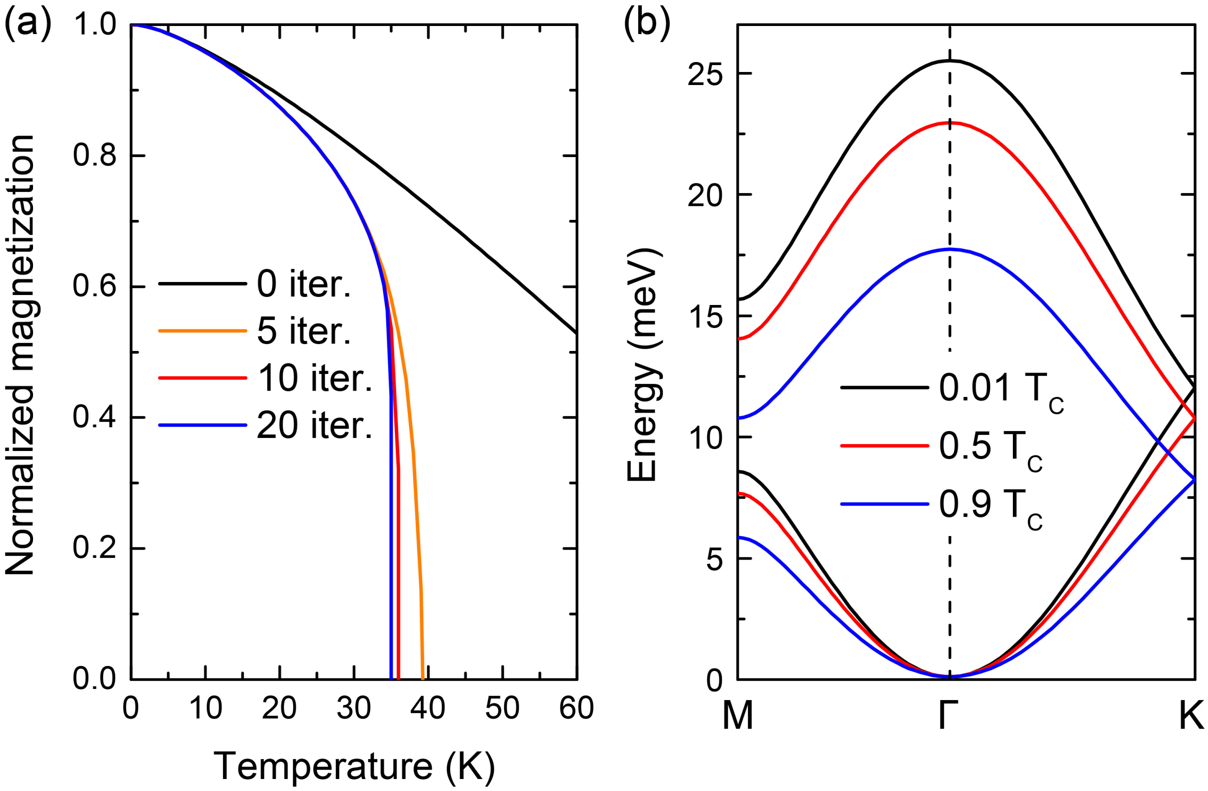

Figure 3.1: Behavior of RSWT method for monolayer $Cr_2Ge_2Te_6$. Example used is set $A = 0$ and $B = 1$ T. (a) Convergence behavior of the self-consistency procedure of RSWT. Plotted are renormalized magnetization $M(T)/M_0$ as a function of temperature T. T_C (or T_C^*, see text) is determined by the value at which the magnetization vanishes. We show the results of 0, 5, 10, and 20 iterations and T_C eventually converge to 35 K. The 0-iteration represents LSWT, and the curve extends beyond the plot to give a $T_C^* = 110$ K, which greatly overestimates the final result. (b) Temperature-dependent spin-wave excitation spectrum. Shown are results at three different temperatures: $0.01T_C^*$ (i.e., 0.35 K), $0.5T_C^*$ (i.e., 17.5 K), and $0.9T_C^*$ (i.e., 31.5 K). Strong temperature renormalization effect is observed.

dependent spin-wave dispersion at $0.01T_C^*$ (i.e. 0.35 K), $0.5T_C^*$ (i.e. 17.5 K), and $0.9T_C^*$ (i.e. 31.5 K). As seen in experiment, the calculated magnon excitation energy decreases as temperature increases, and more rapidly at higher temperature. This is a direct consequence of spin-wave interactions where temperature changes the strength of the scattering. RSWT nicely captures the temperature renormalization effect, which can be directly compared with neutron diffraction experiments quantitatively.

3.4 Conclusion

In this Chapter, we have developed the formalism of RSWT to calculate magnetic properties and the transition temperature of 2D and 3D ferromagnetic materials. We adopt various approximations and cast the interacting spin-wave problem into a self-consistent frame-work with parameters mapped from *ab initio* DFT calculations. This formalism is general by including the sublattice dependence ν and ν' for pairwise interactions. Magnon-magnon interaction is treated at the Hartree-Fock level, and is important for correctly describing the high-temperature behavior of magnetism. The test calculations show that numerical self-consistent calculations are important to get accurate converged value, and the temperature effect in magnon excitation can be very important. This work provides an accurate theoretical method to study 2D (as well as 3D) ferromagnetism in atomic layers of real materials with physical insights. We expect more interesting physical phenomena in 2D magnetism can be understood and predicted using RSWT, employing the theoretical framework and techniques presented here.

Part II

Studies of superconductivity, and electronic and magnetic interactions in novel materials

Chapter 4

Correlation-enhanced electron-phonon coupling and superconductivity in $\mathbf{Ba_{1-x}K_xBiO_3}$ from $\boldsymbol{GW}$PT

Electron-phonon (e-ph) interactions and superconductivity in correlated materials provide exciting opportunities and challenges for theoretical condensed matter studies. With the new theoretical and computational method – GW perturbation theory (GWPT), we are able to investigate a large number of correlated materials where correlation effects in e-ph interactions were missed from standard density-functional perturbation theory (DFPT) based calculations.

In this Chapter, we apply our newly developed GWPT method to systematically study an intriguing superconductor $Ba_{1-x}K_xBiO_3$. We find that many-electron correlations significantly enhance the e-ph coupling strength and the superconducting T_c, nicely agreeing with experimentally measured $T_c = 30 - 32$ K. This application of GWPT showcases the strong capability of the linear-response technique that enables a full investigation of the Brillouin zone and any pairs of transitions for e-ph coupling, and shows its wide application potential in studying the complex e-ph physics in correlated materials.

4.1 Introduction

We have applied our GWPT method (within a one-shot calculation, i.e. G_0W_0PT) to study superconductivity in $Ba_{1-x}K_xBiO_3$ in its cubic perovskite phase as shown in Fig. 4.1(a), which has an experimentally measured superconducting T_c of $30 - 32$ K [68–70]. Previous *ab initio* studies [37, 71] show that the e-ph coupling calculated within DFT-LDA is too weak to account for such a high T_c in this material, and frozen-phonon GW calculations indicate that many-electron self-energy effects may enhance e-ph interactions, however estimated from only a single $\mathbf{q}$-point calculation for one electronic state [37].

In this work, we find that the GW self-energy renormalizes the DFT-LDA e-ph matrix elements non-uniformly across the Brillouin zone, and enhances the e-ph coupling strength λ by a factor of 2.4. The GWPT-calculated $\lambda = 1.14$ is strong enough to account for the high superconducting transition temperature T_c in $Ba_{0.6}K_{0.4}BiO_3$. We show that the doping dependence in the superconductivity is mainly from a density-of-states (DOS) effect.

4.2 Calculation details

In this work, the DFT and DFPT calculations of $Ba_{0.6}K_{0.4}BiO_3$ are performed using Abinit code [44]. We simulate the potassium (K) doping effects by removing electrons and adding a compensating background charge. The generalized gradient approximation [5] and norm-conserving pseudopotentials [72, 73] are used, with a plane-wave cut-off of 100 Ry. The fully relaxed lattice constant of 4.268 $\AA$ is adopted, which is close to the experimental value of 4.283 $\AA$ [70]. In the GW and GWPT calculations using the BerkeleyGW code [45], we take a 15 Ry cut-off for the screened Coulomb interaction, and include 40 empty bands in the construction of the dielectric matrix and the self-energy operator. In metallic systems, the intra-band contributions near the Fermi level are dominant, therefore a relatively small number of empty bands ensures convergence. In the construction of $\Delta_{\mathbf{q}\nu}\Sigma$ operator, the use of Kohn-Sham eigenvalues, DFPT e-ph matrix elements, and the unperturbed and first-order wavefunctions from DFT and DFPT, respectively, makes the current calculations at the level of one-shot G_0W_0PT.

4.3 Electronic structure and verification benchmark against frozen-phonon technique

We first perform standard DFT and DFPT calculations on $Ba_{0.6}K_{0.4}BiO_3$ using the GGA functional [5]. The calculated Fermi surface shows a regular rounded cubic shape (Fig. 4.1(b)), and is strongly nested. We verify our GWPT method by comparing its results against reference frozen-phonon GW results at a selected high symmetry $\mathbf{q}$-vector. We focus on the single band (labeled as n_0 and highlighted in Fig. 4.1(c)) crossing ε_F, which is expected to give the dominant contribution to superconductivity.

We set up the frozen-phonon calculations by displacing a single oxygen atom along the normal of the cubic face (that it centers on) as a finite-difference perturbation. This displacement is modulated from one primitive unit cell to another according to the phonon wavevector $\mathbf{q}$ of interest. We are in particular interested in the $\mathbf{q} = R$ (TRS-equivalent to $\mathbf{q} = R$) phonon perturbation, which is (nearly) a nesting vector of the Fermi surface and scatters carriers between the states $\mathbf{k} = -R/2$ and $\mathbf{k} = R/2$. Inducing the atom-displacement perturbation in the $2 \times 2 \times 2$ supercell splits the degenerate level at the supercell zone boundary (R' point in Fig. 4.1(d)).

The slope in the change in energy with respect to displacement is given by a specific single e-ph matrix element that can be fitted from finite-difference frozen-phonon calculations, or

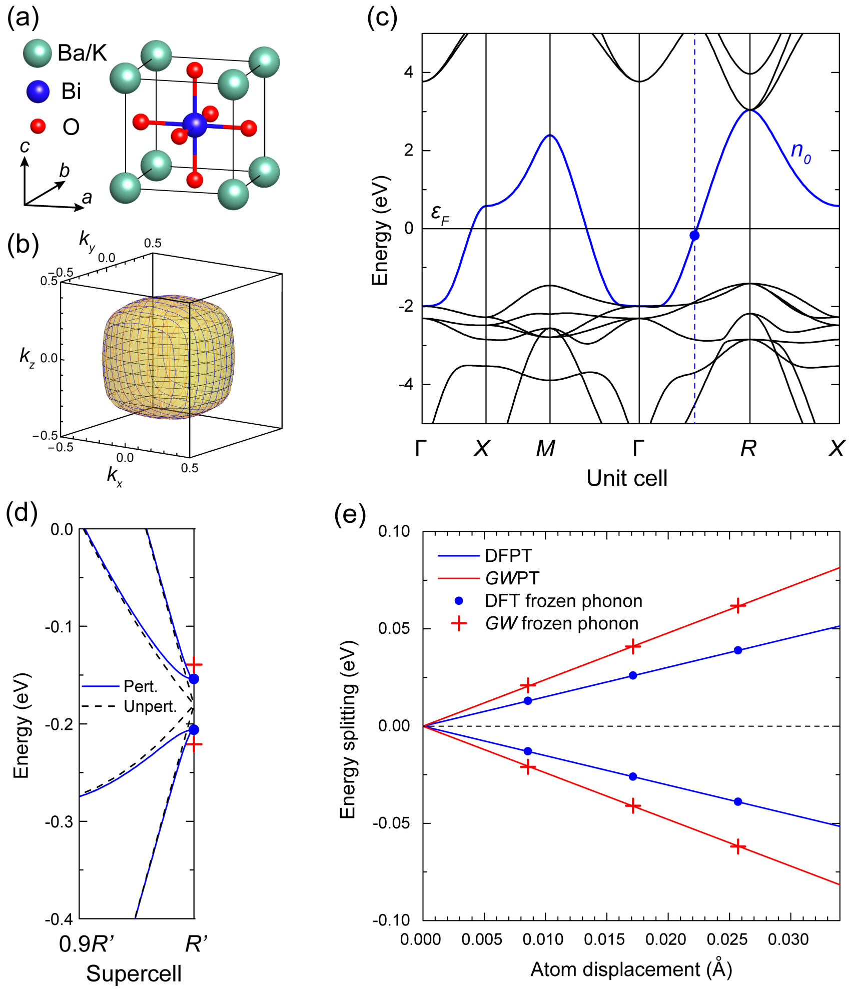

Figure 4.1: (a) Crystal structure of Ba$_{0.6}$K$_{0.4}$BiO$_3$ in the cubic perovskite phase. (b) Calculated Fermi surface of Ba$_{0.6}$K$_{0.4}$BiO$_3$. (c) The DFT-GGA band structure of Ba$_{0.6}$K$_{0.4}$BiO$_3$. The high symmetry points are $\Gamma = (0.0, 0.0, 0.0)$, $X = (0.5, 0.0, 0.0)$, $M = (0.5, 0.5, 0.0)$, and $R = (0.5, 0.5, 0.5)$ in units of primitive reciprocal lattice vectors. The band of interest which crosses the ε_F (set to zero) is highlighted with blue color and labeled as n_0. The state at $\mathbf{k} = R/2$ (blue dashed line) indicated by the blue dot has a band energy slightly below ε_F. (d) The DFT band structure of a $2 \times 2 \times 2$ supercell. The R' point corresponds to the $\mathbf{k} = R/2$ point at the blue dashed line in (c). The degenerate level indicated by the blue dot in (c) splits upon the oxygen-atom-displacement perturbation (see Supplemental Materials [24]) of 0.0171 $\AA$. The corresponding GW quasiparticle energies are indicated by the red crosses. (e) Comparison of energy splitting-versus-displacement curves between perturbation theory and direct frozen-phonon (finite-difference) calculations.

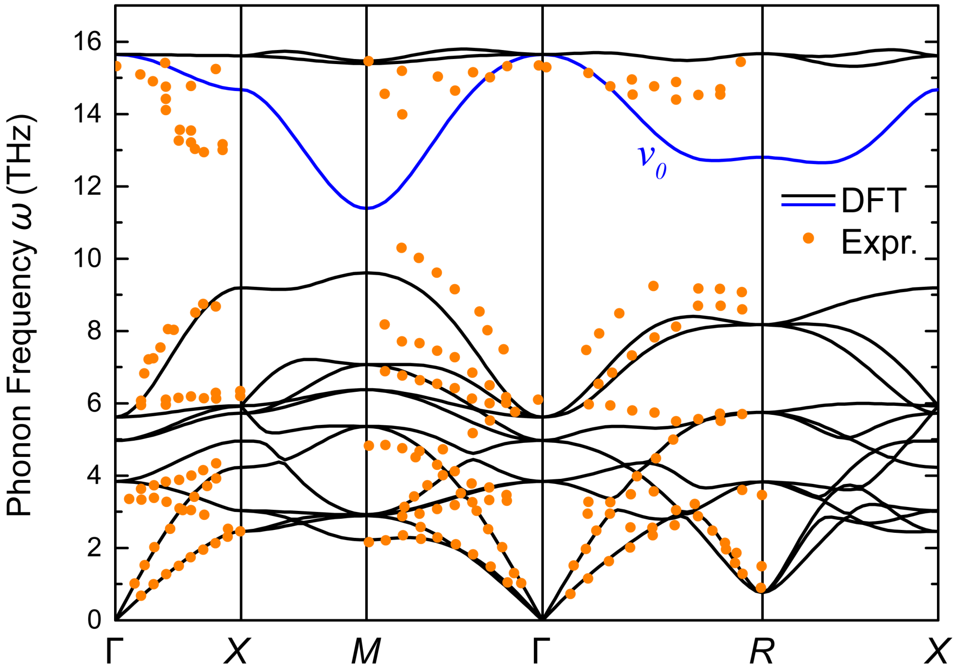

Figure 4.2: Phonon band structure of $Ba_{0.6}K_{0.4}BiO_3$ calculated at the DFT level. The dots are data adopted from neutron diffraction experiments [14]. The oxygen breathing and stretching branch is highlighted with blue color and labeled as ν_0.

directly calculated with the linear-response perturbation theory in a primitive unit cell. As mentioned before, this type of e-ph matrix elements that connect degenerate states is the only one that frozen-phonon GW can relatively accurately calculate by making supercells [16, 37], but GWPT can access all inter-/intra-band e-ph matrix elements across the whole Brillouin zone with equal and high accuracy. As shown in Fig. 4.1(e), we find excellent agreement for this matrix element between frozen-phonon DFT and DFPT, and between frozen-phonon GW and GWPT, nicely verifying our GWPT method. Moreover, the DFPT and GWPT results are significantly different, illustrating the importance of having the correct treatment of the quasiparticle self-energy within the GW approach.

4.4 Phonon spectrum

The phonon band structure directly obtained with DFPT calculations shows imaginary frequencies at R and M points, leading to a structural phase transition involving the oxygen octahedron tilting [37, 70, 71, 74] at low temperature. Such an instability is identified to be physical both theoretically [71] and experimentally [74]. We calculate a finite-temperature phonon spectrum to get rid of the imaginary frequencies [75] so as to proceed with the e-ph calculations. First, a molecular dynamics (MD) sampling at 600 K of a 2 × 2 × 2 supercell is performed using the Vasp code [76], with the projector augmented wave pseudopotentials [77] Then 50 randomly selected structures near the MD equilibrium are recalculated using the Abinit code [44] so that the electronic properties are consistent throughout this work. The calculated forces are then fitted with harmonic potentials to remove the instability, using the Alamode code [78]. Finally the inter-atomic force constants and therefore the dynamical matrices and the phonon spectrum are calculated with the fitted potentials. Our calculated phonon spectrum as plotted in Fig. 4.2 is in good agreement with neutron diffraction experiment [74] , except that the frequencies of the oxygen breathing and stretching branch (highlighted in Fig. 4.2 and we label its mode index as ν_0) are somewhat underestimated.

4.5 Distribution of e-ph matrix elements and the many-electron renormalization across Brillouin zone

To study superconductivity in $\mathrm{Ba_{0.6}K_{0.4}BiO_3}$, we calculate the e-ph matrix elements that scatter quasiparticle states within the n_0 band by performing both DFPT and GWPT calculations on an $8 \times 8 \times 8$ **k**-grid (full grid) and **q**-grid (35 irreducible **q**-points) [24]. These electronic states are coupled most strongly by phonons in the highest three optical branches [17,30]. As an illustration, we pick out one high-frequency oxygen stretching and breathing optical branch (labeled as ν_0 [24]), and plot the distribution of the strength of the e-ph matrix element $|g_{n_0 n_0 \nu_0}(\mathbf{k}, \mathbf{q})|$ varying **k** across the Brillouin zone for selected **q**-points. Fig. 4.3(a-c) show the scatterings for $\mathbf{q} = R$ that are mostly relevant to the superconductivity. For this important phonon mode, GWPT almost uniformly enhances the value of the e-ph matrix elements as compared to DFPT with an enhancement factor of ~ 1.6. This is because the character of the states on the Fermi surface of $\mathrm{Ba_{0.6}K_{0.4}BiO_3}$ is highly isotropic [79]. However, Fig. 4.3(d-i) show strong variances in the distribution (as functions of different **k** and **q** points) of the e-ph matrix elements and also in the enhancement factor of GWPT over DFPT, due to the wavefunction character changing near the Γ point of either the initial or final states. These results, for the first time, systematically reveal the complex nature of many-electron renormalization of the e-ph interactions, demonstrating the strong capability and uniqueness of GWPT.

Note that in the calculation of superconductivity, the DFT band structure is used, as our calculated GW band structure remains quite similar near the Fermi level. The significant enhancement of e-ph coupling is dominantly due to the many-electron renormalization of the e-ph matrix elements.

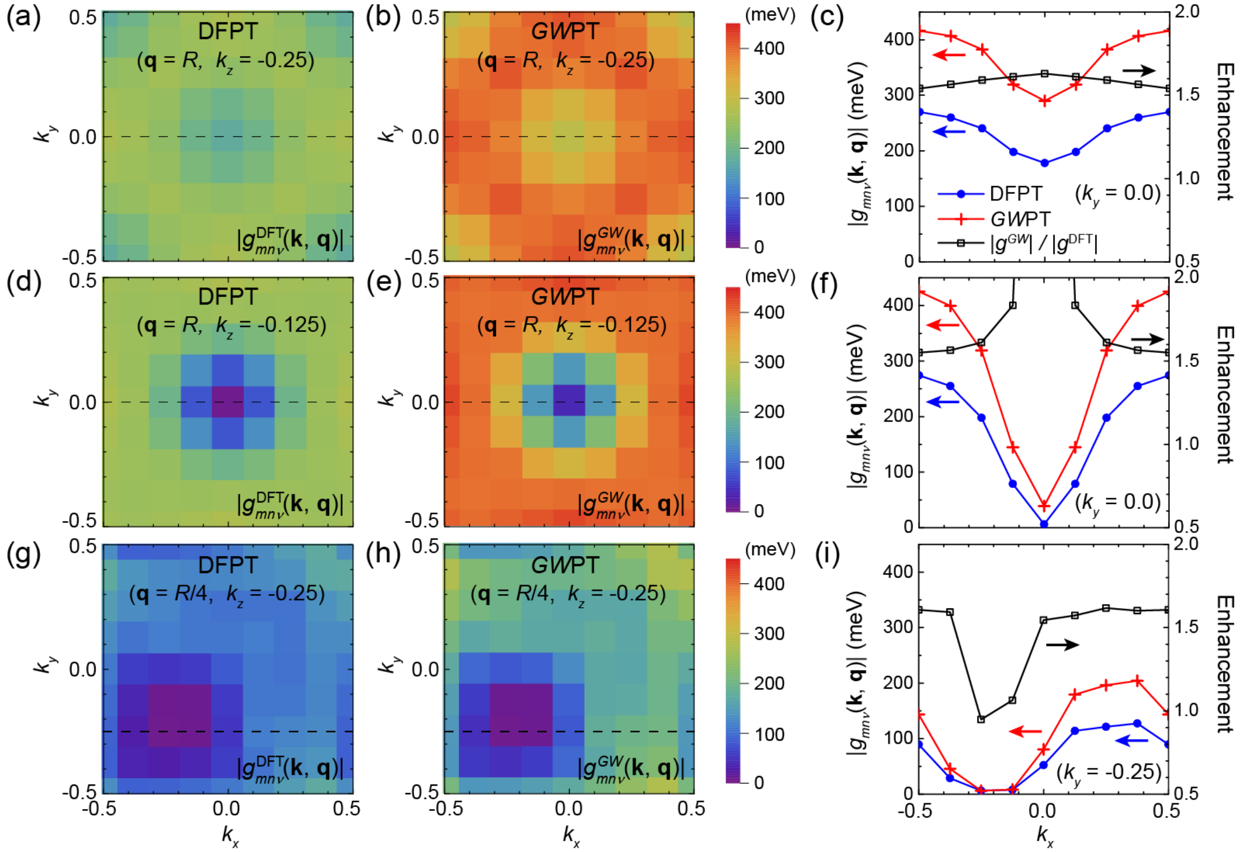

Figure 4.3: Distribution of the absolute value of the e-ph matrix elements $|g_{n_0 n_0 \nu_0}(\mathbf{k}, \mathbf{q} = R)|$ at (a) DFPT and (b) GWPT level, across the k_x–k_y plane at fixed $k_z = -0.25$ of the Brillouin zone. The wavevectors are in units of reciprocal lattice vectors. The data are calculated on an $8 \times 8 \times 8$ $\mathbf{k}$-grid for each $\mathbf{q}$-point. (c) Line profile of (a) and (b) with $k_y = 0.0$, and the path is indicated by the dashed line in (a) and (b). The enhancement factor of $|g^{GW}|/|g^{\mathrm{DFT}}|$ is also plotted. (d-f), (g-i) Similar to (a-c), but with $\mathbf{q} = R$ in the $k_z = -0.125$ plane and $\mathbf{q} = R/4$ in the $k_z = -0.25$ plane, respectively.

	λ	$\omega_{\log}$ (K)	μ^*	T_c (K)
DFPT	0.47	488.2	$0.18-0.08$	$0.61-6.1$
GWPT	1.14	491.3	$0.18-0.08$	$28.5-44.8$

Table 4.1: Calculated e-ph coupling strength λ, logarithmic-averaged phonon frequency $\omega_{\log}$, and superconducting transition temperature T_c (using the McMillan-Allen-Dynes formula [16, 80, 81]) of $Ba_{0.6}K_{0.4}BiO_3$. The effective Coulomb potential parameter μ^* is set to a reasonable physical range, giving the corresponding range of T_c. The experimentally measured T_c is $30-32$ K [68, 69]

4.6 Correlation-enhanced superconductivity in $Ba_{1-x}K_xBiO_3$

We evaluate the superconducting T_c of $Ba_{0.6}K_{0.4}BiO_3$ using the following McMillanAllen-Dynes formula [80, 81],

$$k_{\mathrm{B}}T_c = \frac{\hbar\omega_{\log}}{1.2}\exp\left[-\frac{1.04(1+\lambda)}{\lambda-\mu^*(1+0.62\lambda)}\right], \tag{4.1}$$

where $\omega_{\log}$ is a logarithmic average of the phonon frequencies, λ is the e-ph coupling strength, and μ^* is the effective Coulomb repulsion parameter [16, 80, 81]. The calculation of λ involves an important physical function, called the Eliashberg function $\alpha^2F(\omega)$, which measures the e-ph coupling strength as a function of phonon frequency. The Eliashberg function reads [16, 80, 81],

$$\alpha^2F(\omega) = \frac{1}{N_F}\int\frac{d\mathbf{k}d\mathbf{q}}{\Omega_{\mathrm{BZ}}^2}\sum_{mn\nu}|g_{mn\nu}(\mathbf{k},\mathbf{q})|^2\delta(\varepsilon_{n\mathbf{k}}-\varepsilon_F)\delta(\varepsilon_{m\mathbf{k}+\mathbf{q}}-\varepsilon_F)\delta(\hbar\omega-\hbar\omega_{\mathbf{q}\nu}), \tag{4.2}$$

where N_F is the DOS at the Fermi level, and Ω_{BZ} is the Brillouin zone volume. The e-ph coupling strength λ is then calculated from, [16, 80, 81]

$$\lambda = 2\int_0^\infty \frac{\alpha^2F(\omega)}{\omega}d\omega, \tag{4.3}$$

and the logarithmic average of the phonon frequencies is [16, 80, 81],

$$\omega_{\log} = \exp\left[\frac{2}{\lambda}\int_0^\infty d\omega\frac{\alpha^2F(\omega)}{\omega}\log\omega.\right] \tag{4.4}$$

We can then estimate the superconducting T_c from the above set of equations.

The e-ph coupling strength λ and the characteristic logarithmic-averaged phonon frequency $\omega_{\log}$ [16, 80, 81] are calculated using the e-ph matrix elements that scatter states within the n_0 band for all phonon modes, at both the DFPT and GWPT level (Table 4.1).

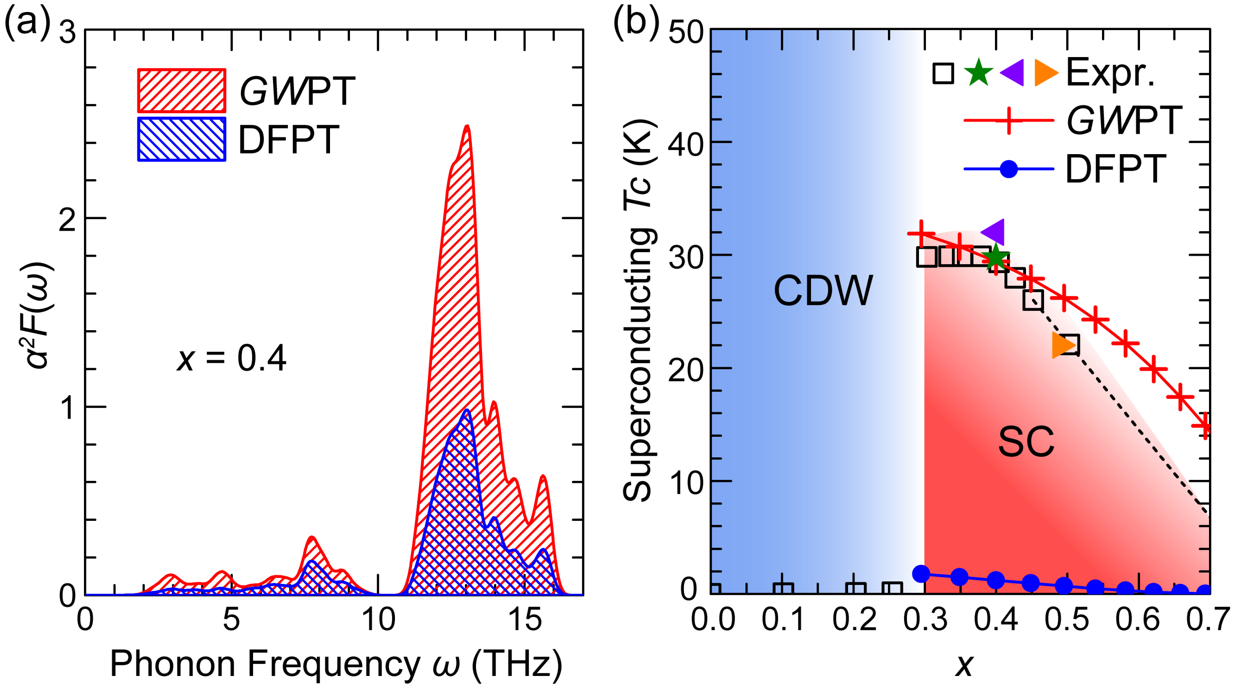

Figure 4.4: (a) Eliashberg function $\alpha^2F(\omega)$ calculated for $\mathrm{Ba_{0.6}K_{0.4}BiO_3}$ (i.e. $x = 0.4$) with e-ph matrix elements from GWPT and DFPT. (b) Phase diagram of $\mathrm{Ba_{1-x}K_xBiO_3}$. Superconducting T_c is calculated with $\mu^* = 0.16$. Experimental data are plotted as squares [70], star [68], left-pointing triangle [69], and right-pointing triangle [79]. Black dashed line represents the extrapolation of experimental data into doping range of $x > 0.5$, which is hard to access experimentally. From the superconducting (SC) phase towards undoped parent composition, for $x < 0.3$, the system undergoes a structural phase transition into the non-superconducting CDW phase.

The correlation-enhanced e-ph interaction strength is directly reflected in the Eliashberg function $\alpha^2F(\omega)$ by comparing the results from DFPT and GWPT in Fig. 4.4(a). The effective Coulomb parameter μ^* [16, 80, 81] is set to a reasonable physical range in Table 4.1. DFPT severely underestimates the superconducting transition temperature, with the calculated T_c in the range of $0.61-6.1$ K for μ^* in the range of $0.18-0.08$. However, GWPT significantly increases T_c to the range of $28.5-44.8$ K for the same range of μ^* (Table I), in good agreement with the experimentally measured T_c of $30-32$ K [21-23]. These results highlight the importance of many-electron correlation effects in e-ph interactions [37] that are well captured by the GWPT method.

We further study the doping dependence of the superconductivity in $\mathrm{Ba_{1-x}K_xBiO_3}$ (superconductivity is observed experimentally for $x > 0.3$) from first principles, calculated using a rigid-band approximation (band energy, phonon spectrum, and e-ph matrix elements are taken from *ab initio* calculations at $x = 0.4$), which is usually a good approximation if the

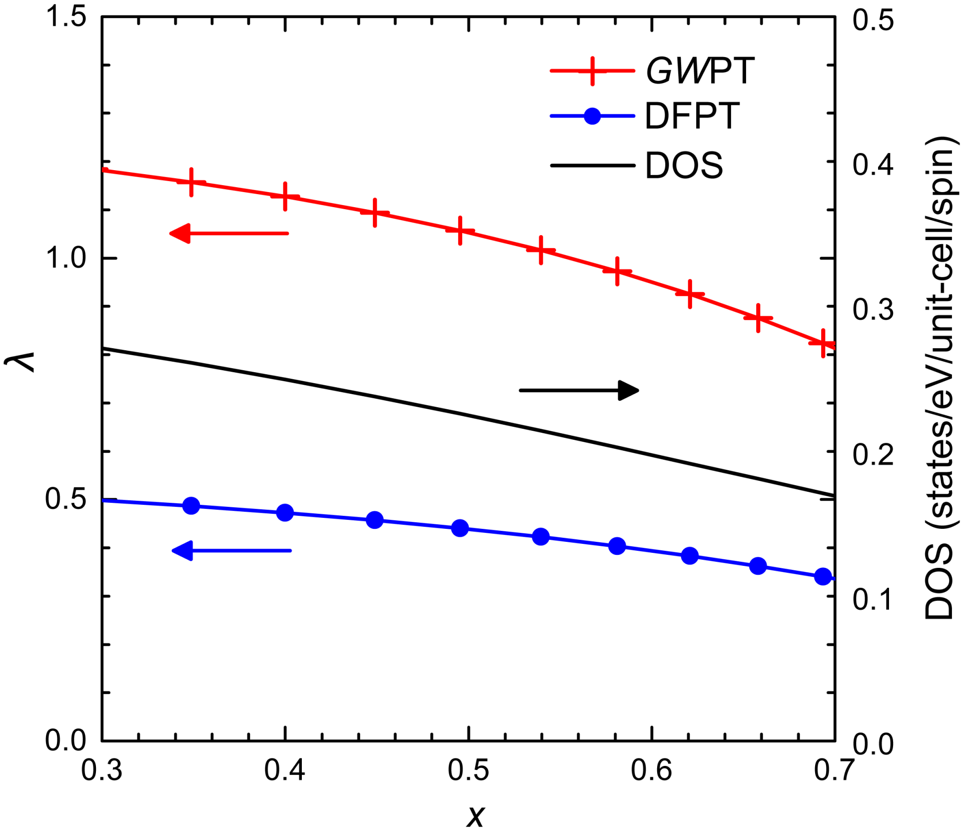

Figure 4.5: Doping-dependent e-ph coupling strength λ from GWPT and DFPT, and density of states (DOS) at Fermi level.

doping range of study is not too large so that the system remains similar. By varying the Fermi level, we recalculate the Eliashberg function $\alpha^2F(\omega)$, the e-ph coupling strength λ, and the superconducting T_c (using $\mu^* = 0.16$).

Fig. 4.4(b) shows that the superconducting transition temperatures from GWPT nicely reproduce the size and shape of the superconducting half dome (however results from DFPT fail significantly) in the phase diagram observed experimentally [68–70]. At doping smaller than $x = 0.3$, the material is in an insulating charge density wave (CDW) phase with strong structural distortions induced by phonon instability and the nested Fermi surface [37, 70, 71, 74, 82, 83] . After $x = 0.4$, an increase in hole doping x suppresses T_c, which is mainly due to a reduced DOS with a shrinking Fermi surface. By looking at the band structure (Fig. 4.1(c)), an increasing hole doping level leads to smaller DOS (Fig. 4.5) from a shrinking Fermi surface. Therefore, the number of allowed scatterings on Fermi surface is reduced, leading to the suppressed T_c. In Fig. 4.5, it is clear that the e-ph coupling strength λ (from both GWPT and DFPT) follows the trend of DOS as a function of doping x. Note that DFPT fails to explain the strength of the superconductivity across the whole range of doping, whereas *GW*PT reproduces the experimental data [70] well. Beyond the rigid-band approximation, other factors involving modifications in crystal structures, quasiparticle bands, phonon spectrum, and e-ph matrix elements may have further influences on the doping dependence. However, within $0.3 < x < 0.5$ ($x > 0.5$ is very hard to access experimentally [70]), the rigid-band approximation based on *GW*PT calculations captures the dominant DOS effects, and provides reliable doping dependence in both magnitude and trend. Our *GW*PT results, along with the recent direct experimental observation of isotropic s-wave superconducting gap [79], strongly support that superconductivity in $\mathrm{Ba}_{1-x}\mathrm{K}_x\mathrm{BiO}_3$ originates from unusually large e-ph interactions, due to many-electron effects.

4.7 Additional discussions on phonons and computational analysis

The phonon branches that couple the most to the electronic states near the Fermi level in $\mathrm{Ba}_{0.6}\mathrm{K}_{0.4}\mathrm{BiO}_3$ are the highest three optical branches, as is evidenced by the dominant spectral weight within the range of $11 - 16$ THz in the Eliashberg function $\alpha^2F(\omega)$ (see Fig. 4.4(a)). We note that our calculated e-ph coupling strength $\lambda = 0.47$ with DFPT is higher than previous ab initio results $\lambda = 0.33$ [37] and $\lambda = 0.29$ [71], and this is largely due to that the previous calculations resulted in higher phonon frequencies (bandwidth ~ 18 THz) than our calculation (bandwidth ~ 16 THz, which agrees better with the experiment). Furthermore, the frequency of the strongly e-ph coupled ν_0 branch is around 13 THz from our calculated $\alpha^2F(\omega)$ function (see Fig. 4.4(a)), whereas in Refs. [37, 71] the dominant peak is localized around 17 THz. The overestimation of phonon frequency will lead to an underestimation of λ, considering which previous calculations [37, 71] and our calculated λ agree nicely. Nevertheless, the calculated T_c at DFPT level is still too low compared with experiments. However, according to our *GW*PT calculations, the absolute values of e-ph matrix elements $|g|$ are increased with an overall enhancement factor ~ 1.6 near the Fermi surface, see Fig. 4.3(c). As is evidenced in the high-frequency range of the $\alpha^2F(\omega)$ function (see Fig. 4.4(a)), the e-ph coupling strength λ is largely enhanced by a factor ~ 2.4 (coming from $|g|^2$) in *GW*PT compared with DFPT. Consequently, the enhancement of T_c due to the many-electron correlation effects is quite significant.

On the computational side, to highlight the significantly high efficiency of linear-response method over frozen-phonon technique, in particular by comparing *GW*PT with frozen-phonon *GW*, we further provide an example of comparison of the computation time (defined as total time spent multiplied by number of central processing units used on Intel Knights Landing architecture) between *GW*PT and frozen-phonon *GW*. For the $\mathbf{q} = (0.5, 0.5, 0.5)$ results plotted in Fig. 4.1(e), the *GW*PT calculation performed in a primitive unit cell takes $\sim 10^4$ seconds, and the frozen-phonon *GW* calculation performed in a $2 \times 2 \times 2$ supercell takes $\times 10^6$ seconds. The efficiency is differed by two orders of magnitude for this relatively

DFPT **k**-grid	λ	$\omega_{\log}$ (K)	μ^*	T_c (K)
$8 \times 8 \times 8$	0.47	488.2	$0.18 - 0.08$	$0.61 - 6.1$
$16 \times 16 \times 16$	0.51	486.4	$0.18 - 0.08$	$1.2 - 8.2$

Table 4.2: Convergence of k-grid for superconductivity properties at DFPT level. The e-ph coupling parameter λ, logarithmic-averaged phonon frequency $\omega_{\log}$, and the superconducting transition temperature T_c are calculated using DFPT on different **k**-grids, but with the same $8 \times 8 \times 8$ phonon **q**-grid. The effective Coulomb potential parameter μ^* is set to the same indicated range.

small supercell. For a fine sampling of the **q**-grid, the frozen-phonon *GW* technique becomes prohibitive, whereas *GW*PT has the capability to access the e-ph physics systematically and efficiently. For example, the **k**- and **q**-grids used in this work for direct calculations is $8 \times 8 \times 8$, and this corresponds to a frozen-phonon supercell containing 2560 atoms, with which the *GW* calculations basically cannot be calculated with any contemporary computational resources. *GW*PT performs all various **k**and **q** calculations using a 5-atom primitive unit cell efficiently, and the scaling is linear in the phonon modes investigated. This highlight the strong necessity of developing the linear-response *GW*PT method to the understanding of many-electron correlation effects in e-ph physics.

In the above discussions, we report the results of both DFPT and *GW*PT calculations on $8 \times 8 \times 8$ **k**- and **q**-grids (phonon **q**-grid is symmetry reduced). We check the convergence of k-grid up to $16 \times 16 \times 16$ with DFPT and list the results in Table 4.2. The superconductivity properties are converged at $8 \times 8 \times 8$ **k** **k**-grid, because of the highly isotropic Fermi surface of $Ba_{0.6}K_{0.4}BiO_3$ [79]. The convergence of phonon **q**-grid is usually faster than that of the **k**-grid; therefore, we use the $8 \times 8 \times 8$ **k**-grid. In fact, *GW*PT can further be combined with Wannier interpolation techniques [30], which we have already successfully developed the interfaces as discussed in Section 2.4.3, to achieve an efficient yet accurate full description of the e-ph coupling properties. For this system, due to the nice isotropic property of the Fermi surface (which is relatively easy to converge) and our convergence test, direct calculation using *GW*PT is enough to give converged results, therefore Wannier interpolation for this system is not necessary here.

4.8 Conclusion

In this Chapter, we have applied the newly developed *GW*PT method from Chapter 2 to study the e-ph coupling and superconductivity in the correlated $Ba_{1-x}K_xBiO_3$. This study demonstrates the capabilities of *GW*PT to efficiently and accurately calculate e-ph matrix elements in full electron and phonon Brillouin zone, and between any pair of electronic states. We find that the correlation effect can have significant role in e-ph coupling. In particular for $Ba_{1-x}K_xBiO_3$, we show that many-electron correlations significantly enhance the e-ph interactions for states near the Fermi surface, and explain the observed high superconductivity transition temperature of the optimally doped $\mathrm{Ba_{0.6}K_{0.4}BiO_3}$. Furthermore, the doping dependence is shown to originate from a DOS effect, and we have nicely reproduced the trend in superconductivity T_c as a function of doping. Our calculations clearly and strongly support the e-ph origin of the high superconductivity T_c, due to the large many-electron correlations enhancement. This application of GWPT in $\mathrm{Ba_{1-x}K_xBiO_3}$ encourages us to study the important role of correlation-incorporated e-ph in various intriguing correlated materials.

Chapter 5

Two-dimensional ferromagnetism in few-layer van der Waals crystals

Atomically thin vdW crystals provide ideal 2D platform for novel physical properties. Before the year of 2017, such achievements have been made mainly for exploring the electric and optical properties. For one of the most important concepts in condensed matter physics − magnetism − the studies were mostly focusing on extrinsically induced magnetism, by defect engineering via vacancies, adatoms, grain boundaries, and edges, by adding magnetic species via intercalation or substitution, or by magnetic proximity effect [84–92].

For intrinsic magnetism, the major hope is to exfoliate intrinsic magnetic 2D materials. In the year of 2017, two ferromagnetic semiconductors $\mathrm{Cr_2Ge_2Te_6}$ [49] and $\mathrm{CrI_3}$ [51] were successfully exfoliated down to few-layer limit. Later, a ferromagnetic metal has been thinned down to the 2D limit and shows the tunability in Curie temperature T_C via electrostatic doping [93]. These works boost both the experimental and theoretical studies in the field of 2D magnet, deepening the understanding of physics and leading to new applicable devices developments.

In this Chapter, we present a collaborative work with experimental groups on the discovery of the 2D vdW crystal $\mathrm{Cr_2Ge_2Te_6}$, published in Ref. [49]. The experimental work was mainly performed by research groups of Prof. Xiang Zhang (University of California at Berkeley) and Prof. Jing Xia (University of California at Irvine). In this project, we developed the renormalized spin-wave theory (RSWT) method described in Refs. [49, 50] and Chapter 3, and applied RSWT to the understanding of the strong dimensionality effect in the layer-dependent magnetism and sensitive magnetic-field dependence that gives rise to an unprecedented magnetic-field control of transition temperature.

5.1 Introduction

Whether or not long-range ferromagnetic order that exists in bulk can persist in 2D regime is a fundamental question, because the strong thermal fluctuations may easily destroy the 2D

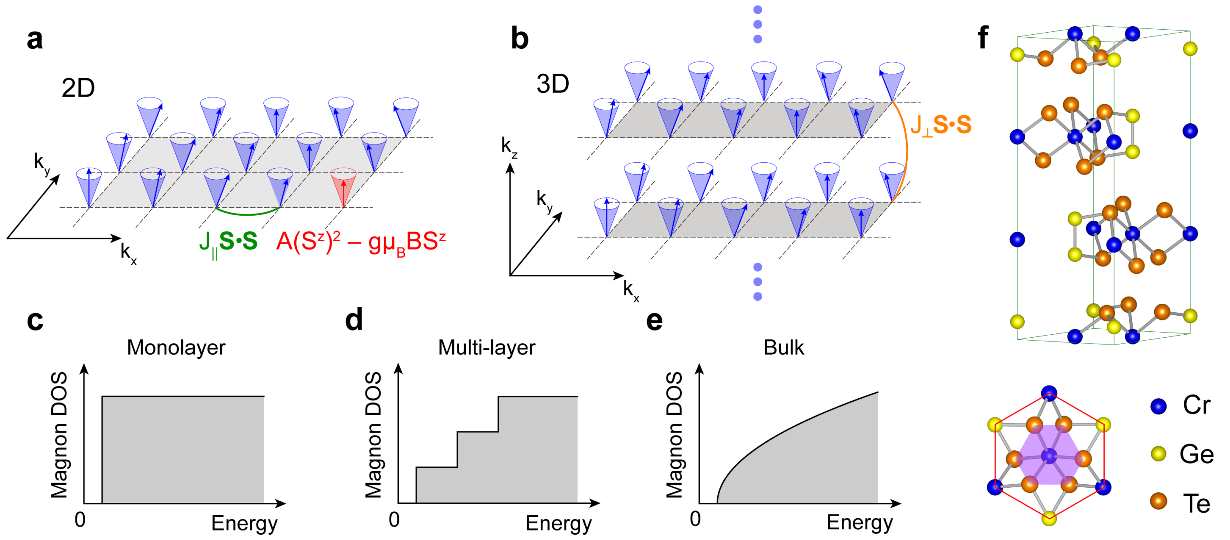

Figure 5.1: Schematics of spin-wave excitations in two and three dimensions. (a) and (b) Ferromagnetic spin-wave excitations in 2D (a) and 3D (b), with intralayer and interlayer exchange interactions $J_{\|}$ (in green) and $J_{\perp}$ (in orange), respectively, single-ion anisotropy A and magnetic field B (in red), as seen in equation (1). The cones are classical precession trajectories of thermally excited spins (arrows). (c)-(e) Schematics of magnon density of states (DOS) per spin around the low-energy band edge of monolayer (c), multi-layer (d) and bulk (e) ferromagnetic materials. The low-energy excitations from the ferromagnetic ground state follow parabolic dispersions; accordingly, the DOS is a step function in 2D and proportional to E in 3D, where E is the excitation energy. Consequently, more magnons are excited by a given thermal energy in 2D than in 3D. In 2D, the ferromagnetic transition temperature T_C is determined primarily by the excitation gap that results from the magnetic anisotropy, whereas in 3D it is determined primarily by exchange interactions. (f) Crystal structure (side and top views) of $Cr_2Ge_2Te_6$. Bulk $Cr_2Ge_2Te_6$ has a layered structure with interlayer vdW spacing of 3.4 $\mathring{A}$.

ferromagnetism, according to the Mermin-Wagner theorem [54]. We illustrate in Fig. 5.1 the physics in the ferromagnetic case. Down to the root of 2D ferromagnetism, the presence of a spin wave excitation gap (a direct result from magnetic anisotropy) is a must for long-range ferromagnetic order at finite temperature.

This work is in collaboration with the experimental groups on the discovery of long-range ferromagnetic order in pristine $Cr_2Ge_2Te_6$ atomic layers by temperature- and magnetic field-dependent Kerr effect via scanning magneto-optic Kerr effect (MOKE) microscopy.

Bulk $Cr_2Ge_2Te_6$ crystal is a ferromagnetic semiconducting material, where Cr is the magnetic ion with a magnetic moment of 3 μ_B ($S = 3/2$) [55, 56], where μ_B is Bohr magneton. The magnetic Cr ions form a honeycomb structure (as graphene) in each 2D layer, and a few-layer sample is formed with an ABC-type stacking, as sketched in Fig. 5.1. Bulk $Cr_2Ge_2Te_6$ has a ferromagnetic phase transition Curie temperature T_C of 66 K, whereas the six-layer sample has $T_C = 10$ K [49], indicative of strong dimensionality effect. Furthermore, it is discovered that T_C can be largely tuned with small external magnetic fields in 2D [49], in strong contrast to 3D. Our theoretical results [56] show quantitative agreement with experimental measurements, demonstrating the reliability of this method and revealing strong dimensionality effect. The very large influence of a small external magnetic field on the ferromagnetic transition temperature in $Cr_2Ge_2Te_6$ owes its origin in having a very small intrinsic anisotropy in the system.

5.2 Experimental results summary

In this project [49], the experimentalists successfully exfoliated the vdW $Cr_2Ge_2Te_6$ to as thin as bilayer limit. Then the long-range ferromagnetic order in pristine $Cr_2Ge_2Te_6$ atomic layers was observed using the temperature- and magnetic-field-dependent Kerr effect study via scanning magneto-optic Kerr microscopy. In the soft, 2D ferromagnetic vdW crystal $Cr_2Ge_2Te_6$, an achieve unprecedented magnetic field control of the ferromagnetic transition temperature using surprisingly small fields (≤ 0.3 T). In this Section, we summarize the main experimental observations in Ref. [49].

Experiment has observed ferromagnetism in few-layer, as shown by the temperature-dependent Kerr image of the bilayer sample in Fig. 5.2(a) with the a small out-of-plane magnetic field of 0.075 T to stabilize the spins. Strictly speaking, upon applying an external magnetic field, we might no longer have a well-defined ferromagnetic phase transition (which is defined at temperature T_C with zero field); therefore, we refer to it as a transition between a ferromagnetic-like state and a paramagnetic-like state, separated by a transition temperature T_C^*. Fig. 5.2(b) and (c) show the emergence of ferromagnetic order in bilayer $Cr_2Ge_2Te_6$ at different temperatures: at 40 K, the Kerr intensity of the scanning area is hardly discernible, except in the region corresponding to thicker flakes (≥ 3 layers); as the temperature decreases, the long strip becomes more easily recognizable, as can be seen at the temperature of liquid helium (~ 4.7 K), the long bilayer strip becomes clearly distinguishable from the surrounding bare substrate, in terms of the intensity of the Kerr rotation angle.

A strong dimensionality effect is observed by a thickness-dependent study under a 0.075-T field. Fig. 5.2 displays a sharp contrast in T_C^* between a bilayer with value of about 30 K to a bulk limit of about 68 K. More layer-dependent data can be found in Ref. [49]. The behaviour of T_C^* from 2D to 3D regimes is similar to the universal trend of many magnetic transition-metal thin films [94, 95], however the interlayer bonding strength in vdW crystals is 2–3 orders of magnitude weaker than that of traditional metals. The strong thickness dependence of T_C^* indicates an essential role of interlayer magnetic coupling in establishing the ferromagnetic order in $Cr_2Ge_2Te_6$ crystals.

Generally speaking, a field that is usually considered to be too small to affect transition temperatures of 3D systems (for example, < 0.5 T) can largely affect the behaviour of soft, 2D

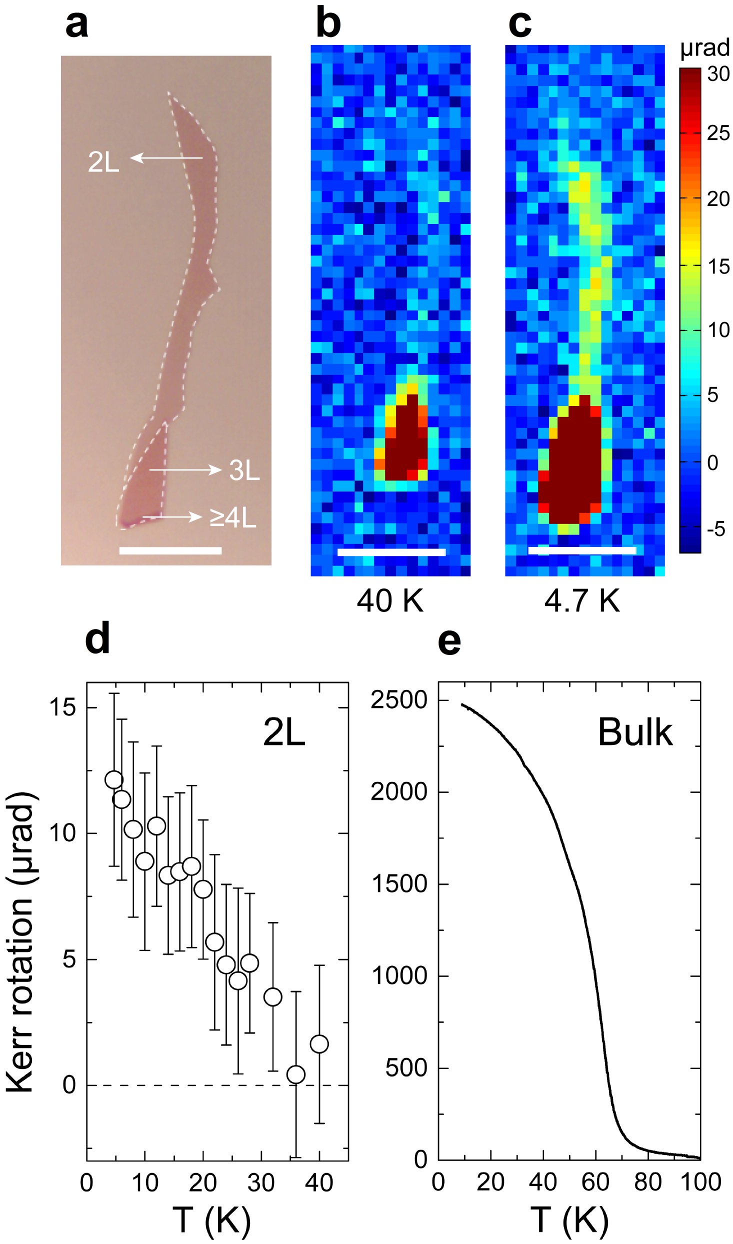

Figure 5.2: Experimental observation [49] of ferromagnetism in bilayer (2L) $Cr_2Ge_2Te_6$ and temperature-dependent Kerr rotation of few-layer and bulk $Cr_2Ge_2Te_6$ crystals. (a) Optical image of exfoliated $Cr_2Ge_2Te_6$ atomic layers, consisting of a 31-μm-long bilayer strip attached to a thicker end ($\geq$ 3 layers). (b) and (c) The emergence of a Kerr rotation signal for the bilayer flake under a magnetic field of 0.075 T, at the temperature 40 K and 4.7 K, respectively. (d) and (e) Temperature dependent Kerr rotation intensities of bilayer and bulk samples under a 0.075-T magnetic field.

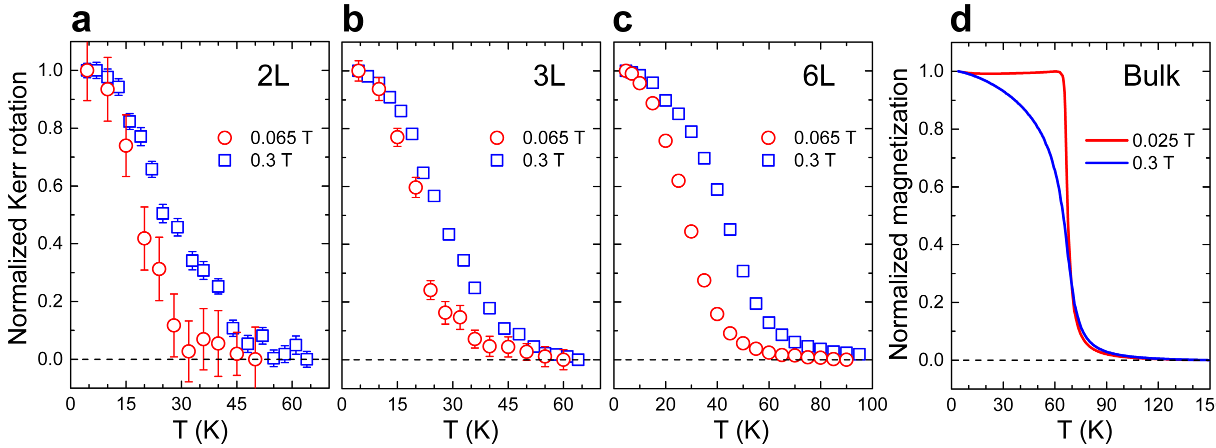

Figure 5.3: Experimental measurement [49] of magnetic field control of the transition temperature of few-layer $Cr_2Ge_2Te_6$ crystals. (a)-(c) Normalized Kerr rotation angle as a function of temperature, under two different magnetic fields: 0.065 T (red circles) and 0.3 T (blue squares), for bilayer (2L) (a) three-layer (3L) (b) and six-layer (6L) (c) flakes. The 0.3-T field shifts the curve markedly with respect to the curve for the 0.065-T field, indicating strong renormalization of T_C^* in few-layer $Cr_2Ge_2Te_6$. Error bars represent the standard deviation of sample signals and are smaller than the plotted point if not shown. d, Temperature-dependent magnetization of the bulk crystal measured by superconducting quantum interference device (SQUID) under fields of 0.025 T (red) and 0.3 T (blue). Compared with the 0.025-T field, the 0.3-T field introduces only a slightly distorted tail above T_C^*. The different behaviours below T_C^* possibly result from domains: under a 0.025-T field, multi-domains are probably formed; under a 0.3-T field, a single-domain was approached.

ferromagnetism by opening the spinwave excitation gap, as sketched in Fig. 5.1(c)–(e). The experiments examine this scenario by performing a temperature-dependent Kerr rotation study on bilayer (2L), three-layer (3L) and six-layer (6L) samples under two contrasting fields: 0.065 T and 0.3 T. As the field increases from 0.065 T to 0.3 T, the T_C^* value of bilayer flakes increases from 28 K to 44 K, that of three-layer flakes increases from 35K to 49K, and that of six-layer flakes increases from 48K to 65K, which is close to the bulk limit, as shown in Fig. 5.3. In sharp contrast, the T_C^* value in the bulk determined from SQUID measurements under fields of 0.025–0.3 T does not show clear change. Under the two magnetic fields (0.065 T and 0.3 T), the overall shift in magnetization temperature curves in 2D layers (Fig. 5.3) is clearly distinguished from the tail effect of a magnetic field on the critical region above T_C^*.

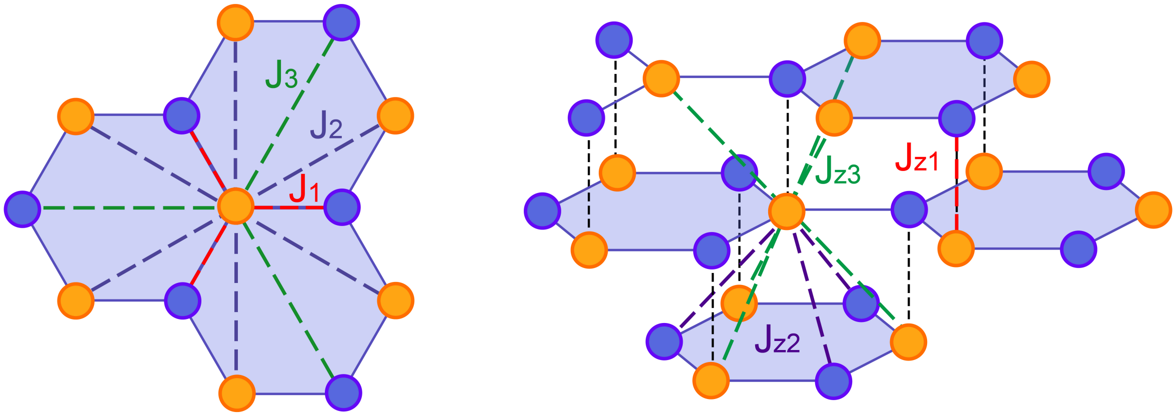

Figure 5.4: Illustration of in-plane (left) and out-of-plane (right) nearest-neighbour exchange interactions. The ABC-type stacked hexagonal lattice is a reduced illustration of atomic arrangement of Cr in bulk $Cr_2Ge_2Te_6$ crystals. J_1, J_2 and J_3 (J_{z1}, J_{z2} and J_{z3}) represent the first, second and third in-plane (out-of-plane) nearest-neighbour spinspin exchange interactions, respectively. Positive and negative J values represent antiferromagnetic and ferromagnetic exchange interactions. Each Cr site carries a spin $S = \frac{3}{2}$ in theoretical simulations. Blue and orange circles represent the Cr ions on A and B sublattices, respectively.

5.3 Theoretical density-functional theory (DFT) computational details

The RSWT formalism we presented in Chapter 3 requires parameters for the spin interactions: the exchange interactions J and single-ion anisotropy A, which come into the following Hamiltonian (resembles Eq. (3.1)),

$$H = \frac{1}{2}\sum_{i,j} J_{ij}\mathbf{S}_i \cdot \mathbf{S}_j + \sum_i A(S_i^z)^2 - g\mu_B \sum_i BS_i^z, \tag{5.1}$$

where $\mathbf{S}_i$ is the spin operator on site i, J_{ij} is the exchange interaction between sites i and j, A is the single-ion anisotropy, g is the Landé g-factor, μ_B is the Bohr magneton and B is the external magnetic field. We include six exchange pairs that are illustrated in Fig. 5.4. These parameters can either be deduced from experiments such as inelastic neutron diffraction, or be evaluated from first-principles calculations using techniques such as those based on density functional theory (DFT). In this work, we perform *ab initio* DFT calculations using the Quantum Espresso package [47] to map out these interactions. We adopt the experimental crystal structure [55] and use the local spin density approximation plus onsite Coulomb repulsion U (LSDA+U) [96] to perform the calculations. A plane-wave basis

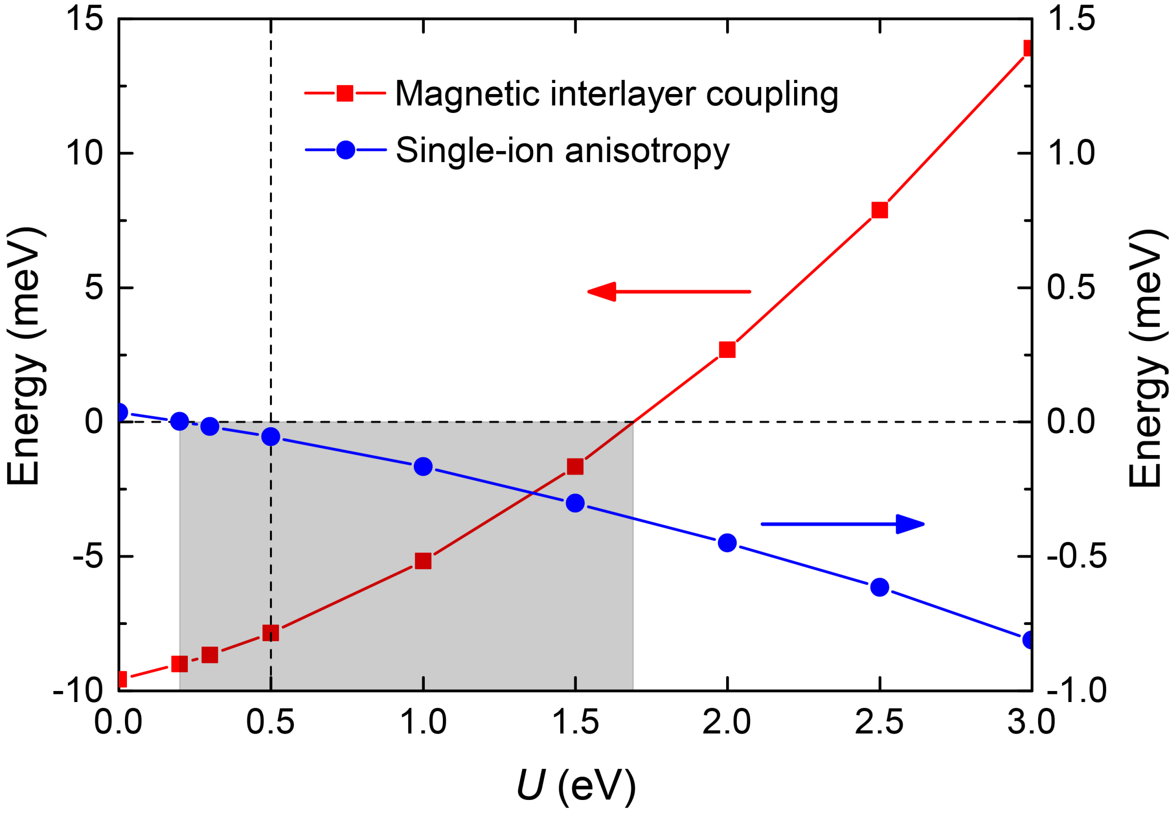

Figure 5.5: Magnetic interlayer coupling and single-ion anisotropy as a function of on-site Hubbard U in bulk $Cr_2Ge_2Te_6$. Blue circles represent single-ion anisotropy, and the magnetic anisotropy is out-of-plane (negative single-ion anisotropy) for $U > 0.2$ eV. Red squares represent interlayer magnetic coupling, which is ferromagnetic (negative) for $U < 1.7$ eV. Bulk $Cr_2Ge_2Te_6$ is a ferromagnet with an out-of-plane easy axis. Therefore, the range $0.2\text{eV} < U < 1.7\text{eV}$ (shaded area) could qualitatively reproduce the bulk magnetic property. We set $U = 0.5$ eV in the subsequent calculations, because the experimentally estimated single-ion anisotropy in the bulk is small.

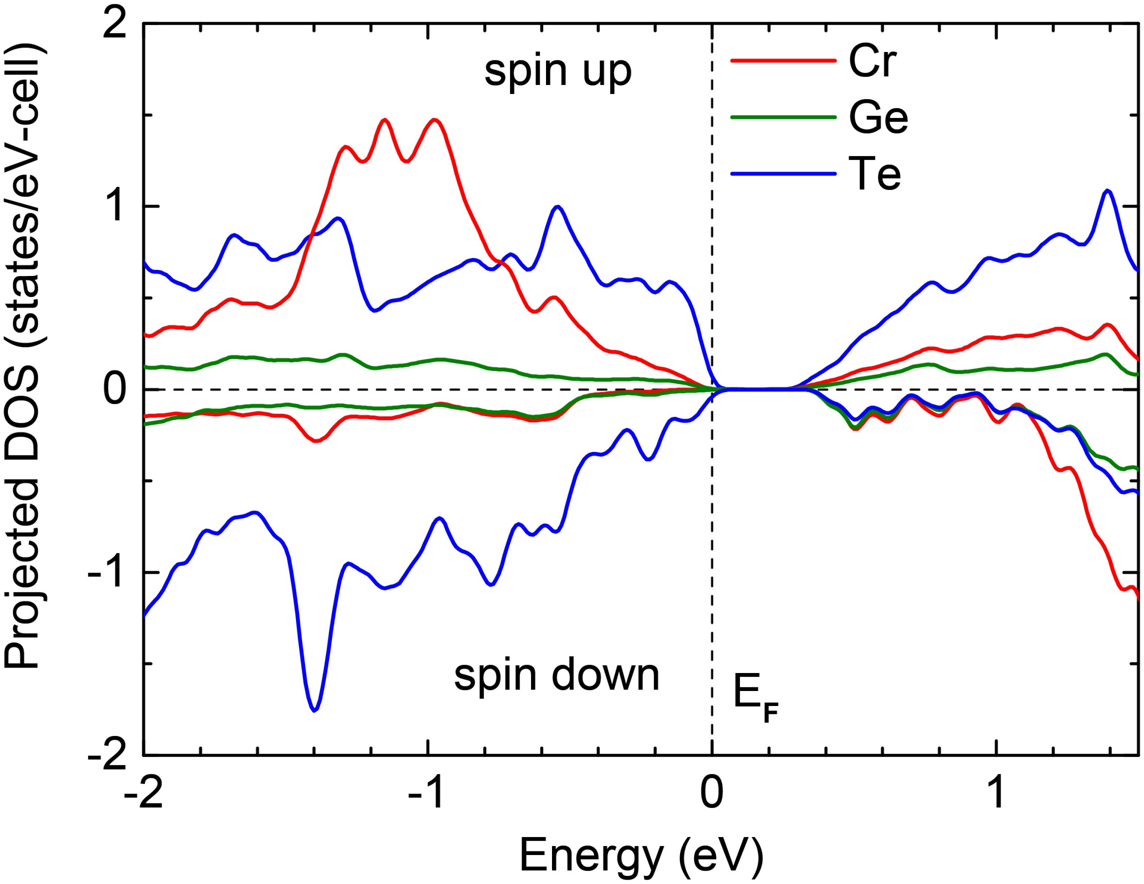

Figure 5.6: Projected and spin resolved density of states (DOS) for the electronic states in bulk $Cr_2Ge_2Te_6$ from first-principles LSDA+U calculations. The top of the valence bands is set to zero energy, indicated by the dashed line. We project the band states onto different atomic orbitals. The results show that bulk $Cr_2Ge_2Te_6$ is a semiconducting ferromagnet, consistent with experiments.

set with 60 Ry energy cutoff in wavefunctions is used. We use projector-augmented-wave (PAW) pseudopotentials from PSLIBRARY [97]. We adopt the four-state mapping analysis to calculate J, taking the advantage that the effects from other sites can be canceled out [98]. For single-ion anisotropy A, we use the full-spinor basis and include fully relativistic effects (i.e., spin-orbit coupling) in the pseudopotentials, and calculate the energy difference per ion between out-of-plane and in-plane spin configurations.

We find that the description of the magnetic properties of $Cr_2Ge_2Te_6$ is sensitive to the choice of the parameter U within LDSA+U. To treat the on-site correlation properly, we adjust the U value to reproduce the correct experimental magnetic ground state of bulk $Cr_2Ge_2Te_6$. Bulk $Cr_2Ge_2Te_6$ is a ferromagnetic insulator with easy-axis anisotropy pointing

Parameters (meV)	J_1	J_2	J_3	J_{z1}	J_{z2}	J_{z3}	A	Bulk T_C
J^{DFT}	-3.76	0.08	-0.16	0.05	-0.12	-0.38	-0.05	90 K
$J = \alpha J^{\mathrm{DFT}}$	-2.71	0.058	-0.115	0.036	-0.086	-0.27	-0.05	65 K

Table 5.1: Exchange parameters and single-ion anisotropy derived from LSDA+U calculations within DFT. The set of parameters J^{DFT} mapped out from the direct LSDA+U calculations give a bulk $T_C = 90$ K. Considering the uncertainties in DFT and for better comparison with experiment, we scale the exchange parameters J DFT with a scaling factor $\alpha = 0.72$ to derive a new set of J, giving a bulk $T_C = 65$ K (as compared to the experimental value of 66 K [49]).

along the (111) direction, perpendicular to the vdW planes. We find that for $U < 0.2$ eV, the system becomes in-plane anisotropic; for $U > 1.7$ eV, the interlayer coupling becomes antiferromagnetic, which means that the bulk crystal becomes an antiferromagnet, see Fig. 5.5. Therefore, we identify the reasonable range of U values to be 0.21.7 eV. In our subsequent calculations, we fix $U = 0.5$ eV, which gives an easy-axis anisotropy with $A = 0.05$ meV (see Section 5.4). For experimental bulk anisotropy, we can estimate the easy-axis single-ion anisotropy A from a mean-field perspective, that the saturation field $B_s \approx 0.5$ T of $B_{\parallel}$ [49], $AS^2 = g\mu_B B_s S/2$, where $S = 3/2$, giving an estimated $A = 0.02$ meV for bulk $Cr_2Ge_2Te_6$. Therefore, our choice of $U = 0.5$ which results from DFT with $A = -0.05$ meV should give relatively reasonable and stable results, considering the level of accuracy of DFT in treating this special case (with a very small energy scale). Fig. 5.6 plots the projected electron density of states (DOS) for bulk $Cr_2Ge_2Te_6$. As we can see from the figure, it is a semiconductor with a band gap of ~ 0.3 eV and is ferromagnetic with the Cr atoms contributing mostly to the magnetic moment.

5.4 Parametrization of spin Hamiltonian for RSWT calculations

We compute a total of six distinct near-neighbor interactions (three in-plane pairs and three interlayer pairs, see Fig. 5.4), from the four-state energy mapping analysis [98]. In particular, J_1, J_{z1}, J_{z2} and J_{z3} are calculated using a $2 \times 2 \times 1$ supercell. For J_2 and J_3, instead of using bulk structures, we used a monolayer $2\sqrt{3} \times 2$ supercell. We find that under the framework of RSWT, the calculated exchange interactions overestimate the bulk transition temperature, giving a bulk $T_C = 90$ K (see Table **??**) whereas the experimental value is 66 K [49]. (Note that in 3D crystalline systems, the effect of small anisotropy to T_C is negligible, as we will show later.) This in fact is a rather satisfactory result, considering the various approximations made and the large uncertainty in DFT in treating this complex system. However, to accurately compare with experimental results, we rescale all the exchange interactions J by a universal factor of 0.72 [49] (see Table **??**) to fit the bulk T_C to the experimental value. The rescaled exchange interactions give a theoretical bulk $T_C = 65$ K. We use this set of rescale exchange parameters (see Table 4.1) fixed through all later calculations, including those for the bulk and the few-layer cases. The single-ion anisotropy A is extracted by calculating the energy difference between all spins along the z direction and along the x or y directions in the bulk. Spinorbit coupling is included by using fully relativistic pseudopotentials. For the value of single-ion anisotropy, DFT calculation gives $A = -0.05$ meV (we do not rescale A), and the experimentally estimated value is $A = -0.02$ meV in the bulk (as discussed above). In 2D atomic layers, however, experiments show an almost negligible anisotropy (yet finite) [49].

5.5 Layer-dependent 2D ferromagnetic transition temperature

Now, we show the results using RSWT of layer-dependent transition temperature T_C^* in the presence of a small external field $B = 0.075$ T in Fig. 5.7. We first set $A = 0$ in the few-layer calculations, and discuss the effects from reasonably small anisotropy afterwards. RSWT quantitatively reproduces and explains the experimental results [49]. The system shows a very strong dimensionality effect with T_C^* dropping significantly from the 3D bulk regime to the quasi 2D regime with a few atomic layers [99], in line with experimental observations. At finite temperatures, spontaneous symmetry breaking (ferromagnetic order) takes place in 3D for the rotationally invariant isotropic Heisenberg system ($A = 0$ and $B = 0$), but is completely suppressed by the thermal fluctuations of the long-wavelength gapless NambuGoldstone modes in 2D. In particular for ferromagnetism, this is due to an increasing value of the available DOS at the onset of the magnon DOS (per spin) as one goes from the bulk to a few layers, as discussed in the schematic DOS in Fig. 5.1. We plot the calculated DOS at $T = 0$ K (at LSWT level) in Fig. 5.8. It indeed shows that the shape at the onset is a step function in the monolayer case (the dispersion of the low-energy excitations is quadratic with the wavevector). In the multilayer case, the step height becomes smaller due to interlayer coupling, and finally in the limit of bulk (infinite number of layers), the value of the DOS at the onset is zero. Therefore, as the number of layers decreases, it requires a lower temperature to excite enough spin-wave excitations to eliminate the net magnetization, resulting in a lower T_C (in general T_C^*) in quasi 2D samples.

Nevertheless, magnetic anisotropy ($A \neq 0$ or $B \neq 0$) could establish ferromagnetic order in 2D at finite temperatures by breaking the continuous rotational symmetry of the Hamiltonian, thereby giving rise to a non-zero excitation gap in the lowest-energy mode of the acoustic magnon branch. Thermal energy at finite T_C^* excites a large number of low-energy but finite-frequency magnon modes, flattening the expectation value of the collective spins. As the number of layers increases (from 2D to 3D), the density of states per spin for the magnon modes near the excitation gap is rapidly reduced (Fig. 5.1(c)–(e) and Fig.

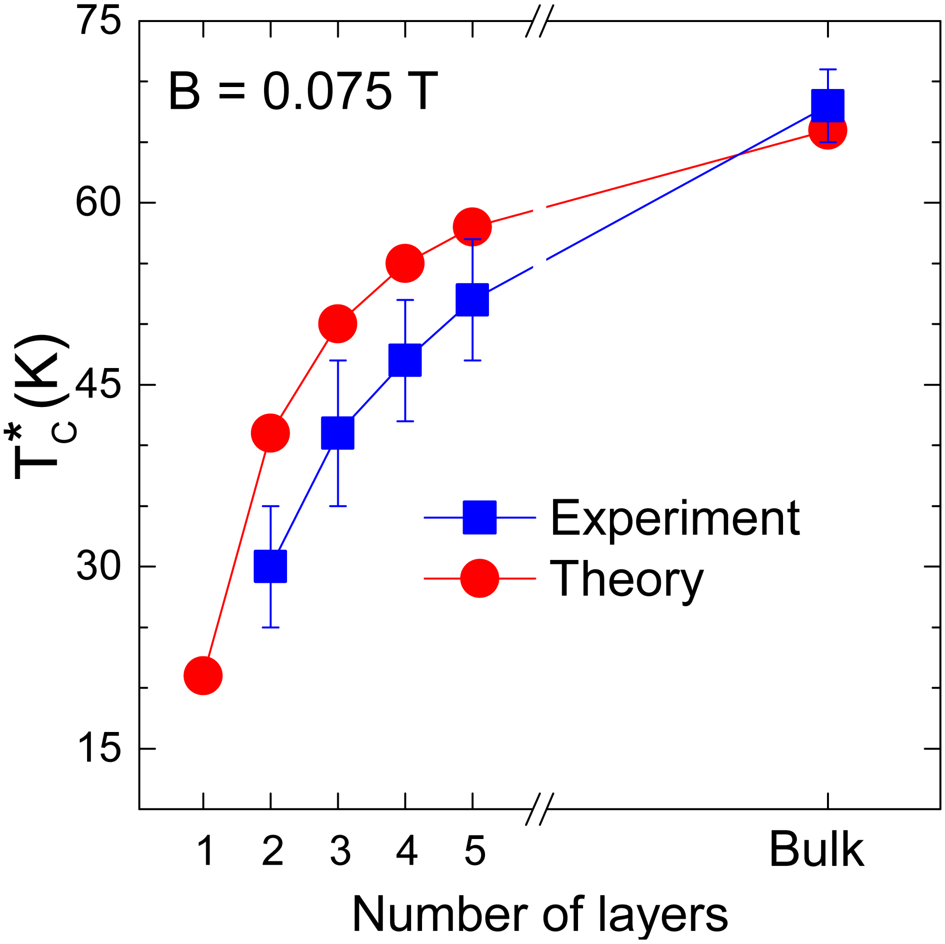

Figure 5.7: Transition temperatures T_C^* (defined as transition temperature under magnetic field) of samples of different thickness, obtained from Kerr measurements (blue squares) and theoretical calculations (red circles), with Zeeman effect from $B = 0.075$ T. A strong dimensionality effect is evident. T_C^* is determined experimentally to be the range (error bars) approximating the paramagnetic tail of the effective Kerr signal, and theoretically by the vanishing net magnetization.

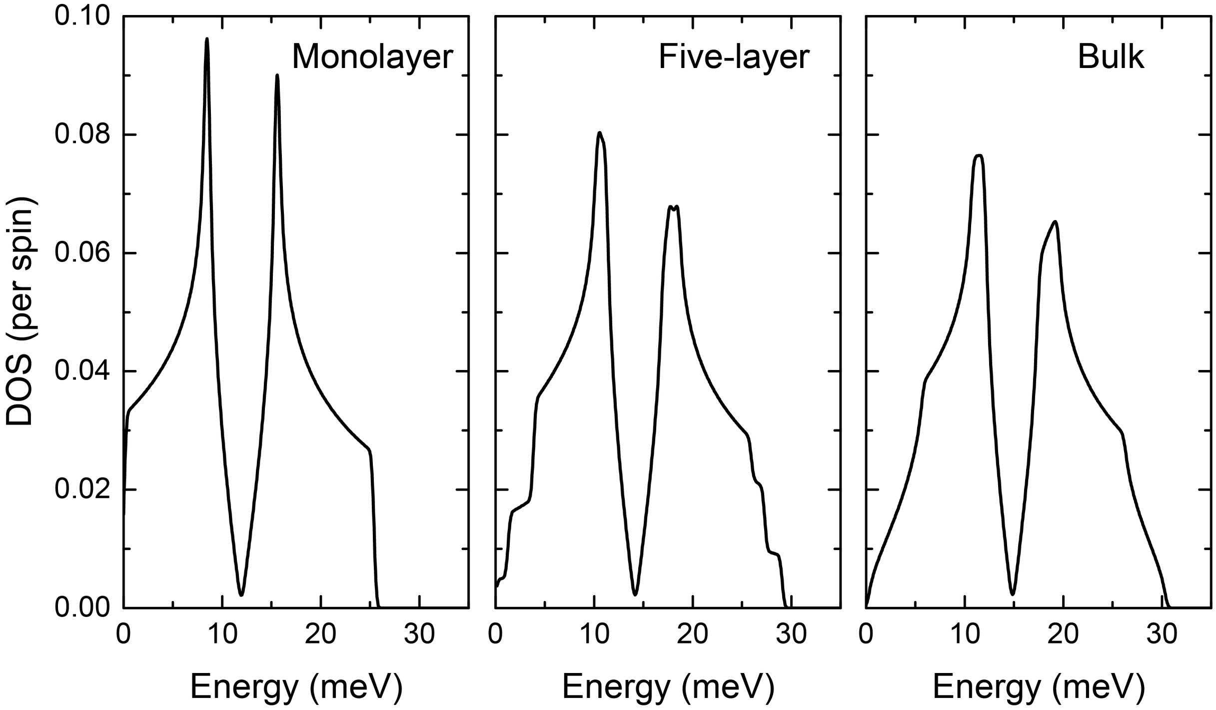

Figure 5.8: Magnon DOS of monolayer, five-layer, and bulk $Cr_2Ge_2Te_6$ at the LSWT level. The low-energy excitation of ferromagnetic spin waves follows a quadratic dispersion, giving a constant and square-root-of-energy dependence of the DOS at the onset in 2D and 3D, respectively. In multilayer case, the large step at the onset in a monolayer splits into several smaller steps at successively higher energy due to interlayer coupling, therefore reducing the value of the DOS at the onset. The different behaviors in the DOS for different number of layers lead to the dimensionality effect in the transition temperature.

5.8), meaning that a higher T_C^* is required to ensure a sufficient population of excitations to destroy the long-range magnetic order, thus leading to a strong dimensionality effect.

5.6 Magnetic field dependence of 2D transition temperature

Next, we look at the effect of external magnetic field. For 3D samples, a small external field is typically used to align the magnetic domains in the experiment to determine the phase transition temperature T_C (more specifically, T_C^*); because the effect of such a small field is negligible, $T_C = T_C^*$. However, for 2D samples, if the intrinsic anisotropy of a material is small compared to the dominant exchange interaction, then a small external field can

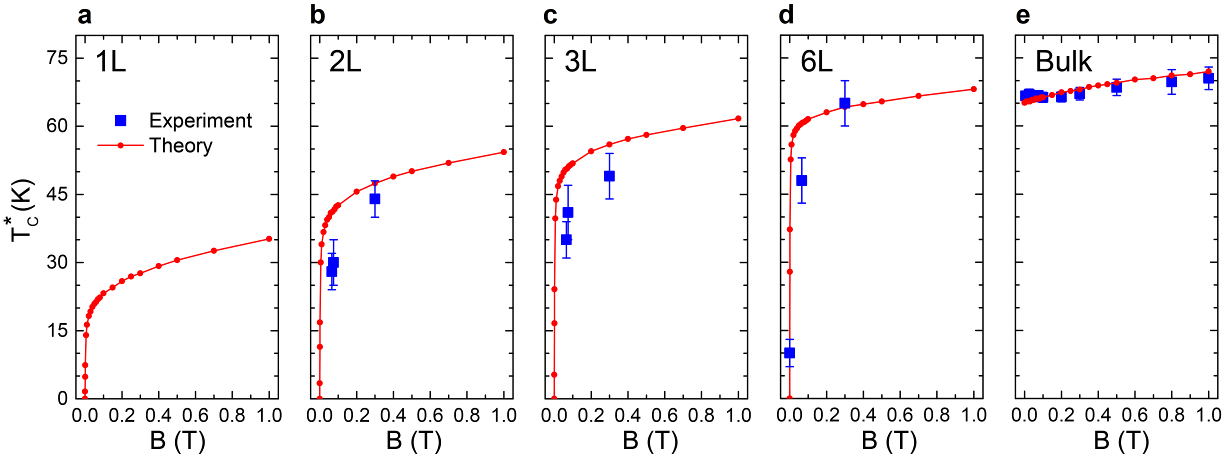

Figure 5.9: Theoretical (red circles) and experimental (blue squares) [49] field dependence of T_C^* in samples of various thickness. T_C^* values for the bulk crystal measured by SQUID are determined at the steepest slop of the magnetizationtemperature characteristic curve. Experimental T_C^* error bars arise from the uncertainty due to the tail effect. In (b–d), results at magnetic fields of 0.065 T and 0.3 T are present. In (b) and (c) for bilayer (2L) and trilayer (3L) samples, experimental results under a 0.075-T field (Fig. 2) are also plotted, and in (d) for six-layer (6L), a zero-field data point is also plotted [49]. In the 2D limit, if the single-ion anisotropy is negligibly small, the transition temperature will be very low, and can easily be tuned with a small magnetic field (for example, $B < 0.5$ T). In the bulk limit, owing to the 3D nature, such tuning is not possible.

make a big difference in establishing magnetic order due to the importance of anisotropy in suppressing magnon excitations in 2D.

By means of the RSWT method, we calculate the magnetic field dependence of T_C^* in few-layer crystals and bulk. The results from the calculations are quantitatively consistent with experimental values. As shown in Fig. 5.9, in the bulk, with applied field ranging from 0 to 1 T, the transition temperature T_C^* is almost unchanged (only varied by $\sim 10\%$) from experimental measurements, and RSWT very nicely reproduces the results [49]. However, in all of the bilayer (2L), three-layer (3L) and six-layer (6L) samples, a remarkable change in T_C^* can be obtained over the magnetic-field range used in experiments (Fig. 5.3). In particular, when we look at the quasi 2D sample of and six-layer thickness experimentally in the absence of an external magnetic field, the measured intrinsic $T_C = 10$ K [49]. In the presence of a small field $B = 0.3$ T, the measured T_C^* dramatically increases to 65 K and reaches the bulk T_C. The change in transition temperature introduced by this small field is 550% [49]! RSWT nicely captures this vast change induced by a small field for the few-layer case. The results demonstrate that the magnetic properties of quasi 2D $Cr_2Ge_2Te_6$ can be very effectively tuned using small magnetic fields, arising from the interplay between magnetic anisotropy and dimensionality in such systems.

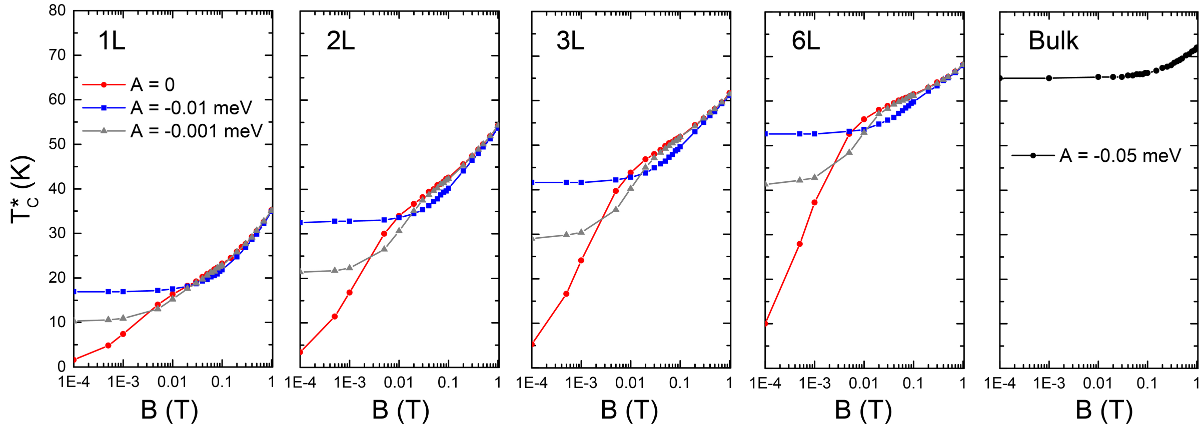

Figure 5.10: Magnetic-field-dependent T_C^* under different values of anisotropy in RSWT. Calculated magnetic-field-dependent T_C^* of monolayer (1L), bilayer (2L), three-layer (3L) and six-layer (6L) samples using $A = 0$, $A = 0.01$ meV and $A = 0.001$ meV in RSWT, as well as that of the bulk, with $A = 0.05$ meV. The x axis (B field) is plotted on a logarithmic scale. The efficient field control of transition temperature is clearly seen for small anisotropies in 2D layers.

5.7 Role of single-ion anisotropy in 2D ferromagnetism

Lastly, we discuss the effects from the single-ion anisotropy term A. In the discussions above, the theoretical results of the magnetic-field dependence for few-layer $Cr_2Ge_2Te_6$ in Fig. 5.9 are calculated with $A = 0$. The RSWT calculated results are consistent with the experimental data in Fig. 5.3 and in Ref. [49]. We consider possible outcomes with a small but finite A in Fig. 5.10, plotting T_C^* versus B assuming different values of the intrinsic single-ion anisotropy: $A = 0$, 0.01 meV, and 0.001 meV. In the absence of an external magnetic field, A will play a dominant role in the determination of T_C, as can be clearly seen in Fig. 5.10; with $A = 0.01$ meV, the 2D layers acquire a sizable T_C. However, when we increase the strength of external field B, the effect from A is weakened, as B will dominate. In this sense, A and B play similar roles in establishing magnetic ordering.

Experimentally, the six-layer sample shows a zero-field transition temperature at 10 K [49]. This temperature, comparing with our numerical results within RSWT, indicates an intrinsic anisotropy in the six-layer sample to be smaller than 0.001 meV in magnitude, because from the calculations $A = 0.001$ meV already gives 41 K in six-layer (see 6L results of Fig. 5.10). Such a tiny intrinsic anisotropy enables an external magnetic field to control T_C^* over a wide range; but its finite value makes the material intrinsically ferromagnetic, evident by the experimentally observed finite remanence at zero field [49].

Looking into the details of Fig. 5.10, we observe a seemingly counterintuitive behavior that comes from the cross over between $A = 0$ and $A \neq 0$ curves: at some finite temperature and at fixed strength of external magnetic field, the T_C^* with a nonzero A is lower than the one with $A = 0$. Naively, we would expect adding a finite A would enhance the total anisotropy and therefore increase T_C^*. This is true when the temperature is very low; that is, the magnon excitation gap is larger with a nonzero A. However, the two types of anisotropies are different in essence – the single-ion anisotropy term is renormalized by temperature (from the $2A \langle n_{\mathbf{k}} \rangle b_{\mathbf{k}_0 \nu_0}$ term), whereas the magnetic field term is not (see Eq. (3.23)). Thus, at higher temperature, the temperature renormalization effect becomes stronger by reducing the excitation gap and hence lowering T_C^*. Therefore, this unexpected behavior can be seen as a consequence of the renormalization of the spin-wave energies by temperature.

5.8 Conclusion

The collaborative work presented in this Chapter reports intrinsic ferromagnetism in the 2D vdW crystal $Cr_2Ge_2Te_6$, in which a strong dimensionality effect arises from weak intrinsic anisotropy. Through the effective engineering of the magnetic anisotropy using small magnetic fields, an unprecedented magnetic-field control of transition temperatures in soft, 2D ferromagnetic vdW crystals is achieved. Our theoretical study based on the RSWT method developed in Chapter 3 corroborates the experimental observations very well. We study the layer-number dependence and the magnetic-field dependence of the transition temperature. Our theoretical results show that the few-layer $Cr_2Ge_2Te_6$ samples behave as 2D ferromagnet with tiny (yet finite) intrinsic anisotropy, and the transition temperature can be tuned over a large range with small external magnetic fields, in nice agreement with the experimental observation. Our work provides an accurate theoretical method to study 2D ferromagnetism in atomic layers of real materials with physical insights.

Chapter 6

Generation of anisotropic Dirac fermions and asymmetric Klein tunneling in black phosphorus superlattices

Artificial lattices have been employed in a broad range of two-dimensional systems, including those with electrons, atoms and photons, in the quest for massless Dirac fermions with high flexibility and controllability. Establishing triangular or hexagonal symmetry, from periodically patterned molecule assembly or electrostatic gating as well as from moiré pattern induced by substrate, has produced electronic states with linear dispersions from two-dimensional electron gas (2DEG) residing in semiconductors, metals and graphene. Different from the commonly studied isotropic host systems, here we demonstrate that massless Dirac fermions with tunable anisotropic characteristics can, in general, be generated in highly anisotropic 2DEG under slowly varying external periodic potentials. In the case of patterned few-layer black phosphorus superlattices, the new chiral quasiparticles exist exclusively in certain isolated energy window and inherit the strong anisotropic properties of pristine black phosphorus. These states exhibit asymmetric Klein tunneling, in which the transmission probability of the wave packets with normal incidence is no longer unity and can be tuned and controlled. In general, the direction of wave packet incidence for perfect transmission and that of the normal incidence are different, and the difference can reach more than 50° under an appropriate barrier orientation in black phosphorus superlattices. Our findings provide insight to the understanding and possible utilization of these novel emergent chiral quasiparticles.

In this chapter, we present the theory of generation of anisotropic Dirac fermions and asymmetric Klein tunneling in few-layer black phosphorus under superlattice potentials. The major results of this project are published in Ref. [100].

6.1 Introduction

The unusual relativistic-like quasiparticles in graphene, known as massless Dirac fermions, intrinsically come from the symmetry of its honeycomb crystal structure [52, 101] and therefore can be designed and manipulated in other systems as first theoretically predicted [102, 103]. Many experimental efforts have been successfully performed to generate Dirac fermions – in common isotropic two-dimensional electron gas (2DEG) on metal or in semiconductor quantum wells under slowly varying periodic potentials formed by molecule assembly or patterned gate [104–109], as well as with ultra-cold atoms trapped in honeycomb optical crystal structures[110, 111] and photons confined in photonic crystals [112, 113]. Similarly, lattice mismatch between graphene and substrate can introduce a long-range moiré superlattice potential, giving rise to new generation of Dirac points in graphene [114–121]. These platforms combine advantages of an extended degree of control and various foreign modifications [107], making much physics accessible such as the interplay between Dirac fermions with strong correlation[105] and spin-orbit coupling [108], and Hofstadter butterfly effect [116–120]. Simple symmetry argument prescribes a trigonally modulating potential to generate massless Dirac fermions from an isotropic system [102, 107]. However, previous studies do not address the questions: (1) whether these novel quasiparticles can arise from highly anisotropic host states, and (2) what new physical insights and phenomena one can obtain from such systems?

In this work, we show that massless Dirac fermions with chiral character can in general be generated from anisotropic 2DEG systems. We moreover propose that semiconducting few-layer black phosphorus [122, 123] is an ideal system to achieve this aim upon electron or hole doping [124, 125]. Electron and hole carriers in black phosphorus exhibit strong anisotropy with their effective masses along the two orthogonal crystal axes (m_x^* and m_y^*) differing by an order of magnitude [126–128], forming highly anisotropic two-dimensional systems. With properly designed periodic potential (that can be realized with patterned molecule assembly or electrostatic gating under laboratory conditions), we predict that highly anisotropic massless Dirac fermions are generated and that the ratio of the group velocities along the two crystalline directions reaches $(m_x^*/m_y^*)^{\frac{1}{2}} \sim 1/3$. These anisotropic quasiparticles moreover exist within an isolated energy window separating from other states, hence are expected to be quite measurable and controllable.

Unlike the Klein tunneling process in graphene with carriers normal incident upon a potential barrier always experiencing perfect transmission [129, 130], we find that the anisotropic massless Dirac fermions generated in black phosphorus superlattices allow for an asymmetric Klein tunneling, in which the normal incident wave packets are no more unimpeded and can be generally tuned and controlled. Specifically, in the case of black phosphorus superlattices, that the directions of normal incidence and perfect transmission can be differed by more than 50°. The anisotropic massless Dirac fermions and the asymmetric Klein tunneling, can be manifested in various experiments including scanning tunneling microscope, electrical transport, quantum oscillations, quantum Hall, magnetoresistance etc. Moreover, these novel tunable features could provide new ingredients in applications [131, 132] such as electron optics, where the system serves as anisotropic crystals and negative refraction and anomalous reflection of the chiral Dirac fermions are expected.

6.2 Generation of anisotropic Dirac fermions in few-layer black phosphorus under superlattice potentials

6.2.1 Effective Hamiltonian perturbative analysis

Monolayer and few-layer black phosphorus is a direct gap semiconductor with a band gap of the order of 1 eV at the Brillouin zone center. We shall use this case as a generic example of an anisotropic massive 2DEG. Let us begin with the effective Hamiltonian for the Γ valley of a few-layer black phosphorus (or monolayer phosphorene) [128, 133],

$$H_{\text{2DEG}}(\mathbf{p}) = \frac{p_x^2}{2m_x^*} + \frac{p_y^2}{2m_y^*}, \tag{6.1}$$

where p_x and p_y are the crystal momenta along the x and y direction, respectively. Eq. (6.1) is valid around both the conduction band minimum and valence band maximum, where the band disperses much faster along the x direction (armchair direction) than along the y direction (zigzag direction) as plotted in Fig. 6.1(a) and 6.1(b). An external periodic potential varying much slower than the interatomic distances would create a superlattice structure, mixing the states in the Γ valley. For an isotropic 2DEG, with $m_x^* = m_y^*$, a triangular potential would strongly mix states near three $\mathbf{K}_j$ ($j = 1, 2, 3$) points (with $\mathbf{K}_j = \mathbf{p}_j/\hbar$, determined by the periodicity and orientation of the external potential) with degenerate energy forming a triangle in reciprocal space, leading to Dirac fermions [102, 108]. In an anisotropic 2DEG, $m_x^* \neq m_y^*$ (for electron-doped few-layer black phosphorus, we take $m_x^* = 0.15m_e$ and $m_y^* = 1.18m_e$, where m_e is the free electron mass [126]); however, in general, one can still choose three $\mathbf{K}_j$ points (on which the states have the same energy) and mix these states with some scattering strength from a designed potential, that is, solving an inverse problem.

As a demonstration, we pick up the following three $\mathbf{K}_j$ points,

$$\begin{aligned} \mathbf{K}_1 &= (K_0, 0), \\ \mathbf{K}_2 &= (-\frac{K_0}{2}, \frac{\sqrt{3}}{2}\sqrt{\frac{m_y^*}{m_x^*}}K_0), \\ \mathbf{K}_3 &= (-\frac{K_0}{2}, -\frac{\sqrt{3}}{2}\sqrt{\frac{m_y^*}{m_x^*}}K_0), \end{aligned} \tag{6.2}$$

where K_0 is a free parameter scaling the periodicity of the external potential $U(\mathbf{r})$. (For concrete illustration, we shall set the periodicity of $U(\mathbf{r})$ as 2.35 nm in plotting the figures

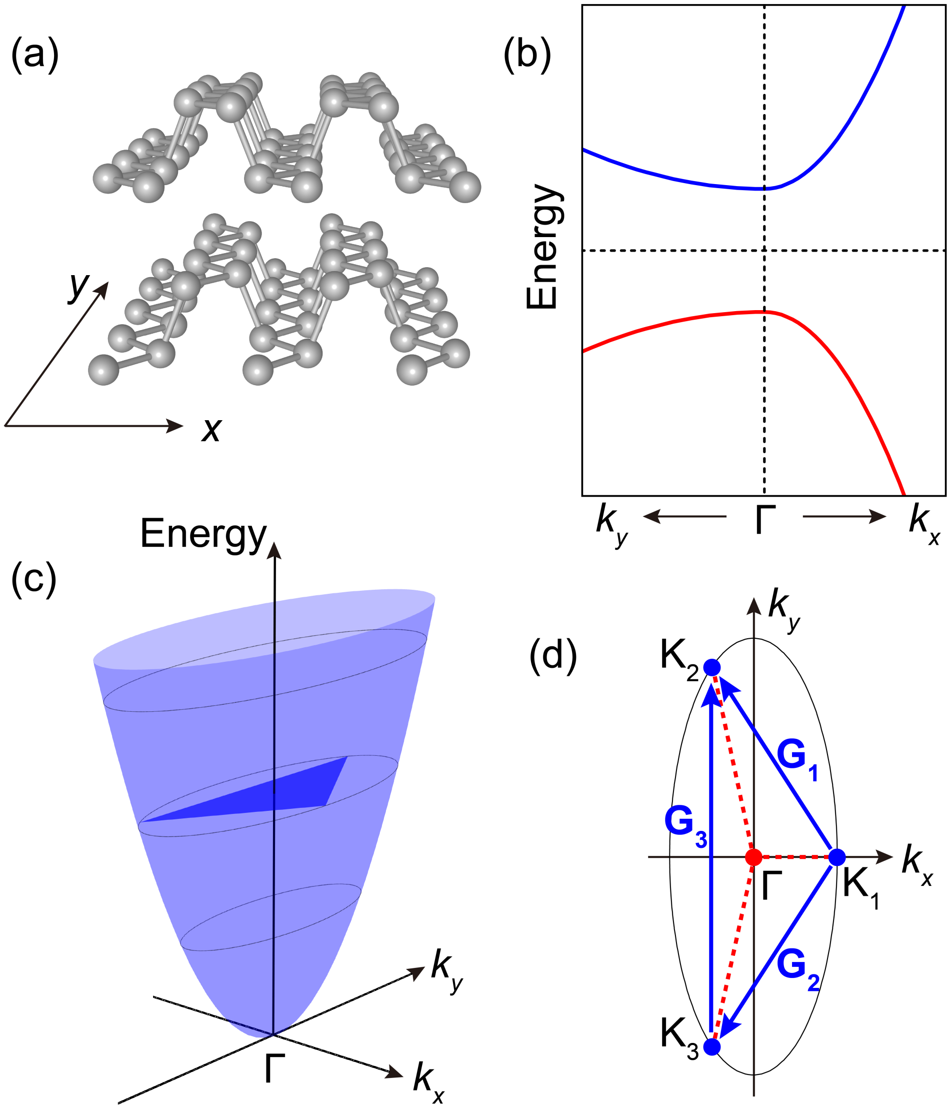

Figure 6.1: Few-layer black phosphorus. (a) Crystal structure and orientation. (b) Schematic band structure near the band gap around the Γ point. Both the lowest conduction band (blue curve) and the highest valence band (red curve) disperse much faster along the k_x direction (armchair direction) than the k_y direction (zigzag direction). (c) Geometry of a triangle defined by three points with degenerate energy in the reciprocal space, fitting into the band dispersion. (d) Elliptical isoenergetic contour. Three $\mathbf{K}_j$ points of degenerate energies on the energy surface define three $\mathbf{G}_j$ vectors. The reciprocal space of the superlattice is spanned by $\mathbf{G}_1$ and $\mathbf{G}_2$.

in the main text.) In generating Eq. (6.2), we demands the $\mathbf{K}_j$ points satisfy the condition, see Fig. 6.1(c) and 6.1(d): $E(\hbar\mathbf{K}_1) = E(\hbar\mathbf{K}_2) = E(\hbar\mathbf{K}_3)$ where $E(\hbar\mathbf{K}_j)$ is the energy eigenvalue of $H_{\text{2DEG}}(\hbar\mathbf{K}_j)$. With the degeneracy condition satisfied, the choice of the three $\mathbf{K}_j$ points still has many possibilities and may lead to rich phases described by a generalized Weyl equation [134] (in general tilted Dirac cones or open Fermi surfaces), and even higher pseudospin fermions can be generated with more $\mathbf{K}_j$ included [135, 136]. Here, we consider a special case with the three $\mathbf{K}_j$ points satisfying $\mathbf{K}_1 + \mathbf{K}_2 + \mathbf{K}_3 = 0$ (Eq. (6.2)) which guarantees electron-hole symmetry in the Dirac cones generated later. The corresponding external potential which would lead to mixing of states near those three $\mathbf{K}$ points then has reciprocal lattice vectors (and superlattice Brillouin zone (Brillouin zone)) given by $\mathbf{G}_1 = \mathbf{K}_2 - \mathbf{K}_1$ and $\mathbf{G}_2 = \mathbf{K}_3 - \mathbf{K}_1$, while $\mathbf{G}_3 = \mathbf{G}_1 - \mathbf{G}_2$ is a dependent vector shown in Fig. 6.1(d).

As a first example, we consider a simple external potential of the following form [102, 108],

$$U(\mathbf{r}) = 2W\left[\cos(\mathbf{G}_1\cdot\mathbf{r}) + \cos(\mathbf{G}_2\cdot\mathbf{r}) + \cos(\mathbf{G}_3\cdot\mathbf{r})\right], \tag{6.3}$$

where W defines the strength of the external potential. The real-space distribution of this potential is plotted in Fig. 6.2a, with $W = 0.1$ eV. We define $\mathbf{k}$ as a small wavevector, i.e. $|k| \ll |K_j|$, and expand the superlattice wavefunction near the three $\mathbf{K}_j$ points as a linear combination of $|j\rangle = e^{i(\mathbf{K}_j+\mathbf{k})\cdot\mathbf{r}}$. In this basis, the Hamiltonian up to the first order in $\mathbf{k}$ is given by $H'_{\text{2DEG}} = H_0 + H_1$, with $\langle j|\, H_0\, |j'\rangle = W(1-\delta_{jj'})$ and $\langle j|\, H_1\, |j'\rangle = \hbar\mathbf{v}_j\cdot\mathbf{k}\delta_{jj'}$, where $\delta_{jj'}$ is the Kronecker delta function and $\mathbf{v}_j = (v_{jx}, v_{jy}) = \left(\frac{\hbar K_{jx}}{m_x^*}, \frac{\hbar K_{jy}}{m_y^*}\right)$. The eigenvalues of the states for $\mathbf{k} = 0$ (Hamiltonian H_0) are $-W$, $-W$, $2W$. The resulting two degenerate states are $\frac{1}{\sqrt{2}}(0, 1, -1)^T$ and $\frac{1}{\sqrt{6}}(2, -1, -1)^T$, and we may construct a 2×2 subspace with them [102, 108]. The Hamiltonian at finite $\mathbf{k}$ in the new basis (setting the zero of energy at $-W$) reads

$$H'(\mathbf{k}) = \frac{\hbar v_0}{2}(-\sigma_z k_x - \gamma_0\sigma_x k_y), \tag{6.4}$$

where $v_0 = \frac{\hbar K_0}{m_x^*}$, $\gamma_0 = \sqrt{\frac{m_x^*}{m_y^*}}$ and $\sigma_j (j = x, y, z)$ are Pauli matrices. With a unitary transform [102, 108], we arrive at

$$H(\mathbf{k}) = \frac{\hbar v_0}{2}(\sigma_x k_x + \gamma_0\sigma_y k_y). \tag{6.5}$$

The Hamiltonian H resembles the low-energy effective Hamiltonian in graphene [52, 101], except that an extra factor γ_0 renormalizes the second term. The energy eigenvalues and the corresponding eigenvectors are therefore given by

$$E(\mathbf{k}) = \frac{\lambda\hbar v_0}{2}\sqrt{k_x^2 + \gamma_0^2 k_y^2}, \tag{6.6}$$

$$\psi_\lambda(\mathbf{k}) = \begin{pmatrix} 1 \\ \lambda e^{i\phi_s} \end{pmatrix}, \tag{6.7}$$

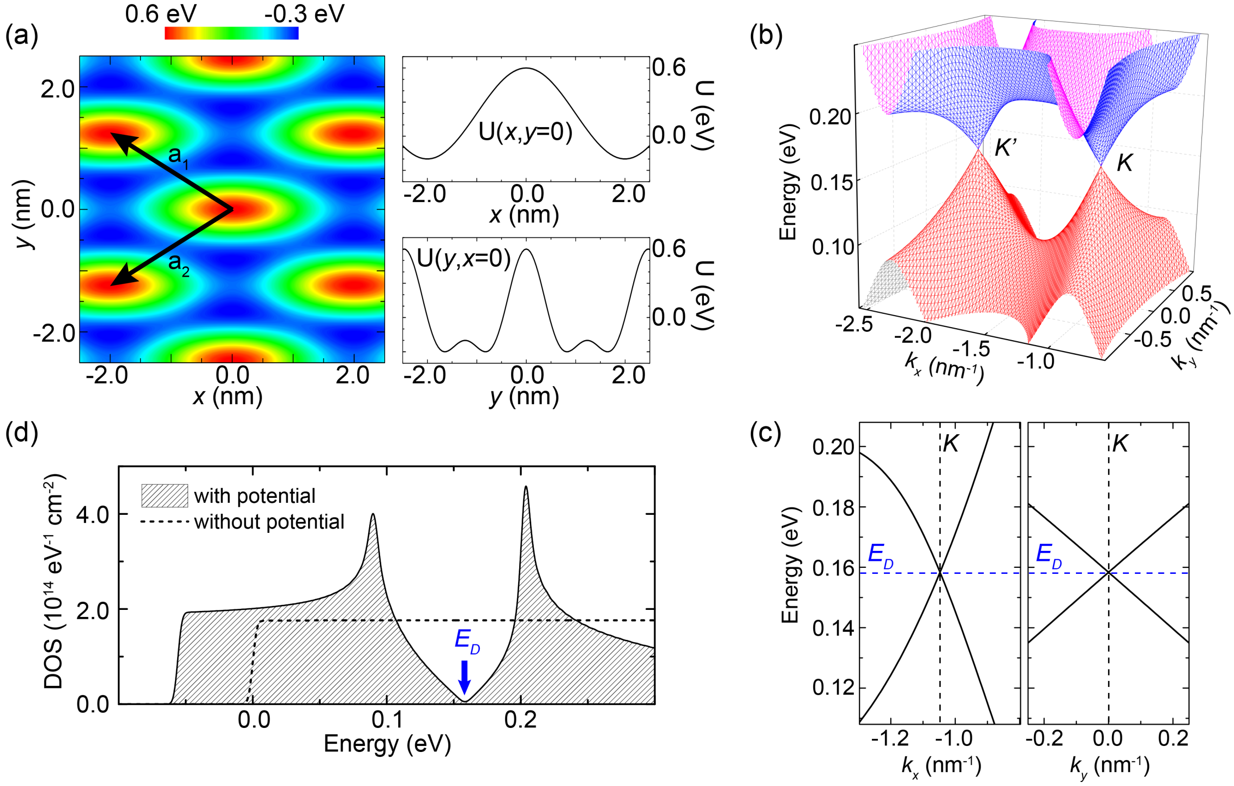

Figure 6.2: Real-space distribution of a superlattice of sinusoidal potentials defined in Eq. (3) with $W = 0.1$ eV. $\mathbf{a}_1$ and $\mathbf{a}_2$ are the superlattice lattice vectors of the external periodic potential. $|a_1| = |a_2| = 2.35$ nm and the angle between the two vectors is 63.4°. The two panels on the right are the line profiles of the external potential. (b) Band structures of the states under the sinusoidal periodic potential from non-perturbative numerical calculations. Anisotropic massless Dirac fermions are generated. (c) Band structures along two directions passing through one Dirac point, showing linearly dispersing features with strongly anisotropic group velocities. (d) DOS of the new Dirac fermions system generated from the external periodic potential, with the real electron spin degree of freedom included. Dashed line represents the DOS from the lowest-energy conduction band in pristine few-layer black phosphorus before applying the external potential.

where $\lambda = \pm 1$ denoting the upper and lower bands, and $\phi_{\mathbf{s}}$ is the polar angle of the pseudospin vector associated with $\psi_\lambda(\mathbf{k})$ with respect to k_x. Here we use the notation $\mathbf{s}$ for pseudospin. The energy dispersion is linear and these quasiparticles behave as massless Dirac fermions under the modulation of the periodic external potential. The group velocity of the linear dispersion varies with the direction of $\mathbf{k}$, resulting in an anisotropic Dirac cone band structure. The anisotropy in the group velocity, $\frac{|v_y|}{|v_x|} = \gamma_0$, originates from the anisotropic effective masses intrinsic to the original anisotropic 2DEG.

6.2.2 Numerical solutions of 2DEG under realistic superlattice potentials

We thus have demonstrated the generation of anisotropic Dirac fermions in general, and in few-layer black phosphorus in particular, for a sinusoidal potential defined in Eq. (6.3) for small $\mathbf{k}$ using a perturbative analysis. We have also numerically solved in the whole superlattice Brillouin zone for the full band structure in Fig. 6.2(b) and 6.2(c), and find that the Dirac cone indeed disperses for a large part of the Brillouin zone much faster along the k_x direction than along the k_y direction.

We now examine more complex shapes for the applied external potential within a unit cell of the superlattice. Considering possible experimental fabrications, we perform calculations on three likely cases: a conical potential of radius 0.3 nm with 2.35 nm periodicity mimicking molecule assembly [106, 109] as shown in Fig. 6.3(a), a rectangular potential of 2.67 nm long and 1.65 nm wide with 7.84 nm periodicity representing patterned electrostatic gating 6.4 [102, 104, 108] as shown in Fig. 6.4(a), and a harmonic potential of radius 0.4 nm with 2.35 nm periodicity as shown in Fig. 6.5(a). We expand all quantities with plane waves to construct the Hamiltonian. The diagonal matrix elements are the unperturbed band energies from Eq. (6.1), and the off-diagonal matrix elements are from the external potential, which is given in a Fourier expansion as

$$U(\mathbf{r}) = \sum_{\mathbf{G}} U(\mathbf{G}) e^{i\mathbf{G}\cdot\mathbf{r}}. \tag{6.8}$$

Numerical diagonalization of the Hamiltonian gives the energy and wavefunction of states in the whole Brillouin zone. Generally, if the potential has small deviations in the corresponding Fourier components given in Eq. (6.3), the generated Dirac points will be slightly shifted away from the three $\mathbf{K}_j$ [111], and the above analysis remains valid approximately.

Since the potential is inversely designed from Eqs. (6.2)–(6.6) in the main text, the lowest $\mathbf{G}$ components, which defines the size and shape of the superlattice, are similar to those defined in Eq. (6.2) in the main text. In the conical potential, a small deviation from Eq. (6.3) exists and effectively shifts the positions of the Dirac points away from the $\mathbf{K}_j$ points. However, if the profile of the potential is very localized (such as in molecule assembly), the analysis in the main text can still be treated as a good approximation to the realistic situation. Therefore, our conclusions on the generation of massless Dirac fermions

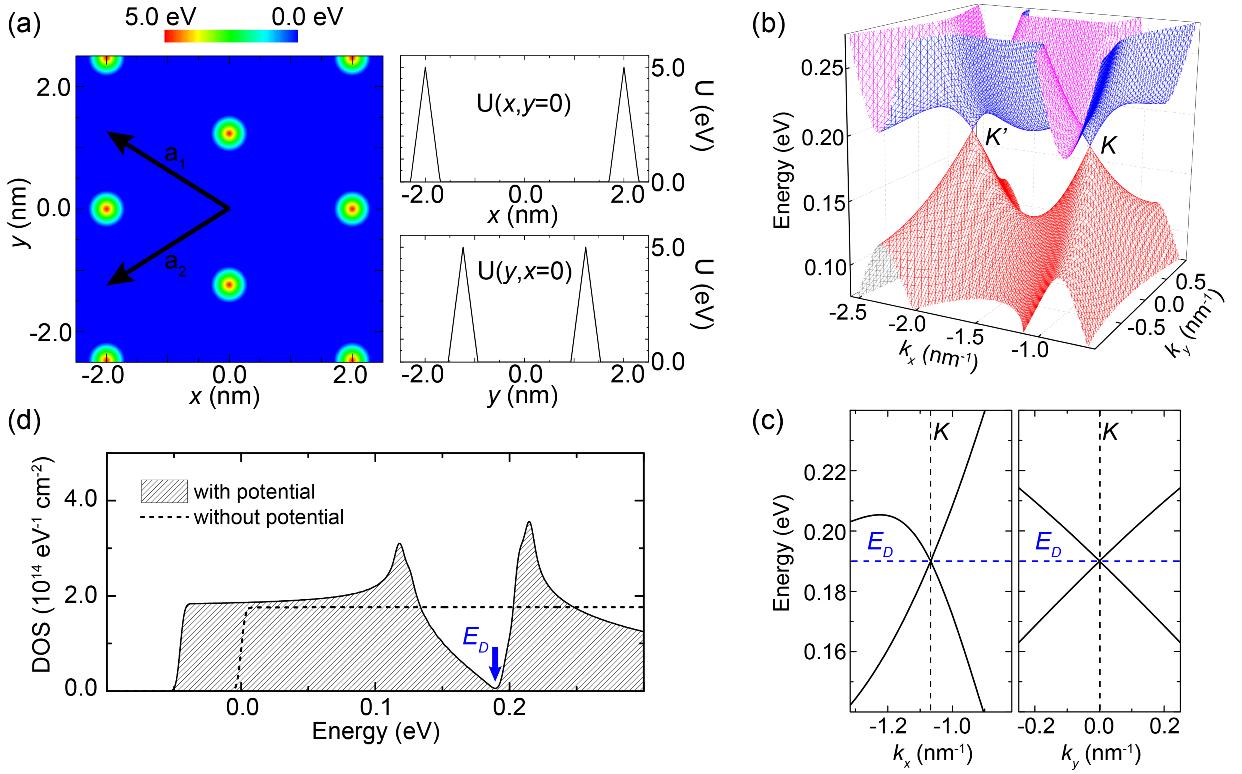

Figure 6.3: Generation of massless Dirac fermions from a superlattice of conical potentials mimicking the molecule assembly. (a) Real-space distribution of the conical potential of radius 0.3 nm with 2.35 nm periodicity (left panel), along with two line profiles plotted (right two panels). (b) Band structures corresponding to the superlattice potential in (a). (c) Band structures along two normal directions passing through one Dirac point. (d) Density of states (DOS) showing V-shape feature with a vanishing point, with the real electron spin degree of freedom included.

around $\mathbf{K}_j$ points are valid, despite the fact that the detailed band structure in other regions away from $\mathbf{k} = 0$ may vary depending on the higher $\mathbf{G}$ components of the potential. The periodicity of the conical potential is 2.35 nm, which is quite accessible in experiments, e.g., the assembly of monoxide molecules on Cu (111) surface takes a periodicity of less than 2 nm. The periodicity of the rectangular potential is 7.84 nm, in the same order of magnitude of what has been achieved with current lithographic techniques that can usually reach 1020 nm. The periodicity of the harmonic potential is 2.35 nm, and we use it to enrich the discussions.

Fig. 6.3(b) and (c), Fig. 6.4(b) and (c) and Fig. 6.5(b) and (c) show the band structures from the molecule assembly, rectangular, and harmonic potentials, respectively. The anisotropic linear dispersions are clearly seen. Note that for all potentials, the density of states (DOS) show the V-shape feature and a vanishing point, as shown in Fig. 6.2(d), Fig. 6.3(d), Fig. 6.4(d), and Fig. 6.5(d).

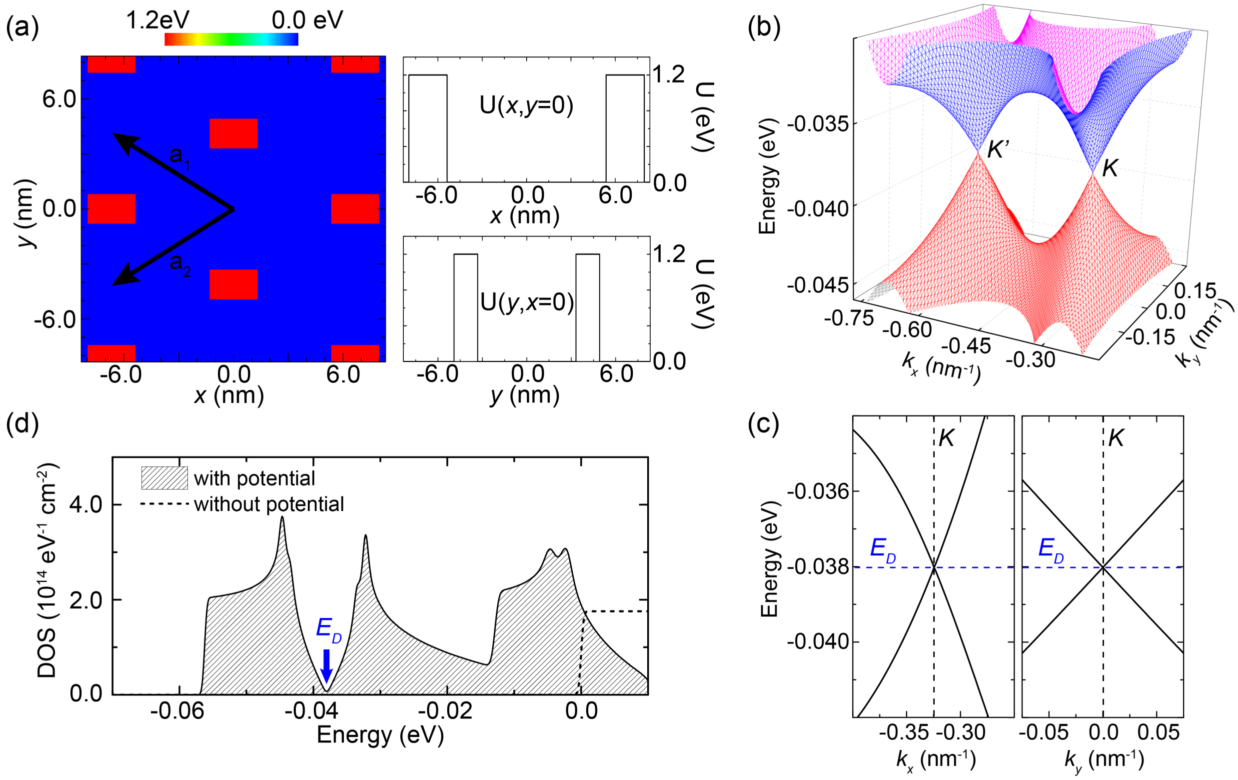

Figure 6.4: Generation of massless Dirac fermions from a superlattice of rectangular potentials mimicking the patterned electrostatic gating. (a) Real-space distribution of a rectangular potential of 2.67 nm long and 1.65 nm wide with 7.84 nm periodicity (left panel), along with two line profiles plotted (right two panels). (b) Band structures corresponding to the superlattice potential in (a). (c) Band structures along two normal directions passing through one Dirac point. (d) Density of states (DOS) showing V-shape feature with a vanishing point, with the real electron spin degree of freedom included.

We further discuss the dependence of superlattice periodicity in the electronic structures. As is shown in Fig. 6.6, we consider the sinusoidal potential (Eq. (6.3)) with different lattice constants. A smaller periodicity gives a wider energy window for the generated massless Dirac fermions, i.e., the range of linearity.

Numerical simulations (i.e., diagonalization of the Hamiltonian matrix) show that all three external potential models produce sizable massless Dirac cones in the superlattice Brillouin zone with quite achievable parameters in experiments. We have therefore given a general scheme for generating massless Dirac fermions in arbitrarily anisotropic 2DEG from the perspective of inverse design.

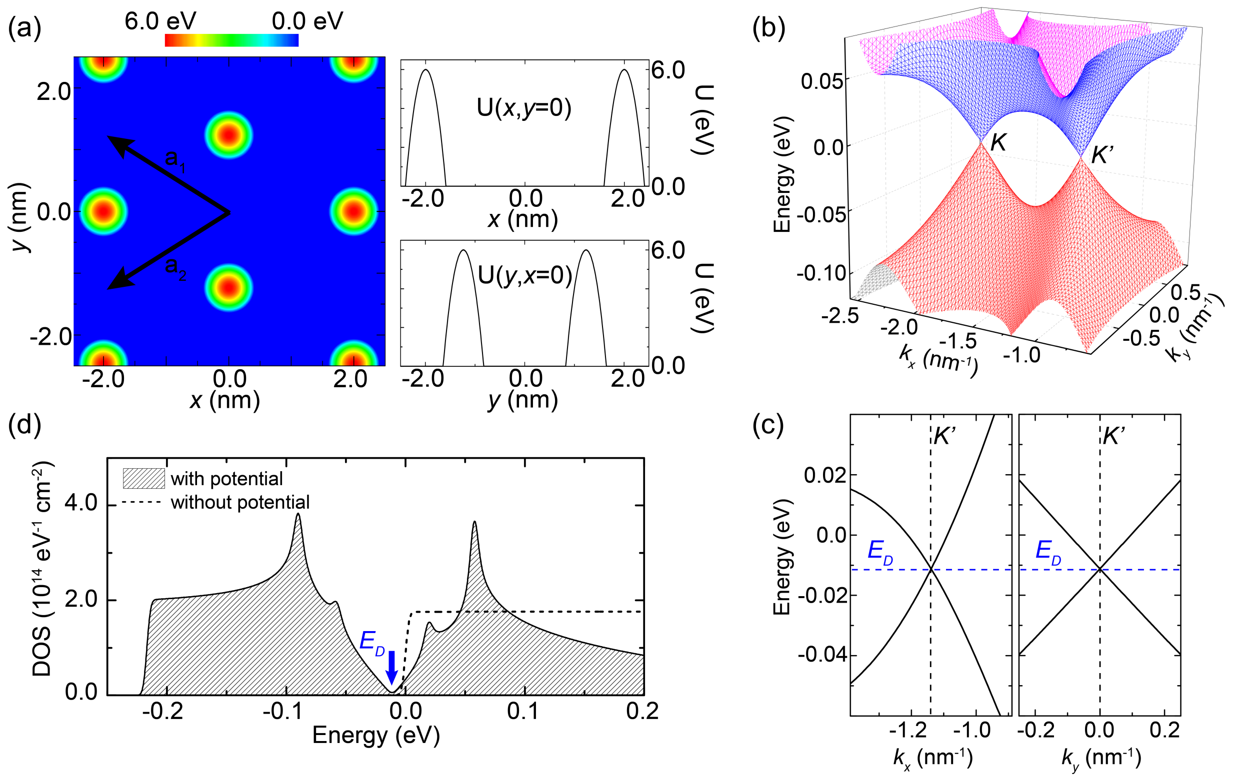

Figure 6.5: Generation of massless Dirac fermions from a superlattice of harmonic potentials. (a) Real-space distribution of a harmonic potential of radius 0.4 nm with 2.35 nm periodicity (left panel), along with two line profiles plotted (right two panels). (b) Band structures corresponding to the superlattice potential in (a). (c) Band structures along two normal directions passing through one Dirac point. (d) Density of states (DOS) showing V-shape feature with a vanishing point, with the real electron spin degree of freedom included.

6.2.3 Isolation of the new Dirac fermions within an energy window

One important prerequisite for ease in probing and utilizing the emergent Dirac fermion states is to have these states well separated from other states. In contrast, tunable and anisotropic Dirac fermions can also be generated in graphene using one-dimensional periodic potential [103, 137] due to the chiral nature of states in the original Dirac cone. However, those new states are typically obscured by other states in the same energy window [103, 137]. In the approach proposed in this work, the resulting density of states (DOS) has a clear V-shape feature and vanishes at the energy where the two new Dirac cones meet, well separated from other states, for example, see Fig. 6.2(d). Thus, they can be more easily characterized in many experiments involving properties of Fermi surface, such as scanning tunneling microscope, electrical transport, quantum oscillations, quantum Hall, magnetore-

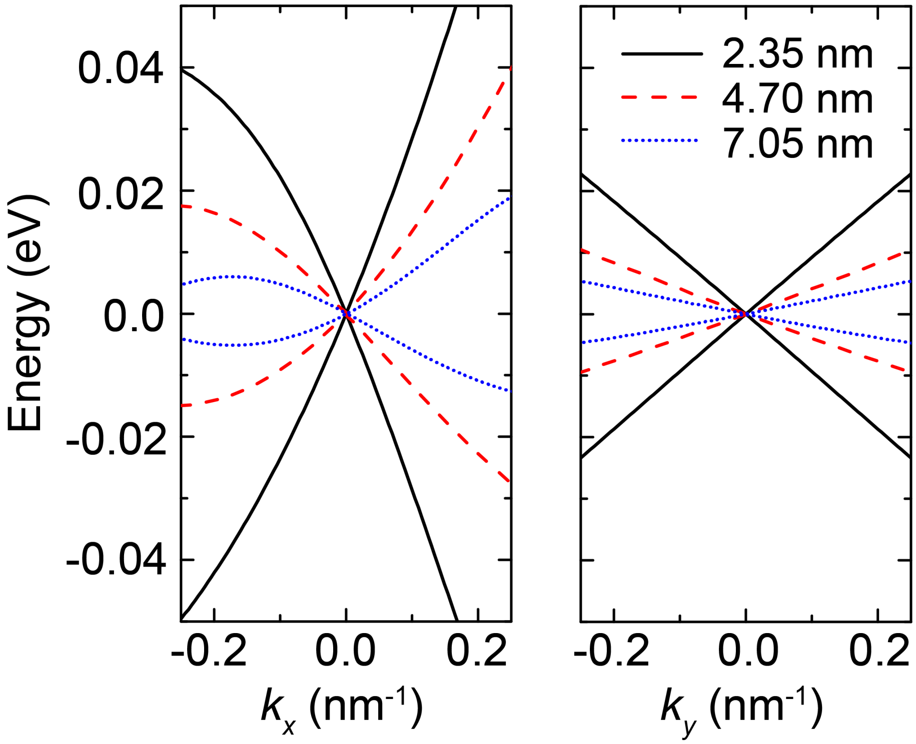

Figure 6.6: Superlattice periodicity dependence of the electronic band structure of the generated massless Dirac fermions. Three lattice constants are considered; smaller periodicity gives a wider energy window hosting the linearly dispersed Dirac states. Note that in this figure, the **k**-space and energy origin have been set to the $\mathbf{K}'$ point and the Dirac point energy, respectively, for a direct comparison of the dispersion.

sistance and electron optics. In the model potential described by Eq. (6.3), the condition $W > \hbar^2 K_0^2/16m_x^*$ should at least be satisfied to avoid obscuring states stemming from original 2DEG dispersions. Comparison with the DOS of the pristine few-layer black phosphorus in Fig. 6.3(d) indicates that the system is only slightly modified by the superlattice potential, and the intrinsic band gap of few-layer black phosphorus ($\sim$ 1 eV) is still well preserved. Therefore, in such a black phosphorus superlattice, the semiconducting phase and the massless Dirac fermion phase can be reversibly switched by tuning the Fermi level within an achievable carrier density range in the order of 10^{12} - 10^{13} cm^{-2}, which is quite achievable experimentally.

Noticing that the DOS plotted in Fig. 6.2(d), Fig. 6.3(d), Fig. 6.4(d), and Fig. 6.5(d) only contain the effects of the external potential on the original lowest-energy conduction band, whereas in a real material, other bands need to be considered as well, because they may appear in the same energy window where the Dirac fermions exist, and therefore affect the experimental measurements. Previous *ab initio* calculations [126] show that in a monolayer black phosphorus, there are other bands $\sim$ 0.2 eV higher in energy than the conduction band minimum and $\sim$ 0.5 eV in energy lower than the valence band maximum. When the number of layers increases to four, other conduction band states almost reach the same energy as conduction band minimum, but those in the valence bands are still at $\sim$ 0.5 eV away from the valence band maximum. Considering the energy of the new Dirac points and potentially interfering states, hole doping in few-layer black phosphorus should be more robust in the generation of isolated-in-energy anisotropic Dirac fermions. However, the ratio of the effective masses along two crystal axes in the hole doping regime varies a lot with number of layers (γ_0 is $\sim 1/6$ in monolayer with hole doping), while it is almost a constant as a function of layer number in conduction bands; therefore, much care is needed in designing the external potential for a sample with a particular number of layers and the choice of electron or hole doping. Nevertheless, our analyses and calculations in the above based on the conduction band states are quite general, and can be easily applied to the valence band and other systems having anisotropic two-dimensional electron gases.

Furthermore, for the purpose of experimental measurement, considering the sensitivity of black phosphorus to air, we note that the encapsulation of the material by using hexagonal boron nitride could be helpful, as it virtually does not affect the electronic structure around the band edges of black phosphorus [138].

6.3 Asymmetric Klein tunneling from anisotropic Dirac fermions

6.3.1 Solving for asymmetric Klein tunneling

Next, we explore one of the most counterintuitive phenomena in graphene, Klein tunneling [129, 130]. That is, normal-incident carriers in graphene experience a perfect transmission through a potential barrier independent of the potential width and height, made possible by the chiral nature of the linearly dispersing Dirac fermions and charge-conjugation symmetry.30 In an anisotropic Dirac cone, for states with wavevector $\mathbf{k}$ (measured from the Dirac point), the wavevector $\mathbf{k}$, the pseudospinor $\mathbf{s}$ and the group velocity $\mathbf{v}$ are all generally noncollinear at a given $\mathbf{k}$ point [103]. In fact, they are given (for the potential in Eq. (6.3)) by $\mathbf{k} = (k_x, k_y)$, $\mathbf{s} = \lambda\hbar(k_x, \gamma_0 k_y)/\sqrt{k_x^2 + \gamma_0^2 k_y^2}$, and $\mathbf{v} = \frac{\lambda v_0}{2}(k_x, \gamma_0^2 k_y)/\sqrt{k_x^2 + \gamma_0^2 k_y^2}$ (with $\lambda = \pm 1$) and related by (Fig. 6.7(a))

$$\tan\phi_{\mathbf{v}} = \gamma_0 \tan\phi_{\mathbf{s}} = \gamma_0^2 \tan\phi_{\mathbf{k}} = \gamma_0^2 \frac{k_y}{k_x}, \tag{6.9}$$

where $\phi_{\mathbf{k}}$, $\phi_{\mathbf{s}}$, $\phi_{\mathbf{v}}$ are referenced to k_x axis (note that $\mathbf{s}$ depends on λ but $\phi_{\mathbf{s}}$ does not, by definition). If the relevant Fourier components of a potential deviate from Eq. (6.3) slightly, Eq. (6.9) is then an approximation.

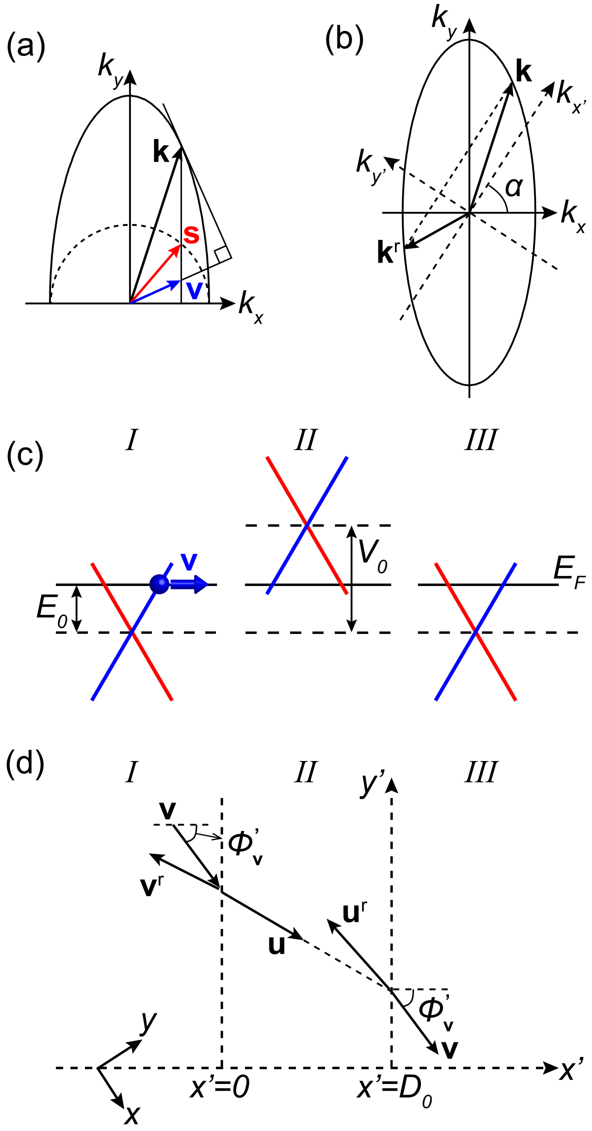

Figure 6.7: (a) Isoenergetic contour of the anisotropic Dirac cone. Wavevector $\mathbf{k}$, pseudospinor $\mathbf{s}$ and group velocity $\mathbf{v}$ are noncollinear in general. (b) With the elliptical Fermi surface plotted in the $x-y$ coordinate system, a potential barrier is created with barrier normal along the x'-direction in the $x'-y'$ coordinate system. Vectors and angles are denoted by the prime symbol if in the $x'-y'$ coordinate system, and by r if representing reflected waves. (c) An $n-p-n$ junction (relative to the Dirac point) created by a potential barrier of height V_0 in region II. (d) Tunneling process through the potential barrier. $\mathbf{u}$, $\mathbf{u}^r$, $\mathbf{q}$, $\mathbf{q}^r$, θ, θ^r, λ' in region II correspond to $\mathbf{v}$, $\mathbf{v}^r$, $\mathbf{k}$, $\mathbf{k}^r$, ϕ, ϕ^r, λ in region I.

Now we consider transmission through a potential barrier with a height of V_0 and width of D_0 ($D_0 \gg |a_1|, |a_2|$) . Let us align the minor and major axes of the elliptical Fermi surface along the k_x and k_y direction, respectively (see Fig. 6.7). We place a potential barrier in the $x' - y'$ coordinate system (rotated at some arbitrary angle α from the $x - y$ coordinate system), and make the barrier axis infinitely long along the y' direction with the width D_0 lying along the x' direction. In the continuum limit, translational symmetry along the y' direction will conserve the $k_{y'}$ wavevector component of a propagating wave impinging on the potential barrier, based on which we look for the solutions in the three regions - before (I), inside (II) and after (III) the barrier. The various notations are illustrated in the caption of Fig. 6.7 and explained in the following context.

With the problem set up in Fig. 6.7, we analytically look for the solutions of this transmission problem. The crystal axes of few-layer black phosphorus and the superlattices are defined in the $x - y$ coordinate system, and the potential barrier is defined in the $x' - y'$ coordinate system. Vectors and angles in the $x' - y'$ coordinate system are denoted with a prime symbol. We consider group velocity $\mathbf{v}$ of the incident wave packet that can hit the potential barrier, so $\phi'_{\mathbf{v}} \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$. With a given $\phi'_{\mathbf{v}}$ ($\phi_{\mathbf{v}} = \phi'_{\mathbf{v}} + \alpha$), because $\phi_{\mathbf{k}} = \frac{k_y}{k_x}$, $\phi_{\mathbf{k}}$ can be determined (depending on $\lambda = \pm 1$), and we can write the incident wavevector in region I as $\mathbf{k} = (k_x, k_y) = (|k| \cos\phi_{\mathbf{k}}, |k| \sin\phi_{\mathbf{k}})$, satisfying,

$$\begin{aligned} E_0 &= \lambda \frac{\hbar}{2} v_0 \sqrt{k_x^2 + \gamma_0^2 k_y^2} \\ &= \lambda \frac{\hbar}{2} v_0 \sqrt{k^2 (\cos^2\phi_{\mathbf{k}} + \gamma_0^2 \sin^2\phi_{\mathbf{k}})}, \end{aligned} \tag{6.10}$$

and we have,

$$|k| = \frac{2|E_0|}{\hbar v_0} (\cos^2\phi_{\mathbf{k}} + \gamma_0^2 \sin^2\phi_{\mathbf{k}})^{-\frac{1}{2}}. \tag{6.11}$$

The same vector is represented in the $x' - y'$ coordinate system as $\mathbf{k} = (|k| \cos\phi'_{\mathbf{k}}, |k| \sin\phi'_{\mathbf{k}})$, where $\phi'_{\mathbf{k}} = \phi_{\mathbf{k}} - \alpha$. In region I, the reflected wavevector takes $\mathbf{k}^r = (k^r_{x'}, k_{y'})$ in the $x' - y'$ coordinate system, because of $k_{y'}$ conservation. In the x-y coordinate system, the reflected wavevector reads,

$$\begin{aligned} k^r_x &= k^r_{x'} \cos\alpha - k_{y'} \sin\alpha, \\ k^r_y &= k^r_{x'} \sin\alpha + k_{y'} \cos\alpha, \end{aligned} \tag{6.12}$$

satisfying,

$$(k^r_x)^2 + \gamma_0^2 (k^r_y)^2 = \frac{4E_0^2}{\hbar^2 v_0^2}, \tag{6.13}$$

and we get the solution of $k^r_{x'}$, as well as the spinor angle $\phi^r_{\mathbf{s}}$. In region III, there is only one wavevector which is the same as $\mathbf{k}$.

In region II, the wavevectors are denoted as $\mathbf{q}$, $\mathbf{q}^r$ with $q_{y'} = q^r_{y'} = k_{y'}$, satisfying,

$$E_0 = \lambda' \frac{\hbar}{2} v_0 \sqrt{q_x^2 + \gamma_0 q_y^2} + V_0, \tag{6.14}$$

where,

$$\begin{aligned} q_x &= q_{x'}\cos\alpha - k_{y'}\sin\alpha, \\ q_y &= q_{x'}\sin\alpha + k_{y'}\cos\alpha, \end{aligned} \tag{6.15}$$

and equivalently, by solving,

$$(q_x^r)^2 + \gamma_0^2 (q_y^r)^2 = \frac{4(V-E_0)^2}{\hbar^2 v_0^2}, \tag{6.16}$$

we can have two solutions for $q_{x'}$, one for incident and the other for reflected wave in region II, depending on the group velocities they associated with. Therefore we obtain the spinor angles $\theta_{\mathbf{s}}$ and $\theta_{\mathbf{s}}^r$. Note that Eq. (6.7) and Eq. (6.9) should be combined (i.e. considering $\lambda(\lambda') = \pm 1$) to determine which quadrant the wavevector and the group velocity lies in, and the spinor direction is determined by spinor angle and $\lambda(\lambda')$ together. In a word, to obtain the solutions, the defining equations of the three vectors should be used, coming from Eq. (6.7) in the main text.

The transmission probability in the ballistic transport limit is obtained by matching the carrier wavefunction at the boundaries [101, 129]. The solutions of the three regions in the $x' - y'$ coordinate system take the form

$$\begin{cases} \psi_{\mathrm{I}}(x',y') = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\ \lambda e^{i\phi_{\mathbf{s}}}\end{pmatrix} e^{i(k_{x'}x'+k_{y'}y')} + \frac{r}{\sqrt{2}}\begin{pmatrix}1\\ \lambda e^{i\phi_{\mathbf{s}}^r}\end{pmatrix} e^{i(k_{x'}^r x'+k_{y'}y')}, & x'<0, \\ \psi_{\mathrm{II}}(x',y') = \frac{a}{\sqrt{2}}\begin{pmatrix}1\\ \lambda' e^{i\theta_{\mathbf{s}}}\end{pmatrix} e^{i(q_{x'}x'+k_{y'}y')} + \frac{b}{\sqrt{2}}\begin{pmatrix}1\\ \lambda' e^{i\theta_{\mathbf{s}}^r}\end{pmatrix} e^{i(q_{x'}^r x'+k_{y'}y')}, & 0<x'<D_0, \\ \psi_{\mathrm{III}}(x',y') = \frac{t}{\sqrt{2}}\begin{pmatrix}1\\ \lambda e^{i\phi_{\mathbf{s}}}\end{pmatrix} e^{i(k_{x'}x'+k_{y'}y')}, & x'>D_0. \end{cases} \tag{6.17}$$

Then we can have the following set of equations from matching the boundary conditions in the above set of equations,

$$\begin{cases} 1 + r = a + b, \\ \lambda e^{i\phi_{\mathbf{s}}} + \lambda r e^{i\phi_{\mathbf{s}}^r} = \lambda' a e^{i\theta_{\mathbf{s}}} + \lambda' b e^{i\theta_{\mathbf{s}}^r}, \\ a e^{iq_{x'}D} + b e^{iq_{x'}^r D} = t e^{ik_{x'}D}, \\ \lambda' a e^{i\theta_{\mathbf{s}} + iq_{x'}D} + \lambda' b e^{i\theta_{\mathbf{s}}^r + iq_{x'}^r D} = \lambda t e^{i\phi_{\mathbf{s}} + ik_{x'}D}. \end{cases} \tag{6.18}$$

With some algebra, we derive the transmission amplitude as,

$$t(\phi_{\mathbf{v}}') = \frac{\lambda\lambda' e^{-ik_{x'}D}(e^{i\theta_{\mathbf{s}}^r} - e^{i\theta_{\mathbf{s}}})(e^{i\phi_{\mathbf{s}}} - e^{i\phi_{\mathbf{s}}^r})}{A}, \tag{6.19}$$

where,

$$\begin{aligned} A = {} & e^{-iq_{x'}D}\left(e^{i\theta_{\mathbf{s}} + i\theta_{\mathbf{s}}^r} + e^{i\phi_{\mathbf{s}} + i\phi_{\mathbf{s}}^r} - \lambda\lambda' e^{i\theta_{\mathbf{s}}^r + i\phi_{\mathbf{s}}^r} - \lambda\lambda' e^{i\theta_{\mathbf{s}} + i\phi_{\mathbf{s}}}\right) \\ & - e^{-iq_{x'}^r D}\left(e^{i\theta_{\mathbf{s}} + i\theta_{\mathbf{s}}^r} + e^{i\phi_{\mathbf{s}} + i\phi_{\mathbf{s}}^r} - \lambda\lambda' e^{i\theta_{\mathbf{s}}^r + i\phi_{\mathbf{s}}} - \lambda\lambda' e^{i\theta_{\mathbf{s}} + i\phi_{\mathbf{s}}^r}\right). \end{aligned} \tag{6.20}$$

The transmission probability through the potential barrier is

$$T(\phi'_{\mathbf{v}}) = tt^*. \tag{6.21}$$

In the case of graphene, besides the resonant unit transmission at specific angles owing to a Fabry-Pérot like effect, normal-incident electron wave packet in graphene always shows perfect transmission because of suppression of backscattering [129] owing to the chiral character of the Dirac fermions. In the anisotropic case, the pseudospin-momentum locking (given by Eq. (6.7) and Eq. (6.9)) gives rise to a complete suppression of the scattering only under the condition $\mathbf{k}' \rightarrow -\mathbf{k}'$. Considering the fact that $k_{y'}$ is conserved, this process will happen only when $k_{y'} = 0$. Consequently, $\phi^r_{\mathbf{s}} = \pi + \phi_{\mathbf{s}}$, $\theta^r_{\mathbf{s}} = \pi + \theta_{\mathbf{s}}$, $\phi_{\mathbf{s}} = \theta_{\mathbf{s}}$, and $q^r_{x'} = -q_{x'}$, resulting in $T = 1$, i.e. perfect transmission independent of the potential barrier height and width.

6.3.2 Numerical results of asymmetric Klein tunneling in patterned few-layer black phosphorus

With an anisotropic Dirac cone, perfect transmission therefore occurs when the incident wavevector $\mathbf{k}$ is along the normal to the potential barrier. However, the fact that the group velocity and the wavevector are generally noncollinear (Eq. (6.9) and Fig. 6.7(a)) leads to a remarkably distinct Klein tunneling behavior compared with graphene. For a wave packet, it is the group velocity, not wavevector, that describes the direction of center-of-mass motion and energy flow. In graphene, normal incident wave packet (or energy flow) is unimpeded; however, in the anisotropic case, the transmission probability of a normal incident wave packet is not unity and can be tuned and controlled. Furthermore, the normal incident direction and the perfect transmission direction are in general different.

In a general asymmetric Klein tunneling process, the normal incidence direction is different from the perfect transmission direction, and this difference can be maximized with given $m^*_x and m^*_y$. The perfect transmission case corresponds to $\mathbf{k} = (k\cos\alpha, k\sin\alpha\ )$, according to the above discussion. In this case, if we limit $0 < \alpha < \frac{\pi}{2}$, we can maximize $(\alpha - \phi_{\mathbf{v}})$, or equivalently $\tan(\alpha - \phi_{\mathbf{v}})$). Together with Eq. (6.9), we may solve for the maximum value (at $\alpha = \alpha_m$) of the following function,

$$\tan(\alpha - \phi_{\mathbf{v}}) = \frac{\tan\alpha - \tan\phi_{\mathbf{v}}}{1 + \tan\alpha\tan\phi_{\mathbf{v}}}. \tag{6.22}$$

The difference between the two directions depends on α and is maximized at

$$\alpha_m = \arctan\frac{1}{\gamma_0}, \tag{6.23}$$

The maxium $(\alpha - \phi_{\mathbf{v}})$ corresponds to $\tan\phi_{\mathbf{v},m} = \gamma_0$, and meanwhile $\tan\phi_{\mathbf{s},m} = 1$. Therefore, we have the following relation,

$$\phi'_{\mathbf{v},m} = \phi_{\mathbf{v},m} - \alpha = \arctan\gamma_0 - \arctan\frac{1}{\gamma_0}. \tag{6.24}$$

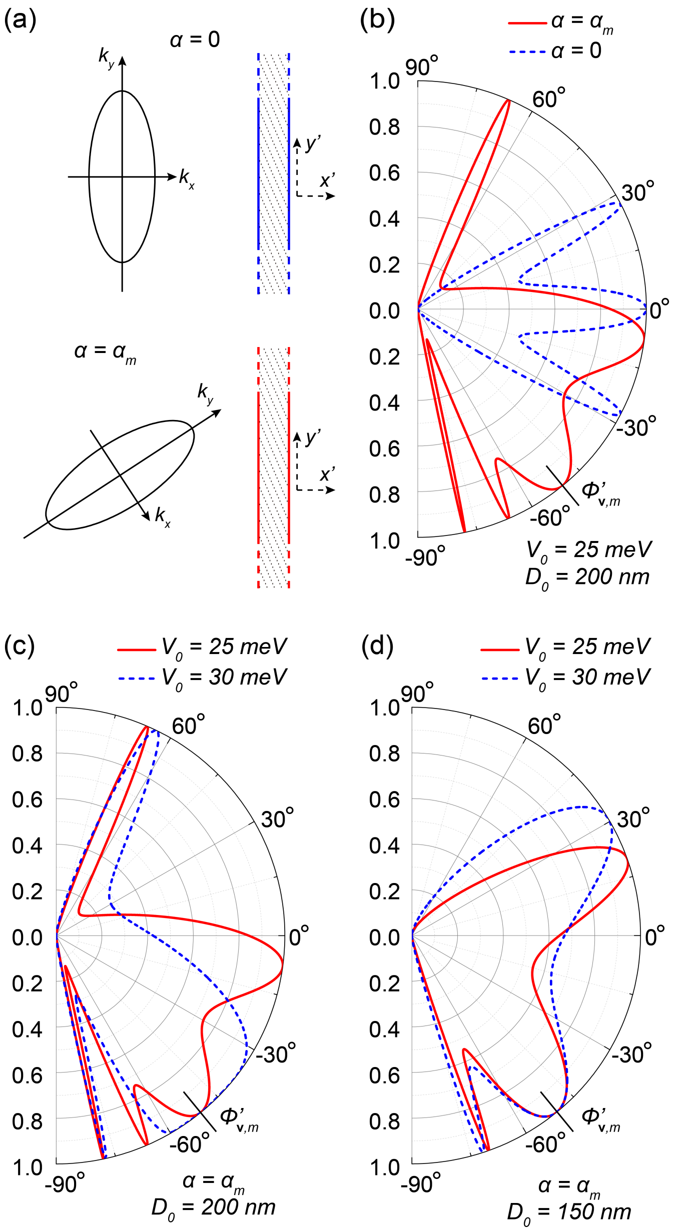

Figure 6.8: Transmission probability T versus the incident angle of group velocity $\phi'_{\mathbf{v}}$ with respect to the potential barrier normal, showing asymmetric Klein tunneling. $E_0 = 10\text{meV}$ is used. (a) Two different potential barrier orientations. At $\alpha = \alpha_m = \arctan\frac{1}{\gamma_0}$, the normal-incidence direction and the perfect-transmission direction are maximally differed by $\phi'_{\mathbf{v},m} = 50.8°$. (b) Symmetric and asymmetric Klein tunneling profiles corresponding to the two geometries in (a). (c, d) Asymmetric Klein tunneling with various parameters for $\alpha = \alpha_m$.

The few-layer black phosphorus under the superlattice potential takes on a value of $\phi'_{\mathbf{v},m} = -50.8°$ with the effective masses considered here for black phosphorus. The directional dependence of the transmission probability $T(\phi'_{\mathbf{v}})$ is plotted in Fig. 6.8, with varying incident angle $\phi'_{\mathbf{v}}$ of the group velocity relative to the potential barrier normal, and in principle is measurable from a directionally dependent nonlocal resistivity experiment [139].

Because of the asymmetric Klein tunneling behavior, the transmission probability of normal incident wave packets can be controlled with different barrier alignment, height and width, whereas as discussed above the perfect transmission direction depends on α only. In the cases where the Fermi surface is smaller inside the potential barrier region than the outside regions, then at certain angles, no propagating wave solutions exist. But evanescent wave solutions are allowed by letting $q_{x'} = i\kappa'$ and $q^r_{x'} = i\kappa'_r$ $(\kappa, \kappa' \in \mathbb{R})$. Therefore, a beam of normal-incident wave packets can be completely turned off by tuning V_0, which is also a missing feature in pristine graphene or any isotropic Dirac fermion system.

6.4 Conclusion

In this work, we have proposed a new platform for the generation of anisotropic Dirac fermion and novel asymmetric Klein tunneling phenomenon. We have adopted effective Hamitonian approach with perturbative analysis and numerical solutions. Our results clearly demonstrate these novel new physics that can be realized in the anisotropic semiconductor few-layer phosphorus. The phenomena found in this work are general and applicable to any anisotropic host systems. Moreover, the emergence of highly tunable and easily accessible anisotropic massless Dirac fermions in few-layer black phosphorus superlattices should provide a range of interesting experimental investigations and a new direction for possible device applications.

Chapter 7

Understanding the formation of electron-irradiation-induced defects in boron nitride

Hexagonal boron nitride (h-BN) has become one of the most important materials in contemporary condensed matter studies due to its various excellent properties. Its defect formation and dynamics are critically important in understanding the physical, mechanical, and chemical properties of h-BN for future harsh-environment applications

In this Chapter, we report a collaborative work with the experimental group of Prof. Alex Zettl (University of California at Berkeley). This work is published as Ref. [140]. In this work, the experimentalists study the defect formation and dynamics of h-BN using aberration-corrected transmission electron microscopy (TEM) at elevated temperatures. The experiment finds particularly interesting phenomena that through the TEM beam, hole defects of triangular and hexagonal shapes are formed with complex dynamics at elevated temperatures. We perform density functional theory (DFT) calculations to understand the stability of such defects at thermal equilibrium, and predict a diagram of preferential defects under different conditions.

7.1 Introduction

Graphene [52] has attracted significant interest over the past decade, which has in turn spurred interest in few-layer and monolayer hexagonal boron nitride (h-BN) [141, 142]. In either its bulk or few-layer forms, h-BN is a wide band gap ($\sim$ 5.5 eV) insulator with comparable mechanical strength to, and higher oxidation resistance than, graphene [142]. These and other characteristics make h-BN a promising candidate in various applications [143–145]. As a substrate for many other 2D materials, h-BN has been shown to reduce charge density fluctuation and surface roughness, thus enhancing electronic properties in overlaid graphene [146, 147] and optoelectronic properties in overlaid few-layer transitional metal dichalcogenides [148].

In this collaborative project, experimentally, the atomic structure, stability, and dynamics of defects in hexagonal h-BN are investigated using an aberration-corrected transmission electron microscope operated at 80 kV between room temperature and 1000 °C. At temperatures above 700 °C, parallelogram- and hexagon-shaped defects with zigzag edges become prominent, in contrast to the triangular defects typically observed at lower temperatures. The appearance of 120° corners at defect vertices indicates the coexistence of both N- and B-terminated zigzag edges in the same defect.

The theoretical efforts which will be mainly discussed in the following, complement the experiments with first-principles calculations which consider the thermal equilibrium formation energy of different defect configurations. We show that, below a critical defect size, hexagonal defects have the lowest formation energy and therefore are the more-stable configuration, and triangular defects are energetically metastable but can be “frozen in” under experimental conditions.

7.2 Experimental results summary

We summarize the experimental observation of electron-irradiation-induced triangular and hexagonal defects in h-BN here [140].

Fig. 7.1(a) and (b) are a representative high-resolution transmission electron microscopy (TEM) image of triangular defects in h-BN, obtained at 500 °C [140]. A corresponding atomic model is also shown (B and N atoms are depicted in pink and blue, respectively) for N-terminated zigzag-edge triangles. Triangular defects of this kind are quite common in h-BN, and many of them have been previously observed at room temperatures [141, 149–151]. The chemical composition of edge-atoms has been identified as N by high-resolution EELS and contrast comparison in scanning transmission electron microscopy (STEM) and high-resolution TEM (N appears to have a slightly higher contrast than B) [141, 150, 152]. The defect formation by a knock-on mechanism has been widely accepted from both experimental and theoretical studies [149, 150, 153, 154].

Fig. 7.1 (c) and (d) show a representative TEM image of hexagonal defects at 900 °C. At 600 and 700 °C, the experiment [140] observes the appearance of hexagonal defects in addition to the conventional triangles; however, the triangular holes become unstable and very quickly (in matter of seconds during the imagining process) transform into a hexagonal configuration. At high temperatures (800, 900, and 1000 °C), the hexagonal defects with distinguishing 120°C corners predominate. This experimental atomic-resolution high-resolution TEM study provides direct, clear, and unambiguous identification of the edge structure of the hexagonal defects [140]. Indeed, at all temperatures studied, the experiment [140] find that all the edges in both cases adopt a zigzag configuration. The simple h-BN atomic models shown in the right panels of Fig. 7.1(d) demonstrate that parallelogram and hexagon-shaped defects (or in general any defects with 120° corners) with zigzag edges must have both N- and B-terminated atoms.

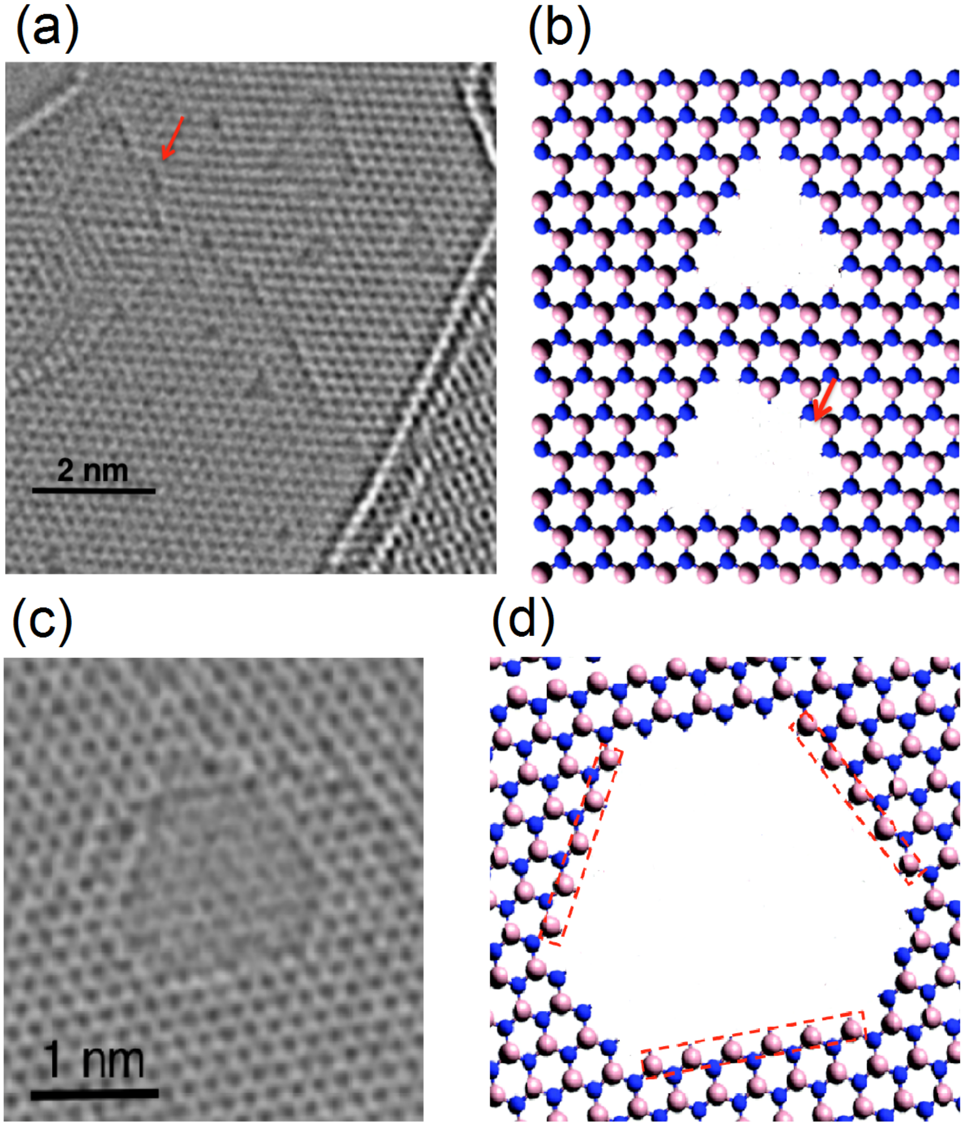

Figure 7.1: Experimental observation [140] of defects in h-BN. (a) and (b) Representative high-resolution transmission electron microscopy (TEM) image of triangle defects at 500 °C (a) and its corresponding atomic model (b). The image and model also show a triangle with an ejected chain of B-N atoms (indicated by red arrows). The blue and pink spheres are N and B atoms, respectively. (c) and (d) High-resolution TEM images of hexagonal defects (at 900 C), showing the presence of both N- and B-terminated zigzag edges in the same defect. Blue and pink spheres represent for N and B atoms, respectively. Long-range B-terminated zigzag edges are indicated by red-dotted boxes.

This key observation [140] apparently contradicts literature wisdom because the presence of a B-terminated zigzag edge has historically been considered possible only as an intermediate transient construct when the triangular defects grow. Defects in h-BN in the literature have either all N-terminated or all B-terminated edges. The experiment also finds that extended B-terminated zigzag edges are stable (at least for several seconds under imaging conditions) and ubiquitous , and both N-terminated and B-terminated edges coexist in the same defect [140].

In the following theoretical investigation, we use DFT calculations to provide insights in the stability of these defects from thermal equilibrium perspective.

7.3 Theoretical DFT computation of formation enthalpy of defect structures, corners, and edges

To gain insight into the defect formation mechanism, we perform first-principles calculations within the density functional theory to compute the formation enthalpy of each defect structure (triangles and hexagons), and we break down to corners and edges.

First-principles calculations are performed within the density functional theory using Quantum Espresso package [47]. We use a plane wave basis set with 45 Ry cutoff, and projector augmented wave pseudopotentials under local density approximation [97, 155] . We fully relax the lattice vectors and internal coordinates of pristine monolayer h-BN with 10 $\mathring{A}$ vacuum size, based on which we construct the defect structures in supercells such that the defect-defect distance in the periodic image is larger than 9 $\mathring{A}$. We have calculated with six defect structures, as shown in Table 7.1, that there are three types of them, i.e., triangular defects with N terminated edges (tri-N), triangular defects with B terminated edges (tri-B), and hexagonal defects with both N and B terminated edges (hex). For each type of defect, we calculate two different sizes, therefore we can extract the edge and corner energies later [156]. In the calculations of defects in supercells, we fix the lattice constant from pristine monolayer h-BN to mimic an environment with relatively low defect density, and fully relax the internal coordinates to let the edges and corners reconstruct.

To calculate the formation enthalpy ΔH of each defect structure, we need the information of chemical potentials, i.e. the chemical environment when the system reaches thermal equilibrium. The chemical potential of monolayer h-BN is denoted as μ_{BN} for 1 B atom and 1 N atom together. The chemical potentials of B and N atoms, are referenced to their elementary substances, chosen as α-boron and N_2 molecule, and denoted as μ_{0B} and μ_{0N} for per atom, respectively. In a thermal equilibrium state with dominating monolayer h-BN, we have a constraint [157],

$$\mu_B + \mu_N = \mu_{BN}. \tag{7.1}$$

Under two extreme chemical environment limits, the chemical potential of B (μ_B) and N

Defect structures	Description	Edge length (number of atoms)	ΔH (Ry), N-rich	ΔH (Ry), B-rich
tri-N-49	Triangle, 3 N-terminated edges, 3 60° N-corners, hole of 28 B and 21 N	7	4.180	5.727
tri-N-64	Triangle, 3 N-terminated edges, 3 60° N-corners, hole of 36 B and 28 N	8	4.816	6.584
tri-B-49	Triangle, 3 B-terminated edges, 3 60° B-corners, hole of 21 B and 28 N	7	5.501	3.955
tri-B-64	Triangle, 3 B-terminated edges, 3 60° B-corners, hole of 28 B and 36 N	8	6.320	4.552
hex-96	Hexagon, 3 N- and 3 B-terminated edges, 6 120° BN-corners, hole of 48 B and 48 N	4	5.566	
hex-150	Hexagon, 3 N- and 3 B-terminated edges, 6 120° BN-corners, hole of 75 B and 75 N	5	7.179	

Table 7.1: Defect structures calculated in this work and their formation enthalpies ΔH under two chemical potential limits.

(μ_N) can therefore be determined as follow,

$$\begin{aligned} &\text{N-rich: } \mu_N = \mu_{0N}, \quad \mu_B = \mu_{BN} - \mu_N, \\ &\text{B-rich: } \mu_B = \mu_{0B}, \quad \mu_N = \mu_{BN} - \mu_B. \end{aligned} \tag{7.2}$$

In reality, the chemical potentials μ_B and μ_N can vary, but would not be out of the range defined by Eq. (7.2) in the thermal equilibrium states. With well-defined chemical potentials under each chemical environment, we can calculate the formation energy for each defect structure by using,

$$\Delta H = E_{\text{supercell}}(\text{defect}) + \mu_B N_{(\text{missing B})} + \mu_N N_{(\text{missing N})} - E_{\text{supercell}}(\text{pristine}), \tag{7.3}$$

where $E_{\text{supercell}}(\text{defect})$ is the total energy of a defect structure in a supercell, $E_{\text{supercell}}(\text{pristine})$ is the total energy of the pristine monolayer BN in the same size supercell, $N_{(\text{missing B})}$ and $N_{(\text{missing N})}$ are the number of missing B and N atoms in the formation of the defect structure (see Table 7.1) from the pristine structure. The calculated formation enthalpy for the six defect structures are listed in Table 7.1. We notice that for the hexagonal defects, the formation enthalpy is independent of the elemental chemical potential B and N, because in such defects, the number of (missing) B and N atoms are always equal, so the chemical potentials simply come from the constraint Eq. (7.1). By calculating two defect structures with

Defects	Edge ΔH (Ry/atom)		Corner ΔH (Ry/Corner)	
	N-rich	B-rich	N-rich	B-rich
tri-N	0.212	0.286	-0.091	-0.091
tri-B	0.273	0.199	-0.077	-0.077
hex	0.269		-0.148	

Table 7.2: Calculated formation enthalpies for edges and corners in different defect types.

different sizes of one type, we can extrapolate the edge and corner energies approximately (the difference in formation enthalpy of two sizes gives the formation enthalpy of the edges, then that of the corners can be derived [156]), as are listed in Table 7.2.

7.4 Defect formation mechanism from thermal equilibrium considerations

As we discussed above, Table 7.1 lists the calculated formation enthalpy for six defect structures with different sizes and shapes, and in two limit conditions, N-rich and B-rich environments. Therefore the edge and corner energies can then be approximately extracted, as shown in Table 7.2. We find that formation of the edges requires more energy, whereas the corners lower the energy due to the appreciable structure reconstructions at the corners, as evidenced by the relaxed structures. More importantly, the 120° BN-corner in the hexagonal defects has lower energy than either the 60° N- or B-corner in the triangular defects. With the extracted edge and corner energies in each type of defect we can extrapolate to obtain the formation enthalpy of defects in different sizes, as plotted in Fig. 7.2, assuming all other effects are negligible.

Fig. 7.2(a) and (b) present the thermal equilibrium formation energy in two limits (N-rich and B-rich environments) as a function of defect circumference (i.e., perimeter). The figures show that when the defect sizes are small ($n < 10$, where n is the number of atoms on all edges of the defect) the hexagonal defects have the lowest formation energy, and when the defects are larger ($n > 10$), triangular defects (either N- or B-terminated edges) are more stable. This is in line with our finding that the formation of a hexagonal defect can lower the total energy by having the 120° BN-corners at the cost of introducing both N- and B-terminated edges. However, when the defect size is large the advantages of the corners lessen compared to the contribution of the edges to the total energy. In this case triangular defects become energetically preferential since they have solely one type of edges.

By varying the chemical potential between two limits (N-rich and B-rich) we derive a defect shape preference diagram in terms of defect circumference and chemical potential using the defect formation energy as shown in Fig. 7.2(c). Apparently under any chemical potential environment, hexagonal defects have the lowest energy when the defect sizes are small ($n < 10$ in most cases). On the other hand triangular defects are more stable at very

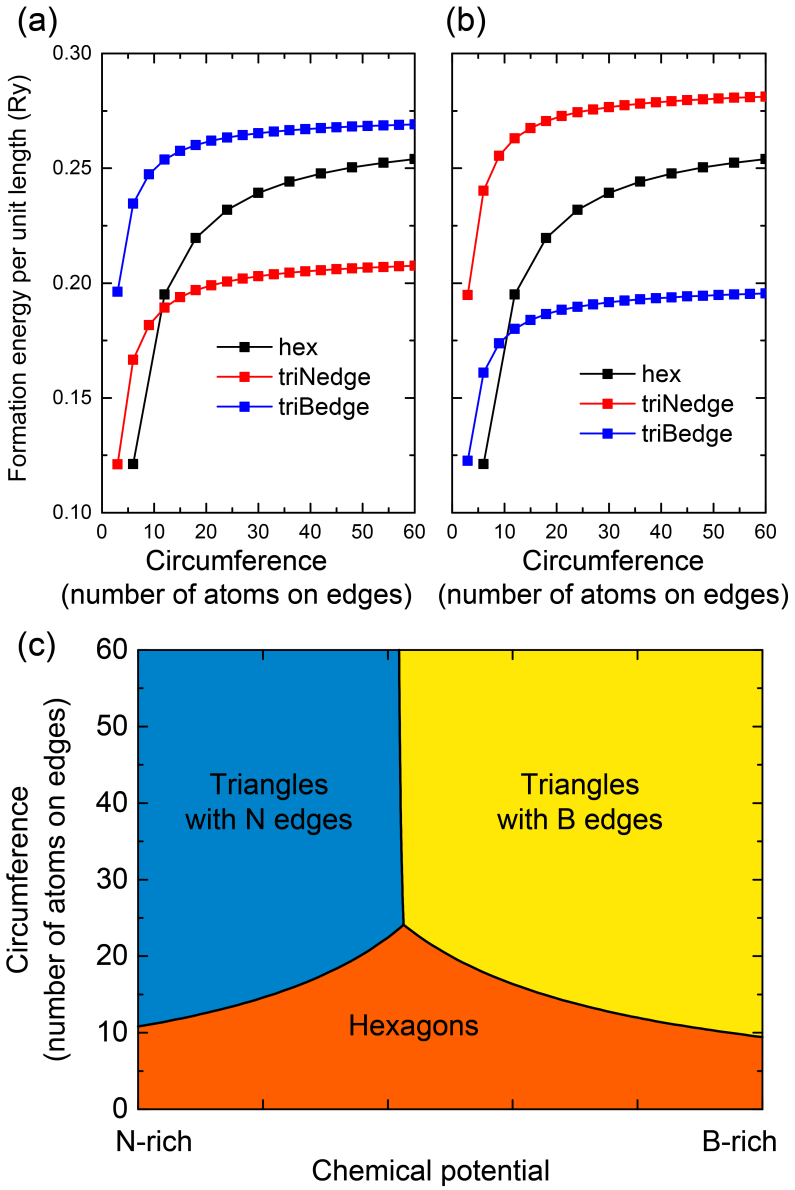

Figure 7.2: (a) Formation energy per unit length (occupied by one atom) as a function of circumference (i.e. number of atoms on all edges) in different types of defects in the N-rich limit. We use the edge and corner energies from Table 7.2 and extrapolate to get the each data point. (b) Similar to (a), but under B-rich limit. (c) Diagram of regions of preferential existence of various defects as a function of circumference versus chemical potential. Triangular and hexagonal defects phases are included. Lines separating different domains are guides for the eyes only.

large sizes ($n > 25$).

This model (considering thermal equilibrium formation energy of defects in the ground state at $T = 0$) provides a physical mechanism to understand, and shows some degree of quantitative agreement with, the experimental observations. Experimentally, the experimentalists [140] observed the presence of hexagonal defects with the size (circumference) in the range of $n = 6$ to $n = 26$ [140], which clearly falls within the hexagonal-defect-dominance space in the diagram (Fig. 7.2(c)). However, for triangular defects, while the theoretical diagram shows their prevalence only at large size ($n > 25$), the experimentalists [140] observe them even with $n = 3$ (single point defect) or $n = 9$ (tetravacancies) at low temperature (< 500°C). While admittedly this theory cannot provide a full explanation to the temperature-dependent defect formation, it gives *significant* insight into the energy landscape of defect stability. Together with the experimentalists, we [140] speculate that the triangular defects are favorable metastable states and at normal conditions (< 500°C) they represent a local attractor in phase space. The lowest energy state and hence globally the most stable one for small sized defects is more likely hexagonal. We conclude that there is an energy barrier between these two states, and indeed by heating the samples to above 700 °C stable hexagonal defects readily emerge experimentally. This sets the energy barrier scale. Overall, the theoretical calculations provide important insights and understandings to the formation of these triangular and hexagonal defects in h-BN.

7.5 Conclusion

This Chapter presents a collaborative work that shows the formation and dynamics of hexagonal defects with both N- and B-terminated zigzag edges in h-BN at high temperatures for the first time. We theoretically performed DFT calculations, which consider only the thermal equilibrium formation enthalpy of static defects. Our calculations provide careful investigation of the detailed electronic and structural properties of these defects. Based on the extrapolated edge and corner energies, we generate a diagram to show the most stable defect structure under different circumference and chemical potential at equilibrium. These results reveal that hexagonal defects are more stable than the triangular defects for small sizes ($n < 10$). In reality, the real experimental conditions involve many other dynamic processes, including chemical etching from gas environment and elastic (knock-on) and inelastic interaction between the electron beam and the material. These factors likely complicate the equilibrium picture and thus give rise to the experimental observation of the temperature-dependent defect formation and dynamics. The insights provided by the theoretical calculations and analysis are important and helpful to further understand of defect physics in h-BN.

Chapter 8

Symmetry rules shaping spin-orbital textures in topological surface states

Topological insulators and topological semimetals have exhibit novel properties and are important research objects in condensed matter physics. In these materials, strong spin-orbit coupling (SOC) often exists, giving rise to exotic electronic properties.

In this Chapter, we discuss a collaborative work with mainly the experimental group of Prof. Alessandra Lanzara (University of California at Berkeley). This work is published as Ref. [158]. The experimentalists have probed the 2D topological surface states residing on Antimony (Sb) (111) surface using spin- and angle-resolved photoemission spectroscopy (spin-ARPES), and discovered the complex behavior of light-matter interactions that are governed by symmetries and SOC. We performed detailed theoretical *ab initio* tight-binding calculations to reproduce and understand such phenomena. Our theory provides significant insights in understanding the interplay between SOC and symmetries in this novel electronic state.

This project [158] discovers that symmetry constrains the way orbital and spin components of a state co-evolve as a function of momentum, and from this, the rules governing how the two degrees of freedom are interwoven are determined. The experiment directly observes this complexity in spin-resolved photoemission, and our *ab initio* calculations of the topological surface states of Sb(111), where the photoelectron spin direction near $\bar{\Gamma}$ is found to have a strong and unusual dependence on photon polarization. This dependence unexpectedly breaks down at large $|k|$, where the surface states mix with other nearby surface states. However, along mirror planes, symmetry protects the distinct spin orientations of different orbitals. The discovery in this project broadens the understanding of surface states with strong SOC, demonstrates the conditions that allow for optical manipulation of photoelectron spin, and will be highly instructive for future spintronics applications.

8.1 Introduction

Materials with strong spin-orbit coupling (SOC) and spin split surface states have garnered significant attention for possible use in spintronic devices, in which the spin degree of freedom would be manipulated electrically [159–162]. In states subject to spin-orbit coupling, in the atomic limit, spin and orbital angular momenta (**S** and **L**) are not good quantum numbers; total angular momentum **J** is instead the conserved quantity [163]. In fact, it has recently been observed in topological surface states that the spin and orbital textures can be "entangled" such that, at a given momentum, there is a mix of orbitals in which each has a distinct spin orientation [163–166]. Thus, fully understanding the wave function of these potentially useful states means characterizing the complex spin-orbital texture.

The dependence of spin texture on wave-function atomic orbital character can give rise to a rich array of physical phenomena. It causes photoelectron spins to point in a direction dependent on photon polarization, allowing for optical control of spin polarization [167, 168]. In fact, the relative weight of $p_{x,y,z}$ orbitals, and hence spin texture of Bi_2Se_3, varies through the atomic layers containing the surface-state wave function [165, 169]. Knowledge of how the spin and orbital degrees of freedom mix is key to interpreting experimental results from spin-orbit materials [163–176], as well as possibly utilizing them technologically.

Spin- and angle-resolved photoemission spectroscopy (spin-ARPES) [177, 178] with tunable photon polarization is uniquely capable of studying spin-orbital texture, as demonstrated with the surface states of Bi_2Se_3, where the spin polarization of photoelectrons was observed to reverse for light polarization rotated 90° [165, 168, 169]. This effect was predicted based on symmetry arguments and a model Hamiltonian [167] and was further discussed microscopically in terms of the constituent atomic orbitals making up the band [163]. With total angular momentum as the conserved quantity, the $J_z = \pm\frac{1}{2}$ basis is used to describe the surface state near $\bar{\Gamma}$. Under this constraint, spin-orbit coupling gives each of the $p_{x,y,z}$ orbitals its own spin texture. Light will select p orbitals oriented along the direction of photon polarization according to the selection rules for the photoemission process [164]. Previous discussion of this phenomenon therefore focused on strong spin-orbit coupling and the symmetries at the $\bar{\Gamma}$ point in Bi_2Se_3: time-reversal, mirror, and C_3 rotational symmetry. Thus far, there have been no tests of how it evolves at high wave vector **k** as the symmetry changes, leaving open questions about the fundamental nature of coupling of orbital textures to distinct spin textures.

Antimony, a topologically nontrivial semimetal, provides an intriguing test case. The Sb(111) surface states have been investigated with ARPES and spin-ARPES, confirming the spin polarization due to strong SOC and nonzero Berrys phase [179–184]. While the (111) surface of Sb has the same symmetries as Bi_2Se_3, its surface states are distinct in their strong **k** dependence. They remain separate from the bulk states out to large $|k|$, allowing for a comparison of the spin-orbital texture near $\bar{\Gamma}$ to that in areas of reduced symmetry, where we will demonstrate that there are significant differences.

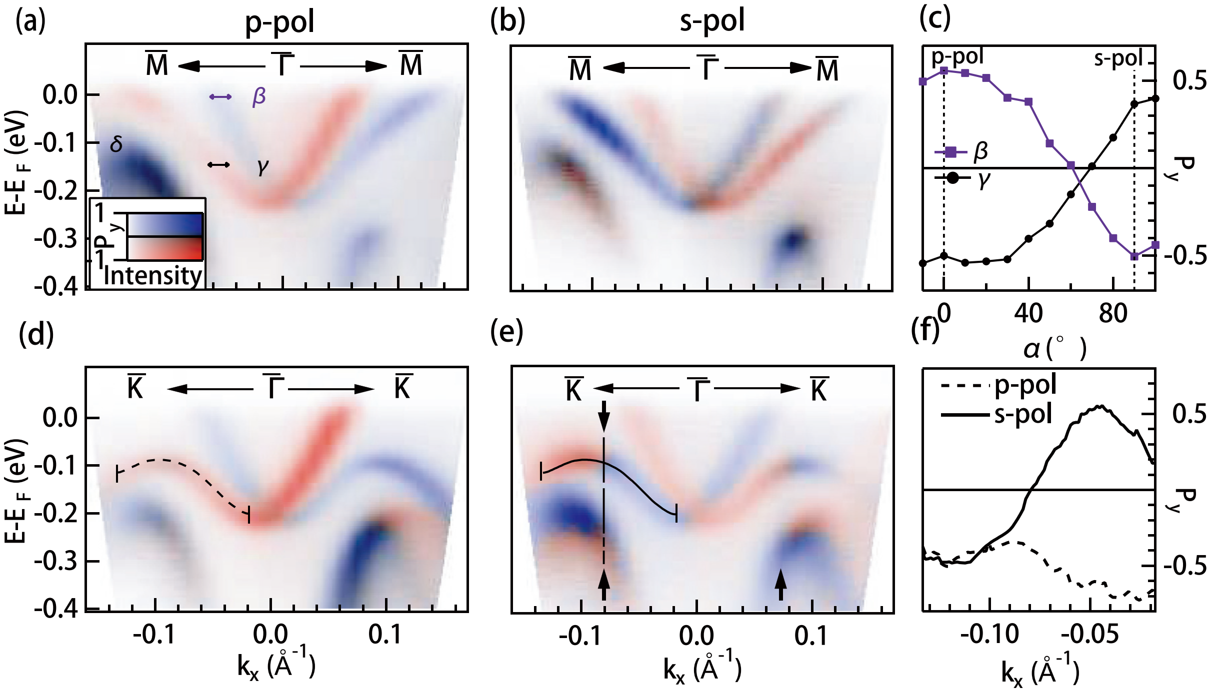

Figure 8.1: Experimental spin-ARPES [158] maps of the $\bar{\Gamma} - \bar{M}$ and $\bar{\Gamma} - \bar{K}$ directions of Sb(111). (a) and (b) Spin-resolved maps of the $\bar{\Gamma} - \bar{M}$ direction, taken with (a) p-polarized and (b) s-polarized light. The two-dimensional color scale displays the total photoelectron intensity by relative darkness and the spin polarization by the balance of red and blue. (c) Spin polarization of the two surface bands as a function of photon polarization angle. The bands are labeled in (a), and spin polarizations were extracted at a fixed **k**, as indicated by the small regions marked with arrows. (d) and (e) Similar to (a) and (b) but with sample azimuth rotated to cut along $\bar{\Gamma} - \bar{K}$. (f) Spin polarization along the left branch of the lower band, as measured with both p- and s-polarized light. The stretch of **k** space plotted here is indicated by dispersive lines in (d) and (e).

8.2 Experimental results summary

Here we summarize the key experimental spin-ARPES measurement of Sb(111) surface. Fig. 8.1 shows the full spin-resolved energy maps for both spin-up and spin-down electrons along $\bar{\Gamma} - \bar{M}$ and $\bar{\Gamma} - \bar{K}$ for both light polarizations. The spin-ARPES maps are plotted in Fig. 8.1 with a color scale in which brightness (from light to dark) corresponds to total photoemission intensity, while color (from red to blue) corresponds to spin polarization.

We first look at the results along the $\bar{\Gamma} - \bar{M}$ direction with p-polarized light in Fig. 8.1(a). The Γ point is enclosed within an electron pocket with positive spin polarization for k_x and negative spin polarization for $+k_x$. The two branches of the surface state meet at The Γ and bend back up to the Fermi level. Furthermore, when the same map is made with s-polarized light as plotted in Fig. 8.1(b), both surface bands show the opposite spin polarization at all momenta from $\bar{\Gamma}$ to the measured range. Furthermore, as in Bi_2Se_3, the spin polarization can be adjusted continuously, as shown in Fig. 8.1(c), by rotating the angle of linear photon polarization.

The $\bar{\Gamma}-\bar{K}$ direction as shown in Figs. 8.1(d)-(f) demonstrates a strong contrast. The same spin dependence on photon polarization is seen near $\bar{\Gamma}$. However, near $k_x \approx \pm 0.08 \mathring{A}^{-1}$, this behavior ceases, and for larger $|k|$, s-polarized light yields the same spin polarization as p-polarized light in the lower surface band.

Besides the rapid change in the spin texture of the lower surface band, an unusual spin polarization appears around the top of the bulk continuum to the left of the dashed line and arrows in Fig. 8.1(e). This is contrary to the common knowledge that in the normal bulk limit of an inversion-symmetric, nonmagnetic crystal, each state is spin degenerate therefore gives zero spin polarization.

Briefly speaking, for well-known topological insulator, e.g. Bi_2Se_3, the observed spin polarization of topological surface states will flip by switching between s- and p- polarized light. This is true in Sb(111) for states around $\bar{\Gamma}$ and along $\bar{\Gamma}-\bar{M}$, but unexpected breaks down at high $|k|$ along $\bar{\Gamma}-\bar{K}$. The arising of some bulk spin polarization is also unexpected. We therefore perform *ab initio* tight-binding calculations to address these unusual phenomena.

8.3 *Ab initio* tight-binding calculation details

To understand these experimental findings, we performed an *ab initio* tight-binding simulation [165]. The basis is chosen to be the Sb p orbitals, and the hopping parameters and on-site energies for the surface and bulk regions were extracted from first-principles calculations within density functional theory (DFT) using the QUANTUM ESPRESSO package [47]. Norm-conserving pseudopotentials with the local-density approximation by Perdew and Zunger [155] were used for Sb in both scalar- and fully relativistic forms. DFT calculations for periodic bulk and a 12-bilayer slab were performed to obtain the hopping parameters within the atomic orbital basis using the WANNIER90 code [48]. The 12-bilayer slab and the bulk DFT calculations are performed to extract the parameters for surface and bulk, respectively. In the tight-binding model, we separate the slab into top and bottom halves and repeat the bulk unit cell in between to fill the two halves. The on-site energies in the bulk region are adjusted to match the middle layers in the 12-bilayer slab. Eventually, a 90-bilayer slab is constructed. All the physical quantities such as band structures, spin textures, orbital projections, and photoemission predictions are calculated from this tight-binding model following the method in Ref. [165]. In the tight-binding model, the SOC strength can be tuned by weighting the hopping parameters between those extracted from scalar- and fully relativistic DFT calculations.

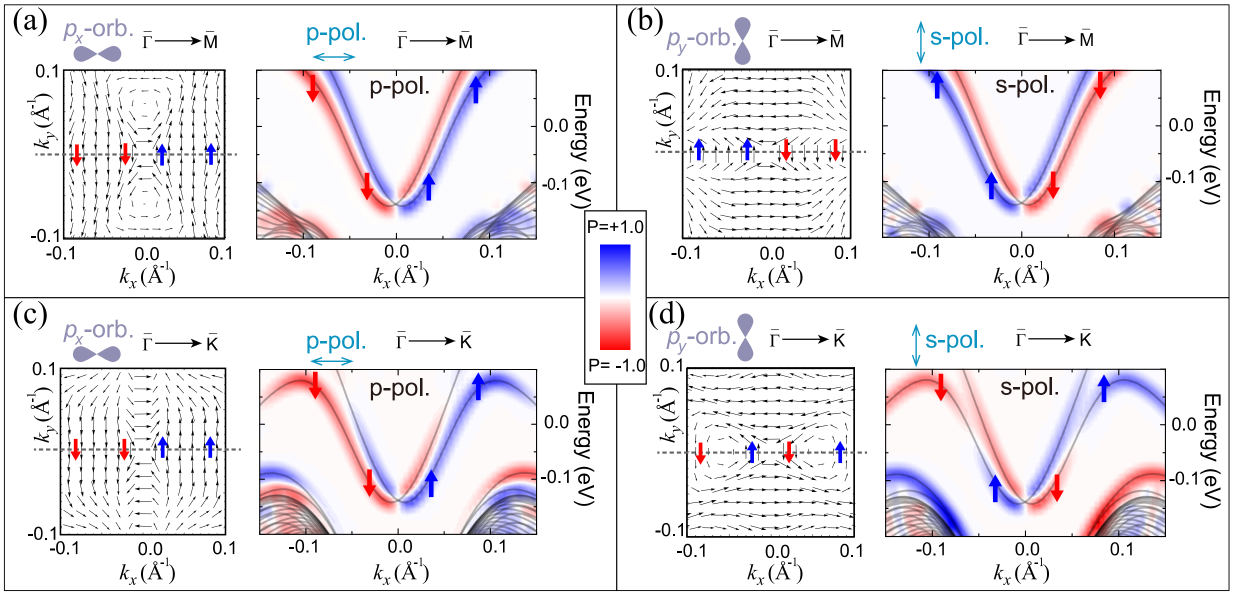

Figure 8.2: Calculated spin-orbital textures and simulated spin-ARPES plots. (a) Left: spin texture of p_x orbitals in the lower surface band, with $\hat{x}$ (horizontal dashed line) oriented along $\bar{\Gamma} - \bar{M}$. These states can be photoemitted by p-polarized light. The black vectors are in-plane expectation values $\langle S_{x,y} \rangle$. Right: simulated spin-ARPES measurement along $\bar{\Gamma} - \bar{M}$ using p-polarized light. Spin polarization of the p_x component of the bands is shown by the color from blue to red, while the band energies are indicated by gray lines. The spin polarization of the lower band (with arrows) is associated with the spin texture in the left panel. (b) Left: spin texture of p_y orbitals in the lower surface band, with $\hat{x}$ oriented along $\bar{\Gamma} - \bar{M}$. These states can be photoemitted by s-polarized light. Right: simulated spin-ARPES measurement along $\bar{\Gamma} - \bar{M}$ using s-polarized light. (c) and (d) Same as (a) and (b), but now with $\hat{x}$ oriented along $\bar{\Gamma} - \bar{K}$ and therefore the simulated measurements along $\bar{\Gamma} - \bar{K}$. Note that in the lower band, the orbital dependence of the spin texture ceases at large $|k|$ for $\bar{\Gamma} - \bar{K}$ but remains in $\bar{\Gamma} - \bar{M}$, consistent with experimental results.

8.4 Simulation of spin textures in topological surface states

Fig. 8.2 shows the calculated p-orbital-dependent spin textures and the simulated spin-ARPES results. The simulated spin measurements utilize the optical selection rule for the dominant p to s transitions, namely, that photons linearly polarized along the i direction ($i = x, y, z$) will allow only a p_i to s transition (if spin-orbit effects in the light-matter interaction are neglected). Although the final states reached in the photoemission process can shape the measured spin polarization [185], the s-wave final states reached in this 6 eV experiment should accept any spin, yielding information about the initial state being probed. Our simulations included p_y orbitals, as probed by s-polarized light, and p_x orbitals, as probed by the x component of p-polarized light. While p_z orbitals also contributed to the measurements with p-polarized light, their spins match those of p_x orbitals along the directions measured, allowing us to focus on a comparison of only p_x and p_y .

In Fig. 8.2, it is clear that in the vicinity of $\bar{\Gamma}$, the spin textures of the lower surface band are the same as those predicted for the Dirac cone in $\mathrm{Bi_2Se_3}$ [163, 165, 167], with p_y orbitals having the opposite spin of px . However, when moving far enough away from $\bar{\Gamma}$ along the $\bar{\Gamma} - \bar{K}$ direction, the spin polarization of p_y orbitals matches that of p_x . We note that the calculations of the upper surface band reveal a similar end to the p-orbital dependence of the spin orientation, albeit once the band is above the Fermi level in measurements. In contrast, the spin of p_y orbitals along $\bar{\Gamma} - \bar{M}$ remains fixed opposite to p_x orbitals for all k on this path.

The simulation of Figs. 8.1(c) and (d) shows another important aspect of the experiment: the apparent spin polarization around the top of the valence band as the surface-state dispersion bends down towards it. From Γ to K, as is evident from the spin polarization, around $|k| = 0.06\mathring{A}^1$ the lower surface band begins to decouple from the upper surface band and starts to pair with another surface-state band that is closer in energy to the bulk valence continuum. Therefore, the experimentally observed spin polarization around the valence band top and below the topological surface bands should be attributed to this newly emerging surface band.

8.5 Theoretical analysis of interplay between spin-orbit coupling and symmetry rules

The results from Sb(111) indicate that the p-orbital dependence of the spin texture breaks down as band mixing alters the basis states for the surface-state wave function. This is highlighted by tuning the strength of SOC α in the tight-binding Hamiltonian with,

$$H_\alpha = H_0 + \alpha \Delta H_{\mathrm{SOC}}, \tag{8.1}$$

where $\Delta H_{\mathrm{SOC}} = H - H_0$, and H is the Hamiltonian with full SOC, and H_0 is the Hamiltonian without SOC but with scalar-relativistic effects. In addition to shrinking the band gap, reducing α reduces the splitting between coupled bands, affording a clearer picture of which states are paired, meaning that they would be degenerate at each k without SOC ($\alpha = 0$).

Fig. 8.3 shows the band structure with varying values of α. Violet is used to highlight the paired surface states. As shown in Fig. 8.3(b), along the $\bar{\Gamma} - \bar{K}$ direction, the lower topological surface band clearly couples with the upper topological surface band around the zone center. However, farther from $\bar{\Gamma}$, the two switch partners: the upper one runs into the conduction-band continuum, and the other couples with a new surface-state band which emerges above the bulk valence band. These two eventually disperse together into the bulk valence-band continuum, as clearly shown in Fig. 8.3(a) with full spin-orbit effects considered. In contrast,

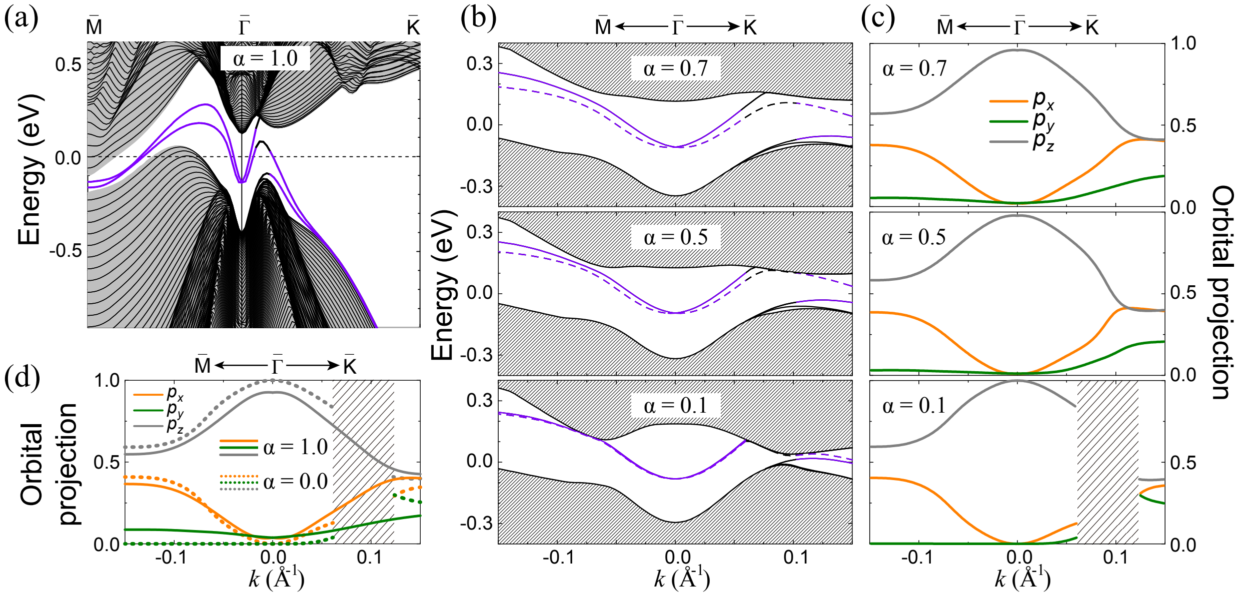

Figure 8.3: Band structure evolution with spin-orbit coupling (SOC) strength α. (a) Calculated 90-bilayer Sb band structures with full SOC ($\alpha = 1$). The darkly shaded area is the projection of bulk states onto the surface Brillouin zone. The violet color indicates the surface states that would be degenerate in the absence of SOC ($\alpha = 1$). (b) Detailed band structures along two directions ($\bar{M} - \bar{\Gamma} - \bar{K}$), with different SOC strengths α. Along $\bar{\Gamma} - \bar{M}$, the two surface bands always couple to each other and stay within the gap; along $\bar{\Gamma} - \bar{K}$, the lower surface band couples to the upper surface band near $k = 0$, then switches to a new surface band closer to the valence bulk continuum at larger $|k|$. (c) Projected p orbital character at various SOC strengths along the lower topological surface band indicated by dashed lines in (b). A rapid change in the orbital character is seen along the $\bar{\Gamma} - \bar{K}$ direction. (d) Similar to (c), showing the two limits of full SOC ($\alpha = 1$, solid lines) and no SOC ($\alpha = 0$, dotted lines). Note that along $\bar{\Gamma} - \bar{K}$, all orbitals have a finite projection even in the absence of SOC, whereas along $\bar{\Gamma} - \bar{M}$, p_y is finite only with SOC. The missing parts of the curves in the $0.06 < k < 0.12$ range (shaded area) shown in (c) and (d) in the small-α cases represent the fact that the band of interest disperses into the bulk continuum, as can be seen in (b).

along $\bar{\Gamma} - \bar{M}$ [Fig. 8.3(b)], the two surface states of interest are always coupled to each other and remain within the gap, maintaining the p-orbital dependence of their spins. The surface bands appear in pairs at any individual $\mathbf{k}$ point due to the degeneracy when SOC is completely turned off. The presence of SOC will split the degenerate bands, highlighting the fact that each of the single-surface bands connects the valence and conduction bulk continua, a result of the topologically nontrivial nature of Sb.

In Sb, as in Bi_2Se_3, the electronic states around the Fermi level are dominated by $p_{x,y,z}$ orbitals, which can take on $S_z = \pm\frac{1}{2}$. In the topological surface states near $\bar{\Gamma}$, $\mathbf{L}$ and $\mathbf{S}$ are coupled in such a way that $J_z = \pm\frac{1}{2}$ [163, 167]. As shown in Fig. 8.3(d), in the absence of SOC, at $\bar{\Gamma}$ there is only p_z character; that is, the orbital projection is zero for p_x and p_y. An eigenstate of J_z will remain such even with SOC turned on. Thus, SOC will mix in $p_{x,y}$ orbitals, giving them a finite projection in Figs. 8.3(c) and (d), while keeping $J_z = \pm\frac{1}{2}$ dominant in the vicinity of $\bar{\Gamma}$. This means that in this region, the states can be described sufficiently by a two-band model [163, 167]. The $J_z = \pm\frac{1}{2}$ requirement determines the spin texture that each p orbital must have. In other words, for in-plane orbitals,

$$|p_\pm\rangle = \frac{1}{\sqrt{2}}\left(\mp|p_x\rangle - i\,|p_y\rangle\right), \text{ with } L_z = \pm 1, \tag{8.2}$$

the surface state can be constructed from the following two basis states,

$$\begin{aligned} |p_+, \downarrow\rangle &= |p_+\rangle \otimes |\downarrow\rangle\,, \text{ carrying } J_z = \frac{1}{2}, \\ |p_-, \uparrow\rangle &= |p_-\rangle \otimes |\uparrow\rangle\,, \text{ carrying } J_z = -\frac{1}{2}. \end{aligned} \tag{8.3}$$

Such states will always show a p-orbital-dependent spin texture, meaning opposite spins will be measured with s-polarized and p-polarized light.

Symmetry provides similar constraints along $\bar{\Gamma} - \bar{M}$. In the absence of SOC, mirror symmetry excludes p_y orbitals along this momentum direction because they cannot mix with $p_{x,z}$ orbitals. This is shown in Fig. 8.3(d), where the $\alpha = 0$ case has a p_y projection of zero along the $\bar{\Gamma} - \bar{M}$ line. Turning on SOC will mix in p_y orbitals [making their contribution finite along $\bar{\Gamma} - \bar{M}$ in Figs. 8.3(c) and (d)] by allowing them to couple to spinors in a way that respects mirror symmetry; that is, p_y will couple to the spinor opposite of that to which $p_{x,z}$ orbitals couple. The $p_{x,y}$ orbitals at $\bar{\Gamma}$ and the p_y orbitals along $\bar{\Gamma} - \bar{M}$ are present only because of SOC and are subject to symmetry rules. Therefore, they are constrained in the spins to which they couple.

However, at high $|k|$ along $\bar{\Gamma} - \bar{K}$, where the same symmetry constraints do not exist, $p_{x,y}$ orbitals contribute appreciably even in the absence of SOC ($\alpha = 0$), as can be seen in Fig. 8.3(d). Therefore, turning on SOC will change the orbital character only slightly here, and the properties detected with any photon polarization are primarily non-SOC effects. At high $|k|$, two basis states are no longer sufficient to describe the surface complexity, as is evident from the presence of extra bands along $\bar{\Gamma} - \bar{K}$ in Fig. 8.3(a). In this region with more states, more basis vectors are needed,

$$\begin{aligned} |p_+, \uparrow\rangle &= |p_+\rangle \otimes |\uparrow\rangle\,, \text{ carrying } J_z = \frac{3}{2}, \\ |p_-, \downarrow\rangle &= |p_-\rangle \otimes |\downarrow\rangle\,, \text{ carrying } J_z = -\frac{3}{2}. \end{aligned} \tag{8.4}$$

Generally, the inclusion of $J_z = \pm\frac{3}{2}$ components without symmetry constraints will alter the phase between the spin-up ($|\uparrow\rangle$) and spin-down ($|\downarrow\rangle$) components of the real spinor wave functions that couple to p orbitals and may lead the spins not to reverse with different photon polarizations (e.g., nonzero linear combinations of $|p_+, \uparrow\rangle$ and $|p_+, \downarrow\rangle$ will never show this effect). This less constrained spin-orbital coupling along $\bar{\Gamma} - \bar{K}$ is in contrast to $\bar{\Gamma}$ and $\bar{\Gamma} - \bar{M}$, where symmetry protects the way that $J_z = \pm\frac{1}{2}$ and $J_z = \pm\frac{3}{2}$ mix, preserving the observation of opposite spins with different photon polarizations. Last, we note that hexagonal warping effects [186] could not be responsible for the sudden change of the spin texture at large $|k|$ because they do not alter the orbital texture and instead just diminish the magnitude of in-plane spin polarizations.

8.6 Conclusion

In this Chapter, we present a collaborative work with experimental groups on the spin textures of the topological surface states of Sb(111). The theoretical *ab initio* tight-binding calculations and analysis of the spin-textures show nice agreement with, and provide deep insights and understandings to various experimental ARPES observations. Our results show that the coupling between the spin textures and orbital textures can vary across the Brillouin zone. Knowledge of the full complexity of a surface state's wave function, including the symmetry rules governing the coupling of spin and orbital degrees of freedom, is a fundamental prerequisite for its application to spintronics or other technologies. One would expect to see orbitals with different spin orientations in the parts of a spin-orbit material's surface Brillouin zone where various symmetries protect it. Away from these momenta, surface states are allowed to mix with other states, ending the requirement that each orbital character couple to a distinct spin texture. This picture provides an understanding of the full complexity of a surface states spin degree of freedom.

Bibliography

[1] R. O. Jones and O. Gunnarsson, “The density functional formalism, its applications and prospects”, Reviews of Modern Physics **61**, 689 (1989).

[2] P. Hohenberg and W. Kohn, “Inhomogeneous electron gas”, Physical Review **136**, B864 (1964).

[3] W. Kohn and L. J. Sham, “Self-consistent equations including exchange and correlation effects”, Physical Review **140**, A1133 (1965).

[4] J. P. Perdew, “Density-functional approximation for the correlation energy of the inhomogeneous electron gas”, Physical Review B **33**, 8822 (1986).

[5] J. P. Perdew, K. Burke, and M. Ernzerhof, “Generalized gradient approximation made simple”, Physical Review Letters **77**, 3865 (1996).

[6] F. Bloch, “Über die quantenmechanik der elektronen in kristallgittern”, Zeitschrift für Physik **52**, 555–600 (1929).

[7] M. S. Hybertsen and S. G. Louie, “First-principles theory of quasiparticles: calculation of band gaps in semiconductors and insulators”, Physical Review Letters **55**, 1418 (1985).

[8] M. S. Hybertsen and S. G. Louie, “Electron correlation in semiconductors and insulators: band gaps and quasiparticle energies”, Physical Review B **34**, 5390 (1986).

[9] L. Hedin, “New method for calculating the one-particle Green’s function with application to the electron-gas problem”, Physical Review **139**, A796 (1965).

[10] L. Hedin and S. Lundqvist, “Effects of electron-electron and electron-phonon interactions on the one-electron states of solids”, Solid State Physics **23**, 1–181 (1970).

[11] G. Onida, L. Reining, and A. Rubio, “Electronic excitations: density-functional versus many-body Green’s-function approaches”, Reviews of Modern Physics **74**, 601 (2002).

[12] F. J. Dyson, “The S matrix in quantum electrodynamics”, Physical Review **75**, 1736 (1949).

[13] M. S. Hybertsen and S. G. Louie, “*Ab initio* static dielectric matrices from the density-functional approach. I. formulation and application to semiconductors and insulators”, Physical Review B **35**, 5585 (1987).

[14] A. Maradudin and S. H. Vosko, “Symmetry properties of the normal vibrations of a crystal”, Reviews of Modern Physics **40**, 1 (1968).

[15] S. Baroni, S. De Gironcoli, A. Dal Corso, and P. Giannozzi, "Phonons and related crystal properties from density-functional perturbation theory", Reviews of Modern Physics **73**, 515 (2001).

[16] F. Giustino, "Electron-phonon interactions from first principles", Reviews of Modern Physics **89**, 015003 (2017).

[17] Z. Li, G. Antonius, M. Wu, F. H. da Jornada, and S. G. Louie, "Electron-phonon coupling from *ab initio* linear-response theory within the *GW* method: correlation-enhanced interactions and superconductivity in $Ba_{1-x}K_xBiO_3$", Physical Review Letters **122**, 186402 (2019).

[18] H. J. Choi, D. Roundy, H. Sun, M. L. Cohen, and S. G. Louie, "The origin of the anomalous superconducting properties of MgB_2", Nature **418**, 758 (2002).

[19] M. Bernardi, D. Vigil-Fowler, J. Lischner, J. B. Neaton, and S. G. Louie, "*Ab initio* study of hot carriers in the first picosecond after sunlight absorption in silicon", Physical Review Letters **112**, 257402 (2014).

[20] J. Noffsinger, E. Kioupakis, C. G. Van de Walle, S. G. Louie, and M. L. Cohen, "Phonon-assisted optical absorption in silicon from first principles", Physical Review Letters **108**, 167402 (2012).

[21] S. Baroni, P. Giannozzi, and A. Testa, "Green's-function approach to linear response in solids", Physical Review Letters **58**, 1861 (1987).

[22] X. Gonze, D. C. Allan, and M. P. Teter, "Dielectric tensor, effective charges, and phonons in α-quartz by variational density-functional perturbation theory", Physical Review Letters **68**, 3603 (1992).

[23] S. Y. Savrasov, "Linear response calculations of lattice dynamics using muffin-tin basis sets", Physical Review Letters **69**, 2819 (1992).

[24] R. Sternheimer, "Electronic polarizabilities of ions from the hartree-fock wave functions", Physical Review **96**, 951 (1954).

[25] X. Gonze, "Adiabatic density-functional perturbation theory", Physical Review A **52**, 1096 (1995).

[26] N. Marzari and D. Vanderbilt, "Maximally localized generalized Wannier functions for composite energy bands", Physical Review B **56**, 12847 (1997).

[27] I. Souza, N. Marzari, and D. Vanderbilt, "Maximally localized Wannier functions for entangled energy bands", Physical Review B **65**, 035109 (2001).

[28] N. Marzari, A. A. Mostofi, J. R. Yates, I. Souza, and D. Vanderbilt, "Maximally localized Wannier functions: theory and applications", Reviews of Modern Physics **84**, 1419 (2012).

[29] F. Giustino, J. R. Yates, I. Souza, M. L. Cohen, and S. G. Louie, "Electron-phonon interaction via electronic and lattice Wannier functions: superconductivity in boron-doped diamond reexamined", Physical Review Letters **98**, 047005 (2007).

[30] F. Giustino, M. L. Cohen, and S. G. Louie, "Electron-phonon interaction using Wannier functions", Physical Review B **76**, 165108 (2007).

[31] Z. Li, G. Antonius, M. Wu, F. H. da Jornada, and S. G. Louie, In preparation.

[32] F. Giustino, M. L. Cohen, and S. G. Louie, "Small phonon contribution to the photoemission kink in the copper oxide superconductors", Nature **452**, 975 (2008).

[33] F. J. Di Salvo, D. Moncton, and J. Waszczak, "Electronic properties and superlattice formation in the semimetal $\mathrm{TiSe_2}$", Physical Review B **14**, 4321 (1976).

[34] M. Lazzeri, C. Attaccalite, L. Wirtz, and F. Mauri, "Impact of the electron-electron correlation on phonon dispersion: failure of LDA and GGA DFT functionals in graphene and graphite", Physical Review B **78**, 081406 (2008).

[35] A. Grüneis, J. Serrano, A. Bosak, M. Lazzeri, S. L. Molodtsov, L. Wirtz, C. Attaccalite, M. Krisch, A. Rubio, F. Mauri, et al., "Phonon surface mapping of graphite: disentangling quasi-degenerate phonon dispersions", Physical Review B **80**, 085423 (2009).

[36] C. Faber, J. L. Janssen, M. Côté, E. Runge, and X. Blase, "Electron-phonon coupling in the $\mathrm{C_{60}}$ fullerene within the many-body GW approach", Physical Review B **84**, 155104 (2011).

[37] Z. Yin, A. Kutepov, and G. Kotliar, "Correlation-enhanced electron-phonon coupling: applications of GW and screened hybrid functional to bismuthates, chloronitrides, and other high-T_c superconductors", Physical Review X **3**, 021011 (2013).

[38] G. Antonius, S. Poncé, P. Boulanger, M. Côté, and X. Gonze, "Many-body effects on the zero-point renormalization of the band structure", Physical Review Letters **112**, 215501 (2014).

[39] C. Faber, P. Boulanger, C. Attaccalite, E. Cannuccia, I. Duchemin, T. Deutsch, and X. Blase, "Exploring approximations to the GW self-energy ionic gradients", Physical Review B **91**, 155109 (2015).

[40] B. Monserrat, "Correlation effects on electron-phonon coupling in semiconductors: many-body theory along thermal lines", Physical Review B **93**, 100301 (2016).

[41] G. Strinati, "Application of the Green's functions method to the study of the optical properties of semiconductors", La Rivista del Nuovo Cimento **11**, 1–86 (1988).

[42] M. Rohlfing and S. G. Louie, "Electron-hole excitations in semiconductors and insulators", Physical Review Letters **81**, 2312 (1998).

[43] M. Rohlfing and S. G. Louie, "Electron-hole excitations and optical spectra from first principles", Physical Review B **62**, 4927 (2000).

[44] X. Gonze, B. Amadon, P.-M. Anglade, J.-M. Beuken, F. Bottin, P. Boulanger, F. Bruneval, D. Caliste, R. Caracas, M. Côté, et al., "ABINIT: first-principles approach to material and nanosystem properties", Computer Physics Communications **180**, 2582–2615 (2009).

[45] J. Deslippe, G. Samsonidze, D. A. Strubbe, M. Jain, M. L. Cohen, and S. G. Louie, "BerkeleyGW: a massively parallel computer package for the calculation of the quasiparticle and optical properties of materials and nanostructures", Computer Physics Communications **183**, 1269–1289 (2012).

[46] S. Poncé, E. R. Margine, C. Verdi, and F. Giustino, "EPW: electron–phonon coupling, transport and superconducting properties using maximally localized Wannier functions", Computer Physics Communications **209**, 116–133 (2016).

[47] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, et al., "QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials", Journal of Physics: Condensed Matter **21**, 395502 (2009).

[48] A. A. Mostofi, J. R. Yates, G. Pizzi, Y.-S. Lee, I. Souza, D. Vanderbilt, and N. Marzari, "An updated version of Wannier90: a tool for obtaining maximally-localised Wannier functions", Computer Physics Communications **185**, 2309–2310 (2014).

[49] C. Gong, L. Li, Z. Li, H. Ji, A. Stern, Y. Xia, T. Cao, W. Bao, C. Wang, Y. Wang, Z. Q. Qiu, R. J. Cava, S. G. Louie, J. Xia, and X. Zhang, "Discovery of intrinsic ferromagnetism in two-dimensional van der Waals crystals", Nature **546**, 265 (2017).

[50] Z. Li, T. Cao, and S. G. Louie, "Two-dimensional ferromagnetism in few-layer van der Waals crystals: renormalized spin-wave theory and calculations", Journal of Magnetism and Magnetic Materials **463**, 28–35 (2018).

[51] B. Huang, G. Clark, E. Navarro-Moratalla, D. R. Klein, R. Cheng, K. L. Seyler, D. Zhong, E. Schmidgall, M. A. McGuire, D. H. Cobden, W. Yao, D. Xiao, P. Jarillo-Herrero, and X. Xu, "Layer-dependent ferromagnetism in a van der Waals crystal down to the monolayer limit", Nature **546**, 270 (2017).

[52] A. K. Geim and K. S. Novoselov, "The rise of graphene", Nature Materials **6**, 183–191 (2007).

[53] G. R. Bhimanapati, Z. Lin, V. Meunier, Y. Jung, J. Cha, S. Das, D. Xiao, Y. Son, M. S. Strano, V. R. Cooper, et al., "Recent advances in two-dimensional materials beyond graphene", ACS Nano **9**, 11509–11539 (2015).

[54] N. D. Mermin and H. Wagner, "Absence of ferromagnetism or antiferromagnetism in one-or two-dimensional isotropic Heisenberg models", Physical Review Letters **17**, 1133 (1966).

[55] V. Carteaux, D. Brunet, G. Ouvrard, and G. Andre, "Crystallographic, magnetic and electronic structures of a new layered ferromagnetic compound $Cr_2Ge_2Te_6$", Journal of Physics: Condensed Matter **7**, 69 (1995).

[56] H. Ji, R. Stokes, L. Alegria, E. Blomberg, M. Tanatar, A. Reijnders, L. M. Schoop, T. Liang, R. Prozorov, K. Burch, et al., "A ferromagnetic insulating substrate for the epitaxial growth of topological insulators", Journal of Applied Physics **114**, 114907 (2013).

[57]X. Li and J. Yang, "CrXTe$_3$ (X = Si, Ge) nanosheets: two dimensional intrinsic ferromagnetic semiconductors", Journal of Materials Chemistry C **2**, 7071–7076 (2014).

[58]N. Sivadas, M. W. Daniels, R. H. Swendsen, S. Okamoto, and D. Xiao, "Magnetic ground state of semiconducting transition-metal trichalcogenide monolayers", Physical Review B **91**, 235425 (2015).

[59]H. L. Zhuang, Y. Xie, P. Kent, and P. Ganesh, "Computational discovery of ferromagnetic semiconducting single-layer CrSnTe$_3$", Physical Review B **92**, 035407 (2015).

[60]M. Bloch, "Magnon renormalization in ferromagnets near the Curie point", Physical Review Letters **9**, 286 (1962).

[61]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).

[62]T. Holstein and H. Primakoff, "Field dependence of the intrinsic domain magnetization of a ferromagnet", Physical Review **58**, 1098 (1940).

[63]F. Bloch, "Zur theorie des ferromagnetismus", Zeitschrift für Physik **61**, 206–219 (1930).

[64]A. Corciovei, "Spin-wave theory of ferromagnetic thin films", Physical Review **130**, 2223 (1963).

[65]E. Rastelli, A. Tassi, and L. Reatto, "Selfconsistently renormalized spin-wave approximation for some two-dimensional magnetic systems", Journal of Physics C: Solid State Physics **7**, 1735 (1974).

[66]R. Birgeneau, W. Yelon, E. Cohen, and J. Makovsky, "Magnetic properties of FeCl$_2$ in zero field. I. excitations", Physical Review B **5**, 2607 (1972).

[67]O. Dietrich, J. Als-Nielsen, and L. Passell, "Neutron scattering from the Heisenberg ferromagnets EuO and EuS. III. spin dynamics of EuO", Physical Review B **14**, 4923 (1976).

[68]R. J. Cava, B. Batlogg, J. Krajewski, R. Farrow, L. Rupp Jr, A. White, K. Short, W. Peck, and T. Kometani, "Superconductivity near 30 K without copper: the Ba$_{0.6}$K$_{0.4}$BiO$_3$ perovskite", Nature **332**, 814 (1988).

[69]P. Kumar, D. Hall, and R. Goodrich, "Thermodynamics of the superconducting phase transition in Ba$_{0.6}$K$_{0.4}$BiO$_3$", Physical Review Letters **82**, 4532 (1999).

[70]S. Pei, J. Jorgensen, B. Dabrowski, D. Hinks, D. Richards, A. Mitchell, J. Newsam, S. Sinha, D. Vaknin, and A. Jacobson, "Structural phase diagram of the Ba$_{1-x}$K$_x$BiO$_3$ system", Physical Review B **41**, 4126 (1990).

[71]V. Meregalli and S. Y. Savrasov, "Electron-phonon coupling and properties of doped BaBiO$_3$", Physical Review B **57**, 14453 (1998).

[72]D. Hamann, "Optimized norm-conserving vanderbilt pseudopotentials", Physical Review B **88**, 085117 (2013).

[73]M. Van Setten, M. Giantomassi, E. Bousquet, M. J. Verstraete, D. R. Hamann, X. Gonze, and G.-M. Rignanese, "The PseudoDojo: training and grading a 85 element optimized norm-conserving pseudopotential table", Computer Physics Communications **226**, 39–54 (2018).

[74]M. Braden, W. Reichardt, W. Schmidbauer, A. Ivanov, and A. Y. Rumiantsev, "Lattice dynamics of (Ba/K)BiO_3", Journal of Superconductivity **8**, 595–598 (1995).

[75]O. Hellman, I. Abrikosov, and S. Simak, "Lattice dynamics of anharmonic solids from first principles", Physical Review B **84**, 180301 (2011).

[76]G. Kresse and J. Hafner, "*Ab initio* molecular dynamics for liquid metals", Physical Review B **47**, 558 (1993).

[77]G. Kresse and D. Joubert, "From ultrasoft pseudopotentials to the projector augmented-wave method", Physical Review B **59**, 1758 (1999).

[78]T. Tadano, Y. Gohda, and S. Tsuneyuki, "Anharmonic force constants extracted from first-principles molecular dynamics: applications to heat transfer simulations", Journal of Physics: Condensed Matter **26**, 225402 (2014).

[79]C. Wen, H. Xu, Q. Yao, R. Peng, X. Niu, Q. Chen, Z. Liu, D. Shen, Q. Song, X. Lou, et al., "Unveiling the superconducting mechanism of $Ba_{0.51}K_{0.49}BiO_3$", Physical Review Letters **121**, 117002 (2018).

[80]W. McMillan, "Transition temperature of strong-coupled superconductors", Physical Review **167**, 331 (1968).

[81]P. B. Allen and R. Dynes, "Transition temperature of strong-coupled superconductors reanalyzed", Physical Review B **12**, 905 (1975).

[82]N. Plumb, D. Gawryluk, Y. Wang, Z. Ristić, J. Park, B. Lv, Z. Wang, C. Matt, N. Xu, T. Shang, et al., "Momentum-resolved electronic structure of the high-T_c superconductor parent compound $BaBiO_3$", Physical Review Letters **117**, 037002 (2016).

[83]A. Khazraie, K. Foyevtsova, I. Elfimov, and G. A. Sawatzky, "Bond versus charge disproportionation in the bismuth perovskites", Physical Review B **98**, 205104 (2018).

[84]H. González-Herrero, J. M. Gómez-Rodríguez, P. Mallet, M. Moaied, J. J. Palacios, C. Salgado, M. M. Ugeda, J.-Y. Veuillen, F. Yndurain, and I. Brihuega, "Atomic-scale control of graphene magnetism by using hydrogen atoms", Science **352**, 437–441 (2016).

[85]R. Nair, M. Sepioni, I.-L. Tsai, O. Lehtinen, J. Keinonen, A. Krasheninnikov, T. Thomson, A. Geim, and I. Grigorieva, "Spin-half paramagnetism in graphene induced by point defects", Nature Physics **8**, 199 (2012).

[86]K. M. McCreary, A. G. Swartz, W. Han, J. Fabian, and R. K. Kawakami, "Magnetic moment formation in graphene detected by scattering of pure spin currents", Physical Review Letters **109**, 186604 (2012).

[87] J. Červenka, M. I. Katsnelson, and C. F. J. Flipse, "Room-temperature ferromagnetism in graphite driven by two-dimensional networks of point defects", Nature Physics **5**, 840 (2009).

[88] B. Uchoa, V. N. Kotov, N. M. R. Peres, and A. H. C. Neto, "Localized magnetic states in graphene", Physical Review Letters **101**, 026805 (2008).

[89] J. Jung, T. Pereg-Barnea, and A. H. MacDonald, "Theory of interedge superexchange in zigzag edge magnetism", Physical Review Letters **102**, 227205 (2009).

[90] Y.-W. Son, M. L. Cohen, and S. G. Louie, "Half-metallic graphene nanoribbons", Nature **444**, 347 (2006).

[91] E. V. Castro, N. M. R. Peres, T. Stauber, and N. A. P. Silva, "Low-density ferromagnetism in biased bilayer graphene", Physical Review Letters **100**, 186803 (2008).

[92] T. Cao, Z. Li, and S. G. Louie, "Tunable magnetism and half-metallicity in hole-doped monolayer GaSe", Physical Review Letters **114**, 236602 (2015).

[93] Y. Deng, Y. Yu, Y. Song, J. Zhang, N. Z. Wang, Z. Sun, Y. Yi, Y. Z. Wu, S. Wu, J. Zhu, et al., "Gate-tunable room-temperature ferromagnetism in two-dimensional Fe_3GeTe_2", Nature **563**, 94 (2018).

[94] Y. Li and K. Baberschke, "Dimensional crossover in ultrathin Ni(111) films on W(110)", Physical Review Letters **68**, 1208 (1992).

[95] R. Zhang and R. F. Willis, "Thickness-dependent Curie temperatures of ultrathin magnetic films: effect of the range of spin-spin interactions", Physical Review Letters **86**, 2665 (2001).

[96] M. Cococcioni and S. De Gironcoli, "Linear response approach to the calculation of the effective interaction parameters in the $\mathrm{LDA} + U$ method", Physical Review B **71**, 035105 (2005).

[97] A. Dal Corso, "Pseudopotentials periodic table: from H to Pu", Computational Materials Science **95**, 337–350 (2014).

[98] H. Xiang, C. Lee, H.-J. Koo, X. Gong, and M.-H. Whangbo, "Magnetic properties and energy-mapping analysis", Dalton Transactions **42**, 823–853 (2013).

[99] P. Bruno, "Magnetization and Curie temperature of ferromagnetic ultrathin films: the influence of magnetic anisotropy and dipolar interactions", MRS Online Proceedings Library Archive **231** (1991).

[100] Z. Li, T. Cao, M. Wu, and S. G. Louie, "Generation of anisotropic massless Dirac fermions and asymmetric Klein tunneling in few-layer black phosphorus superlattices", Nano Letters **17**, 2280–2286 (2017).

[101] A. C. Neto, F. Guinea, N. M. Peres, K. S. Novoselov, and A. K. Geim, "The electronic properties of graphene", Reviews of Modern Physics **81**, 109 (2009).

[102] C.-H. Park and S. G. Louie, "Making massless Dirac fermions from a patterned two-dimensional electron gas", Nano Letters **9**, 1793–1797 (2009).

[103] C.-H. Park, L. Yang, Y.-W. Son, M. L. Cohen, and S. G. Louie, "New generation of massless Dirac fermions in graphene under external periodic potentials", Physical Review Letters **101**, 126804 (2008).

[104] M. Gibertini, A. Singha, V. Pellegrini, M. Polini, G. Vignale, A. Pinczuk, L. N. Pfeiffer, and K. W. West, "Engineering artificial graphene in a two-dimensional electron gas", Physical Review B **79**, 241406 (2009).

[105] A. Singha, M. Gibertini, B. Karmakar, S. Yuan, M. Polini, G. Vignale, M. Katsnelson, A. Pinczuk, L. Pfeiffer, K. West, et al., "Two-dimensional Mott-Hubbard electrons in an artificial honeycomb lattice", Science **332**, 1176–1179 (2011).

[106] K. K. Gomes, W. Mar, W. Ko, F. Guinea, and H. C. Manoharan, "Designer Dirac fermions and topological phases in molecular graphene", Nature **483**, 306 (2012).

[107] M. Polini, F. Guinea, M. Lewenstein, H. C. Manoharan, and V. Pellegrini, "Artificial honeycomb lattices for electrons, atoms and photons", Nature Nanotechnology **8**, 625 (2013).

[108] O. P. Sushkov and A. C. Neto, "Topological insulating states in laterally patterned ordinary semiconductors", Physical Review Letters **110**, 186601 (2013).

[109] S. Wang, L. Z. Tan, W. Wang, S. G. Louie, and N. Lin, "Manipulation and characterization of aperiodical graphene structures created in a two-dimensional electron gas", Physical Review Letters **113**, 196803 (2014).

[110] B. Wunsch, F. Guinea, and F. Sols, "Dirac-point engineering and topological phase transitions in honeycomb optical lattices", New Journal of Physics **10**, 103027 (2008).

[111] L. Tarruell, D. Greif, T. Uehlinger, G. Jotzu, and T. Esslinger, "Creating, moving and merging Dirac points with a Fermi gas in a tunable honeycomb lattice", Nature **483**, 302 (2012).

[112] M. C. Rechtsman, J. M. Zeuner, A. Tünnermann, S. Nolte, M. Segev, and A. Szameit, "Strain-induced pseudomagnetic field and photonic landau levels in dielectric structures", Nature Photonics **7**, 153 (2013).

[113] M. C. Rechtsman, J. M. Zeuner, Y. Plotnik, Y. Lumer, D. Podolsky, F. Dreisow, S. Nolte, M. Segev, and A. Szameit, "Photonic Floquet topological insulators", Nature **496**, 196 (2013).

[114] I. Pletikosić, M. Kralj, P. Pervan, R. Brako, J. Coraux, A. N'diaye, C. Busse, and T. Michely, "Dirac cones and minigaps for graphene on Ir(111)", Physical Review Letters **102**, 056808 (2009).

[115] M. Yankowitz, J. Xue, D. Cormode, J. D. Sanchez-Yamagishi, K. Watanabe, T. Taniguchi, P. Jarillo-Herrero, P. Jacquod, and B. J. LeRoy, "Emergence of superlattice Dirac points in graphene on hexagonal boron nitride", Nature Physics **8**, 382 (2012).

[116] L. Ponomarenko, R. Gorbachev, G. Yu, D. Elias, R. Jalil, A. Patel, A. Mishchenko, A. Mayorov, C. Woods, J. Wallbank, et al., "Cloning of Dirac fermions in graphene superlattices", Nature **497**, 594 (2013).

[117] C. R. Dean, L. Wang, P. Maher, C. Forsythe, F. Ghahari, Y. Gao, J. Katoch, M. Ishigami, P. Moon, M. Koshino, et al., "Hofstadter's butterfly and the fractal quantum hall effect in moiré superlattices", Nature **497**, 598 (2013).

[118] B. Hunt, J. Sanchez-Yamagishi, A. Young, M. Yankowitz, B. J. LeRoy, K. Watanabe, T. Taniguchi, P. Moon, M. Koshino, P. Jarillo-Herrero, et al., "Massive Dirac fermions and Hofstadter butterfly in a van der Waals heterostructure", Science **340**, 1427–1430 (2013).

[119] W. Yang, G. Chen, Z. Shi, C.-C. Liu, L. Zhang, G. Xie, M. Cheng, D. Wang, R. Yang, D. Shi, et al., "Epitaxial growth of single-domain graphene on hexagonal boron nitride", Nature Materials **12**, 792 (2013).

[120] G. Yu, R. Gorbachev, J. Tu, A. Kretinin, Y. Cao, R. Jalil, F. Withers, L. Ponomarenko, B. Piot, M. Potemski, et al., "Hierarchy of Hofstadter states and replica quantum hall ferromagnetism in graphene superlattices", Nature Physics **10**, 525 (2014).

[121] E. Wang, X. Lu, S. Ding, W. Yao, M. Yan, G. Wan, K. Deng, S. Wang, G. Chen, L. Ma, et al., "Gaps induced by inversion symmetry breaking and second-generation Dirac cones in graphene/hexagonal boron nitride", Nature Physics **12**, 1111 (2016).

[122] L. Li, Y. Yu, G. J. Ye, Q. Ge, X. Ou, H. Wu, D. Feng, X. H. Chen, and Y. Zhang, "Black phosphorus field-effect transistors", Nature Nanotechnology **9**, 372 (2014).

[123] H. Liu, A. T. Neal, Z. Zhu, Z. Luo, X. Xu, D. Tománek, and P. D. Ye, "Phosphorene: an unexplored 2D semiconductor with a high hole mobility", ACS Nano **8**, 4033–4041 (2014).

[124] L. Li, G. J. Ye, V. Tran, R. Fei, G. Chen, H. Wang, J. Wang, K. Watanabe, T. Taniguchi, L. Yang, et al., "Quantum oscillations in a two-dimensional electron gas in black phosphorus thin films", Nature Nanotechnology **10**, 608 (2015).

[125] L. Li, F. Yang, G. J. Ye, Z. Zhang, Z. Zhu, W. Lou, X. Zhou, L. Li, K. Watanabe, T. Taniguchi, et al., "Quantum Hall effect in black phosphorus two-dimensional electron system", Nature Nanotechnology **11**, 593 (2016).

[126] J. Qiao, X. Kong, Z.-X. Hu, F. Yang, and W. Ji, "High-mobility transport anisotropy and linear dichroism in few-layer black phosphorus", Nature Communications **5**, 4475 (2014).

[127] R. Fei and L. Yang, "Strain-engineering the anisotropic electrical conductance of few-layer black phosphorus", Nano Letters **14**, 2884–2889 (2014).

[128] A. Rodin, A. Carvalho, and A. C. Neto, "Strain-induced gap modification in black phosphorus", Physical Review Letters **112**, 176801 (2014).

[129] M. Katsnelson, K. Novoselov, and A. Geim, "Chiral tunnelling and the Klein paradox in graphene", Nature Physics **2**, 620 (2006).

[130] A. F. Young and P. Kim, "Quantum interference and Klein tunnelling in graphene heterojunctions", Nature Physics **5**, 222 (2009).

[131] P. Rickhaus, R. Maurand, M.-H. Liu, M. Weiss, K. Richter, and C. Schönenberger, "Ballistic interferences in suspended graphene", Nature Communications **4**, 2342 (2013).

[132] Q. Wilmart, S. Berrada, D. Torrin, V. H. Nguyen, G. Fève, J.-M. Berroir, P. Dollfus, and B. Plaçais, "A Klein-tunneling transistor with ballistic graphene", 2D Materials **1**, 011006 (2014).

[133] J. Pereira Jr and M. Katsnelson, "Landau levels of single-layer and bilayer phosphorene", Physical Review B **92**, 075437 (2015).

[134] M. Goerbig, J.-N. Fuchs, G. Montambaux, and F. Piéchon, "Tilted anisotropic Dirac cones in quinoid-type graphene and α-(BEDT-TTF)$_2$I$_3$", Physical Review B **78**, 045415 (2008).

[135] Z. Liu, J. Wang, and J. Li, "Dirac cones in two-dimensional systems: from hexagonal to square lattices", Physical Chemistry Chemical Physics **15**, 18855–18862 (2013).

[136] B. Bradlyn, J. Cano, Z. Wang, M. Vergniory, C. Felser, R. J. Cava, and B. A. Bernevig, "Beyond Dirac and Weyl fermions: unconventional quasiparticles in conventional crystals", Science **353**, aaf5037 (2016).

[137] C.-H. Park, L. Yang, Y.-W. Son, M. L. Cohen, and S. G. Louie, "Anisotropic behaviours of massless Dirac fermions in graphene under periodic potentials", Nature Physics **4**, 213 (2008).

[138] Y. Cai, G. Zhang, and Y.-W. Zhang, "Electronic properties of phosphorene/graphene and phosphorene/hexagonal boron nitride heterostructures", The Journal of Physical Chemistry C **119**, 13929–13936 (2015).

[139] G.-H. Lee, G.-H. Park, and H.-J. Lee, "Observation of negative refraction of Dirac fermions in graphene", Nature Physics **11**, 925 (2015).

[140] T. Pham, A. L. Gibb, Z. Li, S. M. Gilbert, C. Song, S. G. Louie, and A. Zettl, "Formation and dynamics of electron-irradiation-induced defects in hexagonal boron nitride at elevated temperatures", Nano Letters **16**, 7142–7147 (2016).

[141] N. Alem, R. Erni, C. Kisielowski, M. D. Rossell, W. Gannett, and A. Zettl, "Atomically thin hexagonal boron nitride probed by ultrahigh-resolution transmission electron microscopy", Physical Review B **80**, 155425 (2009).

[142] M. L. Cohen and A. Zettl, "The physics of boron nitride nanotubes", Physics Today **63**, 34–38 (2010).

[143] D. Golberg, Y. Bando, Y. Huang, T. Terao, M. Mitome, C. Tang, and C. Zhi, "Boron nitride nanotubes and nanosheets", ACS Nano **4**, 2979–2993 (2010).

[144] A. Fathalizadeh, T. Pham, W. Mickelson, and A. Zettl, "Scaled synthesis of boron nitride nanotubes, nanoribbons, and nanococoons using direct feedstock injection into an extended-pressure, inductively-coupled thermal plasma", Nano Letters **14**, 4881–4886 (2014).

[145] A. Harley-Trochimczyk, T. Pham, J. Chang, E. Chen, M. A. Worsley, A. Zettl, W. Mickelson, and R. Maboudian, "Gas sensors: platinum nanoparticle loading of boron nitride aerogel and its use as a novel material for low-power catalytic gas sensing", Advanced Functional Materials **26**, 314–314 (2016).

[146] W. Gannett, W. Regan, K. Watanabe, T. Taniguchi, M. Crommie, and A. Zettl, "Boron nitride substrates for high mobility chemical vapor deposited graphene", Applied Physics Letters **98**, 242105 (2011).

[147] C. R. Dean, A. F. Young, I. Meric, C. Lee, L. Wang, S. Sorgenfrei, K. Watanabe, T. Taniguchi, P. Kim, K. L. Shepard, et al., "Boron nitride substrates for high-quality graphene electronics", Nature Nanotechnology **5**, 722 (2010).

[148] G.-H. Lee, X. Cui, Y. D. Kim, G. Arefe, X. Zhang, C.-H. Lee, F. Ye, K. Watanabe, T. Taniguchi, P. Kim, et al., "Highly stable, dual-gated MoS_2 transistors encapsulated by hexagonal boron nitride with gate-controllable contact, resistance, and threshold voltage", ACS Nano **9**, 7019–7026 (2015).

[149] J. C. Meyer, A. Chuvilin, G. Algara-Siller, J. Biskupek, and U. Kaiser, "Selective sputtering and atomic resolution imaging of atomically thin boron nitride membranes", Nano Letters **9**, 2683–2689 (2009).

[150] C. Jin, F. Lin, K. Suenaga, and S. Iijima, "Fabrication of a freestanding boron nitride single layer and its defect assignments", Physical Review Letters **102**, 195505 (2009).

[151] J. H. Warner, M. H. Rummeli, A. Bachmatiuk, and B. Buchner, "Atomic resolution imaging and topography of boron nitride sheets produced by chemical exfoliation", ACS Nano **4**, 1299–1304 (2010).

[152] G. H. Ryu, H. J. Park, J. Ryou, J. Park, J. Lee, G. Kim, H. S. Shin, C. W. Bielawski, R. S. Ruoff, S. Hong, et al., "Atomic-scale dynamics of triangular hole growth in monolayer hexagonal boron nitride under electron irradiation", Nanoscale **7**, 10600–10605 (2015).

[153] J. S. Kim, K. B. Borisenko, V. Nicolosi, and A. I. Kirkland, "Controlled radiation damage and edge structures in boron nitride membranes", ACS Nano **5**, 3977–3986 (2011).

[154] J. Kotakoski, C. Jin, O. Lehtinen, K. Suenaga, and A. Krasheninnikov, "Electron knock-on damage in hexagonal boron nitride monolayers", Physical Review B **82**, 113404 (2010).

[155] J. P. Perdew and A. Zunger, "Self-interaction correction to density-functional approximations for many-electron systems", Physical Review B **23**, 5048 (1981).

[156] S. B. Zhang and S.-H. Wei, "Surface energy and the common dangling bond rule for semiconductors", Physical Review Letters **92**, 086102 (2004).

[157] T. Schmidt, R. Baierle, P. Piquini, and A. Fazzio, "Theoretical study of native defects in BN nanotubes", Physical Review B **67**, 113407 (2003).

[158] K. Gotlieb, Z. Li, C.-Y. Lin, C. Jozwiak, J. H. Ryoo, C.-H. Park, Z. Hussain, S. G. Louie, and A. Lanzara, "Symmetry rules shaping spin-orbital textures in surface states", Physical Review B **95**, 245142 (2017).

[159]S. Datta and B. Das, “Electronic analog of the electro-optic modulator”, Applied Physics Letters **56**, 665–667 (1990).

[160]I. Žutić, J. Fabian, and S. D. Sarma, “Spintronics: fundamentals and applications”, Reviews of Modern Physics **76**, 323 (2004).

[161]D. D. Awschalom and M. E. Flatté, “Challenges for semiconductor spintronics”, Nature Physics **3**, 153 (2007).

[162]J. E. Moore, “The birth of topological insulators”, Nature **464**, 194 (2010).

[163]H. Zhang, C.-X. Liu, and S.-C. Zhang, “Spin-orbital texture in topological insulators”, Physical Review Letters **111**, 066801 (2013).

[164]Y. Cao, J. Waugh, X. Zhang, J.-W. Luo, Q. Wang, T. Reber, S. Mo, Z. Xu, A. Yang, J. Schneeloch, et al., “Mapping the orbital wavefunction of the surface states in three-dimensional topological insulators”, Nature Physics **9**, 499 (2013).

[165]Z.-H. Zhu, C. Veenstra, G. Levy, A. Ubaldini, P. Syers, N. Butch, J. Paglione, M. Haverkort, I. Elfimov, and A. Damascelli, “Layer-by-layer entangled spin-orbital texture of the topological surface state in Bi_2Se_3”, Physical Review Letters **110**, 216401 (2013).

[166]Z. Xie, S. He, C. Chen, Y. Feng, H. Yi, A. Liang, L. Zhao, D. Mou, J. He, Y. Peng, et al., “Orbital-selective spin texture and its manipulation in a topological insulator”, Nature Communications **5**, 3382 (2014).

[167]C.-H. Park and S. G. Louie, “Spin polarization of photoelectrons from topological insulators”, Physical Review Letters **109**, 097601 (2012).

[168]C. Jozwiak, C.-H. Park, K. Gotlieb, C. Hwang, D.-H. Lee, S. G. Louie, J. D. Denlinger, C. R. Rotundu, R. J. Birgeneau, Z. Hussain, and A. Lanzara, “Photoelectron spin-flipping and texture manipulation in a topological insulator”, Nature Physics **9**, 293 (2013).

[169]Z.-H. Zhu, C. Veenstra, S. Zhdanovich, M. Schneider, T. Okuda, K. Miyamoto, S.-Y. Zhu, H. Namatame, M. Taniguchi, M. Haverkort, et al., “Photoelectron spin-polarization control in the topological insulator Bi_2Se_3”, Physical Review Letters **112**, 076802 (2014).

[170]H. Mirhosseini, J. Henk, A. Ernst, S. Ostanin, C.-T. Chiang, P. Yu, A. Winkelmann, and J. Kirschner, “Unconventional spin topology in surface alloys with Rashba-type spin splitting”, Physical Review B **79**, 245428 (2009).

[171]C. Veenstra, Z.-H. Zhu, M. Raichle, B. Ludbrook, A. Nicolaou, B. Slomski, G. Landolt, S. Kittaka, Y. Maeno, J. H. Dil, et al., “Spin-orbital entanglement and the breakdown of singlets and triplets in Sr_2RuO_4 revealed by spin-and angle-resolved photoemission spectroscopy”, Physical Review Letters **112**, 127002 (2014).

[172]S. N. P. Wissing, A. B. Schmidt, H. Mirhosseini, J. Henk, C. R. Ast, and M. Donath, “Ambiguity of experimental spin information from states with mixed orbital symmetries”, Physical Review Letters **113**, 116402 (2014).

[173] H. Mirhosseini, M. Flieger, and J. Henk, "Dirac-cone-like surface state in W(110): dispersion, spin texture and photoemission from first principles", New Journal of Physics **15**, 033019 (2013).

[174] K. Miyamoto, H. Wortelen, H. Mirhosseini, T. Okuda, A. Kimura, H. Iwasawa, K. Shimada, J. Henk, and M. Donath, "Orbital-symmetry-selective spin characterization of Dirac-cone-like state on W(110)", Physical Review B **93**, 161403 (2016).

[175] H. Maaß, H. Bentmann, C. Seibel, C. Tusche, S. V. Eremeev, T. R. Peixoto, O. E. Tereshchenko, K. A. Kokh, E. V. Chulkov, J. Kirschner, et al., "Spin-texture inversion in the giant Rashba semiconductor BiTeI", Nature Communications **7**, 11621 (2016).

[176] J. H. Ryoo and C.-H. Park, "Spin-conserving and reversing photoemission from the surface states of Bi_2Se_3 and Au (111)", Physical Review B **93**, 085419 (2016).

[177] C. Jozwiak, J. Graf, G. Lebedev, N. Andresen, A. Schmid, A. Fedorov, F. El Gabaly, W. Wan, A. Lanzara, and Z. Hussain, "A high-efficiency spin-resolved photoemission spectrometer combining time-of-flight spectroscopy with exchange-scattering polarimetry", Review of Scientific Instruments **81**, 053904 (2010).

[178] K. Gotlieb, Z. Hussain, A. Bostwick, A. Lanzara, and C. Jozwiak, "Rapid high-resolution spin-and angle-resolved photoemission spectroscopy with pulsed laser source and time-of-flight spectrometer", Review of Scientific Instruments **84**, 093904 (2013).

[179] K. Sugawara, T. Sato, S. Souma, T. Takahashi, M. Arai, and T. Sasaki, "Fermi surface and anisotropic spin-orbit coupling of Sb(111) studied by angle-resolved photoemission spectroscopy", Physical Review Letters **96**, 046411 (2006).

[180] T. Kadono, K. Miyamoto, R. Nishimura, K. Kanomaru, S. Qiao, K. Shimada, H. Namatame, A. Kimura, and M. Taniguchi, "Direct evidence of spin-polarized band structure of Sb(111) surface", Applied Physics Letters **93**, 252107 (2008).

[181] D. Hsieh, Y. Xia, L. Wray, D. Qian, A. Pal, J. Dil, J. Osterwalder, F. Meier, G. Bihlmayer, C. Kane, et al., "Observation of unconventional quantum spin textures in topological insulators", Science **323**, 919–922 (2009).

[182] G. Bian, T. Miller, and T.-C. Chiang, "Passage from spin-polarized surface states to unpolarized quantum well states in topologically nontrivial Sb films", Physical Review Letters **107**, 036802 (2011).

[183] M. Bianchi, D. Guan, A. Stróżecka, C. H. Voetmann, S. Bao, J. I. Pascual, A. Eiguren, and P. Hofmann, "Surface states on a topologically nontrivial semimetal: the case of Sb(110)", Physical Review B **85**, 155431 (2012).

[184] A. Takayama, T. Sato, S. Souma, and T. Takahashi, "Rashba effect in antimony and bismuth studied by spin-resolved ARPES", New Journal of Physics **16**, 055004 (2014).

[185] J. Sánchez-Barriga, A. Varykhalov, J. Braun, S.-Y. Xu, N. Alidoust, O. Kornilov, J. Minár, K. Hummer, G. Springholz, G. Bauer, et al., “Photoemission of Bi_2Se_3 with circularly polarized light: probe of spin polarization or means for spin manipulation?”, Physical Review X **4**, 011046 (2014).

[186] L. Fu, “Hexagonal warping effects in the surface states of the topological insulator Bi_2Te_3”, Physical Review Letters **103**, 266801 (2009).